Rulemaking1CEm Resource

From: RulemakingComments Resource
Sent: Monday, November 16, 2015 5:51 PM
To: Rulemaking1CEm Resource
Subject: Comment on PRM-20-28, 20-29 & 20-30
Attachments: NRC-2015-0057-DRAFT-0500.pdf

DOCKETED BY USNRC—OFFICE OF THE SECRETARY

SECY-067
PR#: PRM-20-28, PRM-20-29, and PRM-20-30
FRN#: 80FR35870
NRC DOCKET#: NRC-2015-0057
SECY DOCKET DATE: 11/13/15
TITLE: Linear No-Threshold Model and Standards for Protection Against Radiation
COMMENT#: 501

Hearing Identifier: Secy_RuleMaking_comments_Public
Email Number: 1297

Mail Envelope Properties (2ede25c9825b44779b0bf888ccb65608)

Subject: Comment on PRM-20-28, 20-29 & 20-30
Sent Date: 11/16/2015 5:50:35 PM
Received Date: 11/16/2015 5:51:00 PM
From: RulemakingComments Resource

Created By: RulemakingComments.Resource@nrc.gov

Recipients:
"Rulemaking1CEm Resource" <Rulemaking1CEm.Resource@nrc.gov>
Tracking Status: None

Post Office: HQPWMSMRS02.nrc.gov

Files	**Size**	**Date & Time**
MESSAGE	298	11/16/2015 5:51:00 PM
NRC-2015-0057-DRAFT-0500.pdf	25501466	

Options
Priority: Standard
Return Notification: No
Reply Requested: No
Sensitivity: Normal
Expiration Date:
Recipients Received:

PUBLIC SUBMISSION

As of: 11/16/15 3:25 PM
Received: November 13, 2015
Status: Pending_Post
Tracking No. 1jz-8m8i-sbgv
Comments Due: November 19, 2015
Submission Type: Web

Docket: NRC-2015-0057
Linear No-Threshold Model and Standards for Protection Against Radiation

Comment On: NRC-2015-0057-0086
Linear No-Threshold Model and Standards for Protection Against Radiation; Extension of Comment Period

Document: NRC-2015-0057-DRAFT-0500
Comment on FR Doc # 2015-20722

Submitter Information

Name: Edward Calabrese
Address:
Dept of Environmental Health Sciences
University of Massachusetts
Amherst, MA, 01003
Email: edwardc@schoolph.umass.edu

General Comment

See attached file(s)

Attachments

NRC COMMENTS (11) - NRC

Arch Toxicol_ Abuse of RA - with Supplement

Arch Toxicol Origins of LNT

Arch Toxicol_ Cancer RA Unraveling - with Supplement

32.full

EMM_Key studies_RiskAssessment 2011

Fd Chem Toxicol Calabrese et al 2015

Archiv Toxicol USNAS Misled

HPJ-S-15-00205 Submitted 8-17-15

Hum Exp Toxicol-2000b-Calabrese-32-40

Environ Res_LNT Dogma artful dodges

Hum Exp Toxicol-2000d-Calabrese-76-84

Hum Exp Toxicol-2000e-Calabrese-85-97

Hum Exp Toxicol-2010-Calabrese et al-667-77

Hum Exp Toxicol-2000a-Calabrese-2-31

RTP Calabrese_Blain 2011 Database occurrence

TAP Calabrese_Blain 2005 Database overview

Intern J Toxicol-2008-Calabrese et al-369-78

Toxicol. Sci.-2001-Calabrese-330-8

Toxicol. Sci.-2003-Calabrese-246-50

Hum Exp Toxicol-2000c-Calabrese-41-75

ToxSci NCI Calabrese et al 2006

MODEL UNCERTAINTY VIA THE INTEGRATION OF HORMESIS AND LNT AS THE DEFAULT IN CANCER RISK ASSESSMENT

Edward J. Calabrese, Ph.D.

Department of Environmental Health Sciences

Morrill I, N344

University of Massachusetts

Amherst, MA 01003

Phone: 413-545-3164

Fax: 413-545-4692

E-mail: edwardc@schoolph.umass.edu

Abstract

On June 23, 2015 the U.S. Nuclear Regulatory Commission (NRC) issued a formal notice in the *Federal Register* that it would consider whether "it should amend its 'Standards for Protection Against Radiation' regulations from the linear non-threshold (LNT) model of radiation protection to the hormesis model." The present commentary supports this recommendation based on the:

(1) Flawed and deceptive history of the adoption of LNT by the U.S. National Academy of Sciences (NAS) in 1956.

(2) The documented capacity of hormesis to make more accurate predictions of biological responses for diverse biological endpoints in the low dose zone.

(3) The occurrence of extensive hormetic data from the peer-reviewed biomedical literature that revealed hormetic responses are highly generalizable, being independent of biological model, endpoint measured, inducing agent, level of biological organization and mechanism.

(4) The integration of hormesis and LNT models via a model uncertainty methodology that optimizes public health responses at 10^{-4}. Thus, both LNT and hormesis can be integratively used for risk assessment purposes and this integration defines the so-called "regulatory sweet-spot".

Overview

The comments offered here assess the scientific foundations of the three petitions (Carol Marcus, Michael Miller, and Mohan Doss) to the NRC proposing a change in the use of the LNT for risk assessment to the hormesis dose response. This assessment includes the scientific and

historical foundations of the LNT recommendation by the NAS BEAR I Committee, Genetics Panel in 1956 for regulatory agencies to adopt linearity at low dose for ionizing radiation risk assessment, how this occurred, and what it means today for NRC regulations. The comments also assess the scientific foundations of hormesis, including how accurately it predicts low dose effects and how this model compares with other dose response models such as the LNT and threshold models. Finally, it will be shown how hormesis could be applied to cancer risk assessment and how this may be used to optimize the health of radiation exposed workers and the general public.

The Scientific Foundations of LNT as Adopted by Regulatory Agencies, Including the NRC, are Based on a Fabrication and Falsification of the Research Record By the U.S.NAS BEAR I Committee, Genetics Panel (1956)

The use of the LNT for radiation induced mutation originated in 1928 with a publication by the famous physical chemist Gilbert Lewis in the journal *Nature* (Calabrese, 2013a). The article offered a mechanism for the theory of evolution. While this specific hypothesis of Lewis would not be generally accepted, subsequent research by several students of Herman J. Muller provided support for a linearity response for gonadal mutation in male fruit flies at very high doses (i.e., several hundred thousand fold greater than normal background). Muller would refer to this linear response as the Proportionality Rule. This was the term used throughout the 1930s and 1940s for what would now be called the LNT. The Proportionality Rule (i.e., LNT) became linked to a mechanism in the mid 1930s via the collaboration of leading radiation geneticists and several prominent physicists-yielding the LNT single hit theory. The single hit mechanism was based entirely on “hit theory”. This early history is described and critiqued in detail by Calabrese

(2013a) which is attached. During World War II the U.S. Atomic Energy Commission (AEC) funded research at the University of Rochester to determine the shape of the dose response in the low dose zone. The principal research was done under the direction of Curt Stern. This research and related activities are told in considerable detail by Calabrese (2011) in the attached article. The Stern research is central as it was upon these findings that the LNT would be based and accepted by U.S. regulatory agencies. Thus, a careful assessment of their research is essential for an evaluation of the three petitions to the NRC. Calabrese (2011) has shown that the interpretations of Stern and his manipulations of the publication process led to ideologically-based deliberate distortions of the nature of the dose response in the low dose zone. The history of the LNT and the roles of Stern and Muller are assessed in detailed by Calabrese (2011, 2015a, b) in the attached papers. These findings reflect documented deceptive actions by Muller on multiple occasions in order to ensure acceptance of the LNT. These publications provide a fundamental backdrop for the critical actions of the BEAR I Committee, Genetics Panel, which is now summarized.

Substantial research has recently shown that the NAS BEAR I Committee, Genetics Panel misrepresented the research record in its key technical publication in *Science* (US NAS, 1956) (June) that recommended the switch from threshold to LNT for risk assessment. This scientific misconduct has now been extensively documented in peer-reviewed publications (Calabrese 2013b, 2014, 2015c). As is presented in the Calabrese (2015c) paper the Panel was extremely concerned that their recommendation to switch to the LNT model be accepted. However, there were very strong misgivings amongst the Panelists that their LNT recommendations would not be accepted if the Panel's uncertainties and fundamental scientific disagreements concerning transgenerational genetic risks were made known via their

publications to the scientific community and the general public. These fears are documented in the Calabrese papers (2015a, b, c) via letters and other correspondence of Panel members. In the 1956 *Science* paper (US NAS, 1956) of the Panel it is written that all geneticists on the Panel (i.e. 12) were challenged to estimate the number of adverse reproductive genetic outcomes that would occur over ten generations of U.S. residents at a given level of gonadal radiation exposure. Of the 12, nine provided detailed reports with estimates. All such written documentation are publically available and provide key documentation to support the conclusions of the Calabrese paper (Calabrese 2015c). The evidence shows that the estimates of the expert Panelists wildly varied, revealing great uncertainty both within and between expert geneticists. Such profoundly large inconsistencies and disagreements were disturbing and a non-scientific ideologically-based decision was made to drop the three estimates showing the lowest damage. This significantly reduced the "appearance" of uncertainty. Yet, when the 1956 *Science* paper was published the authors (i.e., U.S. NAS Genetics Panel) stated that of the 12 geneticists on the Panel only six took up the challenge and provided estimates. However, we now know that this was not true and can be shown to be a demonstrably false statement. Dropping of the three lowest genetic damage estimates reduced a significant amount of variation yet excessive uncertainty still remained. For the remaining six estimates the uncertainly range was 750 fold and was still considered too excessive and was feared this could jeopardize acceptance for the LNT recommendation. Thus, the Panel then falsified the *Science* paper by stating their range of uncertainty to be only 100 fold. This falsification of the research record would have been discovered if the data had been published. However, the Panel formally voted not to make the data public and therefore it became impossible to challenge the falsification of the *Science* paper since no Panel member revealed these deceptions. Finally, there were three Panel geneticists who refused to provide

estimates because the process was excessively uncertain and could not be relied upon. These perspectives were also deliberately omitted as well from the *Science* paper, further misleading the *Science* journal readership.

The documentation of these actions is well established within the Calabrese papers. It shows that the key actions of the BEAR I Genetics Panel were dishonest and yet it was upon their recommendation that the linearity paradigm became accepted, adopted and implemented within the U.S. and worldwide. Thus, the foundation of the LNT was based on misrepresentations, intending to mislead regulatory agencies and others. In fact, the NRC publication of 1981 (U.S. NRC, 1981) addressing cancer risk assessment makes note of the 1956 Genetics Panel activity, using this deception based activity as foundational material. As history demonstrates, the Genetics Panel was successful in their deceptions because of the great authority of the NAS and the willingness of the regulatory and scientific communities to accept what they were told without examining the basis for the recommendation. While these accusations seem harsh, the documentation supports each statement. The problem is that it has taken some six decades for these deceptions to be revealed. Thus, the regulatory process was literally taken hostage by leading radiation geneticists acting via the prestigious U.S. National Academy of Sciences much like a highly infectious virus in order to manipulate and direct the actions of regulatory agencies in the U.S and elsewhere to their own ideological viewpoint.

Refusal of the NAS Genetics Panel to Document the Basis of the LNT Recommendation

The BEAR I Genetics Panel deliberately refused to provide any documentation to describe the scientific basis for their recommendation that the LNT be adopted by regulatory agencies. Newly uncovered documents reveal that this decision was made in order not to show profound disagreements on uncertainty in risk estimation and to focus on the identification of self-serving grant funding opportunities. The basis of their decision is given in Calabrese (2015b). More specifically, some six months after publication of their landmark 1956 report (NAS/NRC, 1956), the BEAR Genetics Panel was challenged by a number of distinguished biologists to provide the documentation upon which it based its linearity decision. It should be known that the NAS Genetics Panel had never developed any written basis for the linearity decision. It was simply by proclamation within the Panel as seen by a reading of the Panel transcripts. Now when forced to confront the reality that it had no written basis, the Panel decided that it would not provide one. This outrageous and arrogant decision was shared in writing with the President of the NAS at the time (Dr. Detlev Bronk), thereby making him fully aware of this decision. Yet he would do nothing to reverse it, making him a party to this decision.

Following the acceptance of LNT, cancer risk assessment would become strongly model driven as is seen in the later BEIR Committee reports starting in 1972. Once the LNT concept was accepted as a scientific and inaccessible belief it was transformed into a model-based construct that could not be proven wrong or easily modified. This was the case even after the discovery of DNA repair, apoptosis, adaptive response, hormesis and other new concepts, all of which could profoundly affect the shape of the dose response in the low dose zone.

Hormesis Outcompetes LNT and Threshold

Hormesis, including radiation hormesis, has a long history going back over 100 years. Calabrese and Baldwin (2000a-e) have summarized these early developments in detail (see attached). In fact, as early as 1917 ionizing radiation was shown to significantly enhance the lifespan of the insect model, the confused flour beetle, in an extremely well-designed study that has been repeatedly confirmed.

Thousands of studies have been published over the past several decades on hormesis and show it to be reproducible, generalized and independent of biological model, agent, endpoint and mechanism. In multiple direct head-to-head comparisons, the hormetic model has strikingly outperformed LNT and threshold models for accuracy in low dose predictions Calabrese and Baldwin 2001, 2003; Calabrese et al., 2006, 2008, 2010). It is important to note that the many valid hormesis studies not only clearly show the strengths of hormesis but also demonstrate serious flaws in the LNT model and establish that it cannot be used as a default. That is, if the LNT cannot be shown to provide accurate estimates in so many experimental systems and for a wide range of endpoints, including those affecting the process of cancer, then it is not possible to rely upon it as a default dose response risk assessment model. While it is widely quoted that a single valid study can discredit a powerful theory, LNT has been shown to be invalid in not one, but multiple thousands of peer-reviewed and reproducible studies, affecting a very broad spectrum of biological models and endpoints, including each key stage of the process of carcinogenesis, including tumor formation. With such extensive documentation showing the limitations of the LNT model, it is not scientifically possible to use the LNT as the default model for risk assessment and the basis for regulatory decision making. The LNT model has always

been impossible to prove correct but it could be proven to be incorrect. This is literally what this massive set of published papers on hormesis does.

The Hormesis Database

While the LNT model is being criticized in these comments for its fraudulent origin and integration into U.S. regulatory agencies and its discrediting by a very large number of valid hormesis studies, the proposal that the NRC is considering is to switch to the hormetic dose response model. The NRC should note that an Hormesis Database was created nearly 20 years ago via funding from multiple sources but principally via the U.S. Air Force to the University of Massachusetts at Amherst. This database is being continuously expanded and now there are several different types of hormetic databases which serve differing purposes. In 2005 Calabrese and Blain first published a detailed description of the original Hormesis Database. This paper has been updated on two occasions (2009 and 2011). The Hormesis Database provides detailed information on each hormetic dose experiment that first passes rigorous evaluative criteria. The findings indicate that hormesis is highly generalizable and is independent of biological model, level of biological organization (i.e., cell, organ, and organism), endpoints measured, inducing agent (e.g. chemical class, physical agents such as ionizing radiation, etc.), developmental processes, gender and mechanism. The quantitative features of the hormetic dose response are similar across all of the above parameters, suggesting that the hormetic response is constrained by the limits of biological plasticity (Calabrese, 2013c). Thus, hormesis is fundamental, generalizable, quantifiable and mechanistically explained. Also, unlike the LNT model, it can be tested in the observable range and accepted or rejected for any specific experiment. This is a very valuable feature as one does not have to rely on extrapolative modeling but on empirical data.

In the early 2000s, the most significant concern with the hormesis model was that it needed to be explained in mechanistic terms. Today this is not a concern and is useful only as an historical note. For example, in 2013 Calabrese (2013d) provided specific mechanisms for 400 different hormetic dose responses (see attached paper), where the response was mediated by a specific receptor and/or cell signaling pathway. No other dose response model has had such a plethora of mechanistic documentation to support and explain it. Further, a new hormesis mechanism paper by Calabrese is in its final stages of preparation prior to submittal to a journal. This new paper will contain nearly 600 additional hormetic dose responses with clearly identified molecular mechanisms. Thus, about 1000 dose responses for hormesis are now available with mechanisms.

These developments of the past two decades have provided information on the occurrence of hormetic dose responses, their frequency, generalizability and mechanisms. It provides a sound foundation upon which to build a regulatory program, especially given the fact that its conclusions and predictions are testable. These features make the hormetic dose response a sound choice upon which to base risk assessments upon, including cancer and non-cancer endpoints.

The New Goal: Using Hormesis To Optimize Worker Health and the Public Health

These goals can be achieved best at present via the integration of the LNT and hormesis models via a model uncertainty methodology. Recent papers by Calabrese et al. (2015a, b-see attached) demonstrate that the public health would be optimized at a LNT based risk of 10^{-4}, the dose of the hormetic nadir in animal studies. This integration yields the optimal public health response within the context of both defining and minimizing risk model uncertainty with LNT

providing the upper bound and hormesis the lower bound of risks. Thus, the NRC should change from a LNT-model based risk assessment as a default to the integrated LNT-Hormesis model as described in Calabrese et al. (2015a,b). This model could also be applied to epidemiological data with slight modification.

Funding Acknowledgement

Research activities in the area of dose response have been funded by the United States Air Force and ExxonMobil Foundation over a number of years. However, such funding support has not been used for the present manuscript. This paper will also be submitted to the Nuclear Regulatory Commission as public comments to the petition.

Declaration of Conflicting Interests

The author declares no conflict of interest.

References

Calabrese EJ. Key studies used to support cancer risk assessment questioned. Environ Mol Mut 2011; 52: 595-606.

Calabrese EJ. Origin of the linearity no threshold (LNT) dose-response concept. Arch Toxicol 2013a; 87: 1621-1633.

Calabrese EJ. How the US National Academy of Sciences misled the world community on cancer risk assessment: New findings challenge historical foundations of the linear dose response. Arch. Toxicol. 2013b; 87: 2063-2081.

Calabrese EJ. Biphasic dose responses in biology, toxicology and medicine: Accounting for their generalizability and quantitative features. Environ Poll 2013c; 182:452-460.

Calabrese EJ. Hormetic mechanisms. Crit Rev Toxicol 2013d; 43(7):580-606.

Calabrese EJ. The Genetics Panel of the NAS BEAR I Committee (1956): Epistolary evidence suggests self-interest may have prompted an exaggeration of radiation risks that led to the adoption of the LNT cancer risk assessment model. Arch Toxicol 2014; 88(9): 1631-1634.

Calabrese EJ. An abuse of risk assessment: How regulatory agencies improperly adopted LNT for cancer risk assessment. Arch Toxicol 2015a; 89(4): 647-648.

Calabrese EJ. Cancer risk assessment foundation unraveling: New historical evidence reveals that the US National Academy of Sciences (US NAS), Biological Effects of Atomic Radiation (BEAR) Committee Genetics Panel falsified the research record to promote acceptance of the LNT. Arch Toxicol 2015b; 89(4): 649-650.

Calabrese EJ. On the origins of the linear no-threshold (LNT) dogma by means of untruths, artful dodges and blind faith. Environ Res 2015c; 142: 432-442.

Calabrese EJ, Baldwin LA. Chemical hormesis: Its historical foundations as a biological hypothesis. Hum Exper Toxicol 2000a; 19: 2-31.

Calabrese EJ, Baldwin LA. The marginalization of hormesis. Hum Exper Toxicol 2000b; 19: 32-40.

Calabrese EJ, Baldwin LA. Radiation hormesis: Its historical foundations as a biological hypothesis. Hum Exper Toxicol 2000c; 19: 41-75.

Calabrese EJ, Baldwin LA. Radiation hormesis: The demise of a legitimate hypothesis. Hum Exper Toxicol 2000d; 19: 76-84.

Calabrese EJ, Baldwin LA. Tales of two similar hypotheses: The risk and fall of chemical and radiation hormesis. Hum Exper Toxicol 2000e; 19: 85-97.

Calabrese EJ, Baldwin LA. The frequency of U-shaped dose responses in the toxicological literature. Toxicol Sci 2001; 62(2): 330-338.

Calabrese EJ, Baldwin LA. The hormetic dose-response model is more common than the threshold model in toxicology. Toxicol Sci 2003; 71: 246-250.

Calabrese EJ, Blain R. The occurrence of hormetic dose responses in the toxicological literature, the hormesis database: An overview. Toxicol Appl Pharmacol 2005; 202(3): 289-301.

Calabrese EJ, Blain RB. Hormesis and plant biology. Environ Poll 2009; 157(1): 42-48.

Calabrese EJ, Blain RB. The hormesis database: The occurrence of hormetic dose responses in the toxicological literature. Reg Toxicol Pharmacol 2011; 61(1): 73-81.

Calabrese EJ, Staudenmayer JW, Stanek EJ III, et al. Hormesis outperforms threshold model in National Cancer Institute antitumor drug screening database. Toxicol Sci 2006; 94(2): 368-378.

Calabrese EJ, Stanek EJ III, Nascarella MA, et al. Hormesis predicts low dose-responses better than threshold models. Intl J Toxicol 2008; 27: 369-378.

Calabrese EJ, Hoffmann GR, Stanek EJ III, et al. Hormesis in high-throughput screening of antibacterial compounds in E. coli. Hum Exp Toxicol 2010; 29: 667-677.

Calabrese EJ, Shamoun DY, Hanekamp JC. Cancer risk assessment: Optimizing human health through linear dose-response models. Food Chem Toxicol 2015a; 81: 137-140.

Calabrese EJ, Shamoun DY, Hanekamp JC. The integration of LNT and hormesis for cancer risk assessment optimizes public health protection. Health Physics Journal 2015b; In press.

National Academy of Sciences (NAS)/National Research Council (NRC). The biological effects of atomic radiation (BEAR): a report to the public. NAS/NRC, Washington DC; 1956.

U.S. NAS BEAR I Committee, Genetics Panel. Genetic effects of atomic radiation. Science (1956); 123(3209): 1157-1164.

U.S. Nuclear Regulatory Commission (US NRC). Instruction concerning risks form occupational radiation exposure. Regulatory Guide 8.29 (Task OH 902-4) July 1981, Washington DC; 1981.

Arch Toxicol
DOI 10.1007/s00204-015-1454-4

LETTER TO THE EDITOR, NEWS AND VIEWS

An abuse of risk assessment: how regulatory agencies improperly adopted LNT for cancer risk assessment

Edward J. Calabrese

Received: 15 December 2014 / Accepted: 6 January 2015

Abstract The Genetics Panel of the National Academy of Sciences' Committee on Biological Effects of Atomic Radiation (BEAR) recommended the adoption of the linear dose–response model in 1956, abandoning the threshold dose–response for genetic risk assessments. This recommendation was quickly generalized to include somatic cells for cancer risk assessment and later was instrumental in the adoption of linearity for carcinogen risk assessment by the Environmental Protection Agency. The Genetics Panel failed to provide any scientific assessment to support this recommendation and refused to do so when later challenged by other leading scientists. Thus, the linearity model used in cancer risk assessment was based on ideology rather than science and originated with the recommendation of the NAS BEAR Committee Genetics Panel. Historical documentation in support of these conclusions is provided in the transcripts of the Panel meetings and in previously unexamined correspondence among Panel members.

Keywords Mutation · Linear non-threshold (LNT) · Risk assessment · Carcinogen · Threshold dose response · Ionizing radiation

Electronic supplementary material The online version of this article (doi:10.1007/s00204-015-1454-4) contains supplementary material, which is available to authorized users.

E. J. Calabrese (✉)
Department of Public Health, Environmental Health Sciences, Morrill I, N344, University of Massachusetts, Amherst, MA 01003, USA
e-mail: edwardc@schoolph.umass.edu

The most significant event in the history of environmental risk assessment was the recommendation by the United States National Academy of Sciences (NAS), Biological Effects of Atomic Radiation (BEAR) Committee, Genetics Panel in 1956 to switch from a threshold to a linear dose–response model for the assessment of genomic mutation risk (Anonymous 1956; NAS/NRC 1956). Within a brief period of time, this recommendation became generalized to somatic cells by other governmental advisory committees and was eventually applied to cancer risk assessment. Although this linear dose–response paradigm was originally intended to be used for ionizing radiation, it would later be adopted by the US Environmental Protection Agency and directly applied to chemical carcinogens (Albert 1994; Calabrese 2013a, b), thereby affecting worldwide cancer risk assessment for the past several decades.

Given the significance of this action by the NAS BEAR I Committee, Genetics Panel and the long history of the threshold dose–response model in regulatory practice, I was interested in learning the answers to several key questions: how was this recommendation made, what was the nature of the debate, what were the persuasive and compelling arguments, and what were the roles played by various individuals on the Panel? I therefore obtained transcripts of the BEAR I Committee, Genetics Panel meetings in 1955 and 1956. It was a bit like reading the book after seeing the end of the movie. To my surprise, the BEAR I Committee, Genetics Panel was uniformly confident in their belief that linearity for genomic risk assessment was the correct perspective, while being arrogantly dismissive of both the threshold perspective and those who supported it. So dismissive of the alternative model was the Genetics Panel that it was never viewed as a debatable issue, nor was it ever debated. What a disappointment. I had so looked forward to retrospectively witnessing how the leading thinkers of

Published online: 18 January 2015

their time confronted this seminal issue on dose–response, how they intellectually sparred with one another, and whose logic and facts helped carry the day for the linearity model. The NAS BEAR I Committee, Genetics Panel made the switch from a threshold to a linear dose–response risk assessment model by "proclamation," with no debate and without providing a detailed (or actually even any) evaluation, such as would be expected of any scientific advisory group—most certainly of one at the level of the National Academy of Sciences on such matters of national and international significance. In retrospect, this should not have been too surprising as I had documented in previous publications (Calabrese 2011a, b, 2012, 2013a, b) the inherent intellectual dishonesty of key leaders of the radiation genetics community, such as Curt Stern and Hermann Muller on the issue of threshold versus linear dose–response and how they successfully distorted the scientific record in order to achieve their goal of a linear dose–response for risk assessment. The linear dose–response recommendation by this Genetics Panel would be broadly extolled by leading media outlets on the day of its release as the most extensive assessment ever undertaken on the topic by a most prestigious group of American scientists. The National Academy of Sciences report was literally a front-page story in the New York Times with the linearity risk assessment framework leading the way.

Despite the widely acknowledged success of the BEAR I Committee, Genetics Panel in getting their message out to the scientific community, governmental bodies, and the public, the reports of the BEAR I Committee, Genetics Panel were eventually read by members of the scientific community. This resulted in a number of leading biologists challenging the Genetics Panel, demanding to know the scientific basis of the decision in favor of linearity. However, as noted above, the Genetics Panel had not undertaken such an assessment and was not in a position to explain their actions nor to defend a report that lacked a scientific foundation. Showing its disdain for those challenging this report, the Genetics Panel decided not to provide the information to the scientific community. This decision was rendered to the President of the National Academy of Sciences without any evidence of his objection. The adoption of the linear non-threshold (LNT) dose–response model by the National Academy of Sciences therefore was made without a scientific assessment and, of course, a refusal to provide one when challenged.

The recommendation to switch to a linear dose–response by the NAS BEAR I Committee, Genetics Panel, as announced to the world by leading media outlets, reflects an abdication of societal responsibility on a critical and enduring public health issue. This paper provides the first reporting of these actions in the history of the National Academy of Sciences and in governmental risk assessment practices for cancer. It reveals that current cancer risk assessment practices originated from an ideological set of beliefs from leading scientists rather than a scientific assessment. A fully documented assessment of this story is provided in the Supplementary Data section.

Acknowledgments Research activities in the area of dose–response have been funded by the United States Air Force (FA9550-13-1-0047) and ExxonMobil Foundation over a number of years. However, such funding support has not been used for the present manuscript.

Conflict of interest Author declares no conflict of interest.

References

Albert RE (1994) Carcinogen risk assessment in the U.S. Environmental Protection Agency. Crit Rev Toxicol 24(1):75–85

Anonymous (1956) Genetic effects of atomic radiation. Summary Report of the Committee on Biological Effects of Atomic Radiation by the National Academy of Sciences, BEAR I Genetics Panel (W. Weaver, Chair). Sci 123:1157–1164. [Erratum, Science 124:170]

Calabrese EJ (2011a) Muller's Nobel lecture on dose–response for ionizing radiation: ideology or science? Arch Toxicol 85(12):1495–1498

Calabrese EJ (2011b) Key studies used to support cancer risk assessment questioned. Environ Mol Mutagen 52(8):595–606

Calabrese EJ (2012) Muller's Nobel Prize lecture: when ideology prevailed over science. Toxicol Sci 126:1–4

Calabrese EJ (2013a) Origin of the linearity no threshold (LNT) dose–response concept. Arch Toxicol 87:1621–1633

Calabrese EJ (2013b) How the US National Academy of Sciences misled the world community on cancer risk assessment: new findings challenge historical foundations of the linear dose response. Arch Toxicol 87(12):2063–2081

National Academy of Science (NAS)/National Research Council (NRC) (1956) The biological effects of atomic radiation. A report to the public. Washington

Supplementary Material

LNT'S FAILED HISTORY:
An Abdicated Responsibility - How the US NAS BEAR I Committee Genetics Panel Failed To Assess LNT Prior To Recommending Its Use by US Regulatory Agencies

Edward J. Calabrese, Ph.D.
Professor of Toxicology, Department of Public Health
Environmental Health Sciences, Morrill I, N344
University of Massachusetts, Amherst, MA 01003
Phone: 413-545-3164, Fax: 413-545-4692
E-mail: edwardc@schoolph.umass.edu

Abstract

The U.S. National Academy of Sciences (NAS) Biological Effects of Atomic Radiation (BEAR) I Genetics Panel report recommended a linear dose response to assess the risk of genomic mutation from ionizing radiation. This represented a major change assessing risks which had been based on a threshold dose response model. This recommendation was soon generalized to somatic injury and applied to cancer risk assessment for ionizing radiation and later for chemical carcinogens. An evaluation of the transcriptional records of the Genetics Panel, intra-panel correspondence and work products, reveals that the Panel failed to provide an assessment of which dose response model best characterized the effects of ionizing radiation on the genome. Lacking such an assessment, the recommendation for a linear model was based upon an assumption of the Panel.

The Panel's failure to assess the scientific basis of the dose response for ionizing radiation, while recommending strongly a switch to linearity, represents an abdication of responsibility. It led to a deliberately false public understanding that their risk assessment for ionizing radiation was based on "the most comprehensive effort" ever undertaken in the United States by a committee of outstanding scientists as characterized by a front page New York Times story (Leviero 1956) one day after the release of the Panel report (June 13, 1956) and similarly reported in other scientific and public venues.

Key Words: linearity, threshold, mutation, risk assessment, dose response, cancer

Introduction

The US NAS BEAR I Committee Genetics Panel in 1956 recommended that the risks associated with ionizing radiation to the human genome no longer be evaluated via the use of a threshold dose response model but with a linear at low dose model. This recommendation was quickly adopted by the scientific and regulatory communities and soon generalized to somatic cells for application to cancer risk assessment for ionizing radiation (Taylor 1960, 1963, 1965). Some two decades later the U.S. NAS Safe Drinking Water Committee (NAS 1977) relied upon this linearity at low dose recommendation for assessing risks of chemical carcinogens. In many respects, therefore, the report of the 1956 BEAR I Genetics Panel was the most influential advisory report ever published on risk assessment. The Genetics Panel published two reports, one as part of a general NAS document intended for the media and the general public (NAS/NRC 1956), while the other was a more technical paper published in the journal Science (Anonymous 1956a). The key conceptual conclusion of the Genetics Panel was that ionizing radiation induces genomic mutations which are nearly always harmful and the damage is irreversible, cumulative, and directly proportional to dose, such that there is no safe level of exposure.

NAS Genetics Panel

Since the toxicology, medical and regulatory communities were still being dominated by the threshold dose response model for all endpoints during this time period, the rejection of threshold dose response and its replacement with the linear model constituted no less than a major scientific and regulatory revolution. As such, one would expect that a principal task of the Panel was to document the strengths and limitations of the threshold and linearity dose response models and thoroughly debate this topic during their sessions prior to recommending the retention of the threshold model for genetic risk assessment, a switch to linearity or some other risk assessment approach. In anticipation of reading such an historic debate, yet knowing in advance that the Genetics Panel recommended the rejection of the threshold model and the immediate transition to linearity, the transcripts of the Genetics Panel meetings were obtained from the US NAS Archives. I was surprised to learn that the Panel did not research, assess, nor debate the dose response question. The issue of dose response risk assessment model selection had been "decided" by the closely knit radiation genetics community prior to the creation of the Panel, based on the leadership of Hermann J. Muller and Curt Stern [Calabrese 2013; Crow 1995). In fact, at the first meeting of the Genetics Panel on November 21, 1955 at Princeton University, the well-known geneticist Alfred Sturtevant from California Technical Institute was dismissive of the issue of dose response as he had "no doubt about the correctness of the linear dose response" model and that any effort to further document support for it would only be for "propaganda value," as means to educate and convince the non-geneticists. This dismissive, and indeed arrogant attitude, was pervasive amongst the geneticists on the Panel concerning their unique professional insights on the issue of mutation. In line with this perspective, the key leaders of the genetics community ascribed to a series of firmly held beliefs about radiation and mutations. In fact, at the second meeting of the Panel (February 5, 1956) Tracy Sonneborn, a member of the Panel and colleague of Muller at the University of Indiana, read into the record what amounted to a detailed series of "beliefs", in essence, a geneticist's creed, about dose response, mutation, ionizing radiation and risk assessment (starting on page 81 of the transcript) (i.e. nearly always harmful, irreversible, cumulative and linear) (NAS 1956). Amongst the Panel of 17 members, of which 13 were prominent geneticists, there was no dissent.

The "Debate"

The only attempt at "dissent" was initiated by Bentley Glass on February 5, 1956 (page 108 of the transcript) (NAS 1956). Glass stated that the only challenge to their geneticist creed as articulated by Sonneborn, to which he was aware, concerned the concept of linearity. Glass stated he wanted to explore the question (i.e., the challenge to linearity) within the Panel, "not because I believe personally in the objection that I am going to raise but to play the role of the devil's advocate here." What follows next is the transcript discussion immediately after the comment of Glass:

"DR. CROW: Which assumptions are these?

DR. GLASS: Well, they were in Dr. Weaver's formation too, but they are the two at the beginning of Sonneborn's genetic considerations.

After having made a talk to the physicists at Rutger's recently on this general topic of "The Geneticist Views the Dangers from Atomic Radiations," I was surprised to find that one of the geneticists who dained to come out to hear the talk challenged this particular assumption which I had put out as one of the assumptions that all geneticists are agreed upon, and his line of reasoning – which, of course, is something that the physicists will very eagerly and quickly seize upon I think because most of them want to believe in a threshold effect as at least a possibility, if not demonstrated beyond all question at the moment – his line of reasoning was as follows: that the view that there is no threshold in the response of mutations to dosage is largely based, apart from the experimental data, on the target theory of the effects of radiation, and that the microbial geneticists (and this man was a microbial geneticist) having shown that there is a chemical and indirect mediation between the production of ionizations and the

occurrence of point mutations makes it altogether probable that somewhere or other there is a threshold, and he felt very uncomfortable about the assumption that there is no threshold if you go down to low enough doses. This is heresy in their midst.
DR. WRIGHT: In energy if not in ionization. Isn't your threshold there in energy? Perhaps one electron volt or two does account for the threshold. But ionization is so far above any possible threshold that it does not seem to me that bears on the ionization argument at all.
DR. STURTEVANT: I have met with this objection. They have usually been willing to agree, however, if I worded it that at the moment the best bet is that there is no threshold and we have to proceed on that.
DR. GLASS: That is all right. But I think we have to take some cognizance of this argument.
DR. CROW: Do you know for certain in any area?
DR. WRIGHT: Isn't the experimental evidence practically conclusive there, to the extent that they have been spaced so that from the physicist's standpoint there is no possibility?
DR. CROW: If you have one ionization per hour or whatever.
DR. GLASS: It is convincing me, too.
DR. RUSSELL: There is both the theoretical and the practical viewpoint they have these several orders of magnitude from all the other kinds of things that we are questioning and recommending research on."

Chairman Weaver then refocused the discussion by inviting Panel member Bernard Kaufmann to discuss research of Arnold H. Sparrow from Brookhaven National Laboratory on mutations in plants at low doses. Kaufmann stated that Sparrow and Singleton (Sparrow and Singleton 1953) reported that 0.41 r per day gives a statistically significant mutation effect. Kaufmann failed to note that (on the top of Sparrow & Singleton's page 37) there was actually mutation data for a dose (0.084 r/day) lower than 0.41 r/day and that it had no treatment effect. This finding would have challenged the linearity position if it had not been omitted by Kaufmann. The page 37 statement of Sparrow and Singleton (1953) is as follows:

> "The data in table 2 show that 0.084 r per day caused no significant increase but that 0.41 r per day (or higher) did show a statistically significant effect (table 2). However, the increase was less than twice that of the control. Since 0.41 r per day of radiation is more than one thousand times greater than the naturally occurring intensity these data do not support the theory that the spontaneously occurring micronuclei are produced by naturally occurring ionizing radiation."

After the brief discussion of the Sparrow data and the misrepresentation of his data by Kaufmann all discussion on the issue of linearity vs threshold ended for the BEAR I Genetics Panel.

It is difficult to comprehend that this was the extent to which the Genetics Panel acknowledged the dose-response controversy and discussed the key scientific issues concerning the nature of the dose-response in the low dose zone. This had been a matter of contention for the past two decades with various high level advisory committees in the US and internationally. It was also a critical component of Muller's Nobel Prize lecture (Calabrese 2011a, 2012) and a major component of the health effects research of the Manhattan Project (Calabrese 2012, 2013; Caspari and Stern 1948; Spencer and Stern 1948; Uphoff and Stern 1949) and of the Atomic Energy Commission. In many respects, the principal reason for the creation of the Genetics Panel was to address the issue of how to assess genetic risks at low doses of ionizing radiation. In the end, the Panel provided the scientific community and the public with a statement of beliefs, none of which was researched, documented, assessed, debated and refined as might be expected if a legitimate evaluation process had been followed.

Acknowledgement of the BEAR I Genetics Panel Failure

On November 26, 1956 Bentley Glass wrote to the BEAR II Genetics Panel stating:
"From impressions I have gathered during the course of the past five and a half months since our report [BEAR I Genetics Panel Report] was released to the public [i.e., June 12, 1956], I have come to the

conclusion that there are several matters of some urging for consideration by our Committee." The second of these considerations related to the linearity question as now stated by Glass:

> "II. I have met continuing doubt from well-informed biological scientists in regard to the geneticists' assumption that there is no threshold for mutation. This leads me to believe that there is a need to prepare a statement and exposition of this point that will (A) summarize existing data on the matter, (B) present the physical arguments against the existence of a threshold, and (C) deal with the experimental possibilities of further investigating the question in suitable biological material."

The statement of Glass is significant in light of the report of the Genetics Panel in Science (Anonymous 1956a). It is clear that he received significant push-back to the LNT assumption by some "well informed biologists" such that he now felt it was necessary for the new Genetics Panel (i.e., BEAR II) to provide documentation in support of linearity and against threshold. Now that the Panel's report was challenged, Glass felt the need for an appropriate scientific response. Even in the case of Glass, his written statement indicates bias as he recommends not a search for scientific understanding of the nature of the dose response in the low dose zone for ionizing radiation, but how to make the case for linearity and against threshold. Based on such insights into the actions of NAS BEAR I Genetics Panel, this group was selected based on both high achievement and their unified belief that genetic mutations were considered irreversible, cumulative and linear with respect to dose. So strong was their collective belief that the group failed to provide any scientific justification for their highly influential linear dose response recommendation. Despite this suggestion by Glass now nearly six months after the release of the report, there was no demonstrable attempt to address this most fundamental issue, but rather their first item on the BEAR II Genetics Panel agenda was to propose a funded research program for the genetics community (Memo to Members of the Academy Genetic Committee - i.e., BEAR II) (Beadle 1956a).

This challenge of Glass (1956) would be a continuing one (August 24, Beadle Memo to Genetics Panel) (Beadle 1956b) for the Genetics Panel, even proceeding the letter of Glass (1956) and a finalizing of their internal debate based on a September 11, 1957 letter from the Chairman of BEAR II Genetics Panel (G. Beadle) (Beadle 1957) to Detlev Brock, President of the NAS and copies to Weaver (Chairman of BEAR I Genetics Panel) and the Panel. In this September 11, 1957 letter, Beadle stated that the development of a detailed technical document that would provide the scientific basis for the BEAR I Genetics Panel report was not justified since it would require excessive resources (i.e. one or two geneticists working full time), and there did not appear to be mounting external pressure to do so. Beadle then offered the incomprehensible suggestion that since several published review papers (none were identified) that presumably included some topics addressed in some manner by the Panel, there was no need to consider this issue further. Thus, the request of Glass was finally tabled, and the NAS leadership was fully informed of this decision.

Discussion

So what do these historical insights mean? The switch from threshold to linearity for risk assessment by the US and other governments that followed the NAS report was not based on an assessment of the issue, but rather on a set of pre-conceived beliefs. As demonstrated in a series of previous articles (Calabrese 2011a,b, 2012), these beliefs had been acquired via deliberate misrepresentation of the scientific literature by key leaders of the radiation genetics community, led by the Nobel Prize winner H.J. Muller and Curt Stern (Calabrese 2011b, 2013). It is apparent that the NAS administration, the scientific community and regulatory agencies failed to demand that the Genetics Panel provide a scientifically supported basis for their recommendation of a switch to the linear dose response.

A strong indicator of their public success became evident almost immediately when the New York Times (Leviero 1956) provided a front page story on June 13, 1956 with the title "Scientists Term Radiation A Peril to Future of Man: Even Small Doses Can Prove Harmful to Descendents of Victim". The first paragraph of the article stated that "A committee of outstanding scientists reported today that atomic radiation, no matter how small the dose, harms not only the person receiving it but also all his descendents." The next paragraph would claim that "it was the most comprehensive United States effort to determine how the future of the human race might be affected by the unleashing of nuclear power." Similar reports were also found in the Washington Post (Haseltine 1956), Time Magazine (Anonymous 1956b,c), US News and World Report (Anonymous 1956d), News of Science Section, Science journal (Anonymous 1956e), The Saturday Review (Muller 1956), Challenge Interviews (Weaver 1956), Journal of The Franklin Institute (Weaver 1957a), Bulletin of Atomic Scientists (Weaver 1957b), Public Health Reports (Weaver 1957c), Scientific American (Crow 1959; Beadle 1959), The Lancet (Anonymous 1956f,g) and other leading publications.

As the present paper demonstrates, the Genetics Panel's effort was anything but comprehensive. Rather, it represented an abdication of professional and ethical responsibility, using their outstanding reputations to present a false image of a detailed and objective assessment when it was their ideology that prevailed. While previous articles have captured Muller and Stern's scientific deceptions on the issue of linearity and their impact on the Genetics Panel (Calabrese 2013), and several members of the Panel in serious self-serving comments that undercut the credibility of the Panel (Calabrese 2014), the present paper has captured their silence and illusion as far as an effort to assess the nature of the dose response in the low dose zone.

Policy should be based on facts, not assumptions. In the absence of a factual foundation, the assumptions should be stated, explained, and justified. Not only did the Genetics Panel fail to serve the public, it was permitted to mislead US national policy and cancer risk assessment predictions and that of other countries by a compliant NAS administration, scientific community and press, under the false impression that their recommendation represented an objective and comprehensive assessment. The implications of this deception have been enormous and continue to the present.

Acknowledgement

Research activities in the area of dose response have been funded by the United States Air Force and ExxonMobil Foundation over a number of years. However, such funding support has not been used for the present manuscript. Author declares no conflict of interest.

References

Anonymous (1956a) (Genetics Panel, W. Weaver, Chair). National Academy of Sciences (NAS), Biological Effects of Atomic Radiation (BEAR). Genetic Effects of Atomic Radiation. Science, 123:1157-1164. [Erratum, Science, 124:170]

Anonymous (1956b) X-ray danger. Time Magazine-Medicine, October 10; pp. 67.

Anonymous (1956c) Atomic radiation: The rs are coming. Time Magazine-Science, June 25, 64-65.

Anonymous (1956d) What you should know about danger from x-rays. US News and World Report, June 29; pp. 44-48.

Anonymous (1956e) Biological effects of Atomic Radiations. Science-News of Science 123:1110-1111.

Anonymous (1956f) Biological effects of atomic radiation. The Lancet 167:1007.

Anonymous (1956g) Radiation hazards. The Lancet 167:999-1000.

Beadle GW (1956a) Memo to Members of the Academy Genetics Committee, American Philosophical Society, Philadelphia PA USA. August 6.

Beadle GW (1956b) Memo to Genetics Committee, American Philosophical Society, Philadelphia PA USA. August 24.

Beadle GW (1957) Letter to Detlev Bronk, American Philosophical Society, Philadelphia PA USA. September 11.

Beadle GW (1959) Ionizing radiation and the citizen. Sci Amer 201:219-232.

Calabrese EJ (2011a) Muller's Nobel lecture on dose-response for ionizing radiation: ideology or science? Arch Toxicol 85:1495-1598.

Calabrese EJ (2011b) Key studies used to support cancer risk assessment questioned. Environ Mol Mut 52:595-606.

Calabrese EJ (2012) Muller's Nobel Prize lecture: When ideology prevailed over science. Toxicol Sci 126:1-4.

Calabrese EJ (2013) How the US National Academy of Sciences misled the world community on cancer risk assessment: New findings challenge historical foundations of the linear dose response. Arch Toxicol 87:2063-2081; DOI: 10.1007/s00204-013-1105-6.

Calabrese EJ (2014) The Genetics Panel of the NAS BEAR I Committee (1956): epistolary evidence suggests self-interest may have prompted an exaggeration of radiation risks that led to the adoption of the LNT cancer risk assessment model. Arch Toxicol DOI: 10.1007/s00207-014 1306-7.

Caspari E, Stern C (1948) The influence of chronic irradiation with gamma-rays at low dosages on the mutation rate in Drosophila Melanogaster. Genetics 33:75-95.

Crow JF (1959) Ionizing radiation and evolution. Sci Amer 201:138-160.

Crow JF (1995) Quarreling geneticists and a diplomat. Genetics 140:421-426.

Glass B (1956) Memo to the Committee on Genetic Effects of Atomic Radiation, American Philosophical Society, Philadelphia PA USA. November 26.

Haseltine N (1956) Nation's top scientists call for atomic radiation control; fear shorter life expectancy and mentally deficient babies. The Washington Post, June 13, p. 14.

Leviero A (1956) Scientists term radiation a peril to future of man: even small dose can prove harmful to descendant of victim, report states. The New York Times June 13, page 1.

Muller HJ (1956) Race poisoning by radiation. The Saturday Review, June 9, pp. 9-11, 37-39.

National Academy of Sciences (NAS) (1956) BEAR I Committee Transcript Proceedings Conference on Genetics, Chicago IL, USA, February 5-6.

National Academy of Sciences (NAS) Safe Drinking Water Committee (1977) Drinking water and health. National Academy of Sciences (NAS), Washington DC, USA.

National Academy of Sciences (NAS)/National Research Council (NRC) (1956). The biological effects of atomic radiation (BEAR): A report to the public. NAS/NRC, Washington DC, USA.

Sparrow AH, Singleton WR (1953). The use of radiocobalt as a source of gamma rays and some effects of chronic irradiation on growing plants. The American Naturalist 87:29-48.

Spencer WP, Stern C (1948). Experiments to test the validity of the linear R-dose mutation frequency relation in Drosophila at low dosage. Genetics 33:43-74.

Taylor LS (1960) Radiation protection standards. Radiology 71:824-831.

Taylor LS (1963) Radiation effects on man – Radiation protection standards. Nucleonics 21:58-60.

Taylor LS (1965). Philosophical influences on radiation protection standards. Hlth Phy 11:859-864.

Uphoff DE, Stern C (1949) The genetic effects of low intensity irradiation. Science 109:609-610.

Weaver W (1956) Warren Weaver Interview – How radiation can affect our children. Challenge Interviews 5:54-59.

Weaver W (1957a). Radiations and the genetic threat. J Franklin Inst 263:283-293.

Weaver W (1957b). Science and the citizen. Bull Atomic Sci December, pp. 361-365.

Weaver W (1957c). New dimensions of learning in a free society. Pub Hlth Rep 72:1005-1008.

Arch Toxicol (2013) 87:1621–1633
DOI 10.1007/s00204-013-1104-7

REVIEW ARTICLE

Origin of the linearity no threshold (LNT) dose–response concept

Edward J. Calabrese

Received: 15 March 2013 / Accepted: 11 July 2013 / Published online: 26 July 2013

Abstract This paper identifies the origin of the linearity at low-dose concept [i.e., linear no threshold (LNT)] for ionizing radiation-induced mutation. After the discovery of X-ray-induced mutations, Olson and Lewis (Nature 121(3052):673–674, 1928) proposed that cosmic/terrestrial radiation-induced mutations provide the principal mechanism for the induction of heritable traits, providing the driving force for evolution. For this concept to be general, a LNT dose relationship was assumed, with genetic damage proportional to the energy absorbed. Subsequent studies suggested a linear dose response for ionizing radiation-induced mutations (Hanson and Heys in Am Nat 63(686):201–213, 1929; Oliver in Science 71:44–46, 1930), supporting the evolutionary hypothesis. Based on an evaluation of spontaneous and ionizing radiation-induced mutation with *Drosophila*, Muller argued that background radiation had a negligible impact on spontaneous mutation, discrediting the ionizing radiation-based evolutionary hypothesis. Nonetheless, an expanded set of mutation dose–response observations provided a basis for collaboration between theoretical physicists (Max Delbruck and Gunter Zimmer) and the radiation geneticist Nicolai Timoféeff-Ressovsky. They developed interrelated physical science-based genetics perspectives including a biophysical model of the gene, a radiation-induced gene mutation target theory and the single-hit hypothesis of radiation-induced mutation, which, when integrated, provided the theoretical mechanism and mathematical basis for the LNT model. The LNT concept became accepted by radiation geneticists and recommended by national/international advisory committees for risk assessment of ionizing radiation-induced mutational damage/cancer from the mid-1950s to the present. The LNT concept was later generalized to chemical carcinogen risk assessment and used by public health and regulatory agencies worldwide.

Keywords Ionizing radiation · Linearity · Dose response · Risk assessment · Threshold dose response · Target theory · Eugenics · LNT

E. J. Calabrese (✉)
Department of Public Health, Environmental Health Sciences, University of Massachusetts, Morrill I, N344, Amherst, MA 01003, USA
e-mail: edwardc@schoolph.umass.edu

Introduction

In 1956, the US National Academy of Sciences (NAS) Committee on Biological Effects of Atomic Radiation (BEAR I)/Genetics Panel issued the most far reaching recommendation in the history of risk assessment that genomic risks associated with exposure to ionizing radiation should be evaluated with a linear dose–response model, no longer via the threshold dose–response model that had long been the "gold" standard for medicine and physiology (Calabrese 2005, 2009a, 2011). The Genetics Panel members believed that there was no safe exposure to ionizing radiation for reproductive cells with the mutation risk being increased even with a single ionization (Hamblin 2007). The LNT concept was generalized in 1958 to somatic cells and cancer risk assessment by the National Committee for Radiation Protection and Measurement (NCRPM) (Whittemore 1986). Quickly thereafter, other national and international advisory committees and organizations adopted such judgments for ionizing radiation (Calabrese 2009b). In 1977, the Safe Drinking Water Committee (SDWC) of the US NAS extended the linear dose–response risk assessment model of the BEAR/

Biological Effects of Ionizing Radiation (BEIR) committees to chemical carcinogens, a recommendation that was soon adopted and implemented by the Environmental Protection Agency (EPA). On a parallel track, similar LNT risk assessment procedures were adopted by the Food and Drug Administration (FDA) in 1977 concerning animal carcinogen drug residues.

Despite the fact that the LNT model has been of central importance in chemical and ionizing radiation regulatory risk assessment, its origin is not within the environmental/occupational risk assessment domain. The current paper provides a novel historical assessment of the scientific origin of the LNT. It will show that the LNT was first applied to the field of biology in 1928 to explain the occurrence of genetic variation that would serve as the "biological engine" for evolution. The paper will also demonstrate how the linear dose–response model as proposed by Olson and Lewis (1928), which soon afterward became transformed into a "Proportionality Rule" by Muller (1930), became mechanistically framed within the context of a single-"hit" hypothesis based on the target theory by Timoféeff-Ressovsky et al. (1935) in a unique collaborative effort between leading theoretical physicists and radiation genetics. This paper extends two earlier publications within Archives of Toxicology concerning historical foundations of the LNT concept (Calabrese, 2009b) and threshold/hormetic (Calabrese 2009a) models.

Evolution and LNT

Since the publication of the Origin of Species in 1859 by Darwin and the rediscovery of the works of Mendel on gene inheritance, there was intense interest in the biological community to determine the cause of genetic change or novelty that would be subject to natural selection, thereby providing an important mechanism of evolution. As noted by Patterson (1933), a well-known colleague of Hermann J. Muller at the University of Texas/Austin, "the important question in biology is the problem of evolution" referring to the need to understand the mechanism of evolution at the gene level. Despite the fact that the gene was more of a concept than a physical entity during the early decades of the twentieth century, it was widely believed that the gene was the basic unit of heredity and that the driving force for evolutionary change must be via the induction of heritable genetic changes or mutations at the gene level (Muller 1922). This perspective provided the basis for intense interest by numerous genetics researchers in the second and third decades of the twentieth century to induce alterations in heritable traits by environmental (e.g., temperature) alterations, physiological stressors (e.g., starvation), as well as toxic chemicals and ionizing and non-ionizing radiation.

Given the central importance of evolution in biology and underscoring the intensity of the competition to be the first to demonstrate inducible heritable changes, Muller (1927) provided only an initial "discussion" of his mutagenicity findings with no data in his now famous *Science* paper that led to his Nobel Prize in 1946. This was done in order to secure recognition of being the first to report induction of heritable mutations by an environmental agent (i.e., X-rays). The supporting data were published the next year in a conference proceeding of very limited distribution based on the World Cat database (Muller 1928a) and also within the *Proceedings of the National Academy of Sciences* (PNAS) (Muller 1928b). Not only were the findings of mutation significant so too was the fact that the mutation rate was increased by about 150-fold at the highest dose tested.

Muller speculated that naturally occurring ionizing radiation might be a significant explanatory factor for genetic variation and may drive the evolution process. However, Muller was cautious in making the mutation–evolution link as the doses he had used to induce mutation were extremely high, exceeding background by about 200,000-fold, causing sterility or mortality in a substantial proportion of the fruit flies tested. In addition, the dose response was not linear but closer to a square root function due to a modest decline from linearity at the highest dose (Muller 1927, 1928a). If the true dose response for ionizing radiation-induced gene mutation was linear at low dose, as a general condition, then it may have explanatory implications for an evolution mechanism. Consequently, he soon directed several members in his laboratory to assess the topic of dose response more fully than he did in his groundbreaking mutation discovery. While the follow-up research by Muller's group was being undertaken, Axel R. Olson and the prestigious physical chemist Gilbert N. Lewis (1928) of the University of California/Berkeley published a proposal on April 28, 1928, in *Nature* that natural radioactivity was likely a significant cause of mutation that could generate variability from the parent generation and affect the process of evolution. These authors based this supposition on a report of January 1, 1928, in *PNAS* by Goodspeed and Olson on X-ray-induced heritable changes in tobacco. These authors claimed that the tobacco plant studies were specially planned to facilitate a direct comparison of mutation rates between the artificial X-rays and "naturally occurring radiations." Olson and Lewis (1928) also stated that "since the rays can only be effective when they are absorbed, and this produces ionizations, it seems safe to assume that the various rays will produce biological effects *in proportion to the ionization which they cause*" (emphasis added), a perspective based on the emerging target theory for radiation-induced biological effects proposed by leaders in the physics community (Glocker 1927; Crowther 1924).

Olson and Lewis (1928) then utilized a simple linear mathematical model to derive a mutation estimate at a selected natural background radiation dose. With this method, they estimated the number of variants (mutants) induced per year by natural radiation. These authors concluded that "it seems, therefore not altogether extravagant to assume that such variations as actually occur in nature are due largely to the radioactivity of the environment." The involvement of Gilbert Lewis in this activity, while unexpected, was derived from his research in the 1920s in the area of radiation physics (Coffey 2008). Furthermore, his eclectic research activities had also drawn him toward evolutionary theory, the subject of his major presentation (i.e., Silliman Lecture) at Yale, just preceding the development of the LNT paper in *Nature* (Lewis 1926). This lecture followed that of Thomas Hunt Morgan of Columbia University in 1925, Muller's Ph. D. advisor and 1936 Nobel Prize recipient. The perspective of Olson and Lewis (1928) was also independently advanced by Muller in a paper read before the National Academy of Sciences on April 24, 1928, and published on September 14, 1928. The statement of Muller (1928b) was principally conceptual, lacking the detailed formulation of Olson and Lewis (1928).

The following year, Babcock and Collins (1929a, b) tested the hypothesis of Olson and Lewis (1928). They found a location in which the natural radiation was twice that found in their University of California/Berkeley laboratory. Using the ClB strain sex-linked recessive *Drosophila* assay, they reported an increase in mutation that corresponded in the same proportion as the difference in background radiation, supporting the proportionality hypothesis. Detailed experimental methods including the actual radioactivity levels were never published, although such data were promised to be provided in a subsequent paper. In 1930, Hanson and Heys provided further support for the hypothesis that "natural radiation may be responsible for the mutations that are the grist of the natural selection mill with the resulting evolution of new forms." Their findings were based on a study of fruit fly mutations in an abandoned carnotite (i.e., uranium) mine. Such interpretations were initially supported by commentaries by various authors (Lind 1929; Dixon 1929, 1930).

In 1930 Muller and Rice University physicist, Mott-Smith, challenged this LNT evolution perspective by reporting that natural radiation, which was of such a low-dose rate, could only account for about 1/1,300 of the gene mutations that occurred spontaneously in *Drosophila melanogaster*, assuming a linear dose response. The authors concluded that other causes must explain the origin of most mutations that spontaneously occur. Nonetheless, in his dissertation, under the direction of Muller, Oliver (1931) stated that cosmic and terrestrial radiations must account for some proportion of the spontaneous mutations (see Muller 1930). This conclusion was justified on the belief that the response is linear at low dose, with there being no threshold for a mutation response. This relationship was stated as holding true for all types of high-energy radiation (e.g., gamma, beta, X-rays and probably ultra-violet rays). Thus, Oliver (1931) concluded that "by inference it can be added that the cosmic and the terrestrial radiations also are capable of producing mutations in proportion to their power of ionization." Oliver (1931) also extended the concept of proportionality to chromosomal inversions and translocations further arguing for the support of a background radiation influence. For example, Muller and Altenburg (1930) noted that translocations are induced at a similar frequency as gene mutations. Given these circumstances, Oliver (1931) noted that "one would expect each of the classes of changes considered to occur with the same frequency when the individuals are subjected only to the natural conditions, if natural radiation can account for all mutations…" Despite this interpretation of environmental radiation-induced genetic changes, Oliver (1931) concluded that "some other condition must, therefore, enter in order to explain the difference in non-radiated material, between the frequency of gene mutation and that of the other type of genetic changes." (p. 34)

Even though Muller dismissed natural radiation as providing a quantifiably significant mutational influence to derive genetic novelty for evolutionary change, he still retained his belief in the linear dose–response relationship (p. 238) (Muller 1930) based on the findings of Hanson and Heys (1929, 1930) and Oliver (1930). Even though the hypothesis of Olson and Lewis (1928) did not maintain significant support for long within the scientific community, Muller and other leaders of the radiation genetics community became strong advocates of the LNT model to account for genomic mutations and the occurrence of cancer.

It may seem difficult to understand in retrospect why prominent scientific leaders such as Gilbert N. Lewis, Hermann J. Muller and others so quickly adopted a belief in linearity at low dose. In the case of Muller, he was fully committed to this view after the publication of only three studies (Hanson and Heys 1929, 1930; Oliver 1930) in which the lowest cumulative dose was roughly 285 r, administered in an acute manner, the rough approximation of 1,000 modern chest X-rays in 3.5 min or 5 chest X-rays/s.

In his rather copious publications during this period of "belief"/concept formulation, Muller never addressed contemporary publications that did not support a linear interpretation (Patterson 1928; Weinstein 1928; Stadler 1930, 1931). Yet, he was well aware that the lowest doses in the Hanson and Heys (1929, 1930) and Oliver (1930) papers were acute studies that grossly exceeded background radiation exposure. To think within a linear dose–response term

framework ran counter to pharmacological and chemical toxicological experience at that time. As Zimmer (1966) reflectively wrote, toxic chemicals in the early decades of the twentieth century demonstrated "no effect up to a threshold dose and then climbed steeply up to 100 %." Muller and others argued that the genetic response to ionizing radiation demanded a different evaluative framework.

Target theory and LNT

A likely explanation for Muller's (and possibly Gilbert N. Lewis's) acceptance of the LNT in the absence of convincing dose–response data may be found within the scientific culture at the time. X-ray-induced mutational effects were placed within the context of what was called the radiation target theory. This theory was quantitative and dosimetric, with mathematical calculations related to quantum mechanics, reflecting the leadership of prestigious theoretical physicists (von Schwerin 2010). The formation of a physics-based target theory was established prior to the discovery of inducible mutations by Muller (1927) by medical physicists such as Dessauer (1922), Glocker (1927) and Crowther (1924, 1926, 1927), setting the stage for a novel scientific framing of the mutational data in the 1930s. The mutation findings of Muller (1927) were a major scientific advance that easily fit into the target theory concept while also markedly advancing the scientific standing of target theory itself.

The radiation target theory as applied to mutations was formulated by the detailed interactions and collaborations of leading radiation geneticists and theoretical physicists during the mid-1930s. During this time, radiation geneticists, lead by Nicolai Timoféeff-Ressovsky, and physicists, including Niels Bohr, with a profound interest in the interface of physics and biology, would meet each year, typically in Copenhagen and Belgium for extensive discussions. From these exchanges developed the seminal conceptual paper by Timoféeff-Ressovsky and the physicists Max Delbruck and Kevin Gunter Zimmer (Timoféeff-Ressovsky et al. 1935) that would establish a conceptual framework for gene structure, target theory for the induction of mutations via ionizing radiation, the single-hit mechanism hypothesis to account for the shape of the LNT dose response and the application of this dose–response model for what was to become modern cancer risk assessment. The genetic target theory saw mutation as a purely physical action following an all or none law in which a single ionization or energy absorption produces the mutational effect independent of all other ionizations and energy absorptions.

This linearity feature stands in contrast to normal physiology that invariably deals with large numbers of molecules of each kind, and where the elimination of a single molecule would not result in observable effects (Delbruck 1940). The energy of ionizing radiation was assumed to be essentially transformed into a genetic effect. According to the physicist turned biologist Max Delbruck (1969 Nobel Prize recipient in Biology and Medicine), the proportionality rule that was proposed earlier by Muller, based on the research of Hansen and Heys (1929) and Oliver (1930, 1931) and supported in experimental research by Timoféeff-Ressovsky et al. (1935), provided the basis of the single-hit mechanism interpretation and the calculation of the size of the gene (Delbruck 1940). Table 1 provides a listing of quotes in which the early conceptual framing of the dose–response proportionality concept occurred. The transforming of a dose–response hypothesis based on a very limited amount of data into a biological "Rule" by Muller was done without significant discussion of the concept, its possible mechanisms as well as the recognition of data that may contradict this "Rule."

Although Muller was a geneticist, he was drawn quickly toward the physics-mutation interface, accepting significant elements of target theory for radiation-induced mutational effects, including the important assumptions that damage was proportional to the energy absorbed, linear dose–response modeling and that effects were cumulative and deleterious (Muller et al. 1936). Muller knew Timoféeff-Ressovsky, having met him in the Soviet Union in 1922, encouraging him and his colleagues to transform his laboratory to one of the *Drosophila* genetics. Muller renewed contact with Timoféeff-Ressovsky during the 5th International Congress on Genetics in 1927. From November 1932 to September 1933, Muller researched in Berlin with Timoféeff-Ressovsky. He also participated in the physics-biology/mutation discussions in Copenhagen in 1936, engaging Niels Bohr and other leading physicists. Experiments of radiation geneticists during this period were often designed within the context of this target theory framework. This was also the case for critical studies performed a decade later under the aegis of the Manhattan Project at the University of Rochester under the direction of Curt Stern (with Muller serving as a consultant) (Spencer and Stern 1948; Caspari and Stern 1948).

The hit hypothesis

As noted above, in his Nobel Prize research, Muller reported that the induction of mutations was not directly proportional to the X-ray dose, but rather to the square root of the dose (Muller 1927). Based on discussion with the physicist and future Nobel Prize winner Irving Langmuir (1932 Nobel Prize in Chemistry), Muller (1927) stated that this observation suggested that the induction of mutation was not caused directly by a single quantum of energy.

Table 1 Documentation of the introduction of the proportionality rule concept into the mutation literature, 1929–1960

References	Quote
Hanson and Heys (1929)	"It is only to be expected that the number of mutations be directly *proportional* to the number of rays to which the organisms are exposed." Page 207
Muller (1930)	"Since then Hanson, using radium, and Oliver in our laboratories using X-rays, have both found that the frequency of mutations produced is exactly *proportional* to the energy of the dosage absorbed… There is, then, no trace of a critical or threshold dosage beneath which the treatment is too dilute to work." Page 236
Oliver (1930)	"That is there is a direct *proportionality* between the percent of lethals and the length of time of treatment may be seen more readily by a comparison of the t_1 values calculated from the results for each of the given doses." Page 45
Stadler (1930)	"Mutation frequency increased approximately in direct *proportion* to dosage." Page 13
Hanson et al. (1931)	"Taking the amount of ionization in air as a measure, the mutation rate seems to vary approximately in direct *proportion* to the intensity." Page 142
Oliver (1931)	"By inference it can be added that the cosmic and the terrestrial radiations of higher energy content also are capable of producing mutations in *proportion* to their power of ionization." Page 480
Oliver (1931)	"The relation of *proportionality* to the dosage applies not merely to the lethals in general, but, more specifically, to the lethal gene mutations." Page 485
Oliver (1931)	"…[gene mutations and gene rearrangements] all probably occur in direct *proportion* to the dosage, no matter how small a dose is used." Page 486
Patterson (1931)	"In general their results [i.e., Hanson and Heys 1928 and Oliver 1930] justify the conclusion that the rate is directly *proportional* to the dosage employed." Page 133
Hanson and Heys (1932)	"Further evidence of the *proportionality rule* from a study of the effects of equivalent doses differently applied." Page 335
Hanson and Heys (1932)	"Experiments planned with a view to determining within what limits the *proportionality rule* holds show again a strict correspondence existing between the amount of radium administered and the consequent biological effect, the induced mutation frequency obtained varying directly with the dosage." Page 343
Hanson (1933)	"The rate seems to be directly *proportional* to the dosage. Muller has named this the '*proportionality rule.*' For example, when all other factors are kept constant, doubling the time of exposure also doubles the number of lethal mutations." Page 486
Oliver (1934)	"The frequency of induced mutations is directly *proportional* to the intensity of the treatment." Page 391
Delbruck (1940)	"The *proportionality rule* gave the basis for the single-hit interpretation…" Page 359
Stern (1950)	"The *proportionality rule* has been proven to hold over a wide range. Figure 155 shows that, for Drosophila, the relation is essentially linear over the range from 25 r to several thousand r. It has further been shown that the frequency of induced mutations is independent of the time over which the radiation is applied." Page 433
Stern (1960)	"It has been established for a variety of experimental organisms that the number of mutations induced by radiation is proportional to the dose. This *proportionality* has been proven to hold over a wide range of dosages." Page 491

However, subsequent exposure experiments by Hanson and Heys (1929), Oliver (1930, 1931) and later by Timoféeff-Ressovsky et al. (1935), even though all experiments were at very high dose, supported a proportionality relationship, which was consistent with the "hit" theory of mutation in which the X-ray treatment excites an electron in the target gene. This excitation was proposed to affect a permanent change or mutation to a different molecular structure. Ionizing irradiation was the only effective way to induce mutations; it showed no threshold, suggesting that the absorption of radiation is a quantized and additive process (von Schwerin 2010). A "quantum-jump" was considered to be the physical process caused by a hit on a target, resulting in mutation. Treatment effects induced by a physical agent like ionizing radiation were believed to be caused by one or several discrete biophysical events, that is, hits on a target. Based on hypotheses about what constituted a hit, statistical models were used to construct dose–response relationships. If there was only a single hit on a single target, the dose response was linear. As the number of assumed hits increased, a more threshold like the dose response would appear. In a practical sense, the mathematical model-derived dose response based on an assumed number of hits could be visually matched against the laboratory-obtained dose–response curve. Using this direct and simplified approach, researchers like Muller, Timoféeff-Ressovsky and participating physicists decided the theoretical number of hits. This type of target theory was especially strong in Germany, with support from leaders such as Boris Rajewsky (Director of the KWI for biophysics, 1936), Timoféeff-Ressovsky and others (von Schwerin 2010). This conceptual framework led to the conclusion that mutation was a

single-hit process, proceeding from a single ionization, from a quantum of ionizing radiation in a specific sensitive zone of the gene.

This theoretically based perspective became not only a workable model but a firm belief within the radiation genetics community even though there was no knowledge of the physical nature of the gene. As coauthor of the Timoféeff-Ressovsky et al. (1935) paper, Delbruck subsequently noted in his Nobel Prize lecture that it was thought that genes were very stable and, therefore, showed characteristics of molecules. However, the gene concept at that time was simply that of Mendelian algebraic rates, lacking structural chemistry insight. There was much speculation of gene structure including that of submicroscopic steady-state systems or even an entity not readily analyzable in chemistry as proposed by Bohr (1933).

The paper of Timoféeff-Ressovsky et al. (1935), as noted above, was striking in its collaboration between physics and genetics, its proposed chemical nature of the gene, size of the gene and in the proposal of a "hit" hypothesis as the foundation of the linear dose response for ionizing radiation-induced mutation. While the gene structure and size framework would be bypassed and replaced by the DNA structure of Watson and Crick (1953), the hit theory component of Timoféeff-Ressovsky et al. (1935) was accepted and implemented by the radiation genetics community. The term "hit hypothesis" became commonly used in the lexicon of radiation genetics, including those comprising the BEAR I Committee/Genetics Panel that recommended changing to a linear model from a threshold model for assessing mutation risks from ionizing radiation (Calabrese 2013).

The impact of this 1935 article was facilitated by the actions of Timoféeff-Ressovsky who sent reprints to key researchers. However, the overall immediate impact of the paper was very limited as it was published in an obscure Gottingen journal that was not cited in any leading index with only four issues being printed before ceasing publication. This paper, which provides the origin of the single-hit hypothesis to support a linear dose–response model, was not even cited in the BEAR I report that implemented the concept. Yet, the term "hit" hypothesis and target theory became commonly used, even if credit was not often given to the original paper (Timoféeff-Ressovsky et al. 1935). Nonetheless, this paper did receive a major endorsement in the 1944 book *"What is Life"* by Erwin Schrodinger, a Nobel Prize physicist (1933), raising its visibility in the physics community.

The concept of the gene and its striking stability suggested it must have a unique atomic composition. Delbruck (1970) believed that such stability might be due to each atom of a gene being fixed in its mean position and electron-stable, sunk in an energy well, now seen having stability due to the function of the hydrogen bond. Mutations of such genes could only occur following the absorption of high energies as from ionizing radiation, not from heat under physiological conditions. In fact, a modest increase in vibrational energy was estimated to increase the atomic stability, decreasing mutational risk. Since a transaction in an atom can be affected by a single digit eV and that the initial impact of an X-ray can be several fold greater, it was believed that any gene would be at risk for mutation from radiation. Since the initial energy of impact exceeds a threshold energy of activation, ionizing-radiation should affect not only the induction of a localized mutation but also that of a broad range of gene targets.

The mutation hit theory was challenged by Caspari and Stern (1948) in a chronic, very low-dose rate study, leading to the hypothesis that either a threshold exists or multiple independent primary actions are required for a mutation to occur, or that a recovery or repair effect/process occurred at a very low-dose rate (Howarth et al. 1950; Key 1951). Over the next several decades, the dominance of the physics-based target theory would yield to improved chemical/biological/physiological understandings of the mutation process, including such modified target theory effects of ionizing radiation as DNA repair (in reproductive and somatic cells), adaptive response, the bystander effect as well as the recognition that the biological effects of ionizing radiation are principally due to the generation of hydroxyl radicals/hydrated electrons from cellular water and their migration to cellular targets (Collinson et al. 1962; Czapski and Schwartz 1962; Weiss 1944). In fact, even as the target theory was being applied to mutation by Timoféeff-Ressovsky et al. (1935), the recognition of repair processes, including DNA repair, were emerging (Hanawalt 1994). Such challenges to the hit theory would eventually be brought to the BEAR Committee by Russell (1956, 1963) from Oak Ridge, but only after the BEAR 1 Committee made its linearity recommendation.

Edward Lewis (1957a), another radiation geneticist Nobel Prize (1995) recipient, published a very influential *Science* article in 1957, strongly supporting a linear relationship for cancer, relying on linearity data in the Uphoff and Stern (1949) paper. In subsequent Congressional Testimony, Lewis (1957b) would argue that the dose response was linear, regardless of the mechanism, and should be accepted as such whether or not a mechanism could even be discerned. These comments of Lewis suggested that he recognized the growing mechanistic challenge to the single-hit theory as well as new conceptual problems (e.g., multiple biological processes could yield a linear relationship that did not require a single-hit process) emerging from the physics and genetics communities, including Zimmer (1941), a coauthor of the Timoféeff-Ressovsky et al. (1935) paper and radiation biologists/geneticists (Haas et al. 1950;

Kimball 1952). However, the time period within which Muller's mutation findings were produced was one of the cultural scientific dominance of physics. Association with the leadership of the physics community served to enhance the significance of the mutational findings and its assumed linearity at low dose, as well as providing Muller with an expanded scientific and cultural context that recognized his achievements and enhanced his scientific reputation.

The influence of the hit concept of Timoféeff-Ressovsky et al. (1935) was facilitated via subsequent publications of Lea (1940, 1946), which offered further justification for the target theory-based LNT-single-hit hypothesis for mutation. The publications of Lea were not only authoritative extensions of Timoféeff-Ressovsky et al. (1935) but more readily available than the Timoféef-Ressovsky et al. (1935) paper with its publication in a defunct journal.

Regulatory agency actions

Ionizing radiation

In the radiation risk assessment area, two endpoints were adopted to which linearity was applied: germ cell mutations and cancer. In the case of germ cell mutations, based on several publications in the early 1950s by Muller (1951, 1954), the BEAR I Genetics Panel (1956) proposed to limit exposure to ionizing radiation such that exposure would not exceed doubling of background mutations from conception through the first 30 years of life. The panel assumed that exposure to ionizing radiation could cause mutations to germ cells in a linear manner and had the potential to cause adverse genetic effects in individuals and future generations. The panel derived a risk assessment methodology for application to both first-generation offspring and total genetic risk, including future generations. The panel derived a doubling dose method (i.e., the dose of ionizing radiation, assuming linearity at low dose, that would equal the number of mutations resulting from background exposure), to estimate population-based risks. This doubling dose methodology would predict the number of genetic diseases based on three parameters: the assumed doubling dose, the proposed exposure limit and the background incidence of genetic disease. Based on this risk assessment framework, the panel recommended a "uniform national standard" such that the members of the general population would not receive more than a cumulative dose of 10R from conception through 30 years. This basic method of the BEAR I Committee, using the doubling dose/linear framework, has been refined with recent advances allowing one to integrate between rates of radiation-induced mutation based on mouse studies and the risk of inducible genetic disease in people [Sankaranarayanan and Chakraborty 2000a, b; Sankaranarayanan and Wassom 2008 (see Lyon 2003 for an alternative view)].

In the case of somatic effects, cancer risks were estimated via the use of a linear dose–response model. Assuming linearity to zero, it was estimated that exposure of one rem to one million people each year would cause one to two new cases of leukemia on an annual basis for first decade of life (ICRP 1962; Sowby 1965; UNSCEAR 1962, 1964). As with chemical carcinogenesis risk assessment, therefore, the foundations of the LNT modeling for ionizing radiation-increased cancer risks are directly traced back to Lea, Timoféeff-Ressovsky et al. and ultimately to Muller's proportionality rule.

Chemical carcinogens

Five years after the publication of the BEAR 1 report, Mantel and Bryan (1961) published their influential paper entitled "Safety' Testing of Carcinogenic Agents" based on the probit dose–response model in order to estimate tumor incidence for carcinogens. Biostatistical estimates of cancer risks were first provided by Bryan and Shimkin (1943) when they applied the probit model to estimate the cancer risk of three carcinogenic hydrocarbons (i.e., 20-methylcholanthrene; 1,2,5,6-dibenzanthracene; 3,4-benzpyrene) in strain C_3H male mice.

The motivation for Mantel and Bryan to develop the biostatistical model for predicting carcinogen risk was due to the fact that Mantel, a biostatistician at the US National Cancer Institute (NCI), was asked by the Director of the NCI to develop guidelines for the number of laboratory animals that would be needed to establish the safety of a test agent within the context of a hazard assessment. This response followed a request, after the Thanksgiving cranberry scare of 1959, by the Secretary of the Department of Health, Education and Welfare (HEW) to the NCI. The cranberry scare was a public relations nightmare in which trace residues of a cancer-causing herbicide [i.e., amitrole (3-amino-1,2,4-triazole)] were detected in some sources of cranberries just before the holiday. The secretary of HEW recommended against buying cranberries that year, leading to a consumer panic that threatened the industry. In order to avoid such situations in the future, the secretary of HEW requested the NCI to provide guidance on which cancer-causing substances were "safe" and at what dosage levels.

Mantel and Bryan (1961) noted the generality of their modeling approach and proposed the concept of a virtually safe dose with an estimated risk of 1/100 million. Some 12 years later, the FDA would propose the use of the Mantel-Bryan (1961) model and recommend the 1/100 million safety guide in their July 19, 1973 risk assessment proposal in the Federal Register. When the rule was finalized in

1977, the Mantel-Bryan probit model was retained but with several modifications and with the acceptable (de minimus) risk being reduced to 1/million. This value was considered as the level below which no additional regulatory action would be taken within the context of the safety of animal carcinogen residues. The finalized Mantel-Bryan model of the FDA was the first quantitative risk assessment model approved by a regulatory agency. Two years later, the FDA (1979) significantly revised the cancer risk assessment policy, replacing the modified Mantel-Bryan model with a linear dose–response model based on multiple factors, including its more conservative risk estimation and ease of calculations (Anonymous, 1979). In the low-dose zone, the one-hit model discussed above is closely approximated by a simple linear model.

The US EPA strategy for assessment and regulation of carcinogens displayed a profound evolution during the 1970s. Based on expert testimony during pesticide hearings, EPA attorneys developed a legal brief that embodied "cancer principles" (NAS 1983). These "principles" suggested that carcinogen exposures should be prevented. As the concept of "banning" carcinogenic agents was soon seen as unrealistic, EPA quickly adopted non-regulatory guidelines for a general risk assessment process (EPA 1976). This process advocated the use of quantitative risk assessment as a means to differentiate risks among chemicals and engineering processes. The guidance was very general, being limited to less than a page within the Federal Register. These guidelines were followed by a paper from the EPA Carcinogen Assessment Group (CAG) (Albert et al. 1977), which provided a strong endorsement of the LNT concept, arguing that linearity was supported by human epidemiological studies (e.g., ionizing radiation and cigarette smoking related lung cancer) and mutagenicity studies that were also claimed to follow a linear dose response and believed to be the underlying mechanisms of carcinogenesis. In a March 15, 1979, Federal Register, the EPA Administrator Douglas Castle stated that "Risk assessment from animal data is performed using the 'one-hit' model" based on the 1976 Interim Guidelines (EPA 1976). He went on to state that "the one-hit model was endorsed by the four agencies in the Interagency Regulatory Liaison Group" based on its highly conservative nature and the uncertainties in extrapolating from animal data to human responses and the possibility that humans may be more susceptible than the animal model, because of broad human interindividual variability in exposures and "other unknown factors". The strongly clarifying and underlying statement of the administrator was due in part to the fact that EPA had used other cancer risk assessment models under other regulatory acts and by other US federal agencies.

According to Albert (1994), Chair of the EPA Cancer Assessment Group (CAG) during the 1970s, the EPA adopted the linear no threshold model (LNT) of the Atomic Energy Commission (AEC) that had been applied to estimating risks from fallout from atomic weapon tests. The LNT model was attractive to EPA since it was very simple to apply; all that was needed in a toxicological sense was to identify the lowest dose of agent that induced a statistically significant response and draw a straight line to the origin of the graph for the dose versus cancer incidence. Its biological plausibility was based on the linearity of mutation dose response within the framework of target theory. He noted that "any difference between chemical carcinogens and ionizing radiation could be waived aside as they both cause genetic damage…"

Statisticians would argue that the straight line extrapolation to zero from the lowest statistically significant response ignored data at the high doses. Thus, during a meeting of leading statisticians called by the CAG, a decision was made to change from the single-hit model to the multi-stage model since it used all the data, while retaining linearity at low dose and being compatible with the concept of cancer being a multi-stage process. Consistent with this assessment, the NAS Safe Drinking Water Committee (1977) recommended the adoption of LNT modeling for risk assessment using a multi-stage model. However, in 1982, the Safe Drinking Water Committee (SDWC) was skeptical about LNT modeling for chemicals and rescinded its endorsement of the LNT model noting "…more confidence could be placed in mathematical models for extrapolation if they incorporated biological characteristics of the animal studies… since the users of this volume will be likely to favor different varieties of the conventional extrapolation models or will have access to some of the newer developmental methodologies, it is premature at this stage to recommend any single approach by selecting it for calculations…" (p 8). However, since LNT modeling was already in use by EPA, in 1983, the SDWC again endorsed the LNT model and its subsequent use became the default methodology for chemical cancer risk assessment. According to Albert (1994), none of the possible models (single hit, multi-hit, logit, probit, multi-stage, others) were biologically credible. The agency simply needed one that would be acceptable. The agency applied LNT risk assessment methods using the multi-stage model for the regulation of trihalomethanes in drinking water in a November 29, 1979, notice in the Federal Register (EPA Environmental Protection Agency (US EPA) 1979a, b), a process that would be followed in subsequent EPA cancer risk assessments.

The parallel, yet converging linear dose–response strategies of the EPA and FDA represent the regulatory origin of current cancer risk assessment practices throughout the world. They are directly traced back to the efforts of Lea (1946) and Timoféeff-Ressovsky et al. (1935), all of which stemmed from the "Proportionality Rule" of Muller (1930).

Eugenics

While the LNT concept for mutation was born within the intellectual and scientific framework of the physics-based radiation target theory, its applications also found supportive resonance within the philosophical, ideological and political frameworks of eugenics. German eugenicists expressed considerable concern that ionizing radiation may hurt the German germ plasm (Proctor 1999; Martius 1931). Educational programs based on these concerns cautioned against exposures to ionizing radiation that might adversely affect future generations of Germans. Recommendations as early as 1927 by the Bavarian Society for Pediatrics and Gynecology stated that women receiving excess X-rays during pregnancy should abort their fetuses. Pushing this concept even further, in 1930, Eugene Fisher, director of the Kaiser Wilhelm Institute for Anthropology, argued that women exposed to X-rays should be permanently prevented from having children (Proctor 1999). Muller's own history is replete with his highly visible association with national and international activities advancing eugenics philosophy and agenda. Even as late as 1955, Muller gave a strong eugenics advocacy presentation in Germany, testing such ideas with a large audience of Nobel Prize winners (The Lindau Mediatheque 1955).

The biophysical concept of the gene had important eugenics implications. Since mutations could be induced by ionizing radiation in a linear at low-dose manner, this concept provided the principal foundation that all ionizing radiation—whether via medical diagnosis/treatment or industrially—was a concern for "genetic health". The genetic toxicology studies of Timoféeff-Ressovsky et al. (1935) transformed these above-cited radiation health concerns, providing biophysical models and the LNT-single-hit model risk assessment paradigm. Such actions provided a key vehicle by which eugenics would focus on radiation protection for preventing the occurrence of genetic defects. In fact, the development and activities of the genetics department of the Kaiser Willheim Institute under the direction of Timoféeff-Ressovsky was affected by such perspectives (Gausemeier 2010).

The concept of LNT for ionizing radiation-induced mutation was, therefore, built upon a scientific/cultural framework and applied to a range of health-related policies, especially those of eugenics during the early decades after the discovery of X-ray-induced mutations. In fact, the eugenics area would serve as an intellectual training ground for how ideas such as LNT could be "softened", humanized and successfully integrated within a post-World War II society. Some aspects of eugenics advocacy and the LNT concept would morph into modern regulatory policy for carcinogen regulation, evolving from that of preserving the gene pool of certain racial subgroups or other targeted populations to a humanistic framework that would reduce mutational risks to entire populations.

Evolution and endogenous mutations

The LNT had its start in an attempt to explain evolution, finding other outlets in the world of eugenics and later public health regulatory policies. While Muller was a leader in these activities, he did not abandon his quest to determine those underlying factors that served to provide the novel mutations for natural selection. In fact, prior to his discovery of X-ray-induced mutations in 1927, Muller reported that temperature increases enhanced the mutation rate by about two-fold (Muller 1928c). However, the temperature hypothesis was placed on the research back burner when high doses of X-rays were found to markedly enhance mutation frequency. Muller would return to the temperature–evolution hypothesis some three decades later, completing an intellectual and professional circle, reflected in the comments of Plough and Ives (1934), his former colleagues at Amherst College (1940–1945) who noted that "since Muller and Mott-Smith conclude that natural radiation is inadequate to account for mutations in nature, it seems possible to suggest that ubiquitous temperature variations may play that role". If Muller had lived into the decades of the 1980s (he died in the 1967), he would have begun to appreciate the so-called other conditions suggested by Oliver (1931) as the cause of the overwhelming proportion of spontaneously occurring mutations is now believed to be derived from endogenous metabolism, for which complex and integrative DNA repair processes have been selected for via natural selection (De Bont and van Larebeke 2004; Lindahl 1996).

Summary

The LNT concept was initially proposed to account for evolutionary change and then later applied for the assessment of risks for some genetic diseases and cancer incidence (Table 2). The initial data upon which the LNT concept was based were limited to a few studies of an acute nature and at very high doses. Within a decade, the LNT dose–response model was provided with a mechanistic foundation via the integration of the single-hit concept within target theory. The LNT-single-hit model was then used by radiation geneticists to frame the intellectual debate on low-dose ionizing radiation risk to the human genome. It provided the basis for the recommendations of the US NAS BEAR I Committee in 1956 for

Table 2 LNT history: the temporal sequence leading to the LNT dose–response model for cancer risk assessment

References	Specific temporal events
Muller (1927)	Mutation findings—X-rays induce mutations in fruit flies ⇓
Olson and Lewis (1928)	LNT model proposed to account for evolutionary changes following Muller's discovery that X-rays can induce mutations in fruit fly germ cells ⇓
Muller (1930)	Develops proportionality rule (i.e., linear dose response) for ionizing radiation-induced mutagenicity ⇓
Timoféeff-Ressovsky et al. (1935)	Application of radiation target theory for mutagens. Used target theory to propose a hit theory for ionizing radiation-induced mutation. The hit mechanism was used to explain the LNT dose response ⇓
BEAR I 1956 (Biological Effects of Atomic Radiation Committee, Genetics Panel)	Proposes the use of the linear dose–response model for germ cell mutation, using the "doubling rule" ⇓
Mantel and Bryan (1961)	Develops carcinogen risk assessment model based on the probit model. This activity was undertaken to advise US governmental agencies on chemical risk assessment ⇓
FDA (1973)	Proposes a probit-based quantitative risk assessment method for cancer risk based on the Mantel and Bryan 1961 paper. The proposal stated that an acceptable risk was 1/100 million ⇓
EPA (1976) (see Albert et al. (1977), Anonymous (1979)	Proposed guidelines for carcinogen risk assessment based on quantitative risk assessment. Recommended a linear dose–response model ⇓
FDA (1977)	FDA rule finalized, retaining the Mantel-Bryan model with some modifications. The acceptable risk value was changed to 1/1 million (10^{-6}) ⇓
U.S. NAS Safe Drinking Water Committee (1977)	Recommended that EPA adopt LNT for carcinogen risk assessment. This recommendation was profoundly significant given the widespread multimedia regulatory functions of EPA. Within 2 years of the recommendation, EPA applied the LNT to the regulations of trihalomethanes (e.g., chloroform) in drinking water ⇓
FDA (1979)	Replaced the modified Mantel-Bryan model with the LNT model for carcinogen risk assessment, based on the following reasons: 1. Linear procedure is least likely to underestimate risk. 2. Linear extrapolation does not require complicated mathematical procedures. 3. No arbitrary slope is needed to carry out linear extrapolation. 4. Several significant limitations were found with the application of the Mantel-Bryan model (Anonymous 1979) ⇓
EPA (1979a, b)	EPA established a national drinking water standard for trihalomethanes (including chloroform) based on an LNT methodology as recommended by the US NAS Safe Drinking Water Committee (1977)

the switch from a threshold to a linear dose–response model for estimating ionizing radiation-induced germ cell mutation using the doubling dose concept. The LNT-single-hit model was soon generalized to the process of cancer risk assessment and adopted by national and international committees concerned with ionizing radiation by the late 1950s and early 1960s. Five years later, Mantel and Bryan (1961), researchers at the US National Cancer Institute, proposed a probit model-based cancer risk assessment method. It was the Mantel and Bryan (1961) model that was proposed by the FDA in 1973 for cancer risk assessment procedures, being replaced with a LNT model by the FDA in 1979, the same year that EPA applied the LNT for the regulation of carcinogens (i.e., trihalomethanes) in drinking water. The LNT model and its single-hit explanation/mechanism theory, therefore, can be traced back to the concept of radiation-induced mutation target theory as proposed by Timoféeff-Ressovsky et al. (1935), which was founded on the proportionality rule of Muller (1930) which itself had its origins in the 1928 paper of Olson and Gilbert that created the LNT concept following the seminal findings of Muller (1927) that ionizing radiation could induce mutation in the germ cells of fruit flies.

Acknowledgments The research on the topic of hormesis has been supported by awards from the US Air Force and ExxonMobil Foundation over a number of years. Sponsors had no involvement in study design, collection, analysis, interpretation, writing and decision to submit.

References

Albert RE (1994) Carcinogen risk assessment in the U.S. Environmental Protection Agency. Crit Rev Toxicol 24(1):75–85

Albert E, Train E, Anderson E (1977) Rationale developed by the Environmental Protection Agency for the assessment of carcinogenic risks. J Nat Cancer Inst 58:1537–1541

Anonymous (1979) Scientific bases for identification of potential carcinogens and estimation of risks. J Nat Cancer Inst 63(1):241–268

Babcock EB, Collins JL (1929a) Does natural ionizing radiation control rate of mutation. Proc Nat Acad Sci 15(8):623–628

Babcock EB, Collins JL (1929b) Natural ionising radiation and rate of mutation. Nature 124(3119):227–228

Bohr N (1933) Light and life. Nature 131(421–423):457–459

Bryan WR, Shimkin MB (1943) Quantitative analysis of dose-response data obtained with three carcinogenic hydrocarbons in strain C3H male mice. J Nat Cancer Inst 3(5):503–531

Calabrese EJ (2005) Historical blunders: how toxicology got the dose-response relationship half right. Cell Mol Biol 51:643–654

Calabrese EJ (2009a) Getting the dose response wrong. Why hormesis became marginalized and the threshold model accepted. Arch Toxicol 83:227–247

Calabrese EJ (2009b) The road to linearity: why linearity at low doses became the basis for carcinogen risk assessment. Arch Toxicol 83:203–225

Calabrese EJ (2011) Toxicology rewrites its history and rethinks its future: giving equal focus to both harmful and beneficial effects. Environ Toxicol Chem 30(12):2658–2673

Calabrese EJ (2013) How the US National Academy of Sciences misled the world community on cancer risk assessment: new findings challenge historical fundations of the linear dose response. Arch Toxicol (in press)

Caspari E, Stern C (1948) The influence of chronic irradiation with gamma rays at low dosages on the mutation rate in *Drosophila melanogaster*. Genetics 33:75–95

Coffey P (2008) Cathedrals of Science: The Personalities and Rivalries that Made Modern Science. Oxford University Press, Oxford

Collinson E, Dainton FS, Smith DR, Tazuke S (1962) Evidence for unit negative charge on hydrogen atom formed by action of ionising radiation on aqueous systems. Proc Chem Soc 140–144

Crowther JA (1924) Some consideration relative to the action of x-rays on tissue cells. Proc Roy Soc Sect B 96:207–211

Crowther JA (1926) The action of X-rays on Colpidium colpoda. Proc Roy Soc Ser B 100(704):390–404

Crowther JA (1927) A theory of the action of X-rays on living cells. Proc Camb Philos Soc 23:284–287

Czapski G, Schwarz HA (1962) The nature of the reducing radical in water radiolysis. J Phys Chem 66:471–479

De Bont R, van Larebeke N (2004) Endogenous DNA damage in humans: a review of quantitative data. Mutagenesis 19(3):169–185

Delbruck M (1940) Radiation and the hereditary mechanism. Am Nat 74(753):350–362

Delbruck M (1970) A physicist's renewed look at biology: twenty years later. Science 168(3937):1312–1315

Dessauer F (1922) Uber einige Wirkungen von Strahlen I. Z. Physik 12:38

Dixon HH (1929) Cosmic radiations and evolution. Nature 123(3113): 981

Dixon HH (1930) The mechanism of variation. Nature 125(3165): 992–993

Environmental Protection Agency (US EPA) (1976) Health risk and economic impact assessments of suspected carcinogens. Interim Procedures & Guidelines. [FRL 548-2] Fed Reg 41:21402–21405

Environmental Protection Agency (US EPA) (1979) Water quality criteria. Request for Comments. Notice. [FRL 1062-5] Fed Reg 44(52):15926–15931

Environmental Protection Agency (US EPA) (1979) Control of trihalomethanes in drinking water. National Interim Primary Drinking Water Regulations. [FRL 1312-2] Fed Reg 44(231):68624–68710

FDA (US Food and Drug Administration) (1973) Compounds used in food-producing animals. Procedures for determining acceptability of assay methods used for assuring the absence of residues in edible products of such animals. Proposed rule. Fed Reg 19:19226–19230

FDA (U.S. Food and Drug Administration) (1977) Chapter I—Food and Drug Administration, Department of Health, Education, and Welfare. Subchapter A-General. Subchapter E-Animal Drugs, Feeds, and Related Products. Chemical compounds in food producing animals. Criteria and procedures for evaluating assays for carcinogenic residues in edible products of animals [Docket No. 77 N-0026] Fed Reg 42(35):10412–10437

FDA (U.S. Food and Drug Administration) (1979) Department of Health, Education, and Welfare. Chemical Compounds in Food Producing Animals. Criteria and Procedures for Evaluating Assays for Carcinogenic Residues [Docket No. 77 N-0026] Fed Reg 44(55):17070–17114

Gausemeier B (2010) Genetics as a modernization program: biological research at the Kaiser Wilhelm Institutes and the political economy of the Nazi state. Hist Stud Nat Sci 40(4):429–456

Glocker R (1927) Das Grundgesetz der physikalischen Wirkung von Röntgenstrahlen verschiedener Wellenlänge und seine Beziehung zum biologisched Effekt. Strahlentherapie 26:147–155

Goodspeed TH, Olson AR (1928) The production of variation in nicotiane species by x-ray treatment of sex cells. Proc Nat Acad Sci 14:66–69

Haas Fl, Clark JB, Wyss O, Stone WS (1950) Mutations and mutagenic agents in bacteria. Am Nat 84(817):261–274

Hamblin JD (2007) A dispassionate and objective effort: negotiating the first study on the biological effects of atomic radiation. J Hist Biol 40:147–177

Hanawalt PC (1994) Evolution of concepts in DNA repair. Environ Mol Mut 23(Suppl 24):78–85

Hanson FB (1933) Radiation-genetics. Phys Rev 13(4):466–496

Hanson FB, Heys F (1928) Effects of radium in producing lethal mutations in *Drosophila melanogaster*. Science 68(1753):115–116

Hanson FB, Heys F (1929) An analysis of the effects of the different rays of radium in producing lethal mutations in Drosophila. Am Nat 63(686):201–213

Hanson FB, Heys F (1930) A possible relation between natural (earth) radiation and gene mutations. Science 71(1828):43–44

Hanson FB, Heys F (1932) Radium and lethal mutations in Drosophila. Further evidence of the proportionality rule from a study of the effects of equivalent doses differently applied. Am Nat 66(705):335–345

Hanson FB, Heys F, Stanton E (1931) The effects of increasing x-ray voltages on the production of lethal mutations in *Drosophila melanogaster*. Amer Nat 65(697):134–143

Howarth JL, Miller H, Walter J (1950) Some measurements of gamma-ray doses received by a radiotherapist during radium operations. Br J Radiol 23(268):245–255

International Commission on Radiological Protection (ICRP) (1962) Radiation protection; Recommendations of the Commission (as amended 1959 and revised 1962). Pergamon Press, Inc., New York

Key JM (1951) Neutron and X-ray experiments in barley. Hereditas 37(3):421–464

Kimball RF (1952) Genetic effects of radiation. Ann Rev Nucl Sci 1:479–494

Lea DE (1940) A radiation method for determining the number of genes in the X chromosome of Drosophila. J Gen 39(2):181–188

Lea DE (1946) Actions of radiations on living cells. Cambridge University Press, Cambridge

Lewis GN (1926) The anatomy of science. Silliman Lectures. Yale University Press, New Haven

Lewis EB (1957a) Leukemia and ionizing radiation. Science 125(3255):965–972

Lewis EB (1957b) Testimony: Statement of Dr. Edward Lewis, California Institute of Technology. Hearings before the Special Subcommittee on Radiation of the Joint Committee on Atomic Energy, Congress of the United States. 85th Congress, 1st session, Part 1. United States Government Printing Office, Washington

Lind SC (1929) The chemical effects of radium radiation. Am J Roentgen Rad Ther 21:480–483

Lindahl T (1996) The Croonian lecture, 1996: endogenous damage to DNA. Phil Trans Biol Sci 351(1347):1529–1538

Lyon MF (2003) James Neel and the doubling dose concept. Mut Res 543:115–120

Mantel N, Bryan WR (1961) Safety testing of carcinogenic agents. J Nat Cancer Inst 27(2):455–470

Martius H (1931) Keimschadigung durch Rontgenstrahlen. Strahlentherapie 41:47–66

Muller HJ (1922) Variation due to change in the individual gene. Am Nat 56:32–50

Muller HJ (1927) Artificial transmutation of the gene. Science 66(1699):84–87

Muller HJ (1928a) The problem of genic modification. *Supplementband l der Zeitschrift fur Induktive Abstammungs und Vererbungslehre* Manuscript Department, Lilly Library. Indiana University, Bloomington, pp 234–260

Muller HJ (1928b) The production of mutations by x-rays. Proc Nat Acad Sci 14(9):714–726

Muller HJ (1928c) The measurement of gene mutation rate in Drosophila, its high variability, and its dependence upon temperature. Genetics 13:279–357

Muller HJ (1930) Radiation and genetics. Am Nat 64(692):220–251

Muller HJ (1951) Radiation damage to the genetic material. In: Baitsell GA (ed) Science in progress, vol 7. Yale University Press, New Haven, pp 93–177

Muller HJ (1954) The manner of dependence on the permissible dose of radiation on the amount of genetic damage. Acta Radiol 41:5–20

Muller HJ, Altenburg E (1930) The frequency of translocations produced by x-rays in Drosophila. Genetics 15:283–331

Muller HJ, Mott-Smith LM (1930) Evidence that natural radioactivity is inadequate to explain the frequency of "natural" mutations. Proc Nat Acad Sci 16:277–285

Muller HJ, Timoféeff-Ressovsky N, Delbruck M, Bohr N (1936) Summary of discussions on mutations. Copenhagen 28–29. Manuscripts Department-Muller manuscripts, The Lilly Library, Indiana University, Bloomington

National Academy of Sciences (1983) Risk Assessment in the Federal Government. Managing the process. NAS Press, Washington

National Academy of Sciences (NAS)/National Research Council (NRC) (1956) The biological effects of atomic radiation (BEAR): a report to the public. NAS/NRC, Washington DC

National Academy of Sciences Safe Drinking Water Committee (NAS SDWC) (1977) Drinking water and health, vol 1. National Academy of Sciences, Washington

National Academy of Sciences Safe Drinking Water Committee (NAS SDWC) (1982) Drinking water and health, vol 2. National Academy of Sciences, Washington

Oliver CP (1930) The effect of varying the duration of x-ray treatment upon the frequency of mutation. Science 71:44–46

Oliver CP (1931) An analysis of the effect of varying the duration of x-ray treatment upon the frequency of mutations. Ph.D. Thesis. University of Texas, Austin

Oliver CP (1934) Radiation genetics. Quart Rev Biol 9(4):381–408

Olson AR, Lewis GN (1928) Natural reactivity and the origin of species. Nature 121(3052):673–674

Patterson JT (1928) The effects of x-rays in producing mutations in the somatic cells of Drosophila. Science 68:41–43

Patterson JT (1931) Continuous versus interrupted irradiation and the rate of mutation in Drosophila. Biol Bull 61(2):133–138

Patterson JT (1933) The mechanism of mosaic formation in Drosophila. Genetics 18(1):0032–0052

Plough HH, Ives PT (1934) Heat induced mutations in Drosophila. Proc Nat Acad Sci 20:268–273

Proctor RN (1999) The Nazi War on cancer. Princeton University Press, Princeton

Russell WL (1956) Comparison of x-ray-induced mutation rates in *Drosophila* and mice. Am Nat 90(suppl):69–80

Russell WL (1963) Genetic hazards of radiation. Proc Am Phil Soc 107(1):11–17

Sankaranarayanan K, Chakraborty R (2000a) Ionizing radiation and genetic risks XI. The doubling dose estimates from the mid-1950s to the present and the conceptual change to the use of human data on spontaneous mutation rates and mouse data on induced mutation rates for doubling dose calculations. Mut Res 453:107–127

Sankaranarayanan K, Chakraborty R (2000b) Ionizing radiation and genetic risks. XII. The concept of "potential recoverability correction factor" (PRCF) and its use for predicting the risk of radiation-inducible genetic disease in human live births. Mut Res 453:129–179

Sankaranarayanan K, Wassom JS (2008) Reflections on the impact of advances in the assessment of genetic risks of exposure to ionizing radiation on international radiation protection recommendations between the mid-1950s and the present. Mut Res 658:1–27

Sowby FD (1965) Radiation protection in Canada. Part VII. Setting standards for radiation protection. Can Med Assoc J 92:505–507

Spencer WP, Stern C (1948) Experiments to test the validity of the linear R-dose/mutation at low dosage. Genetics 33:43–74

Stadler LJ (1930) Some genetic effects of x-rays in plants. J Heredity 21:3–19

Stadler LJ (1931) Chromosome number and the mutation rule in avena and triticum. Proc Nat Acad Sci 15:876–881

Stern C (1950) Principles of human genetics. WH Freeman and Company, San Francisco

Stern C (1960) Principles of human genetics, 2nd edn. WH Freeman and Company, San Francisco

The Lindau Mediatheque (1955) 5th Lindau Nobel Laureate meeting. The effect of radiation and other present day influences upon the human genetic constitution. http://www.mediatheque.lindau-nobel.org/#/Video?id=102

Timoféeff-Ressovsky NW, Zimmer KG, Delbruck M (1935) Uber die Natur der Genmutation und der Genstruktur. Nachrichten von der Gesellschaft der Wissenschaften zu Gottingen:Mathematische-Physikalische Klass, Fachgruppe VI, Biologie 1(13):189–245. [English translation: On the Nature

of Gene Mutation and Gene Structure. Reprinted in Sloan PR, Fogel B (editors). (2011). Creating a Physical Biology. The three-man paper and early molecular biology. The University of Chicago Press, Chicago.]

United Nations Scientific Committee on the Effects of Atomic Radiation (UNSCEAR) (1962) Report, 17th session, Supplement No. 16, New York

United Nations Scientific Committee on the Effects of Atomic Radiation (UNSCEAR) (1964) Report, 19th session, Supplement No. 14, New York

Uphoff DE, Stern C (1949) The genetic effects of low intensity in irradiation. Science 109:609–610

Von Schwerin A (2010) Medical physicists, biology, and the physiology of the cell (1920–1940). Ident Mut 231–258

Watson JD, Crick FHC (1953) Molecular structure of nucleic acids. A structure for deoxyribose nucleic acid. Nature 171(4356):737738

Weinstein A (1928) The production of mutations and rearrangements of genes by x-rays. Science 67:376–377

Weiss J (1944) Radiochemistry of aqueous solutions. Nature 153:748–750

Whittemore GF (1986) The national committee on radiation protection, 1928–1960: From professional guidelines to government regulation. Ph.D. Dissertation. Harvard University, Cambridge, Massachusetts

Zimmer KG (1941) Ergebnisse und Grenzen der treffertheoretischen Deutung von strahlenbiologischen Dosis-Effekt-Kurven. Biol Zentral 63:78

Zimmer KG (1966) The target theory. In: Cairns J, Stent GS, Watson JD (eds) Phage and the origins of molecular biology. Cold Spring Harbor Laboratory Press, New York, pp 33–42

Arch Toxicol
DOI 10.1007/s00204-015-1455-3

LETTER TO THE EDITOR, NEWS AND VIEWS

Cancer risk assessment foundation unraveling: New historical evidence reveals that the US National Academy of Sciences (US NAS), Biological Effects of Atomic Radiation (BEAR) Committee Genetics Panel falsified the research record to promote acceptance of the LNT

Edward J. Calabrese

Received: 15 December 2014 / Accepted: 6 January 2015

Abstract The NAS Genetics Panel (1956) recommended a switch from a threshold to a linear dose response for radiation risk assessment. To support this recommendation, geneticists on the panel provided individual estimates of the number of children in subsequent generations (one to ten) that would be adversely affected due to transgenerational reproductive cell mutations. It was hoped that there would be close agreement among the individual risk estimates. However, extremely large ranges of variability and uncertainty characterized the wildly divergent expert estimates. The panel members believed that sharing these estimates with the scientific community and general public would strongly undercut their linearity recommendation, as it would have only highlighted their own substantial uncertainties. Essentially, their technical report in the journal Science omitted and misrepresented key adverse reproductive findings in an effort to ensure support for their linearity recommendation. These omissions and misrepresentations not only belie the notion of an impartial and independent appraisal by the NAS Panel, but also amount to falsification and fabrication of the research record at the highest possible level, leading ultimately to the adoption of LNT by governments worldwide. Based on previously unexamined correspondence among panel members and Genetics Panel meeting transcripts, this paper provides the first documentation of these historical developments.

Keywords Mutation · Cancer · Risk assessment · Linear no-threshold (LNT) · Threshold dose response

In 1956, the US National Academy of Sciences (NAS) published their long-awaited reports addressing national concerns about how ionizing radiation may affect such entities as oceans/fisheries, agriculture/food supply, meteorology/atmosphere, medicine/pathology, genetics and disposal of radioactive wastes. As it turns out, the report that dominated the attention of the scientific community and media was that of the Genetics Panel. It proclaimed there was no safe level of exposure to ionizing radiation and offered dire warnings about severe adverse biological effects occurring in present and future generations. Societies, world governments and medical communities needed to heed the mutational risks that could persist across generations as a result of exposures to even low doses of ionizing radiation. The panel emphasized that the then extant threshold dose–response model was wrong and misled society on the hazards of low doses of ionizing radiation. To better protect the public health and to provide more accurate predictions, the report urged the risk assessment community to adopt a linear dose–response model. This recommendation represented no less than a paradigm shift that would alter the courses of both international environmental policy and cancer risk assessment to the present time. The LNT dose response was soon generalized from assessing the radiation risk of mutation to the radiation risk of cancer and then generalized once again by the US EPA to assessing the chemical risk of cancer. In retrospect, the road to linearity can be directly traced back to the BEAR Committee, Genetics Panel (Calabrese 2009, 2013).

Despite their tidal wave of success in 1956 and in the years following, the radiation genetics community had

Electronic supplementary material The online version of this article (doi:10.1007/s00204-015-1455-3) contains supplementary material, which is available to authorized users.

E. J. Calabrese (✉)
Department of Public Health, Environmental Health Sciences, Morrill I, N344, University of Massachusetts, Amherst, MA 01003, USA
e-mail: edwardc@schoolph.umass.edu

Published online: 20 January 2015

already been seeking a switch from the threshold to the linear dose–response model for nearly 30 years (Calabrese 2013), i.e., starting from a time soon after Muller's famous Nobel Prize winning discovery in 1927 that X-rays can induce mutations in the sperm of male fruit flies. Muller, Curt Stern and other prominent researchers from the radiation genetics community had long challenged the risk assessment methods for ionizing radiation and proposed using the far more conservative linear dose–response model. However, at each turn in the road, another similarly recalcitrant medical committee opposed their challenges and supported the more lenient threshold dose–response model instead. This frustrated Muller and his kindred radiation geneticist colleagues. In all major advisory committees to that point, the cards were "stacked" against them. However, with the creation of the NAS Committee, which was funded by the Rockefeller Foundation, the political tide turned their way. The decision to create an NAS Genetics Panel meant that Muller and his group would no longer be token geneticists on a committee oriented toward and dominated by the medical community; they would now be the dominant force on a BEAR I Committee whose 17 members included 13 notable geneticists. This may have seemed like a dream come true as the panel would now have no opposition to the big issue of the day: that is, finally getting linearity to drive the mutation risk assessment. The panel would soon proclaim that LNT was the new risk assessment "law" of the land, with little, if any, need for discussion, debate or evidence-based examination via scientific assessments. Thus, the panel moved to other challenges. Instead of debating the merits of the threshold vs LNT, the Chair of the Panel requested that all the geneticists on the panel provide their best estimates with upper and lower confidence intervals for the number of adversely affected children born to parents' whose gonads were exposed to a certain dose of radiation.

Despite the fact that there was a wide range of geneticists (e.g., human, fruit fly, bacterial, etc.) comprising the panel, it was hoped that there would be a high degree of agreement/consensus on what the specific population risks might be. If the panel members could independently come to a convergent agreement on risks, it would strongly support their risk assessment judgment and the linearity dose–response paradigm that they wanted society to adopt.

It is here where the story gets interesting. Through a variety of unexpected discoveries, it was possible to determine that the panel of geneticist experts wildly differed among themselves on the estimates of population risks, and, in fact, felt very uncertain about their own estimates of mutation frequency in future generations. The emergence of such uncertainties rattled the leaders of the panel and eventually led the Genetics Panel to omit key data from the research record, all in an effort to disguise the vast uncertainty that existed for the projected human risks. These factors and issues were known by the panel and are evident in the numerous letters that were exchanged between them and the Panel Chair; the panel even voted to hide the uncertainty from the scientific community by omitting key data and misrepresenting the predicted risks. In effect, the NAS BEAR Committee, Genetics Panel committed scientific misconduct in their publication in the journal Science in June, 1956 (Anonymous 1956). By omitting and misrepresenting the actual data, the panel hoped to convince the scientific community and the public to adopt their linear dose–response model in the assessment of risks associated with exposures to ionizing radiation, especially at low doses. These falsifications and fabrications are detailed and presented for the first time in the supplemental data section; they expose the fraudulent actions of the Genetics Panel and call attention to the vast impact they have had on cancer risk assessment.

Acknowledgments Research activities in the area of dose response have been funded by the United States Air Force (FA9550-13-1-0047) and ExxonMobil Foundation over a number of years. However, such funding support has not been used for the present manuscript.

Conflict of interest Author declares no conflict of interest.

References

Anonymous (1956) Genetic effects of atomic radiation. Summary Report of the Committee on Biological Effects of Atomic Radiation by the National Academy of Sciences, BEAR I Genetics Panel (W. Weaver, Chair). Science 123:1157–1164 (1956). [Erratum, Science 124:170]

Calabrese EJ (2009) The road to linearity at low doses became the basis for carcinogen risk assessment. Arch Toxicol 83(3):203–225

Calabrese EJ (2013) Origin of the linearity no threshold (LNT) dose–response concept. Arch Toxicol 87:1621–1633

Supplementary Material

SCIENTIFIC MISCONDUCT BY THE U.S. NATIONAL ACADEMY OF SCIENCES IN RECOMMENDING LNT FOR RISK ASSESSMENT

Edward J. Calabrese, Ph.D.
Professor of Toxicology, Department of Public Health
Environmental Health Sciences, Morrill I, N344
University of Massachusetts, Amherst, MA 01003
Phone: 413-545-3164, Fax: 413-545-4692
E-mail: edwardc@schoolph.umass.edu

ABSTRACT

The National Academy of Sciences Biological Effects of Atomic Radiation I (NAS BEAR I) Committee Genetics Panel (1956) recommended that regulatory agencies switch from a threshold to a linear model for radiation risk assessment of genomic endpoints. This recommendation was generalized to somatic cells for cancer risk assessment and later applied to chemical carcinogens. At the heart of the recommendation were independent estimates by panel geneticists of adverse reproductive effects in humans exposed to a given radiation dose. This paper reveals both an enormous variance and uncertainty in the independent estimates of genetic damage in reports submitted by six geneticists. These estimates were far greater than reported in Science, substantially misrepresenting expert disagreement. While the Science paper indicated that only six (of the 12) geneticists provided estimates, nine did. The three excluded estimates indicated markedly lower damage than did the six. The censoring of the three estimates and the mischaracterization of the uncertainty of the remaining six geneticists deliberately and substantially overstated the degree of confidence and agreement in the estimates of genetic harm. Based on internal correspondence, these actions were designed to ensure that linearity at low dose would be adopted for risk assessment. The omission and misrepresentation of data intended to represent the independent appraisal of panel experts is consistent with definitions of falsification and fabrication of the research record. The principal scientific document therefore supporting the adoption of the LNT was the product of scientific misconduct and that its adoption by regulatory agencies worldwide was based on a process involving falsification and fabrication.

Key Words: linearity, dose response, risk assessment, threshold, mutation, cancer

Introduction

Cancer risk assessment policy in the United States (US) originated from the US National Academy of Sciences (NAS) Biological Effects of Atomic Radiation (BEAR) I Committee, Genetics Panel, in 1956 when they recommended a shift from a threshold to a linear dose response model for the effects of ionizing radiation on genomic mutations. This linearity recommendation was soon generalized to somatic cells by the National Committee on Radiation Protection and Measurement (NCRPM) and applied to cancer risk assessment (Whittemore 1986 – Page 608, Footnote 260). This perspective was widely adopted in numerous countries by various regulatory and public health agencies (e.g. International Committee for Radiation Protection-ICRP) and continues to the present (Calabrese 2009, 2011a, 2013)

Historical reassessment of the BEAR I Genetics Panel report and its scientific foundations has revealed a series of key scientific concerns, flaws, and misrepresentations that challenge the basis of the original linearity recommendation (Calabrese 2011b, 2013). These findings provide substantial evidence

that the switch from threshold to linearity occurred due to ideologically-based deceptions of key papers on ionizing radiation induced mutation by leaders in the radiation genetic community, most notably Hermann J. Muller, Nobel Prize winner, and Curt Stern, University of California-Berkeley professor.

The current paper extends this historical re-evaluation into the activities of the BEAR I Genetics Panel with particular focus on the key publication of that panel in Science, entitled "Genetic Effects of Atomic Radiation". The Science paper represents the "technical" publication (Anonymous 1956) of the panel whereas a "popular" version (NAS/NRC 1956) was written for the media and public and included summaries of each of the NAS BEAR I Committee Expert Panels. Each of the technical Panels published a separate article in the journal Science in 1956. The present paper originated as a result of an unexpected observation that the Science paper of the Genetics Panel reported genetic risk estimates of only six of the 12 geneticist members of the Panel. Yet, an investigation of the personal correspondence and unpublished technical writings of the Genetics Panel uncovered detailed assessments from nine members of the Panel. This previously unrecognized discrepancy led to a further investigation resulting in the present paper.

Estimating radiation induced genetic risk

A key finalizing activity of the Genetics Panel was to provide an estimate of the total number of offspring (including embryonic/fetal deaths/stillbirths and those born but unable to reproduce) that would be adversely affected by mutations (i.e. often referred to as "genetic deaths" –adversely affected offspring that could not reproduce) assuming the entire adult reproducing population received a single dose of 10 roentgens (R) (0.1 Gy) to the gonads. The specific charge to the Genetics Panel was given by the chairman, Warren Weaver of the Rockefeller Foundation, the organization which funded the assessment. This finalizing topic was extensively discussed on Feb. 6, 1956 in the Panel meeting, occupying about 40 transcript pages (NAS 1956a). On Feb. 8, 1956 a memo was sent by Weaver to the Genetics Panel with the subject heading of "Reminder" (Weaver 1956a). He stated the following:

"At the Chicago meeting it was agreed that every geneticist on the panel was invited and indeed urged to undertake an estimation of the expressed damage due to detrimental mutations. The estimation was to apply to the total number of children (say 160 million ?) which will in the future be born to persons now alive in the U.S.; and to their children and so on. At least three estimates are desired: 1) expressed damage to the first set of 160 million children due to a dose of 10r to the persons now living; 2) expressed damage to the F1 through F10 generations of children due to a single dose of 10r to the persons now living; 3) expressed damage to F1 through F10 due to a dose of 10r per generation.

It would be most helpful also to have comparable estimates for 2) and 3) above when the dose is 50r. Linearity applies to 1), so that it is trivially easy to adjust that estimate to any dose.

It was agreed that each would state the range which applies to the estimates entering the calculation, and the range applying to the final result; and would state his reasoning in enough detail so that the other members could intelligently weigh, criticise (criticize), and compare the various estimates.

May I urge you that you undertake this <u>as promptly as possible</u>, and that you duplicate your report and send it to every member of the panel?"

The basis for the charge conferred upon the geneticists by Weaver was a discussion of a written position statement of geneticist Panel member Tracy Sonneborn that was read into the transcription record on Feb. 5, 1956. In addition to a statement of genetic risk assessment principles, Sonneborn raised the question as what would be the risk communication message for the general public based on their Panel deliberations. This led Weaver to trace three possible options with the most desirable, being the most explicit. That is, he wanted the Panel to provide reproductive risk estimates to the general

public that were specific and quantitative, estimates that would build upon the belief that all doses of radiation were harmful, cumulative, and irreversible.

On page 244 of the Feb. 6, 1956 transcribed proceedings of the Panel (NAS 1956a), Weaver stated his desire to obtain the independent views of the geneticists for the prediction of radiation induced genetic damage for subsequent generations of humans, using the US population as the working example. This was viewed as a means to define the range of informed judgment and to ascertain how much the Panel geneticists were in agreement, a key factor needed to support policy recommendations of the Panel. On page 257 of that same transcript Sonneborn stated **"that the thing of most value in all this calculation would be to show how one can use different methods to make estimates, and see to what extent methods, if possible, variations in approach, lead to different answers. So that if they converge, or tend to converge, then we might have more willingness to put them forth."** Weaver and Sonneborn were, of course, mindful of the fact that the geneticists brought differing technical abilities, research experiences and judgments to the Panel. Panel geneticists were experts with Drosophila, bacteria, paramecia, mice and human population genetics. Some of the Panel members were far more mathematically oriented than others. Each was expected to independently consider the problem and then document in writing their methods, assumptions and their estimated population-based radiation induced genetic harm. These estimates were to be collectively summarized and would define the range of expert agreement/disagreement using the different but complementary approaches of each contributing expert geneticist.

The Genetics Panel

There were 13 *bona fide* geneticists who were members of the Genetics Panel. These included: George W. Beadle – California Institute of Technology; James Cotterman – Baylor University; James F. Crow – University of Wisconsin; Milislav Demerec – Carnegie Institution of Washington; H. Bentley Glass – Johns Hopkins University; Berwind P. Kaufmann – Carnegie Institution of Washington; Clarence C. Little – Roscoe B. Jackson Memorial Laboratory; Hermann J. Muller – Indiana University; James V. Neel – University of Michigan; William L Russell – Oak Ridge National Laboratory; Tracy Sonneborn – Indiana University; Alfred Sturtevant – California Institute of Technology; and Sewall Wright – University of Wisconsin. Other members of the committee were the chairman Warren Weaver, Rockefeller Foundation, a mathematician; C. Failla, Columbia University, a health physicist; Alexander Hollaender, a physical chemist and administrator of Oak Ridge National Labs; and Shields Warren, New England Deaconess Hospital, a physician, who was also a member of the NAS BEAR I Pathology Panel. After the meeting in Chicago (February 5/6, 1956), Cotterman resigned from the Panel due to academic work load issues, leaving 12 geneticists on the committee. Of the 12 geneticist members, estimates of nine members have been obtained via searching past committee communications and letter exchanges from all members of the committee. The three members not providing estimates were Little, Neel, and Sonneborn. While there is no record of Little's actions on this matter, Sonneborn and Neel refused the request of Weaver to provide estimates of genetic damage since they did not believe that they could be reliably made. In fact, Neel argued that the scientific basis for such estimates were so uncertain that providing them would be a violation of his obligation as a scientist to society, an unethical condition. On April 6, 1956 he specifically addressed his concern with Chairman Weaver (Neel 1956a):

"The geneticist has social responsibilities, but he also has responsibilities as a scientist. One is that in an area as critical as this one is, he must beware of letting his conjectures get too far in advance of his facts. It is to me an exceedingly tenable position, having stated the general genetic argument, to say flatly that we know so little about the quantitative aspects." In fact, Neel was so adamantly opposed to the decision to develop and provide such genetic estimates of damage that he wrote Weaver on March 8, 1956 stating that he would **"go down with flags flying and guns booming to the last"** (Neel 1956b).

Organizing the estimates

The nine estimates were provided by the end of February, 1956. At a March 1, 1956 meeting of the Genetics Panel, a decision was made on how the estimates would be processed. In a March 2, 1956 brief summary (Weaver 1956a) (Muller manuscripts (mss), Lilly Library) of the March 1 meeting, Weaver indicated that **"Jim Crow was made the chairman of a sub-committee (he can commandeer help on this if he wishes) to go through all the damage estimates, compare them, and display assumptions, methods, input, and results in some sort of chart or graphic form."** This description did not indicate that Crow was authorized to exclude individual contributions, nor to make judgments as to which estimates were the most/least credible. Crow's role as sub-committee Chair was to organize and integrate the submitted estimates in a coherent manner so that the entire Panel could intelligibly, efficiently and objectively view the submitted estimates.

The Science journal article

In the Genetics Panel Science paper it was stated that **"Six of the geneticists of this committee considered the following problem: suppose the whole population of the United States received one dose of 10 roentgens of radiation to the gonads. What is the estimate of the total number of mutants which would be induced by this radiation dose and passed on to the next total generation of about 200 million children? Each geneticist calculated what he considered to be the most probable estimate, and then bracketed this by his minimum and maximum estimates. Each thus said, in effect: I feel reasonably confident that the true value is greater than my minimum estimate and less than my maximum. My best judgment, as stated in a single figure, is what I have labeled the most probable estimate."** Note that authors published an errata one month later indicating that the 200 million number of children should have been 100 million (Errata-Science, Volume 124, page 170, 1956). The paper goes on to state that **"the most probable estimates as thus calculated by the six gentcists (geneticists) do not differ widely. They bunch rather closely around the figure 5 million. Four of the six estimates are very close to that figure, and the other two differ only by a factor of 2.**

These six geneticists concluded, moreover, that the uncertainty in their estimation of the most probable value was about a factor of 10. That is to say, their minimum estimates were about 1/10, and their maximum estimates about 10 times the most probable estimate."

Inconsistencies emerge from the Science paper

A detailed evaluation of the Genetics Panel article in Science reveals a number of anomalies and concerns.

1. The invitation to provide estimates was made with some urgency to all geneticists on the panel. The reader is informed that only six provided estimates. The specific six geneticists were not identified in the Science paper. The Science paper did not provide any description of the biological models and methods used by the six geneticists to obtain their respective estimates. While only six geneticists were stated as having **"considered the …problem"**, all 12 geneticists were urged to take on this task. We also know that the number of geneticists providing written assessments was nine. Based on a letter from James Crow to Warren Weaver on May 21, 1956 (Crow 1956a), he listed the names of six geneticists on the Panel who provided estimates. These geneticists included Beadle, Crow, Glass, Muller, Russell and Sturtevant. It is this group of six that the Science article values were based upon. However, detailed mutational estimates were provided in a professional and timely manner consistent with the request of Weaver by three other Panel geneticists [Demerec (2/14/56 letter; 2/11/56 document; and stamped received 2/16/56) (Demerec 1956); Kaufmann (2/27/56 letter to Weaver and stamped received 2/29/56) (Kaufmann 1956); and Wright (2/22/56 letter to Weaver and stamped received on 2/24/56 (Wright 1956)]. Each of these three documents (Demerec, Kaufmann and Wright) were obtained from the Lilly Library, papers of Hermann J. Muller at the University of Indiana as well from the

files of multiple other members of the Panel. It therefore seems that Weaver received the assessments and that they were distributed to Panel members, as expressly recommended in the Weaver memo (Weaver 1956b). Thus, it is likely that all or most Genetics Panel members saw the three assessments that were excluded from the reporting in the Science publication. The question needs to be raised as to why didn't the Science article incorporate the independent findings from the other three Panel geneticists since they were professionally addressed and completed in a timely manner?

2. The estimates of the six Panel geneticists were summarized by James Crow in three letters [March 12, 1956 (Crow 1956b); March 29, 1956 (Crow 1956c); and May 21, 1956 (Crow 1956a)] to Weaver. The March 12, 1956 letter provided a graph showing the minimum, maximum and best estimates for total genetic damage expected from all descendants of the first generation of parents exposed to 10r, the cumulative damage through ten generations for a single generation exposed to 10r or the damage at the 10th generation from 10r per generation up to that time. In this letter, Crow stated that he made a decision to exclude the estimates of Drs. Wright and Demerec while the non-cited report of Kaufmann was not mentioned. However, the estimates of Wright were included in the March 12, 1956 graph. More specifically, in the final paragraph of this letter, Crow writes that he did not provide **"Dr. Wright's methods which are greatly different, but clearly given in his letter."** In the next sentence he then writes that **"I haven't included Dr. Demerec's estimate on the graph for it, too, is based on quite different assumptions that lead to a greatly different value than the others obtained".** The only specific comment that Crow offered for the Demerec estimation was that it **"was based on bacterial mutation rates"**. As for the Wright estimates, Crow stated that he **"counted only mutations causing conspicuous effects in postnatal life"**. However, the Wright document states that his analysis included the **"conspicuous detrimental effect on viability or fecundity"**. Thus, the effects estimated by the Wright analysis would not have been restricted to only those occurring in postnatal life, thereby contradicting the statement of Crow. Furthermore, the postnatal effects estimated by Wright were not restricted to an age of the offspring to display the harmful effect. However, Crow accepted the analysis of William Russell (1956) which restricted the age of expressing damage to only three weeks, a limitation that Russell noted. The actions of Crow to exclude the Wright and Demerec estimates are at variance with the instruction that he present **"all the damage estimates, compare them, and display assumptions, methods, input, and results in some sort of chart or graphic form."**

3. In the March 29, 1956 letter to Weaver, (Crow 1956c), Crow stated: **"The limits presented on our estimates of genetic damage are so wide that the reader will, I believe, not have any confidence in them at all."** He then makes the statement that **"I recommend one of two things: omit the estimates entirely, or b) give a single best estimate of the number of mutations, or a narrow range of estimate, based on direct extrapolations from mouse and Drosophila."** Crow then inexplicably states that **"We then state that these are based on mouse data and let the reader add his own uncertainty factor."** Basing estimates on Drosophila and mice and then telling the reader that the estimates are based only on mice is deceptive. Asking the reader to construct their own uncertainty factor with what is highly censored data lacking upper and lower bounds is disingenuous. This recommendation followed from his earlier comment in the March 12, 1956 letter (Crow 1956b) that **"...the groups differ widely in their confidence in the best estimate, as indicated by their grossly discrepant minimum and maximum estimates."** In the May 21, 1956 letter, he again provided mutational estimates of the six geneticists, including their best estimates and ranges (Crow 1956a). He stated in the letter

that he did not support publishing the table. This issue was eventually considered for a vote by the entire committee and the table was not included, only the summary statement as given above.

4. Table 1 reveals that the best (i.e., most probable) estimates of affected children after 10 generations ranged from 2 to 10 million, a five-fold range. The range of uncertainty varied considerably amongst the geneticists. Beadle and Glass, for example, gave a 2,000 fold uncertainty range while Muller's estimate was 10-fold. In the Science paper, it was stated that the uncertainty estimates were within 10-fold of the best estimates (i.e., 100-fold range). However, in the case of Beadle, the bounded values range from 20-fold below to 100-fold above his best estimate. For Glass, his range was from 40-fold below to 50-fold above his best estimate. In the case of Muller, his bounded values were 4-fold below and 2.5-fold above his best estimate. Similarly large variation in uncertainty estimates was also presented by Crow in his March 12, 1956 (Crow 1956b) letter to Weaver for the damage expected in the first offspring generation. In fact, the estimated number of affected individuals ranged from 5000- 20,000,000, a 4000 fold value. It is not possible to discern how the minimum and maximum values were derived in some cases. In the case of Beadle, for example, the values are provided without any explanation. As for Russell the best estimate was derived but lower and upper bound estimates are not included. It is not known how Crow derived such values. The case of Crow's own estimate may also be instructive. He stated on March 29, 1956c (Crow 1956c) **"I shall use as a minimum estimate a direct extrapolation from Drosophila and as a maximum some calculation from the sex-ratio data in the Japanese cities. An estimate from mouse data turns out to be just about half way between these, so I shall use it as the most probable estimate."** With such non-sequitur biological reasoning guiding the genetic risk estimates as well as with estimates lacking any documentation, it is not surprising that the committee members did not want to share their procedures with others. Furthermore, there was an absence of criticism of the bizarre approach offered by Crow as well as the lack of documentation for upper and lower bound estimates by Beadle, Russell and others despite the high priority placed on such values. Even Sturtevant, whose lower and upper bound estimates were exactly ± 10-fold, stated on February 20, 1956 to Weaver (Sturtevant 1956), **"After going through these calculations I come out with a feeling that they are rather futile. At almost every step it has been necessary to make a guess, often with little to go on, and with no real basis for setting limits within which the true value probably lies".** This statement raises the question that Sturtevant may even have been open to the possibility that the true estimate was not even within his derived lower and upper bounds. The statement of Sturtevant contradicts the earlier statement in the Science article which claimed that **"Each thus said, in effect, I feel reasonably confident that the true value is greater than my minimum estimate and less than my maximum."** In fact, the term guess was also used by other reporting geneticists. For example, Russell referred to his extrapolation process also as a "guess" (Russell 1956).

5. According to Neel the closeness of the best estimates of the six geneticists selected by Crow was due to the fact that they used essentially the same assumptions for gene number, mutation rates, and other parameters (Neel 1956c). Neel thought that their scientific "agreement" was illusionary since there was little independence of thought. In fact, once each of the six had to think far more independently on the problem as with the case of estimating upper and lower uncertainty bounds their "apparent" agreement strikingly disappeared, supporting the Neel perspective. However, knowing that he had been overruled on genetic risk estimations, Neel

asked to be dissociated with any aspect of the report that provided quantitative estimates of genetic damage or a recommended permissible dose (Jolly 2003 - page 359).

6. The 100-fold range reported in the Science article to characterize the uncertainty about the best estimate strikingly conflicted with and misrepresented the range of uncertainty of this group of geneticists which had a mean value of 745 (uncertainty range 10-2000 – see Table 1) for the 10 generation estimate (median 180) and a mean of 756 (uncertainty range 100-2857) for the first generation (median 312.5) (data from March 12, 1956 Crow Letter to Weaver) (Crow 1956b).

Draft BEAR I Genetics Panel Report of March 19, 1956

In this draft report (NAS 1956b), Weaver stated that the **"estimates have been independently furnished by seven of the geneticists of this Panel. Each estimated the total damage (that is to say, the number of genetic deaths or extinctions) which would occur among our 100 million children. Each of the seven geneticists stated his result in terms of a range of values, giving what he considered to be a reasonable lowest figure and also a reasonable highest figure. Thus there were in all seven low estimates and seven high estimates. The lowest of the low estimates was 5,000 damaged individuals among our direct children. The highest of the high estimates was 20 million. This extreme range simply reflects our lamentable lack of information on human radiation genetics."** This range of high to low would indicate a 4000-fold range and probably reflected damage estimated in the first offspring generation. This draft differed from the May 21, 1956 table of Crow which showed a maximum 2,000 fold range, reflecting damage after 10 generations. There is also no explanation why there were seven geneticist estimates stated in this draft but only six were mentioned in the Science paper.

Chairman Weaver also raised the question that **"The public may well ask, why were there only seven such estimates when this Panel includes a larger number of geneticists? The answer is that many highly competent geneticists would not wish to undertake a calculation of this sort, either because they are not specifically experienced in the more mathematical aspects of genetics, or because they doubt that such a calculation is very useful at the present state of knowledge."** The response of Weaver failed to indicate that nine of the 12 geneticists provided estimates. He referred to seven Genetics Panel members as providing estimates. At some point one of the seven reports was reduced to six without explanation.

There were substantial differences as to what the trans-generational mutation risk might be, even amongst the six geneticist values that were used. The extent of the variation and the acknowledgment of the "futility" in making the given estimates undercut the confidence in the process and in its conclusions. Furthermore, this high degree of uncertainty would have been considerably greater if these geneticists were not constrained by Weaver (in his memo directions to the geneticists) to employ a linearity dose response assumption. The variability characterization estimates in the Science article provide a distinctly different impression than the actual estimates of the six geneticists, reflecting data censoring, and leading to false conclusions concerning the degree of uncertainty among the participating geneticists and the extent to which the experts agreed/disagreed amongst themselves on this matter.

On June 5, 1956 Chairman Weaver (1956c) wrote to the Genetics Panel informing them of the results of the voting on an unresolved aspect of the final report. One of the key questions was whether to include the table of Crow which showed the variation in genetic damage estimates between and within subjects. This decision was accomplished by a vote of 15 members of the Panel. The results were 7 against including the table, 6 for publishing it, with two designated as "indifferent". Thus, the recommendation of Crow prevailed and the table was not included. Based on incomplete obtained correspondence, it is known that Crow, Muller, Glass and Sonneborn voted against publishing the table.

The Missing Expert Geneticist Estimates

1) **What information did the three missing geneticist estimates provide?**
   Kaufmann: His report was six pages, providing a detailed mathematical derivation of mutation frequency based on human mutation rates, providing all assumptions used in his calculations. Three pages were text, two pages were tables and one page contained two summary figures. He provided genetic damage estimates over 10 generations and expressed the number of affected children per 200 million. Adjusting for 100 million, he reported 195, 000 affected children in generation #1. In his February 27, 1956 letter (Kaufmann 1956) to Weaver, Kaufmann stated that his **"calculations show that under the defined conditions the visible genetic damage resulting from chronic exposure to 10 r per generation is small in comparison with that of spontaneous origin."** Kaufman noted that his estimates, based on research of the geneticist Herman Moser, were quite complicated and he was not able to provide upper and lower confidence intervals. Note that Crow (1956b) included estimates of "Moser" which were subsequently dropped. These appear to be the estimates of Kaufmann.
   Wright: His report provided a nine page detailed assessment and mathematical derivation of genetic damage over 10 generations (Wright 1956). He also provided two estimates of damage based on assumptions related to the proportion of dominant mutations. The estimates provided in the report range from 34,000 to 67,000 (i.e. ~50,000 average) children affected in generation #1 based on the assumptions used. Upper and lower bound estimates were provided.
   Demerec: He provided a detailed, single spaced, three page assessment using *E. coli* as his model. He presented evidence for spontaneous mutation rates for 26 genes and X-ray induced mutation rates for the same genes. He also indicates that similar data existed for multiple chemical mutagens (Demerec 1956). The information that Demerec had developed on the *E. coli* model far exceeded that of any of the other models presented including Drosophila, mice and humans. His risk estimates were based on 160,000,000 as was originally proposed by Weaver. He estimated that there would be 14,200 affected children in generation #1 due to spontaneous mutation; in contrast, 8,320 additional children were estimated to be affected by the 10 r exposure to their parents. This value was adjusted to 5,200 for a 100 million population. Demerec stated that since he did not include "lethals" his estimate would need to be adjusted to some degree upward. He also stated that **"I do not wish to venture into speculations about the genetic damage that would be sustained by subsequent generations, for such speculations could only be based on assumptions not supported by experimental evidence."**
2) The three best estimates were 195,000 (Kaufman), 50,000 (Wright), and 5,200 (Demerec) (83,000 mean) for the first generation. On balance these are lower (~70 %) than for the other five geneticists (275,000 mean) (note that Beadle did not provide an estimate for generation #1). If all geneticist estimates for generation #1 were included, the generation #1 mean estimates would range from a low of ~ 5000 to a high of ~350,000, a ~70 fold difference, a value reflecting considerable variation amongst the geneticists.

Discussion

The present assessment indicates that the US NAS BEAR I Committee Genetics Panel Science paper included a series of significant misrepresentations relating to the central charge of the Panel, that is, predicting the public health risks of ionizing radiation from all sources, such as medical, fall-out and other means of exposure. The Panel knowingly reported that only six of the 12 geneticists provided estimates, when nine did. While reasons for this omission of data are speculative, two (i.e. Demerec and Wright) of the three excluded views offered notably lower estimates of risk than the six geneticists presented in the paper. If all geneticists' estimates were presented, it would have markedly altered the range of the best estimates; the range of the ~2-fold mean difference would increase to ~70-fold for

generation #1damage. If the experts were shown not to be in agreement, it would affect the credibility of their report and most likely weaken support for its policy recommendations.

The decision by Crow to exclude the estimates of Demerec and Wright overstepped his authority as given by Chairman Weaver. There was no documentation to support these actions. There is no indication of why the estimates of Kaufmann were omitted. Despite such actions of Crow to exclude these three, there is no evidence that such actions were disputed. This suggests that the affected members and the entire Panel accepted this decision. If this is the case, then it is even more problematic as the incorrect statement that only six geneticists provided estimates rather than the nine was known and agreed to by all. Whether this was due to internal pressure to conform to the goal of displaying highly consistent estimates is not known. Yet, the efforts of the three excluded geneticists were substantial, with no evidence that they withdrew their estimates.

The Panel also stated that the uncertainty range (i.e., upper to lower bound) was about 100-fold for the six geneticists used while the table of Crow had Beadle and Glass with ranges of 2,000 and a mean of 745, and uncertainty values that were even higher after the first generation. This misrepresentation, though different from omitted data, is consistent with the intention of creating the appearance of greater agreement amongst the geneticists than actually existed. Furthermore, the Science article failed to provide the reasons why three individuals (i.e., Little, Neel, and Sonneborn) did not provide estimates. Given the critical significance for Panel agreement with respect to the acceptance of policy recommendations, it was necessary to present the views of the entire group. In fact, the views of Sonneborn and Neel were that such estimates could not be reliably done, with Neel being particularly strident on this issue as quoted above. It is not clear what the opinions of Little were. Thus, if the broad spectrum of views and risk estimates had been summarized in the Science article, it would have undermined the conclusions and policy recommendations of the Panel even further. On this later matter, it is worthy of note that in his May 21, 1956 letter to the Panel (Crow 1956a), Crow states that **"Once again, I urge that we not include the table at all..."** He then displayed a table of the values of the six geneticists and stated that **"I include these values for Committee members' inspection because I believe the only thing worse than publishing the estimates at all is to publish the wrong values".** This recommendation was adopted and the table showing the extremely wide differences between the experts was dropped and was replaced with the above noted misrepresentation.

While the estimates of the panel members were clearly efforts in risk assessment speculation, they contributed to the appearance of confident conclusions throughout the Science article. In retrospect it appears that the exclusion of the Demerec and Wright estimates occurred because the derived values were “different”, that is, far lower than the other six. This was assumed to be due to the use of other methods, models or approaches. However, the charge of the Panel was to present the range of independent estimates and to describe them along with their assumptions. If complementary methods and models could be shown to have considerable agreement, it would bring added confidence to the recommendations of the Panel as noted above by Sonneborn. However, if the expert estimates were widely or wildly divergent, then it was feared that it would undermine acceptance and use of their report. In many ways, the future course of action would be affected by the degree of concordance of their estimates. When the estimates were found to be extremely variable, Crow and the other members of the Panel refused to show this variability/uncertainty and in fact acted to exclude low estimates and mischaracterize the estimates retained.

When placed in perspective what does this story reveal? The geneticists of the Panel firmly believed their genetics mutation-credo of nearly always harmful, cumulative, irreversible, and linear at low dose. However, when challenged by Weaver to translate this credo into independent estimates of societal harm, using their own research methods and experience, the results were unanticipated. That is, the estimates revealed much quantitative variability and uncertainty between and within the experts. So great was the range of estimates that actions were taken to prevent this from being exposed,

deliberately creating the false impression of agreement that was not warranted, as seen in the omission and mischaracterization of estimates. Comparing the message of the Science paper with the internal correspondence of the Panel members, reveals a striking dichotomy and dishonesty of the Panel. The falsification of the research record provided the vehicle for the acceptance of the Panel's goal, that is, the use of the LNT for risk assessment and ultimately its acceptance worldwide by regulatory and public health agencies and within the legal system. In a highly ironic twist to the present story, it should be noted that the US National Academy of Sciences building in Washington, D.C. has a statute of Albert Einstein, which is accompanied by a series of his famous quotes. One of these captures of the essence of the present paper and the intolerable actions of the Genetics Panel. "The right to search for truth implies also a duty; one must not conceal any part of what one has recognized to be true." In a retrospective statement some 40 years later concerning the actions of the BEAR I Genetics Panel, Crow (1995) confessed that he and especially Muller exaggerated the dangers of low level radiation and should accept significant blame for an irrational emphasis on such matters by the public and regulatory agencies.

Acknowledgement

Research activities in the area of dose response have been funded by the United States Air Force and ExxonMobil Foundation over a number of years. However, such funding support has not been used for the present manuscript. Author declares no conflict of interest.

Table 1. Trans-generational estimations of the NAS BEAR I Genetics Panel – Subcommittee –
The values represent "the total genetic damage expected among all descendants (through 10 generations) of a single (i.e., first) generation exposure to 10r". Source: Crow, 1956b - March 12, 1956 of Crow to Weaver; Letter of Crow provided the four columns to the right.

Min to Max Range (uncertainty range)	Minimum	Most Probable	Maximum	Author
2,000-fold	100,000	2,000,000	200,000,000	Beadle
2,000-fold	100,000[a]	4,000,000	200,000,000	Glass
260-fold	250,000	5,000,000	72,000,000	Crow
100-fold	600,000	6,000,000	60,000,000	Sturtevant
100-fold	700,000	7,000,000	70,000,000	Russell
10-fold	2,500,000	10,000,000	25,000,000	Muller

[a] On May 21, 1956 Glass wrote to Weaver indicating that his minimum estimate was in error and should be 200,000 rather than 100,000. He stated that the reason for the error was that he based his estimate on 1 r rather than 10 r in his calculations. He stated that his most probable and maximum values were correct. Coincidently, Crow (1956d) wrote to Muller on that same day (May 21, 1956) with copies to Weaver and Glass indicating that Glass's minimum estimate needed to be adjusted to a 10 r exposure (rather than 1 r), which when normalized to a population of 10^8, his value was 100,000, the same as Glass originally reported. No changes were needed for his most probable and maximum values. Thus, the memo of Crow did not agree with the recommended change in damage estimate of Glass. This technical point was not resolved since the Panel decided not to provide the Table in their final report/ Science paper based on a vote by the Panel on May 29, 1956. Note that Glass received the copy of Crow's letter as it was in his files. Glass's May 21, 1956 (Glass 1956) letter explanation seems unlikely as his original genetic damage estimates included values for both 1 and 10 r, with the expected 10-fold damage difference.

References

Anonymous (1956) Genetic effects of atomic radiation. Summary Report of the Committee on Biological Effects of Atomic Radiation by the National Academy of Sciences, BEAR I Genetics Panel (W. Weaver, Chair). Science 123:1157-1164 (1956). [Erratum, Science 124:170].

Calabrese EJ (2009) The road to linearity: Why linearity at low doses became the basis for carcinogen risk assessment. Arch Toxicol 83:203-225.

Calabrese EJ (2011a) Toxicology rewrites its history and rethinks its future: Giving equal focus to both harmful and beneficial effects. Environ Toxicol Chem 30:2658-2673.

Calabrese EJ (2011b) Key studies used to support cancer risk assessment questioned. Environ Mol Mut 52:595-606.

Calabrese EJ (2013) How the US National Academy of Sciences misled the world community on cancer risk assessment: New findings challenge historical foundations of the linear dose response. Arch Toxicol 87:2063-2081.

Crow JF (1956a) Crow, Letter to Weaver, Lilly Library, (Muller mss), Manuscript Department, Indiana University, Bloomington IN, USA. May 21.

Crow JF (1956b) Letter to Weaver, Lilly Library, Manuscript Department, Indiana University, Bloomington IN, USA. March 12.

Crow JF (1956c) Letter to Weaver, Lilly Library (Sonneborn mss), Manuscript Department, Indiana University, Bloomington IN, USA. March 29.

Crow JF (1956d) Letter to Muller, American Philosophical Society (APS), Philadelphia PA, USA. May 21.

Crow JF (1995) Quarreling geneticists and a diplomat. Genetics 140:421-426.

Demerec M (1956) Letter to Weaver, Lilly Library, (Muller mss), Manuscript Department, Indiana University, Bloomington IN, USA. February 14 with February 11 document.

Glass HB (1956) Letter to Weaver, American Philosophical Society (APS), Philadelphia PA, USA. May 21.

Jolly JC (2003) Thresholds of uncertainty: Radiation and responsibility in the fallout controversy, Ph.D. thesis. Oregon State University, USA.

Kaufmann BP (1956) Letter to Weaver, Lilly Library (Muller mss), Manuscript Department, Indiana University, Bloomington IN, USA. February 27.

National Academy of Sciences (NAS) (1956a) BEAR I Committee Transcript Proceedings Conference on Genetics, February 5-6, 1956, Chicago IL, USA.

National Academy of Sciences (NAS) (1956b) Unpublished Draft Report of the Genetics Panel of the National Academy of Sciences, Study of the Biological Effects of Atomic Radiation (BEAR I), American Philosophical Society (APS), Philadelphia PA, USA. March 19.

National Academy of Science (NAS)/National Research Council (NRC) (1956) The Biological Effects of Atomic Radiation. A report to the public. Washington DC, USA.

Neel JV (1956a) Letter to Chairman Weaver. American Philosophical Society (APS), Philadelphia PA, USA. University of Wisconsin. April 6, (Jolly, 2003-page 398).

Neel JV (1956b) Letter to Weaver. American Philosophical Society (APS), Philadelphia PA, SUA. University of Wisconsin. March 8, (Jolly, 2003-page 359).

Neel JV (1956c) Letter to Weaver, American Philosophical Society (APS), Philadelphia PA, USA. March 16 (Jolly, 2003-page 159).

Russell WL (1956) Report to Weaver, Lilly Library. Estimate of expressed damage in the offspring of individual exposed to 10r. Lilly Library, (Muller mss) Manuscript Department, Indiana University, Bloomington IN, USA. February 24.

Sturtevant A (1956) Letter to Weaver, American Philosophical Society (APS), Philadelphia PA, USA. February 20.

Weaver W (1956a) Memo to Genetics Panel, brief report, Lilly Library (Muller mss), Manuscript Department, Indiana University, Bloomington IN, USA. March 2.

Weaver W (1956b) Memo to Members of the Genetics Panel, Lilly Library, Manuscript Department, Indiana University, Bloomington IN, USA. February 8.

Weaver W (1956c) Memo to Genetics Committee, Lilly Library, Manuscript Department, Indiana University, Bloomington IN, USA. June 5.

Whittemore GF (1986) The National Committee on Radiation Protection, 1928-1960: From Professional Guidelines to Government Regulation. Ph.D. thesis, Department of the History of Science. Harvard University, Cambridge MA, USA.

Wright S (1956) Letter to Weaver, Lilly Library (Muller mss), Manuscript Department, Indiana University, Bloomington IN, USA. February 22.

Human & Experimental Toxicology (2000) 19, 32–40

www.nature.com/het

The marginalization of hormesis

EJ Calabrese*,[1] and LA Baldwin[1]

[1]*Department of Environmental Health Sciences, School of Public Health, University of Massachusetts, Amherst, MA 01003, USA*

Despite the substantial development and publication of highly reproducible toxicological data, the concept of hormetic dose-response relationships was never integrated into the mainstream of toxicological thought. Review of the historical foundations of the interpretation of the bioassay and assessment of competitive theories of dose-response relationships lead to the conclusion that multiple factors contributed to the marginalization of hormesis during the middle and subsequent decades of the 20th century. These factors include: (a) the close-association of hormesis with homeopathy lead to the hostility of modern medicine toward homeopathy thereby creating a guilt by association framework, and the carry-over influence of that hostility in the judgements of medically-based pharmacologists/toxicologists toward hormesis; (b) the emphasis of high dose effects linked with a lack of appreciation of the significance of the implications of low dose stimulatory effects; (c) the lack of an evolutionary-based mechanism(s) to account for hormetic effects; and (d) the lack of appropriate scientific advocates to counter aggressive and intellectually powerful critics of the hormetic perspective.
Human & Experimental Toxicology (2000) **19**, 32–40

Keywords: hormesis; low dose; stimulation; β-curve

Introduction

A recent detailed review of the toxicological literature from the late 19th century to the mid 1930s has revealed the publication of numerous papers indicating that the shape of the dose response curve often displayed a low dose stimulatory response followed by an inhibitory response at higher doses.[1] This was especially the case with respect to the responses of plants, bacteria and fungi. The endpoints measured typically involved growth rate, colony number, germination rate, time to germination, and physiological responses such as carbon dioxide production, sugar production and utilization, waste product generation, ammonification, nitrification and nitrogen-fixation. These low dose stimulatory responses became viewed as reproducible and broadly generalizable. The papers were published in leading scientific journals by investigators with notable reputations from outstanding institutions and universities. In fact, several of the leading researchers in the area of low dose stimulatory research involved Nobel prize winners and/or their students.

By 1905 the fundamental shape of the dose-response curve had been proposed[2] and is remarkably consistent with the modern hormetic β-curve for both the range and magnitude of stimulatory responses (Figure 1).[3] Furthermore, by the mid 1920s a scientific journal, *Zell Stimulations Forschungen* [i.e., Cell Stimulation Research[4]], was published in Germany to assess the biological significance of chemically-induced stimulatory responses typically of a low dose nature.

Despite this substantial development and publication of reproducible scientific data the concept of hormetic dose-response relationships became rapidly marginalized during the mid-decades of the 20th century in the field of toxicology and its related disciplines involving assessment of the effects of pesticides, disinfectants, environmental toxins and radioactivity. Such has been the case of its marginalization that the concept of hormesis rarely, if ever, merits even an historical note in any leading toxicological textbook. Thus, several generations of toxicologists have never been formally exposed to the extensive history of the hormetic response, including how it relates to the broader dose-response continuum, what its potential biological significance may be or how such a phenomenon may be studied. In fact, by ignoring this concept and adhering strongly to the regulatory belief that there are no biologically relevant effects below the no observed adverse effect level (NOAEL), a common and unmistakable conclusion that hormesis does not exist has pervaded the toxicology community.

*Correspondence: EJ Calabrese
Received 15 October 1999; Revised 15 October 1999; Accepted 15 October 1999

Within this framework it is easy to see how fair minded toxicologists of the 1990s could dismiss the concept of hormesis.[3,5] First, hormetic effects are generally of a limited nature, usually being only about 30–60% greater than the values of the unexposed controls at their maximum stimulatory response level. Such modest increases over controls may be readily dismissed as manifestations of background variability. Second, most mammalian toxicity studies involve very few doses. In fact, it appears that only 2% of toxicology studies published in the open literature have utilized six or more doses with only 10% of this two percentage number having three or more doses below the NOAEL (i.e., ~2 of each 1000 studies).[5] Given the low magnitude of stimulatory responses and the limited number of studies with an adequate number of doses especially in the low dose area, the modern mammalian toxicologist is not routinely exposed to studies providing an opportunity to evaluate hormesis as a biological hypothesis.

Therefore, the intention of this paper therefore is to assess the following: (1) how the toxicology community came to establish its concepts of dose-response relationships; (2) how such dose-response relationships were interpreted; (3) why the numerous observations of hormetic dose-response effects became marginalized by the toxicology community; and (4) why the concept of hormesis played no role in the development of the current tenets of toxicology and pharmacology with respect to the bioassay including its design, conduct and interpretation and its relationship to risk assessment principles and procedures.

Interpreting dose-response relationships

Background

The development of conceptual understandings of the dose-response curve has had a long history with its earliest, most definitive and sustained articulation derived from studies on chemical disinfectants [see review[6]]. Pasteur's work with disinfectants first definitively established that preservatives act by means of their toxic effects on microorganisms. These findings were extended by Lister in England to disinfectants and human disease causing microbes. Such scientific and medical milestones were followed by a vast amount of work which began a more formal characterization of the disinfectant properties of numerous agents and the development of the concept of a disinfectant ranking system using comparative potency rankings. However, it was not until 1886 that Robert Koch published the first systematic studies on disinfectants employing pure cultures of bacteria. He assessed the effects on anthrax spores of the then popular disinfectants carbolic acid (i.e., phenol) and sulfur dioxide, plus many other previously unstudied agents including mercuric chloride. Soon these findings were extended to practical applications by establishing a series of emulsified disinfectants thereby making use of soaps as effective agents to improve hygiene.[6] Thus, the waning years of the 19th century brought forth remarkable advances in the understanding of the biology of microorganisms, the development of methods for their evaluation, and their implications for human health and disease. Theories for the intrepretation of dose-response relationships were subsequently proposed and debated.

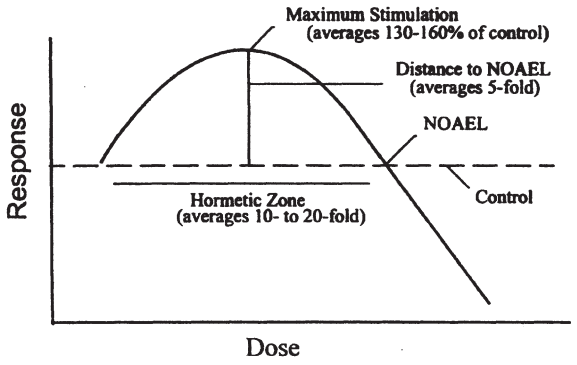


Figure 1 Dose-response curve depicting the characteristics of the chemical hormetic zone [modified from Calabrese and Baldwin[3]]. Abbreviations: NOAEL=no observed adverse effect level; LOAEL=lowest observed adverse effect level; ZEP=zero equivalent point

The unimolecular theory

In 1897 Kronig and Paul published an assessment of disinfectant properties using definite species of micro-organisms and disinfectants of every chemical class in widely varying concentrations.[7] This extremely comprehensive study also employed improved methods for more reliably quantifying bacteria. Such work set the stage for subsequent researchers to develop the means for standardization and comparison of various disinfectants (e.g., carbolic acid coefficient comparisons). The Kronig and Paul paper[7] was also important for the concept of hormesis since it established the importance of assessing biological effects of disinfectants over a broad concentration range and over time, thereby providing a foundation to establish dose-effect and dose-time-effect relationships.

The findings of Kronig and Paul[7] suggested a logarithmic relationship between the numbers of surviving bacteria and survival times. This logarithmic relationship was also reported by Madsen

and Nyman[8] and Chick[6] in the disinfection of anthrax spores with mercuric chloride and by Chick (1908) with phenol. Of significance is that Chick[6] observed that the dose-time-effect curve was quite similar to that expressing the course of a 'unimolecular reaction' (i.e., first order reaction). This follows directly from the Law of Mass Action where the velocity of any reaction is proportional to the active mass of reacting substance present at that moment. Therefore, Chick[6] concluded that the velocity of disinfection at any instant is proportional to the number (or weight) of living bacteria present.

Challenges to the unimolecular theory

Chick extended her original findings[6] of experiments with anthrax spores and *B. paratyphosus* to several additional bacterial models to enhance their generalizability.[9] Her striking claims, however, elicited an enormous debate in the microbiological literature over the next three decades and beyond, involving some of the most luminary scientists of the day [e.g., the Nobel laureate Arrhenius[10]]. The first wave of criticism was offered by Eijkman,[11] Hewlett[12] and Reichel,[13] followed by Loeb and Northrop,[14] Brooks,[15] Peters,[16] Smith,[17,18] Shackell,[19] Shackell *et al*,[20] Buchanan and Fulmer,[21] and Clark.[22]

The duration and intensity of the debate over how to interpret the shape of the dose-response curve reflects how pivotal this issue was in establishing the fundamental tenets of pharmacology, toxicology and even more recently with respect to current practices in risk assessment. In fact, the scope of the debate over the theory of unimolecular reactions was soon applied to the newly emerging field of radiation biology. More specifically, Blau and Altenburger,[23] supporting the view of Chick,[6,7] reported that the destruction of micro-organisms by radiation yielded unimolecular dose-response curves, leading those authors to conclude that the death of cells was produced by one or at most a few quanta of energy. Such an interpretation was subsequently adopted by early radiation response theoreticians.[24–27] Thus, the theoretical constructs of radiation dose-response evolved from the earlier debate sparked over the interpretation of disinfection dose-response issues. But, in fact, as pointed out by Packard[28] and Clark[22] the respective chemical and radiation dose-response interpretational controversies are strikingly similar and essentially hinge on the resolution of the same biological/toxicological points.

By the 1930s the developing consensus had emerged that strict acceptance of the unimolecular theory (i.e., the logarithmic death rate) should be abandoned based on the following arguments originally put forth by Brooks[15] and refined by Buchanan and Fulmer[21]: (a) Therewa a lack of correspondence of the experimental curve with the theoretical curve at the beginning of the experiment. The experimental studies of Chick[6,9] had inadequate data at short intervals at the beginning of the studies. In fact, in numerous instances more than half of the cells had been destroyed by the end of the first time interval reported. Such experimental data precluded an adequate assessment of the distribution of susceptibility in the population of bacteria and is believed to have been a major factor leading to erroneous conclusions supporting the unimolecular response theory. (b) There was a lack of correspondence of actual survivors' curve with the logarithmic curve in the latter part of response curve. When the values of the velocity constant are small there is a tendency for them to decrease rather than remain constant as predicted. (c) There were difficulties in assumption of uniform susceptibility. The overwhelming data for biological studies indicate that not all cells or organisms are equally susceptible; in addition the ratio of susceptible to non-susceptible cells is not a constant (i.e., in equilibrium) as is assumed in the unimolecular theory. Variation in response to toxic substances is widespread rather than the exception. Finally, (d) there was difficulty associated with securing a theoretical basis for interpreting cell death as a unimolecular reaction.

The characteristic curve theory

As a result of the above arguments, the unimolecular interpretation gradually receded and was essentially replaced by what was initially called the 'characteristic curve' which estimates the distribution of individual variation within populations with respect to responses to xenobiotics. Numerous studies of a substantial nature were published which attempted to characterize the extent of variability with animal models in response to toxic substances. For example, in 1924 van Wijngaarden determined the lethal dose of digitalis in nearly 600 cats, which resulted in a symmetrical bell-shaped distribution of variation.[29] Similarly, Morrell and Chapman reported the response of a population of 1331 rats to a series of neoarsphenamine doses, which also showed a symmetrical distribution of variation over a moderate dose range.[30,31]

Perhaps the most influential assessment of the characteristic curve was the work of Trevan[32] which remains the principal cited reference in the major toxicology texts of the 1990s [e.g., *Cassarett and Doull's Toxicology*[33]]. In setting forth to determine the best manner to estimate quantal effects, Trevan

concluded that the preoccupation of previous attempts to estimate the minimal effective (or lethal) doses should be abandoned in favor of estimating the central tendency of the group response.[32] In fact, Trevan established the concept and terminology of the lethal dose 50 (LD50).[32] He also utilized the term 'characteristic' to define the curve expressing the percentage of mortality or other limiting biological effect induced by varying the dose of a drug on animals of a certain species. He went on to provide a statistical foundation for sample size and statistical power. The proposal to estimate a response descriptor that could be accurately measured provided a strong stimulus for further work. While Trevan published a number of follow-up papers, it was the in-depth and extensive writings of Bliss and, to a lesser but still important extent, Gaddum, who applied the fundamental concepts laid out by Trevan to the major biological disciplines. Bliss published statistical methods for interpreting dose-response functions separately for entomology, microbiology, physiology, pharmacology and toxicology.[34–43] His widely cited papers, along with those of Gaddum[44–46] became a powerful standard for dose-response modeling and interpretation that is widely used today.

The influence of these biostatistical leaders of the 1920s and 1930s on the subsequent generations of toxicologists cannot be underestimated. They also influenced the further development of probit analysis by Finney[47–49] with its subsequent widespread application to the field of toxicology. By 1940, such statistical thinking had been integrated into the assessment of toxicology studies including the cancer bioassay as well evidenced in the classic paper of Bryan and Shimkin.[50] This paper focused on assessing the nature of the dose-response curve for chemical carcinogens over a wide range of exposures. Furthermore, some 21 years later probit analysis became the foundation of the Mantel-Bryan method[51] for low dose cancer risk estimation based on extrapolation beyond the observable range of the experimental data.

Therefore, the two predominant theories for interpreting dose-response relationships, the unimolecular and the characteristic curve, dominated the toxicological landscape in the early decades of the 20th century. While the unimolecular theory was losing out to the characteristic curve, there were other influential attempts to reinterpret the characteristic curve as one expressing that of a chemical process rather than a manifestation of interindividual variation. In fact, this concept in different mathematical forms continues to the present time as seen in examples of various hit and stage theories of carcinogenesis. Other biostatistical transformations were also proposed, including the logistic method.[52–54] The use of logits with quantal data used the assumption that the logarithms of the individual doses are distributed in a complex curve slightly different from the normal curve used in the probit method. In contrast to the characteristic curve of Trevan, with its incorporation into the probit concept, Emmens[55] sought to account for the dose-mortality curve by chance alone using the logistic method. He suggested that if the concept of tolerance is rejected, the law of chance supports the use of logits. This concept was developed earlier by Yule[56] in a more defensible statistical framework via a random hit theory dose-response approach which offers a similar dose response curve to that of the probit method.

In the 1930s the chemical process alternative interpretation to the characteristic curve was proposed in the case of the kinetics of drug action. This was based on the assumption that individuals varied symmetrically with respect to the lethal dose of xenobiotic required to produce a given response [see[22] for a review] and that the rate of xenobiotic entrance into the cell depended on diffusion according to a unimolecular formula. Such an hypothesis offered an explanation for the observation that the time-action curve expressing the rate of destruction of a population of small organisms frequently approximates the curve describing a unimolecular reaction.

Role of hormesis in the dose-response debate

So significant has been the need to resolve the above interpretational issues that the widely cited text, *Handbook of Experimental Pharmacology* by AJ Clark devoted approximately 10% of its pages to critiquing the unimolecular theory and another 15% to that of the characteristic curve and its possible integration with kinetic processes.[22]

While there has been much debate over the unimolecular theory, the characteristic curve and chemical process/kinetic-mechanism-based curves, there was almost no discussion concerning the role of hormetic responses during the early decades of the 20th century when the bioassay was being formalized and its interpretation refined. The only explicit discussion of hormetic responses in relationship to generalizable dose-response functions was that of Clark[22] who refers to this phenomenon as the Arndt-Schulz Law and/or diphasic responses [the term hormesis, from the Greek word meaning 'to excite', was not proposed until 1943[59]]. Rahn[57,58] modeled the hormetic response, which he acknowledged as a

widespread and generalizable phenomenon. He attempted to offer an enzymatic explanation for the low dose stimulatory response by developing an analogy between the actions of temperature on chemical metabolism. It was assumed that the toxic compound is a catalyst enhancing enzyme activity as well as enzyme deterioration. The model was built upon the concept that there was a shifting of the optimum enzyme activity with time from higher to lower concentration of the poison. While Rahn[57,58] cited the example of zymase activity enhancement by arsenate as a striking example of this phenomenon, he concluded there were no sufficiently complete data to adequately evaluate this phenomenon in a modeling sense. It appears that this attempt by Rahn had no demonstrable impact on the subsequent interpretation of the bioassay.

The effort made by Clark[22] to assess the significance of hormetic responses was principally judgmental providing extremely limited information with respect to theoretical foundation, historical basis and presentation of relevant data. The Arndt-Schulz Law was characterized as being 'in accordance with homeopathic doctrines and hence has maintained a certain popularity', thereby suggesting that it drew its strength from a homeopathic perspective rather than a biological/toxicological tradition. Clark further claims that the chief objection to the Arndt-Schulz Law is that 'it is obviously untrue in the case of most drugs that have been studied carefully' without providing adequate documentation to support this statement. He also states that 'many of the effects which appear to support this law were found to have simple explanations', again without providing sufficient information to support this broad conclusion. Based on the information presented on the historical foundations of hormesis by Calabrese and Baldwin,[3] it is apparent that the conclusions of Clark[22] were not only unsubstantiated but, in fact, were inconsistent with a large and generally available database that preceded his book. In addition, Clark's concept of hormesis was lacking in toxicological refinement in the sense that he collectively grouped all low dose stimulatory responses together including stimulatory responses of purified enzymes independent of their presence in a *bona fide* biological system. Despite the lack of an adequate review of the underlying foundation of the concept and data relating to the Arndt-Schulz Law, the strident judgments and highly critical characterizations of Clark are believed to have contributed in an important way to facilitate the marginalization of the hormetic concept within the toxicological community. Clark was a pharmacologist of exceptional reputation in the 1930s. He is credited with novel discoveries in the area of receptor mediated mechanisms. In fact, even as late as 1981 his original findings[60–62] were still being referred to as 'landmark' [see[63]]. In his historical foundations of the receptor theory, Parascandola[64] states that interaction between drugs and receptors was not treated quantitatively by pharmacologists until the research of AJ Clark and JH Gaddum in the 1920s and 1930s. Furthermore, in his foreward to the book '*Towards Understanding Receptors*' Robison[65] refers to the 1937 text of Clark[32] as the 'now classic monograph on General Pharmacology, a book that had great influence on a number of individuals.' Under normal circumstances such a negative statement would be challenged. However, none apparently did. Likewise, no alterative biostatistical models were presented to either challenge or refine those of Trevan, Bliss, and Gaddum which did not take the hormetic perspective into account.

The marginalization of hormesis

The marginalization of hormesis is therefore the result of a complex interactive web of circumstances which may be summarized as follows. First, the early identification of hormetic responses with the concept and practice of homeopathy created a 'guilt by association' situation. In fact, this concern was explicitly articulated by Hueppe in his 1896 book.[66] Even 50 years after the original association as stated by Arndt and Schulz [see[67] for review], Clark[22] continued to promote the close association of hormesis with homeopathy.

The strong association between homeopathy and hormesis (i.e., Arndt-Schulz Law) was based on the need of homeopathy to have an underlying biological foundation for this treatment's practices. The capacity for very low quantities of some poisons to act as stimulants to certain life functions as seen in some of the early research activities of Hugo Schulz[68] provided the scientific foundation for this relationship. While the low dose stimulatory actions of hormesis appear to be consistent with the tenets of homeopathy, this relationship is more apparent than real. More specifically, the stimulatory responses in homeopathy are alleged to become heightened as the solutions to be used in treatment became more dilute. In contrast, the stimulatory action seen in hormetic responses are of a very limited dose-range (usually not greater than a factor of ten, with a maximum stimulatory response of about 4–5-fold from the NOAEL. In addition, the most substantial theory of hormetic responses is that it represents an overcompensation to a limited disruption in homeostasis. Thus, if the experiment involves a properly timed sequence a clear decrement in response will

be observed followed by the hormetic overcompensation. Further, as a medical practice homeopathy treats patients with low doses of agents that elicit similar effects of their diseased condition (i.e., the so-called 'law of similars'). Hormesis has a different temporal arrangement such that an hormetic dose administered prior to a more massive dose as in the case of chemical and radiation adaptive responses may provide protection. Thus, there are important conceptual differences between homeopathy and hormesis that should have become progressively apparent over the decades since Arndt and Schulz sought to use hormesis as an underlying feature to this medical practice. In fact, as early as the late 1890s the temporal features of hormetic responses were being established by Townsend,[69–71] later refined by Branham[72] and Smith[73] and developed into a modern toxicological theory by Stebbing in the late 1970s and early 1980s [see[74] for a recent review]. The bias, therefore, against hormesis because of its relationship with the concept of homeopathy reflects (1) how strongly the homeopathic founders sought its association, (2) how intense the tension was between homeopathy and modern medicine, and (3) the lack of toxicological insight into the fundamental differences between the theoretical basis of homeopathy and the experimentally derived toxicological manifestations of hormetic responses.

Second, during the late 19th century up through at least the first half of the current century, the principal practical interests in dose-response curves of have focused on the upper end of the curve. This has been the case for disinfection biology, for the eradication of pests such as insects, weeds, etc., as well as the need to characterize hazardous situations especially for occupational exposures to workers. Even Winslow, who played such a critical role in establishing hormetic responses in bacteria demonstrably involved himself in the debate over the assertions of Chick and the unimolecular theory of dose response relationships[75] even though Falk,[76] a colleague of Winslow, emphasized the high significance of the low dose stimulatory responses of Hotchkiss,[77] Winslow's PhD student. Winslow never articulated its (i.e., the hormetic response) implications for the generalized dose-response function. Unfortunately, Hotchkiss did not follow up on her findings but redirected her research to other bacteriological questions. Despite the lack of follow through by Winslow and Hotchkiss it should be noted that a number of subsequent leading microbiological textbooks [e.g.,[78–80]] and significant research papers [e.g.,[81]] highlighted these findings and emphasized the hormetic nature of disinfectant dose-response relationships. Nonetheless, these multiple leading microbiological publications had little apparent impact on other related fields in relationship to establishing hormetic effects in the dose-response continuum.

Even though such references as noted above emphasized the regular occurrence of low dose stimulatory responses, the widely cited paper by Marshall and Hrenoff[81] explicitly stated that the stimulatory response 'is frequently of no practical value'. The inclusion of the stimulatory range in the dose-response continuum at least as far as disinfectants was involved, was principally for completeness rather than relevance. Such interpretations would have contributed to a further lowering of the concept of hormesis as a research priority, thereby perhaps inadvertently further trivializing its status.

Third, the lack of a sound evolutionary-based mechanism providing a credible conceptual framework for both evaluating and interpreting new experimental data was a serious limitation affecting the incorporation of the hormetic concept into the mainstream of biological/toxicological theory. This lack of a coherent framework is evident in the limited and confused presentation on this topic in 1937 by Clark, a highly respected pharmacologist.[22] If the evolutionary based mechanism of Stebbing[74] that hormesis may represent an overcorrection of a disruption in homeostasis had been available, it is likely that hormesis would have been seen more as a manifestation of an evolutionary based adaptive strategy than as part of the rationale for the practice of homeopathy.

Fourth, hormetic responses needed not only a leading toxicological advocate to develop its evolutionary basis but also a biostatistical/dose-response modeling one. For example, when Winslow was at Yale University (New Haven, CT, USA), the outstanding biostatistician Bliss was at the Connecticut Agricultural Research Station also in New Haven, Connecticut. If Winslow and Bliss had collaborated on this issue, it is likely that the course of low dose modeling of biological effects would have changed. However, Bliss did not arrive in New Haven until approximately 1943, some 20 years after the ground breaking work of Hotchkiss[77] and toward the end of Winslow's career. It should be mentioned that Bliss[38] was clearly aware of the 1937 book by Clark[22] which was so critical of the Arndt-Schulz Law. In addition, in 1935 Bliss acknowledged the assistance of Clark in his paper on estimating the dosage-mortality curve.[34] Furthermore, in 1936 Clark provided the opening address at a discussion of the chemical and physical foundations of pharmacological actions where one of the leading presenters was Gaddum [see[82]]. In addition, Clark and Gaddum were professors at the same university (University of Edinburgh) and in the same academic

department! Gaddum[44] also acknowledged the assistance of Trevan in reviewing his manuscript for publication. Such a close interaction of Clark with the leading biostatisticians of that era and the acknowledged respect for his professional advice would seem to have precluded the necessary biostatistical interest in the concept of hormesis.

The only notable figure in the biostatistical area, who potentially could have readily explored the statistical features of hormetic responses would have been the American Frank Wilcoxon, famous for the Wilcoxon Signed Rank Test, who explicitly reported hormetic responses in several fungal species exposed to low levels of hydrogen sulphide.[83] Nonetheless, McCallan and Wilcoxon[83] went on to emphasize the upper end of the dose-response curve with no further consideration of the significance of these observations. Note that Clark[22] discussed the report of McCallan and Wilcoxon[83] in considerable detail with no mention of the acknowledged and obvious low dose stimulatory responses.

Discussion

This paper described the historical unfoldings of the concept of dose-response relationships (i.e., dose-effect, dose-time-effect, etc.) and the intellectual struggle over how to interpret their biological significance. Resolution of such complex issues is rarely completed overnight as the debate over the unimolecular theory was very active for over three decades. This is often the nature of scientific debate that depends upon the conduct of new experiments, the subsequent publication of the data and the further refinement of the debate question.

The unimolecular theory had long staying power because of its theoretical linkage with a well-defined chemical process, its formulation in an established statistical model, and its strong advocates including leading researchers (e.g., Chick, Martin) at the world famous Lister Institute, the nobel laureate Svante Arrhenius and the highly regarded Cornell University Professor, Otto Rahn. The appeal to authority, even in scientific circles, can be substantial as Clark[22] noted that 'it is obvious that a physico-chemical theory (i.e., unimolecular theory) regarding the mode of action of drugs, which has received the support of Arrhenius must be considered carefully.'

While this constellation of factors can help to explain the long and continuing debate over the unimolecular theory of dose-response relationships, a more complex interaction of factors was likely at work which served to undermine the successful integration of the concept of hormesis into the mainstream of toxicological thought. Even though hormesis has been an unintended victim of the continuing struggle between traditional medicine and homeopathy, even though society has seen the upper end of the response curve as rightfully more significant than low dose responses, even though many decades of reproducible hormetic responses lacked coherent underlying mechanisms and even though mammalian toxicology studies have generally been relegated to the use of such small numbers of doses that hormetic effects could not be legitimately assessed, there still existed numerous cumulative *bona fide* examples in the literature leading up to the 1930s that should have accorded hormesis the status of a mainstream and an adequately supported biologically-based hypothesis. While any one of the above factors should not have been sufficient to marginalize this concept, their independent, yet collective, force is believed to have achieved that result.

Nonetheless, as has been shown, hormetic responses continued to be reported by numerous investigators over the following decades in various fields of biology. The biological basis therefore of hormesis as a credible hypothesis has been immeasurably strengthened over the past 50 years. Equally important, various toxicological explanations have begun to be offered which provide not only a sound theoretical foundation for hormesis, the hypothesis, but also a means of experimentally evaluating these proposed mechanisms [see[74,84]]. In addition, the issue of low dose cancer risks based on biostatistical models has clearly elevated the significance of hormesis in the risk assessment debate. Likewise, advances in molecular biology are also revealing that hormetic mechanisms may have important biomedical implications with respect to affecting adaptive responses of patients. These collective and relatively recent advances provide a foundation upon which hormesis as a biological hypothesis may now receive a more objective and comprehensive assessment.

Acknowledgements

This article was previously published in *Toxicologic Pathology* 1999; **27**:187–194 and has been reproduced in *Human and Experimental Toxicology* with the approval of *Toxicologic Pathology* and the Society of Toxicologic Pathologists. This work has been sponsored in part by an award to the University of Massachusetts (Edward J Calabrese, PhD, Principal Investigator) by the Texas Institute for Advancement of Chemical Technology, Inc. (TIACT). The text is based upon Appendix 8 of *Hormesis: The Stimulatory Action of Sub-Inhibitory Amounts of Toxins*, a report prepared for TIACT.

References

1 Calabrese EJ, Baldwin LA. Chemical hormesis: Its historical foundations as a biological hypothesis. *Toxicologic Pathology* 1999; **27:** 195–216.
2 True RH, Oglevee CS. The effect of the presence of insoluble substances on the toxic action of poisons. *Botanical Gazette* 1905; **39:** 1–21.
3 Calabrese EJ, Baldwin LA. A quantitatively-based methodology for the evaluation of chemical hormesis. *Human and Ecological Risk Assessment* 1997; **3:** 545–554.
4 *Zell-Stimulations-Forschungen.* **Vol. 1–3.** M Popoff and W Gleisberg (eds.). Paul Parey: Berlin, 1924–1930.
5 Calabrese EJ, Baldwin LA. The dose determines the stimulation (and poison): Development of a chemical hormesis database. *International Journal of Toxicology* 1997; **16:** 545–559.
6 Chick H. An investigation of the laws of disinfection. *Journal of Hygiene* 1908; **8:** 92–158.
7 Kronig B, Paul T. Die chemischen Grudlagen der Lehre von der Ciftwirkung und Disinfection. *Zeitschrift fur Hygiene und Infectionskrankenheiten* 1897; **25:** 1–112.
8 Madsen T, Nyman M. Zur Theories der Desinfektion. I. *Zeitschrift fur Hygiene und Infectionskrankenheiten* 1907; **57:** 388–404.
9 Chick H. The process of disinfection by chemical agents and hot water. *Journal of Hygiene* 1910; **10:** 237–286.
10 Arrhenius S. Quantitative Laws in Biological Chemistry. London, 1915.
11 Eijkman C. Die Ueberlebungskurve bei Abstotung der Bakterien durch Hitze. *Biochemische Zeitschrift* 1908; **11:** 12–20.
12 Hewlett RT. The Milroy Lectures of disinfection and disinfectants. *The Lancet* 1909; March **13:** 20–27.
13 Reichel H. Zur theorie der desinfektion. I, II, and III. *Biochemische Zeitschrift* 1909; **22:** 149–177, 177–199, 201–232.
14 Loeb J, Northrop JH. On the influence of food and temperature upon the duration of life. *Journal of Biological Chemistry* 1917; **32:** 103–121.
15 Brooks SC. A theory of the mechanism of disinfection, hemolysis and similar processes. *Journal of General Physiology* 1918; **1:** 61–80.
16 Peters RA. Variations in the resistance of protozoon organisms to toxic agents. *Journal of Physiology* 1920; **54:** 260–266.
17 Smith JH. The killing of Botrytis spores by phenol. *Annals of Applied Biology* 1921; **8:** 27–50.
18 Smith JH. The killing of Botrytis cinerea by heat, with a note on the determination of temperature coefficients. *Annals of Applied Biology* 1923; **10:** 335–347.
19 Shackell LF. The relation of dosage to effect. II. *Journal of Pharmacology and Experimental Therapeutics* 1925; **25:** 275–288.
20 Shackell LF, Williamson W, Deitchman MM, Katzman GM, Kleinman BS. The relation of dosage to effect. *Journal of Pharmacology and Experimental Therapeutics* 1924/1925; **23/24:** 53–65.
21 Buchanan RE, Fulmer EI. Physiology and Biochemistry of Bacteria. Vol. Effects of Environment Upon Microorganisms. The Williams and Wilkins Co.: Baltimore, MD, 1930.
22 Clark AJ. Handbook of Experimental Pharmacology. Verlig Von Julius Springer: Berlin, 1937.
23 Blau M, Altenburger K. Uber einige Wirkungen von Strahlen. II. *Zeitschrift fur Physik* 1922; **12:** 315–329.
24 Crowther JA. The action of x-rays on *Colpidium colpoda*. *Proceedings of the Royal Society of London B* 1926; **100:** 396.
25 Holweek F, Lacassagne A. Sur le mecanisme de l'action cytogaustique des radiations. *Comptes Rendus Societe Biologie de Paris* 1930; **103:** 766–768.
26 Wyckoff RWG. The killing of certain bacteria by x-rays. *Journal of Experimental Medicine* 1930; **52:** 435–447.
27 Pugsley AT, Oddie TH, Eddy CE. The action of X-rays on certain bacteria. *Proceedings of the Royal Society of London B* 1935; **118:** 276–298.
28 Parkard C. The biological effects of short radiations. *Quarterly Review of Biology* 1931; **6:** 253–280.
29 van Wijngaarden L. *Archiv fur Experimentelle Pathologie und Pharmakologie* 1924; **113:** 40 [cited in 22].
30 Morrell CA, Chapman CW. On the determination of the toxicity of neoarsphenamine I. The increase in toxicity on exposure to air. *Journal of Pharmacology and Experimental Therapeutics* 1933; **48:** 375–390.
31 Morrell CA, Chapman CW. On the determination of the toxicity of neoarsphenamine II. The determination of the characteristic curve for rats. *Journal of Pharmacology and Experimental Therapeutics* 1933; **48:** 391–409.
32 Trevan JW. The error of determination of toxicity. *Proceedings of the Royal Society of London B* 1927; **101:** 483–514.
33 Klaassen C (editor). Casserett and Doull's Toxicology: The Basic Science of Poisons, Fifth Edition, McGraw-Hill: New York, 1996.
34 Bliss CI. Estimating the dosage-mortality curve. *Journal of Economic Entomology* 1935; **25:** 646–647.
35 Bliss CI. The calculation of the dosage-mortality curve. *Annals of Applied Biology* 1935; **22:** 134–167.
36 Bliss CI. The comparison of dosage-mortality data. *Annals of Applied Biology* 1935; **22:** 307–333.
37 Bliss CI. The toxicity of poisons applied jointly. *Annals of Applied Biology* 1939; **26:** 585–615.
38 Bliss CI. The relation between exposure time, concentration and toxicity in experiments on insecticides. *Annals of the Entomological Society of America* 1940; **33:** 721–766.
39 Bliss CI. Biometry in the service of biological assay. *Industrial Engineering and Chemistry* 1941; **13:** 84–88.
40 Bliss CI. The calculation of microbial assays. *Bacteriological Reviews* 1956; **20:** 243–258.
41 Bliss CI. Some principles of bioassay. *American Scientist* 1957; **45:** 449–466.
42 Bliss CI, Cattell M. Biological Assay. *Annual Review of Physiology* 1943; **5:** 479–539.
43 Bliss CI, Packard C. Stability of the standard dosage-effect curve for radiation. *American Journal of Roentgenology and Radium Therapy* 1941; **46:** 400–404.

44 Gaddum JH. The action of adrenalin and ergotamine on the uterus of the rabbit. *Journal of Physiology* 1926; **61:** 141–150.
45 Gaddum JH. Lognormal distributions. *Nature* 1945; **156:** 463–466.
46 Gaddum JH. Bioassays and mathematics. *Pharmacological Reviews* 1953; **5:** 87–134.
47 Finney DJ. The application of the probit method to toxicity test data adjusted for mortality in the controls. *Annals of Applied Biology* 1943/1944; **30/32:** 68–74.
48 Finney DJ. The choice of a response metameter in bio-assay. *Biometrics* 1949; **5:** 261–272.
49 Finney DJ. Probit Analysis, A Statistical Treatment of the Sigmoid Response Curve. 2nd. Edition, Cambridge Univ. Press: London, 1952.
50 Bryan WR, Shimkin MB. Quantitative analysis of dose-response data obtained with carcinogenic hydrocarbons. *Journal of the National Cancer Institute* 1940; **1:** 807–833.
51 Mantel N, Bryan WR. 'Safety' testing of carcinogenic agents. *Journal of the National Cancer Institute* 1961; **27:** 455–470.
52 Reed LJ, Berkson J. The application of the logistic function to experimental data. *Journal of Physical Chemistry* 1929; **33:** 760–779.
53 Berkson J. Application of the logistic function to bio-assay. *Journal of the American Statistical Association* 1944; **39:** 357–365.
54 Berkson J. Why I prefer logics to probate. *Biometrics* 1951; **7:** 329–339.
55 Emmens CW. The dose/response relation for certain principles of the pituitary gland, and of the serum and urine of pregnancy. *Journal of Endocrinology* 1940/1941; **2:** 194–225.
56 Yule GU. On the distribution of deaths with age when the causes of death act cumulatively, and similar frequency distributions. *Journal of the Royal Statistical Society* 1910; **73:** 26–38.
57 Rahn O. Der Einfluss der Temperature und der Gifte aluf Enzymwirkung, Garung und Wachstrum. *Biochemische Zeitschrift* 1915; **72:** 351.
58 Rahn O. Physiology of Bacteria. P. Blakiston's Son and Co.: Philadelphia, 1932.
59 Southam CM, Ehrlich J. Effects of extracts of western red-cedar heartwood on certain wood-decaying fungi in culture. *Phytopathology* 1943; **33:** 517–524.
60 Clark AJ. The reaction between acetyl choline and muscle cells. *Journal of Physiology* 1926; **61:** 530–546.
61 Clark AJ. The antagonism of acetyl choline by atropine. *Journal of Physiology* 1926; **61:** 547–556.
62 Clark AJ. Mode of Action of Drugs on Cells. Arnold: London, 1933.
63 Hollenberg MD, NexΦ E. Receptor binding assays. In: Jacobs S and Cuatrecasas P (eds). *Membrane Receptors. Series B*, Vol. 11. Chapman and Hall: London, 1981, pp. 1–22.
64 Parascandola J. Origins of the receptor theory. In: Lamble JW (ed). *Toward Understanding Receptors.* Elsevier/North-Holland: New York, 1981, pp. 1–7.
65 Robison GA. Forward. In: Lamble JW (ed). *Towards Understanding Receptors.* Elsevier/North-Holland: New York, 1981, pp. v–x.
66 Hueppe F. Principles of Bacteriology. Translated from the German by E.O. Jordon. The Open Court Publishing Company: Chicago, 1896.
67 Boyd LJ. A Study of the Stimile in Medicine. Boericke and Tafel: Philadelphia, 1936.
68 Schulz H. Uber Hefegifte. *Pfluger's Archiv fur die gesemmte Physiologie* 1888; **42:** 517–541.
69 Townsend CO. The effects of ether upon the germination of seeds and spores. *Botanical Gazette* 1899; **27:** 458–466.
70 Townsend CO. The effect of hydrocyanic acid gas upon the germination of seeds. *Scientific American Supplement* 1899; **48:** 20010–20011.
71 Townsend CO. The effect of hydrocyanic acid gas upon grains and other seeds. *Botanical Gazette* 1901; **31:** 241–264.
72 Branham SE. The effects of certain chemical compounds upon the course of gas production by Baker's yeast. *Journal of Bacteriology* 1929; **18:** 247–284.
73 Smith EC. Effects of ultra-violet radiation and temperature on Fusarium. II. Stimulation. *Bulletin of the Torrey Botanical Club* 1935; **62:** 151–164.
74 Stebbing ARD. A theory for growth hormesis. *BELLE Newsletter* 1997; **6:** 1–11.
75 Watkins JH, Winslow C-EA. Factors determining the rate of mortality of bacteria exposed to alkalinity and heat. *Journal of Bacteriology* 1932; **24:** 243–265.
76 Falk IS. The role of certain ions in bacterial physiology. A Review. (Studies of salt action. VII). *Abstract of Bacteriology* 1923; **7:** 33–50; 87–105; 133–147.
77 Hotchkiss M. Studies on salt action. VI. The stimulating and inhibitive effect of certain cations upon bacteria growth. *Journal of Bacteriology* 1923; **8:** 141–162.
78 Salle AJ. Influence of Environment Upon Bacteria. In: *Fundamental Principles of Bacteriology with Laboratory Exercises.* McGraw-Hill Book Company, Inc.: New York, 1939.
79 Clifton CE. Death of bacteria. In: *Introduction to Bacterial Physiology.* McGraw-Hill Book Company, Inc.: New York, 1957.
80 Lamanna C, Mallette MF. Basic Bacteriology: Its Biological and Chemical Background. 3rd Ed. The Williams & Wilkins Co.: Baltimore, 1965.
81 Marshall MS, Hrenoff AK. Bacteriostasis. *Journal of Infectious Diseases* 1937; **61:** 42–54.
82 Clark AJ. Discussion on the chemical and physical basis of pharmacological action, 12 November, 1936. *Proceedings of the Royal Society of London B* 1936/1937; **121:** 580–609.
83 McCallan SEA, Wilcoxon F. The fungicidal action of sulphur. II. The production of hydrogen sulphide by sulphured leaves and spores and its toxicity to spores. *Contributions to the Boyce Thompson Institute* 1931; **3:** 13–38.
84 Hart RW, Frame LT. Toxicological defense mechanisms and how they may affect the nature of dose-response relationships. *BELLE Newsletter* 1996; **5:** 1–16.

This article has been reproduced in *Human & Experimental Toxicology* with the approval of *Toxicologic Pathology* and the Society of Toxicologic Pathologiests

Environmental and Molecular Mutagenesis 52:595–606 (2011)

Commentary

Key Studies Used to Support Cancer Risk Assessment Questioned

Edward J. Calabrese*

Environmental Health Sciences Division, Department of Public Health, School of Public Health and Health Sciences, University of Massachusetts, Amherst, Massachusetts

This paper reassessed studies conducted under the leadership of Drosophila geneticist Curt Stern which played a pivotal role in the acceptance of the linear dose-response model by the U.S. National Academy of Sciences Biological Effects of Atomic Radiation (BEAR) I Committee and the subsequent generalization of their recommendations on the linearity dose-response paradigm for ionizing radiation and chemically induced cancer. The analysis finds serious concerns and flaws in important aspects of these experiments, their assessment, and interpretation. Of particular concern was the failure of Stern's group to provide the necessary and promised experimental documentation to support the findings of three critical summarized experiments published as a brief technical note in *Science*. While this analysis questions the validity of the reported findings and their interpretations, it raises an even more serious concern about the process by which leaders in the radiation genetics community accepted such findings without requiring the necessary documentation and then used this information to support the acceptance of the linear dose-response in public policy matters as affected by risk assessment practices that have continued to the present. Environ. Mol. Mutagen. 52:595–606, 2011.

Key words: linearity; threshold; mutagenicity; ionizing radiation; risk assessment; dose–response

INTRODUCTION

The most important publication in the history of risk assessment was the 1956 report of the U.S. National Academy of Sciences (NAS), called the BEAR I report [U.S. National Academy of Sciences, 1956]. It achieved this distinction since it directly lead to a dose-response revolution, convincing governments worldwide to replace the threshold dose-response model for assessing the risks of ionizing radiation on germ cells with the linear dose-response model. The key conclusions of the BEAR I report that changed the dose–response default status from threshold to linear at low dose are embodied in the following two quotes on page 17 of that document:

''Any radiation dose, however small, can induce some mutations. There is no minimum amount of radiation dose which might be exceeded before any harmful mutations occur.''

''. . .if we increase the radiation that reaches the reproductive glands by *X* percent, the number of mutations caused by radiation will also be increased by *X* percent.''

These dose-response conclusions proved not to be restricted to a narrow biological question; they were generalized to radiation-induced cancer one year later by the National Committee on Radiation Protection (NCRP) (and soon thereafter most international advisory organizations) and then generalized again for all genotoxic chemical carcinogens by the U.S. NAS Safe Drinking Water Committee [National Academy of Sciences, 1977; Calabrese, 2009]. The key societal element in the linearity transformation was that the presumed safety of a threshold dose-response model was replaced with an ''acceptable risk'' concept of a linear dose-response model. With linearity as the guide for cancer risk assessment, no exposure to a

Grant sponsor: Air Force Office of Scientific Research, Air Force Material Command, USAF; Grant Number FA9550-07-1-0248.

*Correspondence to: Edward J. Calabrese, Environmental Health Sciences Division Department of Public Health, School of Public Health and Health Sciences, University of Massachusetts, Amherst, MA 01003, USA. E-mail: edwardc@schoolph.umass.edu

Received 27 May 2011; and in final form 1 June 2011

DOI 10.1002/em.20662

Published online 22 July 2011 in Wiley Online Library (wileyonlinelibrary.com).

carcinogen would be without risk, no matter how low or seemingly inconsequential. The present paper re-examines key publications directed by the eminent geneticist Curt Stern, upon which the linearity decision was, in large measure, based. The present analysis will demonstrate the presence of serious scientific and evaluation flaws in these papers along with a failure of the investigators to provide critical methodological and complementary research findings to validate their conclusions. This paper will also demonstrate that these flawed and unsupported findings were broadly accepted by key leaders in the genetics community and played an important role in the acceptance of linearity at low dose by the BEAR I Committee for radiation induced germ cell mutation and its generalization to cancer risk assessment for radiation and chemical carcinogens for use by federal and state regulatory agencies in the context of legislative requirements.

THE BEAR I COMMITTEE

The key issue for the BEAR I committee as it was created in April 1955 was making a scientifically based decision on whether radiation-induced germ cell mutation followed a threshold or linear dose-response model. As the BEAR I Committee began its deliberations, data from two biological models (i.e., fruit fly and mouse models) would have its primary interest. Considerable mutational data also existed on plants, some of which was directly related to the nature of the dose response. However, plant data were not given the same high priority as the animal findings in committee deliberations.

GERM CELL MUTATION: DROSOPHILA AND MICE

The fruit fly has long been a staple of mutational research, starting in 1910 when Morgan reported on naturally occurring mutations and the later application of this concept by Muller, who in 1927, showed that high doses of ionizing radiation induced germ cell mutations [Muller, 1927]. Research on the effects of ionizing radiation on the mouse genome, especially in the low dose zone, occurred later. For example, the U.S. federal government, under the Manhattan Engineering District (MED), initiated large scale studies during World War II, starting in 1943 under Donald Charles at the University of Rochester. What came out of this massive effort was disappointing as Charles published only a single 3.5 page note (lacking study methods) [Charles, 1950]. Former members of his group published a similarly brief review of the overall work in 1961, six years after Charles died of leukemia [Charles et al., 1961]. This substantial investment on mouse germ cell mutation research, involving more than 250,000 mice, had no noticeable impact on the BEAR I Committee. Starting much later was mouse research under Russell at Oakridge National Laboratories. While Russell generated data in the early 1950s, the dosages were high, in part due to model insensitivity, causing frustration about whether new and more sensitive models were needed to provide insight on the nature of the dose response in the low dose zone for radiation-induced germ cell mutation [Jolly, 2003]. According to Rader [2004], Russell's mouse studies by themselves were ''literally unusable for determining acceptable risk'' at the BEAR I meetings.

The focus of the scientific debate at this critical time over whether ionizing radiation induced germ cell mutations in a threshold or linear manner would be contested with fruit fly data obtained largely following treatment of mature sperm [Oftedal, 1964]. With its three decade head start Drosophila research had the most data and numerous geneticists studying ionizing radiation induced mutations. Thus, there was considerable insight into the nature of the animal model, its advantages and limits, how studies were designed, what they might yield, how they could be analyzed, and their relevance to humans. Based upon research with this Drosophila mature sperm model system, the genetics community would conclude and promote the belief that genetic damage caused by exposure to ionizing radiation was linear, cumulative, and deleterious. This perspective was confirmed, consolidated, and significantly extended in a series of highly influential studies under the direction of Curt Stern.

THE STERN STUDIES

The center of dose-response study was the laboratory of Curt Stern, a professor of genetics at the University of Rochester since 1939 before moving to the University of California at Berkeley in August 1947. The Stern studies were of considerable importance because they were part of a major funded activity of the MED/Atomic Energy Commission (AEC), were directed by a leader in the genetics community, were assessing responses at the lowest doses of ionizing radiation yet tested and offered the best opportunity to clarify the nature of the dose response for germ cell mutation. They would be key studies for not only the scientific community but governmental advisory and regulatory organizations in the U.S. and worldwide for occupational, medical, and environmental risk assessment and for broader atomic policy considerations.

There were three general research projects to assess the nature of the dose response for ionizing radiation under the direction of Stern, each lead by a different person. The first project was lead by Warren P. Spencer, a professor on leave from the College of Wooster, with nearly 20 years research experience with Drosophila. The second project was directed by a senior entomological behavioral geneticist researcher, Ernst Caspari. The third project was not originally planned but created after unexpected findings from the Caspari study which challenged a linearity dose-response interpretation. The third project was given

to a recent graduate of Russell Sage College, Delta E. Uphoff, a new master's student at the University of Rochester. The data collection of the three projects ran sequentially: Spencer's from December 1944 to June 1945, Caspari's from October 1945 to August 1946, while Uphoff's initial experiment, a partial replication of the Caspari experiment, ran from September 1946 to April 1947. During the summer of 1947, Uphoff [Uphoff and Stern, 1947] conducted another experiment at the University of Rochester, a ''chronic'' (i.e., 21 day) exposure to gamma rays. The final Uphoff experiment was performed at the University of California at Berkeley in the first half of 1948. Muller was an official consultant to the series of projects, providing the Muller-5 strain flies which were not susceptible to crossing-over genetic alternations. He also guided the group on breeding practices, data interpretation, and manuscript refinement.

SUPPORT FOR A THRESHOLD DOSE RESPONSE: THE CASPARI AND STERN STUDY

Caspari and Stern wrote a manuscript describing their investigations for the AEC [Caspari and Stern, 1947]. This manuscript was classified by the U.S. government until August 12, 1947, after which it was submitted in November 1947 to the journal *Genetics*, where Stern was the editor and published [Caspari and Stern, 1948]. While these two manuscripts of Caspari and Stern [1947 1948] are nearly identical, with no differences in the data tables, there are some important changes. In the Caspari and Stern [1947] paper, Hermann J. Muller is not included in the Acknowledgments but he is included in the Acknowledgments of Caspari and Stern [1948]. The most significant change between the two manuscript versions is that a key sentence in the Conclusion of the 1947-AEC [Caspari and Stern, 1947] version was dropped in the 1948 version [Caspari and Stern, 1948]. The sentence is a follows: ''*From the practical viewpoint, the results presented open up the possibility that a tolerance dose for radiation may be found, as far as the production of mutation is concerned*'' *(page 15)*. This change suggested support for a threshold dose response and will be discussed later. The Spencer research was published [Spencer and Stern, 1948] in *Genetics* following its declassification (i.e., March 5, 1947).

Spencer's research under the direction of Stern assessed the effects of X-rays on sex-linked recessive lethality in Drosophila males from short term (2–40 min) exposures (i.e., 10–96 r/hr). The cumulative doses ranged from a high of 4,000 r to the lowest then yet tested of 25 r. The study indicated a dose-response relationship that supported a linearity interpretation.

Caspari's research under the direction of Stern assessed the effects of gamma rays (i.e., radium needle) on Drosophila sex-linked recessive lethality. In contrast to the Spencer and Stern [1948] study, females were mated and exposed to radiation (2.5 r/day) for 21 days with sperm stored in the female's spermatheca. Cumulative doses were similar between the studies, both ~50 r. The females were fed a diet that suppressed egg laying during the 21 day irradiation exposure and then placed on a diet and altered environmental conditions to facilitate egg laying and development. The Caspari and Stern study using the aged sperm showed no statistically significant treatment effect, findings supporting a threshold rather than linear model. This finding created a significant problem for Stern and Muller, both strong advocates of a linear at low dose risk assessment model.

SUPPORT FOR THE THRESHOLD MODEL IS A "PROBLEM"

The findings of Caspari and Stern [1948] were of considerable interest since they provided support for the threshold dose-response model at the lowest dose rate of ionizing radiation yet tested. A draft manuscript was shared with Muller in early November 1946. Muller immediately recognized their significance, challenging the linear perspective; within a week he suggested that this experiment be replicated even though he noted that he had no reason to dispute the work of Caspari [American Philosophical Society, 1946]. Despite his knowledge of these novel findings, the credibility of the investigators and his own role as a consultant to Stern's research, one month later, on December 12, 1946, Muller would deliver his Nobel Prize Lecture, unequivocally affirming the validity of the linear dose-response model and claiming still further that there was no excuse any longer to accept a threshold perspective [Muller, 1946]. During this same time period, Caspari also sent a copy of his findings to the Milisav Demerec, the influential head of genetics at the Cold Spring Harbor. According to Caspari, Demerec was not pleased that his data challenged a linearity perspective, even suggesting ways to circumvent this problem and to ''save the hit theory'' [American Philosophical Society, 1947a].

So strong were mounting concerns over the challenge to linearity that even Stern suggested to Caspari that his data must be in error, due to spuriously high control group values [American Philosophical Society, 1947b]. High control group responses would likely preclude the detection of the radiation induced mutation effect at such low dose rates. However, a search of the published literature by Caspari indicated that his mutation frequency was in agreement with the observations of others, including very experienced Drosophila geneticists. Stern was finally forced to withdraw his control group mutation frequency criticism after Muller provided him with a large body of control group data for aging Drosophila sperm, confirming the observations of Caspari.

The ''final'' version of the Caspari and Stern manuscript [1948], as noted above, removed the sentence in the Conclusions which had suggested a tolerance or

TABLE I. Differences Between Spencer/Stern (1948) and Caspari/Stern (1948)

Spencer/Stern	Caspari/Stern
Exposure: X-rays	Exposure: gamma rays (radium needle)
Animal Model: males exposed prior to mating	Animal Model: females exposed after mating
Exposure Duration: acute exposure (minutes)	Exposure Duration: chronic exposure (21 days)
Dose Rate: ~15,000-fold greater than Caspari	Dose Rate: ~1/15,000 of Spencer
Plastic vials to hold flies	Glass vials to hold flies
Temperature: 24°C	Temperature: 18°C
Diet: cornmeal molasses	Diet: honey yeast agar
Age (males): ≤7 days, most 2–4-days old	Age (males): ≥5 days
Controls poorly matched with treatment exposure period.	Controls closely matched with treatment exposure period.
Temperature Control: poor, highly variable based on external conditions.	Temperature Control: good
50-r treatment group: 2 groups with different dose rates and exposure periods all combined.	A single 50-r treatment group; all treated similarly.
Mold Control: used Moldex throughout study.	Possibly less Moldex used in the 21 day radiation exposure period due to the lower temperature (18°C vs. 25°C).
Not corrected for lethal clusters. If so, the treatment group (50 r) used would have had its mutation rate decrease by −8% versus 4% for controls.	Corrected for lethal clusters. No differences between control and treatment.
Control radiation exposure not given.	Control radiation exposure reported as 0.6 r.
50-r treatment group had 20,400 less flies than the Caspari experiment.	
The study was not designed to affect the occurrence of lethal clusters.	The study was designed to minimize the possibility of lethal clusters.
F_0 Breeding Protocol Differed: 40 females/40 males; females – 2-days old	F_0 Breeding Protocol Differed: 50 females/100 males; females ≤16-hours old.
Radiation Exposure Condition Differed: 20 males/capsule; no food in capsule	Radiation Exposure Condition Differed: 50 females/capsule; food in capsule.
Lethal Designation Protocol Differed: Used 6 heterozygote females in F_2 generation to identify lethality.	Lethal Designation Protocol Differed: Used 2 female heterozygotes in F_2 to identify lethality.
A single wild-type male offspring lead to a designation of a viable culture.	A single wild-type male offspring lead to a designation of a semi-lethal.

threshold for the production of ionizing radiation induced sex-linked recessive lethal mutations. In addition, the entire discussion strangely, but strategically, centered on why there was no radiation treatment effect in the Caspari and Stern study [1948], whereas there was no questioning the Spencer and Stern findings [1948] even though both studies used the same cumulative dose. While this is a legitimate area of inquiry, the authors knew from the start that these two studies were fundamentally different and not directly comparable. Despite the use of profoundly different research methodologies between the two studies, Caspari and Stern [1948] concluded by asserting that before their findings could be accepted it would be necessary to exclude all factors that differed between the Spencer and Stern [1948] and Caspari and Stern [1948] studies that could have lead to the lack of a treatment effect in the Caspari and Stern paper [1948]. This was an extraordinary statement. To notify the scientific community not to accept the findings of your research unless and until future research would convincingly demonstrate that none of the methodological differences between the two studies could account for the observation was highly unusual. It should be noted that more than six decades later most of the methodological differences remain unexplained. The extensive experimental differences between these studies are documented in Table I.

DIFFERENCES BETWEEN THE SPENCER AND CASPARI STUDIES

Spencer and Stern exposed male fruit flies to x-rays, with the entire dose administered over a few minutes, with the flies held in plastic vials and reared at 24°C. In contrast, the Caspari and Stern study exposed females with stored/aging sperm to the same cumulative dose over 21 days in stored glass vials at 18°C at a dose rate about 1/6,000 to 1/15,000 of the Spencer and Stern treated flies. The control data in the Spencer and Stern study [1948] were poorly matched for their 50-r treatment group (i.e., the key treatment group comparison with the Caspari and Stern study [1948]. In fact, Spencer and Stern [1948] had two 50-r treatment subgroups. The control mutation rate was averaged over 70 weeks of observation while the 50-r treatment subgroup 1 was only tested over 11 weeks, about 15.7% of the control group's experimental duration. The 50-r subgroup 2 was exposed over 45 weeks (i.e., 61% of the control group experimental duration). The two 50-r subgroups were administered the x-rays with a dose rate that inexplicably differed by about 1.7−2.5-fold. The mutation data from both 50-r subgroups were then combined into one 50-r treatment group for all group effects without providing any information about the responses of the two subgroups. Considerable data from Caspari and

Stern [1948] and Uphoff and Stern [1947] had demonstrated large monthly variation in control group mutation rates. Such a lack of matching of control and low treatment groups as well as the lack of matching by the two 50-r subgroups themselves calls into question the validity of the Spencer and Stern [1948] study for the low dose groups. These critical methodological limitations were never discussed by Spencer and Stern [1948] nor in any of the subsequent papers that have assessed the dose-response findings of Spencer and Stern [1948]. Despite a lack of an explicit discussion of how such methodological weaknesses affected their interpretations, Spencer and Stern [1948] nevertheless noted that ''*at low dosages in the range of 25 r and 50 r . . .the control mutations may equal or exceed in number those produced by the radiation* [by normal variation]. *It is therefore important to collect a large body of control data and to collect these data in so far as possible at the same time and with the same environmental and genetic conditions as the radiation data to reduce errors from control fluctuation to a minimum.*'' The comparison of the control and 50-r treatment group data by Spencer and Stern [1948] did not satisfy their own methodological criteria. In contrast to the Spencer and Stern study [1948], the control and treatment group data collections were properly matched in the Caspari and Stern study [1948]. Consistent with the above-quoted perspective of Spencer and Stern [1948], in the Caspari and Stern study [1948], the control group displayed higher mutation rates than the 52.5-r treatment group during three of the eight treatment months, indicating that natural background variation for this mutational endpoint can be larger than a possible treatment effect at low dose.

The Spencer and Stern study [1948] also used a fruit fly diet that was intentionally different than that used by Caspari and Stern [1948]. The Caspari and Stern diet 1948] was honey yeast agar while that of Spencer and Stern [1948] was cornmeal molasses agar. The reason for the change to honey yeast agar was to suppress egg laying during the 21-day radiation treatment. In addition to egg-laying suppression this diet also affected other parameters with differences in the percent of sterile females (31.0 $\pm$ 1.7 honey yeast agar vs. 42.2 $\pm$ 2.7 cornmeal molasses) and the average number of F1 females per culture (12.1 $\pm$ 0.42 honey yeast agar vs. 19.27 $\pm$ 1.27 cornmeal molasses). Furthermore, in the 11 weeks of collecting data on the 50-r treatment subgroup the x-ray machine was checked only once, yet at most other times in the study it was checked on a weekly basis. Since the authors mention the occurrence of errors in dosimetry, it is not possible to detect the degree of error over the 11-week exposure period for one of their 50-r treatment subgroups.

Other differences between the two studies should also be noted. The studies used opposing strategies for sperm selection. Caspari and Stern [1948] selected males that were $\geq$5-days old whereas Spencer and Stern [1948] selected for males that were $\leq$7-days old. The younger male selection would tend to yield sperm with higher mutation rates. Spencer and Stern [1948] noted that there was inadequate temperature control at different seasons of the year. The experimental test cultures used for the detection of lethals were held on open shelves in a laboratory cabinet. The authors noted that there were wide fluctuations in room temperatures at different times of the year (December–June). These temperature differences were said to have affected a number of processes, including, for example, the time for emergence of the flies in the culture. The temperature differences also clearly had important mutation implications as certain mutants can behave as semilethals at one temperature and as lethals at another temperature. According to these authors, failure to control temperature may adversely affect the accuracy of the mutation scoring. They noted that if the cultures had been reared at a constant temperature a larger group of semilethal and delayed emergence mutants with visible effects could have been included in the totals, increasing mutation totals by 10% or more. In contrast to this concern in the Spencer and Stern study [1948], temperature control was tightly maintained in the Caspari and Stern experiment [1948].

Another difference between these two experiments was that lethal clusters were identified and removed in the statistical analysis of the Caspari and Stern experiment [1948] as shown in [Uphoff and Stern, 1947, 1949], whereas this was not done in the Spencer and Stern study [1948] or even in the later reporting of its data [see Uphoff and Stern, 1949]. While adjusting for the lethal clusters had no effect on the outcome of the Caspari and Stern study [1948], it would have had an impact to reduce the treatment response in the Spencer and Stern study [1948] by several percent. However, the authors never indicated how such an adjustment would have affected hypothesis testing or its impact dose-response modeling. It should be noted that Caspari did raise the question of why his data were being adjusted for the occurrence of lethal clusters by Uphoff but those of Spencer were not in a letter to Stern [American Philosophical Society, 1949]. Stern never addressed the pointed question of why Spencer's research was treated differently only that it was a needed adjustment for his work. Finally, it is also possible that the two experiments used different amounts of mold suppressant (Moldex) since Spencer and Stern [1948] conducted their experiment at 24°C, whereas the Caspari and Stern [1948] study was performed at 18°C during the 21 days of irradiation.

Given the large number of differences in experimental protocol between the two studies as well as missing, mismatched, and recombined data in the Spencer and Stern [1948] paper, there is little justification to speculate on why the results in one study were different than the other,

while holding acceptance of the Caspari and Stern [1948] study in abeyance. Yet the discussion of Caspari and Stern [1948] created in effect a type of scientific ''Strawman'' yielding impossible to resolve questions before dose-response findings supporting a threshold could be considered ''valid.'' In retrospect, the discussion was misdirected and inappropriate. However, with Stern as the editor of *Genetics*, this manuscript, with its obfuscated discussion, was fast tracked into publication, all with the encouragement of Muller [American Philosophical Society, 1947c]. It also had the benefit (whether by design or accident) of ''buying time'' for Uphoff and Stern to complete a replication of the Caspari and Stern study and providing a diversion for Muller's Nobel Prize Lecture comments concerning the impossibility of a threshold response when he knew that data existed that could refute his statements.

UPHOFF'S REPLICATION OF THE CASPARI AND STERN STUDY

The findings of the Uphoff experiment that was designed to replicate the study of Caspari and Stern [1947] were included along with the summary data of two other experiments in a brief note in *Science* [Uphoff and Stern, 1949]. However, an insight into the thinking that motivated this research as well as how the authors viewed the data is seen in their 1947 publication for the AEC files [Uphoff and Stern, 1947]. Following the Caspari and Stern study [1947], Uphoff and Stern [1947] replicated that work as closely as possible except that the exposure to the gamma rays was over 24 hr rather than the 21 days of the Caspari and Stern [1947] study. In their data evaluation, Uphoff and Stern [1947] adjusted the control and treatment group responses of Caspari and Stern [1947] for lethal clusters. This resulted in a modestly decreased control group mutation rate from 0.2738 to 0.2489 and similarly so for the radiation treatment group mutation rate (i.e., 0.3118–0.2848). Such changes did not affect the statistical analysis/conclusions of the Caspari study but were performed to make the two studies as comparable as possible. In contrast to the Caspari and Stern experiment [1947, 1948], a significant treatment effect was reported by Uphoff and Stern [1947].

What could account for these two studies yielding a different finding besides the obvious differences in exposure rate? First, the mutation rates of the radiation treatment groups between the two experiments were similar (Caspari, 0.2848; and Uphoff, 0.2542), being within 11–12% of each other. The principal difference between the two studies was the control mutation rates (Caspari and Stern, 0.2438; Uphoff and Stern, 0.1682). The Uphoff and Stern [1947] control was nearly 40% lower than the Caspari and Stern control [1947].

Uphoff and Stern [1947] addressed these differences by first refloating the proposition that the controls of Caspari and Stern [1947] displayed an abnormally/spuriously high mutation rate. However, as discussed above, this was not supported in several studies [Rajewsky and Timofeeff-Ressovsky, 1939; Ray-Chaudhuri, 1944; Kaufmann, 1947] showing that as sperm age they develop more sex-linked recessive lethal mutations, with similar mutation rates as were observed by Caspari and Stern [1947]. Furthermore, since Muller sent Stern substantial data which supported the findings of Caspari and Stern [1947], the Uphoff and Stern criticism [1947] of Caspari and Stern [1948] was withdrawn. The data indicated that it wasn't that Caspari and Stern [1947] had an unusually high control group mutation rate but rather that Uphoff and Stern's control [1947] was unusually, perhaps aberrantly, low. Statistical testing revealed that it was significantly different from the Caspari and Stern findings [1948]. In their conclusion, Uphoff and Stern [1947] stated that ''*in view of the former results on chronic irradiation (i.e., the Caspari and Stern experiment* [1947], *as well as the fact that the control rate of the present report is unexpectedly low, a final interpretation of the results cannot be offered.*'' Of further interest in the report was the statement of potential observer bias (i.e., presumably Delta E. Uphoff and not Curt Stern). On page 3 (bottom), it is stated that ''*the earliest control value is particularly low, and the question may be raised whether at this initial stage of the project it may reflect a personal bias of the experimenter.*'' In the final analysis there was no clear answer whether there was bias to see a lower control to detect a treatment effect, lack of research experience or simply chance variation. However, in an ironic twist, it should be noted that in 1928 Muller reported his own failure to establish a control mutation rate for the same lethal recessive sex-linked trait, attributing it to ''inexperienced persons'' in his laboratory [Muller, 1928].

THE *SCIENCE* JOURNAL PUBLICATION: STERN'S META-ANALYSIS

The findings of Uphoff's three experiments were published as slightly more than a one-page technical note in *Science* [Uphoff and Stern, 1949]. The authors promised to provide a detailed follow-up paper with the study methods, analysis procedures, and other data. However, no follow-up paper was ever published despite Stern's continued leadership in the field. In contrast to Uphoff's initial experiment, the first of Uphoff's two remaining experiments attempted to closely replicate the earlier Caspari study, now using aged sperm within a so-called chronic exposure of 21 days. The treated flies showed a 0. 2834 mutation rate which was similar to the other 52.5-r exposures, whether the exposure was acute or

chronic. However, the control group mutation rate was again problematic, being inexplicably low at 0.1765. No manuscript has been found that was submitted to the AEC on experiment two during the fall of 1947 as was done with the acute experiment of Uphoff. The third and final Uphoff experiment followed the Caspari design except that the dose was about double (100 r vs. 52.5 r) that used by Caspari. This experiment displayed a control mutation rate of 0.2352, a value similar to the original Caspari and Stern mutation rate [1947]. It also reported a significant treatment effect consistent with a linear interpretation. Before we address the issue of Uphoff's aberrant controls, let us consider the third experiment.

Uphoff and Stern [1949] made a key assumption in the assessment of their 100-r follow-up experiment. They decided that its results could be most reliably compared against the observed mutation rate per r based on the Spencer and Stern study [1948]. These authors reported that such a mutation rate per r was about 0.002%. If this mutation rate were applied to the 100-r experiment of Uphoff and Stern [1949], the control mutation rate of 0.2352 would increase by about 0.20% to 0.4352, a value close to the observed 0.4658, a finding consistent with a linear model. This interpretation played a pivotal role in the conclusions offered in the Uphoff and Stern paper [1949].

The critical flaw in this interpretation of Uphoff and Stern [1949] was due to the numerous experimental differences between it and the Spencer and Stern study [1947] that could affect mutation rates. The best choice for a standardized comparison for the Uphoff and Stern [1949] 100-r experiment would have been the Caspari and Stern paper [1948] since it was not only essentially identical to that of the 100-r experiment of the Uphoff and Stern study [1949] it also had considerably greater statistical power.

While supporting most strongly a threshold interpretation, the data of Caspari and Stern [1948] was not inconsistent with a linear model. That is, there was a nonstatistically significant increase in the mutation rate of about 11.5% (based on the seven months of direct matching of control and treatment groups). If the 11.5% increase was assumed to reflect a linear dose-response relationship (although not detectable as a significant treatment effect in hypothesis testing) and applied to the 100-r study, a mutation rate of 0.3100 would have been predicted based on its control value of 0.2352. These estimates are far below the 0.4352 mutation rate of Uphoff and Stern [1949]. In fact, the rate of increase of Uphoff and Stern is 2.66-fold greater than that predicted for a linear response using the Caspari and Stern study [1948], an increase that would be highly unlikely at this dosage [Edington, 1956]. Such an aberrant response could have reflected a third case of an ''unexpected'' response by Uphoff. Two earlier experiments resulted in ''unexpectedly'' aberrant control group values, enough so that both Uphoff and Stern [1947] even discounted their attempt to challenge the Caspari experiment. These data also need to be viewed within a framework that their documentation was never published. Thus, there has never been a bonafide basis for relying on it.

Despite these experimental inconsistencies, Uphoff and Stern [1949] concluded that ''*it appears*'' that irradiation at low dosage induces mutations in fruit fly sperm. One sentence later they reached the general conclusion that there is ''*no threshold below which irradiation fails to cause mutation*.'' Given the importance of this conclusion the scientific bases were too limited and inadequate. Just how did Uphoff and Stern [1949] come to this conclusion.

To achieve this goal, Stern had to marginalize Caspari and Stern [1948] even further than attempted in its own constraining discussion while at the same time repackaging the discredited Uphoff and Stern [1947] paper. This latter point was important for two reasons. First, a revitalized Uphoff and Stern paper [1947] would discredit Caspari [Caspari and Stern, 1948]. Secondly, if Uphoff and Stern [1949] could convince their peers that the aberrantly low control group was normal variation, it would also provide support for the second experiment of Uphoff, again with an aberrantly low control group. Uphoff and Stern [1949] failed to disclose or cite their Uphoff and Stern [1947] AEC paper which concluded that these data were un-interpretable due to the aberrantly low control mutation rates. This 1947 AEC paper also had asserted the validity of the Caspari and Stern [1947] control group mutation rate. Their 1947 AEC paper [Caspari and Stern, 1947] was truncated in AEC archives with negligible or no circulation within the scientific community. In fact, their 1947 AEC paper has never been cited until now. Of importance is that the mutation rate with the aging sperm was estimated to rise by 0.05–0.08%/week [Uphoff and Stern, 1947]. Over three weeks this would yield an increase of 0.15–0.24% on top of a background of ~0.10%. This would lead to an estimated mutation rate of 0.25–0.34%, a value consistent with the Caspari data. These findings and the interpretation of their 1947 AEC paper were disregarded by Uphoff and Stern [1949]. They then used the *Science* paper to assess whether the Caspari and Stern [1948] paper failed to detect a treatment effect due to ''errors of sampling,'' a nuanced version of spuriously high control values. Thus, the same ''challenge the control group'' strategy was used again, although it had been previously asked and addressed in detail about a year before with no new findings having emerged to challenge it. Next came the major change in their perspective. Uphoff and Stern [1949] now asserted that their controls were not aberrant but part of the norm. In this way, their uninterpretable data became interpretable. They were now able to minimize and even discredit further the findings

of Caspari and Stern [1948] while promoting the acceptance of their discredited experiment and enhancing the linearity interpretation.

The question was whether Stern would be able to pull off this scientific slight of hand without any debate or controversy. It is a bit like science's version of the famous ''follow the pea game.'' In this case we are trying to follow the studies. In fact, he did pull it off. Subsequently published papers by leading researchers indicated that he was indeed successful in this strategy. For example, Higgins [1951] stated that low levels of radiation produced mutations in fruit fly sperm and that the apparent inconsistency of previous results (i.e., Caspari and Stern [1948] were due to different experimental techniques and ''errors in sampling'' by Caspari and Stern [1948]. No evidence was produced to support this now accepted conclusion. The use of the term ''errors in sampling'' was lifted straight from Uphoff and Stern [1949]. It is not clear where the basis of the statement of errors in experimental techniques came from as there was no documentation to support it. Likewise, Singleton [1954] stated that Uphoff and Stern [1949] demonstrated that the controls of Caspari and Stern [1948] were spuriously high and that there actually was a treatment effect at 2.5 r/day. This statement of Singleton [1954] was of considerable value as he had been the leading opponent of the linearity perspective within the geneticist community. In effect, this statement by Singleton [1954] cleared any remaining path of resistance to the acceptance of Uphoff and Stern [1949] just prior to the convening of the BEAR I Committee. The data however clearly contradicted the conclusion of Caspari and Stern [1948]. In fact, it is odd that Caspari never responded with a letter to the editor to present the data that would have easily countered this view. Uphoff and Stern [1949] had now, in effect, rid themselves of the Caspari ''problem,'' as Stern referred to the Caspari and Stern [1948] findings, in a letter to Edward Noviski [American Philosophical Society, 1948]. They made it ''disappear'' by wrapping it in a 1949 version of a meta-analysis of only their data, redefining what was the norm for control variation and never providing data to support this position.

The continued success of the Stern strategy may be seen in publications after the completion of the BEAR I Report. The future Nobelist EB Lewis in 1957 published a profoundly influential article in the journal *Science* arguing that ionizing radiation induced leukemia with a linear dose response. In this paper, Lewis [1957] used the findings of Stern and his colleagues to reinforce his position as follows:

> Gene mutation has long been known to show a linear relationship with respect to dose of ionizing radiation from studies with Drosophila. This linearity has been extended by Spencer and Stern [1948] to doses of 50 and 25 roentgens. Gene mutation is also known to be directly proportional to the accumulated dose of radiation, even when the radiation is chronically administered at a relatively low dose rate, as in the studies of Uphoff and Stern [1949].

A further comment was published by James V. Neal [1958], member of the BEAR I Committee. He also used the Stern findings to assert his belief in the linearity dose-response paradigm. He stated that ''*In 1927, Muller for the first time clearly demonstrated the mutagenic effects of ionizing radiation. The first work on this subject was done with rather considerable doses of radiation, but, during the past decade, it has been demonstrated to the satisfaction of most geneticists that in the fruit fly, Drosophila, the mutagenic effects of radiation extend to doses as low as 25 to 50 r Spencer and Stern* [1948]; *Uphoff and Stern* [1949]. *Inevitably in work of this type the question of a 'threshold' arises. For technical reasons, it is rather difficult and extremely laborious to study the genetic effects of x-ray doses very much lower than 25 r in higher organisms, and no one has clearly demonstrated the mutagenic effects of doses below this level. On the other hand, in the face of all the evidence concerning the straight line relationship between dosage and mutation production, to me today the burden of proof is clearly on him who assumes that there is a threshold as regards the mutagenic property of x-rays.*''

Similar perspectives have been expressed by other leaders in the radiation research community. For example, writing in *JAMA*, Norwood [1958] noted that ''*Spencer and Stern, using more than 50 million flies, showed that genetic damage was proportional to dosage in the important range of 25 to 50 r.*'' It is also notable that none of these authors who highlighted the Spencer and Stern [1948] and Uphoff and Stern [1949] research ever acknowledged the findings of Caspari and Stern [1948] in accordance with the recommendations of the discussion of Caspari and Stern [1948] that their findings will not be accepted until differences between their findings and Spencer and Stern [1948] be resolved.

CONCLUSIONS

General

(1). Curt Stern assumed that the linear dose-response model was accurate and critically important for public policy. Consequently, Stern directed his scientific efforts to ensure that experiments challenging a linear at low dose perspective would need a higher degree of scientific proof, being subjected to greater efforts at replication and more scrutiny than results that supported a linearity perspective.

(2). This conclusion is supported by (a) the decision to only replicate the findings of Caspari and Stern [1948] and not the Spencer and Stern paper [1948] which supported the linearity perspective, (b) the assertion that the Caspari and Stern findings could not be accepted until it could be determined why they differed from that of Spencer and Stern; (c) the repeated attempts to challenge the findings of Caspari under the assumption that the control group data was spuriously high despite substantial data to the contrary; (d) attempts to enhance the credibility, mask the criticism and further the acceptance of the series of Uphoff experiments; and (e) failure to adjust the Spencer and Stern study [1948] for lethal clusters as was the case for the Caspari and Stern research [1948].

(3). The actions displayed by Stern raise questions about whether and to what extent philosphopical/ideological perspectives may have influenced his science. The present analysis suggests that he used his very elevated reputation, his associations with other leaders in the genetics field, his relationship with key journals such as *Science*, and the complexity of his research to mask his intentions and activities. He was successful in achieving his goal of ensuring acceptance of the linear model via these multiple manipulations and obfuscations as they reinforced similar biases within the genetics community.

(4). The data from Spencer and Stern [1948] and Caspari and Stern [1948] were actually in close agreement on the nature of the dose response in the low dose zone, even though one more strongly supported linearity and the other a threshold interpretation. In both studies, it was clear that at the low doses tested they were close to the limits of detection of a treatment effect. In fact, Spencer and Stern [1948] noted that it was not uncommon for control mutation rates to exceed those seen at 25 and 50 r, due to background variation. In a similar fashion, in three of the eight months of the Caspari study, the controls displayed a higher mutation rate than the treatment group. These observations indicate that in this low dose area both studies found it difficult to distinguish treatments from controls. A treatment effect could become statistically significant when a control group yielded an uncharacteristically low value, something that could happen by chance. This possibly happened in the Uphoff replication of Caspari. The control response was about 40% lowered than expected, leading to the significant treatment effect. Although the control mutation rate was so low in the Uphoff replication experiment, Stern was initially committed to using it. However, the literature research of Caspari which disputed the Stern position and the surprising and copious data of Muller forced him to back down, even though only temporarily.

(5). While the evidence is circumstantial, it appears that Stern was determined to suppress the acceptance of the Caspari study. The discussion of Caspari and Stern [1948] was, in retrospect, a professional oddity despite its scholarliness, yet this discussion was endorsed by Muller, another strong proponent of linearity. When viewed within the framework of promoting the acceptance of linearity at low dose, the decision was another example of Stern placing a road block in the path of acceptance of the Caspari data while trying to appear reasonable and objective.

(6). Even the case of Spencer and Stern [1948], a study that most geneticists of that era could support, had serious methodological issues that challenge the validity of its low dose findings. Nonetheless, both the authors themselves and the genetics community failed to note weaknesses that are obvious in retrospect.

(7). The Stern papers represent a case study for assessing scientific findings within a broader societal context. Stern was an accomplished scientist but his actions suggest strong ideological tendencies. While this historical reassessment has academic interest, the principal significance is that the actions of Stern manipulated the scientific appraisal and the quality of the scientific record on the issue of the dose-response default model, the results of which were to change the course of risk assessment throughout the world for the next 60 years.

Specific

(1). The Caspari and Stern study [1948] supported a threshold dose response. The changing of its conclusion from its initial publication by the AEC [Caspari and Stern, 1947], and the misdirecting of its discussion appears to have been designed to prevent a fair-minded consideration of this possibility.

(2). In Uphoff and Stern's attempt [1947] to replicate the Caspari and Stern study [1947, 1948], the control group was ''unexpectedly'' low, leading the authors to deem the study findings as uninterpretable. This rejection of the study data by Uphoff and Stern [1947] was never revealed in subsequent publications of Stern and colleagues. It was to a large extent hidden in the archives of the AEC.

(3). The Uphoff and Stern technical note [1949] in *Science* included three experiments—two with aberrant control values and one with an aberrantly high treatment group response. The authors never explicitly addressed the aberrant responses and their earlier documentation (see Uphoff and Stern, 1947) which supported the aberrant control group interpretation. They simply reversed their position, without ever mentioning this action, incorporated the discredited data into their 1949 meta-analysis to support a linear-

ity interpretation [Uphoff and Stern, 1947]. Promised documentation to support these studies was never provided.

(4). Uphoff and Stern [1949] also incorrectly used the Spencer and Stern [1948] publication as the gold standard comparison for the Uphoff and Stern 100-r experiment to validate a linearity prediction. There was no justification for its use in this manner given the extensive protocol differences between the two studies.

(5). Important new limitations have been identified in the Spencer and Stern paper [1948]. This study failed to properly match control exposure periods with those of the key low dose comparison group (e.g., 50 r). It employed two 50-r subgroups with differing matching periods with the control and with different dose rates. Yet these authors combined the data of the two subgroups to make a single 50-r treatment group without providing any data to justify this course of action. Detailed control mutation rate data of Caspari and Stern [1948] and Uphoff and Stern [1947] demonstrated wide monthly variability and strongly supported the need to closely match control and treatment exposure periods. The Spencer and Stern study [1948] also explicitly reported poor temperature control which they indicated adversely affected the accuracy of mutation rates. The scoring criteria of Spencer and Stern [1947] for lethality was to some extent admittedly subjective. Yet, the scoring was not single or double blinded. Since there was a strong belief in the linear model and its affirmation, the possibility of a bias may be raised. These findings raise important questions concerning the validity of conclusions concerning low dose responses in the Spencer and Stern [1947] paper, yet they were ignored in the meta-analysis of Uphoff and Stern [1949]. The limitations of the Spencer and Stern paper [1948] may also reflect the high likelihood that this manuscript was published without peer review. This 68-page, highly detailed manuscript, was formally received by *Genetics* on November 25, 1947, and published in January 1948(6) The conclusion of Uphoff and Stern [1949] that ''*there is no threshold below which radiation fails to induce mutations*'' was insufficiently supported and therefore not justified.

Lingering Questions

Stern's (and Uphoff's) broad and unequivocal conclusion [Uphoff and Stern 1949] was as surprising as it was unjustified, possibly due to the significance of the scientific and societal implications. Trying to make sense out of the Stern conclusion is disquieting and may lead to discussion which can be both speculative and judgmental. There are some questions to consider:

- Why did Uphoff and Stern [1949] transition from a very tentative statement that low doses of radiation ''may'' increase mutation rate but then offer such an unequivocal general conclusion supporting linearity regardless of how low the dose was?
- Why did Stern determine it was best to publish a one page technical note in *Science*, knowing the experimental details that lay beneath a brief conclusionary oriented summary?
- How did Stern go from dismissing the findings of Uphoff and Stern [1947] in the replication of Caspari and Stern [1948] to including it positively in his weight of evidence perspective of Uphoff and Stern [1949]?
- Why didn't Uphoff and Stern [1949] inform the *Science* readership that the key data used in their analysis had been recently discredited by them after a detailed documentation of control group mutation incidence?
- Did Uphoff really agree with Stern on the key general conclusion? Or was she somehow forced to agree?
- Why didn't Caspari challenge the de facto dismissal/ignoring of his findings?
- Why didn't Stern raise any concern with the spuriously high treatment response in the 100-r Uphoff study?
- Why did Stern use the Spencer and Stern [1948] study as the gold standard for comparison with the Uphoff and Stern [1949] 100-r study, knowing the long list of experimental protocol differences between the two studies?
- Why did the ''replication'' study of Uphoff and Stern [1947] reduce the sample size by about 40% and expect to detect a treatment effect at low dose?
- Why didn't Uphoff and Stern [1949] follow through with their stated commitment to provide the detailed paper documenting their methods, materials, and other relevant data?
- Was there ever a real commitment to publish the follow-up detailed paper or was this simply part of a broader plan of floating the conclusion in the world's most visible scientific journal?
- Why did *Science* decide to publish a one page note and no methods, with only a promise by the authors that they would provide such information in a subsequent publication?
- Why did Stern apparently approve the highly unusual criticism of possible experimental bias by Uphoff possibly leading to the low control values [Uphoff and Stern, 1947]?
- Why didn't Spencer and Stern [1948] display their weekly/monthly mutation rates for the controls and treatments as did Caspari and Stern [1948] and Uphoff and Stern [1949]?
- Did Stern misuse his role of journal editor in the publication of the Caspari and Stern [1948] paper?
- Why did the genetics community accept the undocumented findings and general conclusions without demanding that Stern follow through with his commitment to provide the detailed paper?

- Why didn't Caspari and Stern [1948] list even more potentially important differences between the studies that they would claim have to be ''explained.''
- Wouldn't the threshold model still be unacceptable to Stern if the Uphoff and Stern study [1947] had supported the Caspari and Stern study [1948]? This would not have answered why the Spencer and Stern [1948] and the Caspari and Stern [1948] studies differed.
- Why would Stern apparently not submit the Spencer and Stern and the Caspari and Stern manuscripts for peer-review?

Final Thoughts

The scientific community had a major stake in the assessment of studies published by Stern and colleagues concerning the nature of the dose response of ionizing radiation in the low dose zone yet remained strikingly silent over this matter. Despite the centrality of these findings and their broad acceptance by the scientific community, there were numerous concerns with the goals of the experiments, the capacity of the study designs to resolve key scientific questions, how the data were and were not reported, the scientific basis of conclusions, the publication of these papers, and finally how such findings should affect policy and risk assessment activities. The Stern Affair deserves a more detailed reassessment based on its central importance in the development of risk assessment policies and procedures and whether these may have been affected by deceptive actions by esteemed scientific leaders of the genetics community.

While this article questions several important scientific decisions by Curt Stern, such questioning is not intended to challenge his ethics, as he was highly regarded by his peers for his scientific and personal integrity [Bern et al., 1985]. Furthermore, there are also important limitations in the historical record that lead to informed but, nonetheless, speculative interpretations about the basis of some of Stern's actions. This is especially the case since it is not possible for Stern to explain his actions/decisions. Nonetheless, the important and challenging questions remain about the judgments of Stern, raising the broader question of whether this series of scientific decisions were the result of what happens when one's science becomes affected by transcience concepts which in this case is what we now call the ''precautionary principle.'' This issue raises a further serious and general concern as ideology-driven science represents a type of ''intellectual'' virus that can undercut the integrity of data-driven processes needed to guide critical societal decisions, and it can do so very effectively in a disguised and difficult to discern manner, as appears to be the case with the history of low dose linearity.

REFERENCES

American Philosophical Society. 1946. Muller letter to Stern. Stern Papers, Muller File, November 12.

American Philosophical Society. 1947a. Caspari letter to Stern. Stern Papers, Caspari File, September 25.

American Philosophical Society. 1947b. Stern letter to Caspari. Stern Papers, Caspari File, Fall 1947, undated.

American Philosophical Society. 1947c. Muller letter to Stern. Stern Papers, Muller File, January 14.

American Philosophical Society. 1948. Stern letter to Edward Noviski. Stern Papers, Noviski File, March 19.

American Philosophical Society. 1949. Caspari letter to Stern. Caspari Papers, Stern File, January 27.

Bern HA, Constance L, Demspter M, Fristrom JW, Stern C. 1985. Zoology, genetics, in memoriam. California Digital Library.University of California, Berkeley.

Calabrese EJ. 2009. The road to linearity: Why linearity at low doses became the basis for carcinogen risk assessment. Arch Toxicol 83:203–225.

Caspari E, Stern C. 1947. The influence of chronic irradiation with gamma-rays at low dosages on the mutation rate in Drosophila melanogaster. MDDC-1200, U.S. Atomic Energy Commission. pp. 1–18. Hathi Trust Digital Library. Available athttp://www.hathitrust.org.

Caspari E, Stern C. 1948. The influence of chronic irradiation with gamma-rays at low dosages on the mutation rate in *Drosophila melaogaster*. Genetics 33:75–95.

Charles DR. 1950. Radiation induced mutations in mammals. Radiology 55:579–581.

Charles DR, Tihen JA, Otis EM, Grobman A. 1961. Genetic effects of chronic X-irradiation exposure in mice. Genetics 46:1

Edington CW. 1956. The induction of recessive lethal in *Drosophila melanogaster* by radiation of different ion density. Genetics 41:814–821.

Higgins E. 1951. Atomic radiation hazards for fish. J Wild Management 15:1–12.

Jolly JC. 2003. Thresholds of uncertainty: Radiation and responsibility on the fallout controversy. Dissertation. Oregon State University, Corvalis, Oregon.

Kaufmann BP. 1947. Spontaneous mutation rate in Drosophila. Am Nat 81:77–80.

Lewis EB. 1957. Leukemia and ionizing radiation. Science 125:965–972.

Muller HJ. 1927. Artificial transmutation of the gene. Science 66:84–87.

Muller HJ. 1928. The measurement of gene mutation rate in Drosophila, its high variability, and its dependence upon temperature. Genetics 13:279–357.

Muller HJ. 1946. Nobel Prize Lecture. Stockholm, Sweden, December 12.

National Academy of Sciences Safe Drinking Water Committee. 1977. Drinking Water and Health, National Academy of Sciences, Washington DC.

Neal JV. 1958. The delays effect of ionizing radiation. JAMA 166:908–916.

Norwood WD. 1958. Common sense approach to the problem of genetic hazard due to diagnostic radiology. JAMA 167:1928–1935.

Oftedal P. 1964. Radiosensitivity of *Drosophila spermatognia*. II. Protracted doses. Hereditas 51:13–30.

Rader KA. 2004. Making mice: Standardizing Animals for American Biomedical Research, 1900–1950. Princeton, NJ: Princeton University Press.

Rajewsky BN, Timofeeff-Ressovsky NW. 1939. Holen-stahlungund die Mutationsrate von *Drosophila melanogaster*, ZIAV 77:488–500.

Ray-Chaudhuri SP. 1944. The validity of the Bunsen-Roscoe law in the production of mutations by radiation of extremely low intensity. Proc Royal Soc Edinburgh 62:66–72.

Singleton WR. 1954. The effect of chronic gamma radiation on endosperm mutations in maize. Genetics 39:587–603.

Spencer WP, Stern C. 1947. Experiments to test the validity of the linear R-dose/mutation frequency at low dosage. Atomic Energy Commission (MDDC-765). pp. 68.

Spencer WP, Stern C. 1948. Experiments to test the validity of the linear R-dose/mutation at low dosage. Genetics 33:43–74.

Uphoff DE, Stern C. 1947. Influence of 24-hour gamma-ray irradiation at low dosage on the mutation rate in Drosophila. MDDC-1492, U.S. Atomic Energy Commission. pp. 1–6. Hathi Trust Digital Library. Available athttp://www.hathitrust. org.

Uphoff DE, Stern C. 1949. The genetic effects of low intensity in irradiation. Science 109:609–610.

U.S. National Academy of Sciences. 1956. Biological effects of atomic radiation. Washington, DC.

Food and Chemical Toxicology 81 (2015) 137–140

Contents lists available at ScienceDirect

Food and Chemical Toxicology

journal homepage: www.elsevier.com/locate/foodchemtox

Review

Cancer risk assessment: Optimizing human health through linear dose–response models

Edward J. Calabrese [a,*], Dima Yazji Shamoun [b], Jaap C. Hanekamp [c]

[a] School of Public Health & Health Sciences, Department of Environmental Health Sciences, Morrill I N344, University of Massachusetts, Amherst, MA 01003, USA
[b] Mercatus Center, George Mason University, 3434 Washington Blvd., Arlington, VA 22201, USA
[c] Roosevelt Academy, Lange Noordstraat 1, NL-4331 CB Middelburg, The Netherlands

ARTICLE INFO

Article history:
Received 27 February 2015
Accepted 18 April 2015
Available online 24 April 2015

Keywords:
Risk assessment
Dose–response
Hormesis
Linearity
LNT
Precautionary principle

ABSTRACT

This paper proposes that generic cancer risk assessments be based on the integration of the Linear Non-Threshold (LNT) and hormetic dose–responses since optimal hormetic beneficial responses are estimated to occur at the dose associated with a 10^{-4} risk level based on the use of a LNT model as applied to animal cancer studies. The adoption of the 10^{-4} risk estimate provides a theoretical and practical integration of two competing risk assessment models whose predictions cannot be validated in human population studies or with standard chronic animal bioassay data. This model-integration reveals both substantial protection of the population from cancer effects (i.e. functional utility of the LNT model) while offering the possibility of significant reductions in cancer incidence should the hormetic dose–response model predictions be correct. The dose yielding the 10^{-4} cancer risk therefore yields the optimized toxicologically based "regulatory sweet spot".

1. Introduction

The assessment of cancer risks from exposure to ionizing radiation and chemical carcinogens by regulatory agencies worldwide is typically performed via the use of linear at low dose modeling. The linear non-threshold (LNT) approach for cancer risk assessment was first proposed for cancer risk assessment by the U.S. National Committee for Radiation Protection and Measurement (NCRPM) in 1958, following the recommendation of the U.S. National Academy of Sciences (NAS) Biological Effects of Atomic Radiation (BEAR) I Genetics Panel to switch from a threshold to a linear model for assessing genomic risk from ionizing radiation in 1956 (Jolly, 2003; Whitemore, 1986).

The LNT approach was later adopted by regulatory agencies starting in the late 1970s assessing risks for chemical carcinogens in all media (e.g. air, water, food and soil) (National Academy of Sciences (NAS), 1977). The initial transition from the threshold to the LNT approach in the mid 1950s was made prior to the discovery of DNA repair, adaptive responses with chemical mutagens and ionizing radiation, apoptosis, pre-conditioning and the resurgence of the hormetic concept, all of which could affect the shape of the dose response in the low-dose zone. The clarification of different mechanisms of action for carcinogens has encouraged the development of cancer risk assessment methods that incorporate knowledge of species specificity and threshold. These approaches are often employed by the U.S. EPA and FDA and most European authorities for non-genotoxic carcinogens (Page et al., 1997; Whysner and Williams, 1992; Williams, 2001; Williams et al., 2012).

These developments have challenged the theoretical and mechanistic basis of the LNT, along with the recognition that epidemiological methods are in effect not capable of detecting risks below twice the normal background (Taubes, 1995). Furthermore, the massive mega-mouse study that used 24,000 animals was only able to estimate risk at the 1% level (ED01 study) (Bruce et al., 1981). Similar limitations were reported for a cancer bioassay study with >40,000 trout (Bailey et al., 2009). These methodological limitations along with the more recent developmental insights on the plethora of adaptive mechanisms that act at low doses have revealed limitations of the LNT model.

2. Developments

The dose–response model that has been shown to have biological plausibility, especially in the low dose zone, is hormesis, a biphasic dose–response. Current interest in hormesis can be traced back to the research of Thomas Luckey on radiation hormesis (Luckey, 1980) and on chemical hormesis by Tony Stebbing (Stebbing, 1982). These researchers stimulated the electric power utilities of Japan

* Corresponding author. School of Public Health & Health Sciences, Department of Environmental Health Sciences, Morrill I N344, University of Massachusetts, Amherst, MA 01003, USA. Tel: +1 413 545 3164; fax: +1 413 545 4692.
E-mail address: edwardc@schoolph.umass.edu (E.J. Calabrese).

http://dx.doi.org/10.1016/j.fct.2015.04.023
0278-6915/

and the U.S. to conduct the first hormesis conference in August, 1985. These three events reactivated interest in the hormesis concept.

Since the initial hormesis conference mentioned here, multiple books have been published on hormesis (Calabrese, 1992, 1994; Costantini, 2014; Elliott, 2008; Luckey, 1992; Mattson and Calabrese, 2010; Rattan and LeBourg, 2014; Sanders, 2010; Stebbing, 2011). Also, many chapters on hormesis in toxicology and pharmacology texts have been produced; hormesis has been the focus of more than a dozen conferences; multiple symposia at major society meetings have addressed hormesis. It is the subject of more than 2000 scientific publications in peer-reviewed journals, and the object of more than 30,000 citations in the Web of Science/Knowledge. Extensive documentations of hormetic dose responses have been summarized from a large and continuously updated database (Calabrese and Blain, 2005, 2009, 2011).

The hormetic dose–response was also found to make more accurate predictions than the LNT or threshold dose–response models in head-to-head comparisons using large, independent data sets (Calabrese and Baldwin, 2003; Calabrese et al., 2006, 2008). Detailed mechanisms of 400 hormetic dose responses have recently been summarized (Calabrese, 2013). Additionally, the hormetic dose response therefore has been demonstrated to be highly generalizable, being independent of biological model (i.e., phylogenetically diverse – from bacteria to humans; in vitro and in vivo), level of biological organization (i.e., cell, organ and organism), endpoint, inducing agent and mechanism.

3. Objective – Integration

Based on these features, it has been proposed that the hormetic dose–response should become the default model for risk assessment for both carcinogens and non-carcinogens. The hormesis database provides strong evidence that dose–response relationships for carcinogens (e.g., DDT, dioxin, multiple PAHs, ionizing radiation) and non-carcinogens typically display hormetic dose response patterns with similar quantitative features. While this line of argument has been made (Calabrese, 2004), this is not the purpose of this paper. The present paper proposes a "practical" and straightforward harmonization of both the LNT and hormetic models for cancer risk assessment. As is customary in such convergences, common ground is sought by various entities (e.g., regulatory agencies and regulated industries), while differences are still recognized and will remain unresolved for now.

We see the following reasons why integration of both models would be beneficial. First, if hormesis describes low-dose exposure impacts of chemicals/ionizing radiation more accurately than the LNT-model does, then the regulatory authorities should apply the best that the toxicological sciences have to offer. The hormetic dose response requires rigorous study designs in order to be properly evaluated, with large numbers of doses, with proper dose spacing, and often within a dose–time framework. When such data are available, the hormetic dose response has far outperformed the threshold and linearity dose response model for accuracy in estimating low dose effects (Calabrese and Baldwin, 2003; Calabrese et al., 2006, 2008).

Second, considering the developments in analytical chemistry, increasingly lower levels of chemicals can be detected. We have entered the realm of atto- (part per quintillion; 10^{-18}) and zeptomoles (part per sextillion; 10^{-21}) of detectable analytes (Pagnotti et al., 2011). Consequently, the unspoken 'logic' of the LNT-model infers that a 'clean bill of health' can never be truly given (Hanekamp et al., 2012). The technology-driven stringency of regulation in the context of the LNT-model can be attenuated with the aid of the biphasic dose–response model. As a result, regulatory expenditures will be reduced along with benefit optimization (Keeney, 1997).

Third, the biphasic dose–response model underscores the beneficial adaptability of organisms' responses to chemical exposure, whereby regulation that expresses the functional integration of both the LNT and hormetic models is better able to address society's fears of carcinogen exposure.

4. Integration – Roadmap

How then do we envision this integration, that is, the harmonization of the hormesis and LNT dose response models for cancer risk assessment? The reconciliation of these two divergent models can surprisingly be made in a direct and uncomplicated fashion.

1) The key aspect of the hormesis/LNT convergence is that when risks are based on chronic animal bioassay studies, the optimal protective effects (i.e., reduction in tumor incidence for the affected below the control group) is predicted to occur at the same dose at which the LNT predicts 10^{-4} risk.
2) To achieve this value, the hormetic-based approach would first estimate a 1% response from the animal bioassay via a BMD-type methodology. When this derived-dose is divided by factor of 100, it yields slightly less than a risk of 10^{-4}. This was shown to be the case for ten highly diverse data sets by Gaylor (1989). The hormetic risk assessment methodology of Calabrese and Cook (2005), which is optimized at the same dose that the LNT estimates a 10^{-4} risk level, predicts benefit while the LNT estimates enhanced cancer risk.
3) We propose that cancer risk assessment adopt an acceptable risk of 10^{-4} using the LNT model since this dose would also yield the optimal hormesis dose response benefit. This dose is the so-called regulatory "sweet-spot" that provides substantial protection against theoretical low dose risks that are far below the detection of even the most demanding epidemiological and toxicological studies/methods, while including benefits predicted by the hormetic dose response model (Fig. 1). This approach would also have the significant societal benefit of affecting a profound reduction in costs (i.e., financial and predicted adverse health), markedly affecting cost/benefit analyses.
4) In a population of one million people, the 10^{-4} risk predicts 100 people (i.e., 10^6 people $\times$ 10^{-4} risk = 100) affected with an organ-specific cancer (e.g., lung, kidney, bladder, etc.) by some deleterious agent that is added to the background for cancer of that organ (Fig. 1). Assuming a 25% tumor background

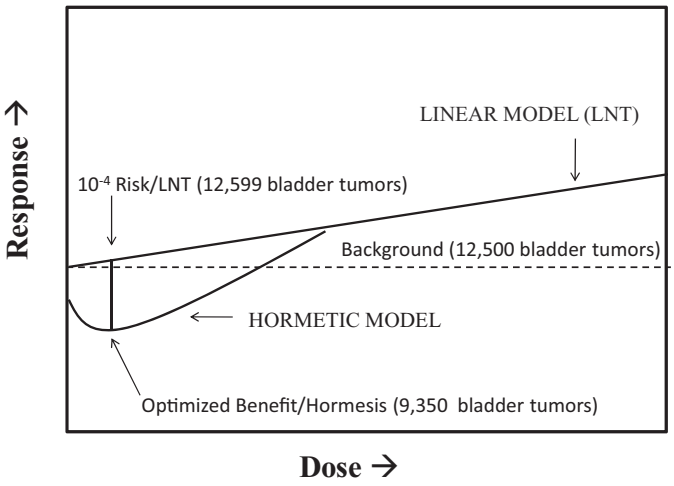


Fig. 1. Functional integration of hormesis and LNT for carcinogen risk assessment; derivation of the optimal regulatory strategy.

incidence, 250,000 of the one million people would be predicted to develop tumors. If the organ in question was responsible for 5% (e.g., bladder) of the above 25% (i.e., 250,000 people), it would represent 12,500 of the 250,000 people with cancer (i.e., $0.25 \times 0.05 = 0.0125$) ($0.0125 \times 10^6 = 12{,}500$). Many organ-specific tumors, including the bladder, affect about 3.5 to 6.0% of the tumor occurrence (National Cancer Institute (NCI), 2014), thus the use of 5% for an organ like the bladder would be a reasonable expectation. Organs affecting a notably higher proportion of people (e.g., about 16–18% per cancer type) are those cancers of the lung, breast and prostate. The 100 newly affected people with *chemically* induced bladder cancer are then randomly distributed among the entire population of one million. This suggests that 25% of the 100 will already be in the process of developing a background tumor, with about 5% of those already targeted for a "spontaneous" bladder tumor ($0.25 \times 0.05 = 1.25\%$). The net result of background (i.e. spontaneous) and tumor-induction via a chemical carcinogen at 10^{-4} is 12,500 ("background") plus 100 new chemically induced cases (i.e., $12{,}500 + 100 = 12{,}600$) minus 1 due to spontaneous and induced bladder tumors in the same individuals. This would yield a total of 12,599 individuals with bladder cancer. The hormetic benefit is likely to affect both background and induced tumor incidence, reducing their incidence by roughly 25% (Calabrese and Blain, 2011), lowering the predicted total number of affected people (12,500) by about 3150. There can be other situations in which the chemical may affect multiple organs with different tumor backgrounds and induced tumor incidence, affecting the nature and complexity of the assessment. For example, in the case of dioxin, it was shown in the Kociba et al. study (Kociba et al., 1978) that has been widely used for cancer risk assessment that hormetic effects appear to occur in multiple organs (i.e., Females: liver, ovary, uterus, cervix/vagina, mammary, pituitary and adrenal; Males: liver, pulmonary, pituitary, pancreas and adrenal). In such cases it may be possible to select that dose which displays the lowest overall tumor incidence for risk assessment purposes. In theory, this type of situation may be predicted to have a greater beneficial effect than described for the bladder cancer. However, it would not be unexpected for the optimal effect to vary by organ. Using a financial metaphor, the convergence of the LNT/10^{-4} risk and hormesis methodologies permits the protection of one's "principle" (i.e., impossible to detect chemically-induced increase in cancer risk) while adding considerable benefit (i.e., large reduction in cancer risk for those affected organs). This compromise strikes an optimized balance in which there is a very low theoretical risk increase and a very high theoretical benefit. Choosing a 10^{-6} acceptable risk would reduce 99 of the 100 theoretically affected people while eliminating the possible hormetic benefit. This type of strategy would prevent the possibility of beneficial effects, which could be substantial.

5) The example presented above addresses the risk of a single complete carcinogen. However, humans live in a highly complex environment involving exposure to a vast array of complete carcinogens, tumor promoters, chemoprotective chemicals and physical agents, all superimposed on dynamic metabolic processes, numerous adaptive mechanisms and complex exposure dynamics. Predicting cancer incidence of complex mixtures from experimental and epidemiological studies is problematic, if not impossible. A very limited, simplified and yet mechanistically oriented approach to assess complex carcinogenic mixtures is the toxic equivalent factor (TEF) that assumes additive processes that act identically (e.g. same receptor) for similarly grouped agents (e.g. dioxins, PAHs and PCBs). The TEF concept was integrated within a LNT perspective. Epidemiological evaluations of complex mixtures reveal the failure of predictions of animal studies to predict human responses. For example, a cup of coffee contains >1000 chemicals of which approximately 30 have been tested for cancer. Of these the majority were carcinogenic in standard rodent model testing. Each cup of coffee contains >10 mg of rodent carcinogens, with American adults drinking three cups per day (Ames and Gold, 2000; Gold et al., 1992). The situation gets more complex as more carcinogens are added via the roasting process. However, despite such exposures to natural and roasted process-related carcinogens, comprehensive epidemiological studies reveal neutral or beneficial effects from lifetime coffee drinking depending on the organ (Bohn et al., 2014; Crippa et al., 2014). Thirty-two occupational epidemiological studies (i.e. case–control – 19 studies; cohort – 13 studies) of gasoline exposure which is a highly complex and variable mixture of >500 saturated/unsaturated hydrocarbons revealed no pattern or clear association between gasoline and any cancer (Keenan et al., 2010). Furthermore, dose responses of complex mixtures [e.g. petroleum (Laughlin et al., 1981), waste-water treatment effluents (De Nicola et al., 2004; Mendoza-Figueroa, 1973; Walsh et al., 1982), complex organochlorine mixtures (Aube et al., 2011)] over a broad dose response often conform to an hormetic dose response. These findings support the conclusion that complex mixtures can induce hormetic dose responses and can be evaluated within the framework proposed here.

6) An important implication of model uncertainty is that it has the potential to undermine and challenge the use of LNT in toxic tort litigation cases. The acknowledgement of substantial and unresolved uncertainty in risk assessment may preclude causation judgments with low dose exposures. In fact, the use of LNT in toxic tort cases in the United States has been successfully challenged in numerous litigations affecting ionizing radiation, asbestos as well as chemical carcinogens, principally due to its lack of validation capacity, inconsistency with published findings and the recognition of substantial adaptive mechanisms that undermine an LNT interpretation (Milward v. Acuity Specialty Products Groups, Inc, 2013; Sutera v. Perrier Group of America Inc, 1997; Whiting v. Boston Edison Co, 1995).

5. Discussion

The search for public health common ground via the integration of opposing risk assessment models is a new approach in the process of risk assessment harmonization. It permits the strengths of opposing perspectives to be incorporated into a unified risk assessment approach. It is recognized that estimates of low risk is a speculative activity, especially when the data are derived from high dose toxicology studies and that there is no current practical way around this limitation. The present recommendation is viewed as substantially conservative, creating the opportunity to benefit from the induction of adaptive responses while recognizing and incorporating model uncertainty into the risk assessment process. We believe that this is a sound foundation upon which to base environmental public health policy.

The precautionary principle, which is at the core of modern governmental environmental health policies, is founded on a toxicological assumption that lower is always safer/better and that zero exposure, especially for carcinogens, is the goal [maximum contaminant level goal (MCLg)] as seen for EPA drinking water standards. The precautionary principle was strongly influenced during its formative development by belief in LNT predictions. Harmonizing of the LNT and hormesis dose response models can provide a vehicle not only for cancer risk assessment but also a novel means, along

with a more biologically based foundation, to guide a broad range of precautionary principle applications.

Conflict of interest

The authors declare that there are no conflicts of interest.

Transparency document

The Transparency document associated with this article can be found in the online version.

Acknowledgement

Long-term research activities in the area of dose response have been supported by awards (to EJC) from the U.S. Air Force (FA9550-13-1-0047) and ExxonMobil Foundation (S18200000000256) over a number of years. The U.S. Government is authorized to reproduce and distribute for governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing policies or endorsement, either expressed or implied. Sponsors had no involvement in study design, collection, analysis, interpretation, writing and decision to submit.

References

Ames, B.N., Gold, L.S., 2000. Paracelsus to parascience: the environmental cancer distraction. Mut. Res. 447, 3–13.
Aube, M., Larochelle, C., Ayotte, P., 2011. Differential effects of a complex organochlorine mixture on the proliferation of breast cancer cell lines. Environ. Res. 111, 337–347.
Bailey, G.S., Reddy, A.P., Pereira, C.B., Harttig, U., Baird, W., et al., 2009. Non-linear cancer response at ultra-low dose: a 40,800-animal ED001 tumor and biomarker study. Chem. Res. Toxicol. 22, 1264–1276.
Bohn, S.K., Blomhoff, R., Paur, I., 2014. Coffee and cancer risk, epidemiological evidence, and molecular mechanisms. Mol. Nutr. Food Res. 58, 915–930.
Bruce, R.D., Carlton, W.W., Ferber, K.H., Hughes, D.H., Quast, J.F., et al. (Members of the Society of Toxicology ED01 Task Force), 1981. Re-examination of the ED01 study: adjusting for time on study. Fund. Appl. Toxicol. 1, 67–80.
Calabrese, E.J., 1992. Biological Effects of Low Level Exposures to Chemicals and Radiation. Lewis Publishers, Michigan.
Calabrese, E.J., 1994. Biological Effects of Low Level Exposures: Dose-Response Relationships. CRC Press, Florida.
Calabrese, E.J., 2004. Hormesis: from marginalization to mainstream: a case for hormesis as the default dose-response model in risk assessment. Toxicol. Appl. Pharmacol. 197, 125–136.
Calabrese, E.J., 2013. Hormetic mechanisms. Crit. Rev. Toxicol. 43, 580–606.
Calabrese, E.J., Baldwin, L.A., 2003. The hormetic dose response model is more common than the threshold model in toxicology. Toxic. Sci. 71, 246–250.
Calabrese, E.J., Blain, R., 2005. The occurrence of hormetic dose responses in the toxicological literature, the hormesis database: an overview. Toxicol. Appl. Pharmacol. 202, 289–301.
Calabrese, E.J., Blain, R.B., 2009. Hormesis and plant biology. Environ. Poll. 157, 42–48.
Calabrese, E.J., Blain, R.B., 2011. The hormesis database: the occurrence of hormetic dose responses in the toxicological literature. Reg. Toxicol. Pharmacol. 61, 73–81.
Calabrese, E.J., Cook, R.R., 2005. Hormesis: how it could affect the risk assessment process. Hum. Exper. Toxicol. 24, 265–270.
Calabrese, E.J., Staudenmayer, J.W., Stanek, E.J., Hoffmann, G.R., 2006. Hormesis outperforms threshold model in National Cancer Institute antitumor drug screening database. Toxicol. Sci. 94, 368–378.
Calabrese, E.J., Stanek, E.J., Nascarella, M.A., Hoffmann, G.R., 2008. Hormesis predicts low-dose responses better than threshold model. Intern. J. Toxicol. 27, 369–378.
Costantini, D., 2014. Oxidative Stress and Hormesis in Evolutionary Ecology and Physiology: A Marriage Between Mechanistic and Evolutionary Approaches. Springer, New York.
Crippa, A., Discacciati, A., Larsson, S.C., Wolk, A., Orsini, N., 2014. Coffee consumption and mortality from all causes, cardiovascular disease, and cancer: a dose-response meta-analysis. Am. J. Epidemiol. 180, 763–775.
De Nicola, E., Gallo, M., Iaccarino, M., Meric, S., Oral, R., et al., 2004. Hormetic versus toxic effects of vegetable tannin in a multitest study. Arch. Environ. Contam. Toxicol. 46, 336–344.
Elliott, K., 2008. Is a Little Pollution Good for You? Oxford University Press, New York. 264 p.
Gaylor, D.W., 1989. Quantitative risk analysis for quantal reproductive and developmental effects. Environ. Health Perspect. 79, 243–246.
Gold, L.S., Slone, T.H., Stern, B.R., Manley, N.B., Ames, B.N., 1992. Rodent carcinogens: setting priorities. Science 258, 261–265.
Hanekamp, J.C., Kwakman, J., Pieterman, R., Ricci, P.F., 2012. The administrative ordering of nature and society – precaution and food safety at the molecular and global level. Eur. J. Risk Reg. 3, 313–325.
Jolly, J.C., 2003. Thresholds of Uncertainty: Radiation and Responsibility in the Fallout Controversy. Dissertation. Oregon State University, Oregon.
Keenan, J.J., Gaffney, S.H., Galbraith, D.A., Beatty, P., Paustenbach, D.J., 2010. Gasoline: a complex chemical mixture, or a dangerous vehicle for benzene exposure? Chemico-Biolog. Inter. 184, 293–295.
Keeney, R.L., 1997. Estimating fatalities induced by the economic costs of regulations. J. Risk Uncert. 14, 5–23.
Kociba, R.J., Keyes, D.G., Beyer, J.E., Carreon, C.E., Wade, D.A., et al., 1978. Results of a two-year chronic toxicity and oncogenicity study of 2,3,7,8-tetrachlorodibenzo-p-dioxin in rats. Toxicol. Appl. Pharmacol. 46, 279.
Laughlin, R.B., Ng, J., Guard, H.E., 1981. Hormesis – a response to low environmental concentrations of petroleum-hydrocarbons. Science 211, 705–707.
Luckey, T.D., 1980. Ionizing Radiation and Hormesis. CRC Press, Florida.
Luckey, T.D., 1992. Radiation Hormesis. Oxidative Stress and Hormesis in Evolutionary Ecology and Physiology: A Marriage Between Mechanistic and Evolutionary Approach. CRC Press, Florida.
Mattson, M.P., Calabrese, E.J., 2010. Hormesis: A Revolution in Biology, Toxicology and Medicine. Humana Press Inc, New York.
Mendoza-Figueroa, T., 1973. Growth and toxicity tests on the marine nonoplanktonic algae Platymonas astetrathele G.S. West in the presence of crude oil and emulsifiers. Environ. Poll. 4, 261–282.
Milward v. Acuity Specialty Products Groups, Inc, 2013. Civil action no. 07-11944-DPW. WL 4812425 (D. Mass. Sept. 6, 2013).
National Academy of Sciences (NAS), 1977. Drinking Water and Health. NAS, Washington DC, p. 939.
National Cancer Institute (NCI), 2014. SEER stat fact sheets. <http://seer.cancer.gov/statfacts/html/urinb/html>.
Page, N.P., Singh, D.V., Farland, W., Goodman, J.L., Conolly, R.B., et al., 1997. Implementation of EPA revised cancer assessment guidelines: incorporation of mechanistic and pharmacokinetic data. Fundam. Appl. Toxicol. 37, 16–36.
Pagnotti, V.S., Chubatyi, N.D., McEwen, C.N., 2011. Solvent assisted inlet ionization: an ultrasensitive new liquid introduction ionization method for mass spectrometry. Anal. Chem. 83, 3981–3985.
Rattan, S.I.S., LeBourg, E., 2014. Hormesis in Health and Disease. CRC Press, Florida.
Sanders, C.L., 2010. Radiation Hormesis and the Linear-No-Threshold Assumption. Springer, New York.
Stebbing, A.R.D., 1982. Hormesis – the stimulation of growth by low levels of inhibitors. Sci. Total Environ. 22, 213–234.
Stebbing, T., 2011. A Cybernetic View of Biological Growth: The Maia Hypothesis. Cambridge University Press, New York.
Sutera v. Perrier Group of America Inc, 1997. 986 F. Supp. 655, 666 (D. Mass.).
Taubes, G., 1995. Epidemiology faces its limits. Science 269, 164–169.
Walsh, G.E., Duke, K.M., Foster, R.B., 1982. Algae and crustaceans as indicators of bioactivity of industrial wastes. Water Res. 16, 879–883.
Whitemore, G.F., 1986. The National Committee on Radiation Protection, 1928–1960: From Professional Guidelines to Government Regulation. Department of the History of Science, thesis dissertation. Harvard University, Cambridge MA.
Whiting v. Boston Edison Co, 1995. 891 F, Supp. 12, 25 (D. Mass.).
Whysner, J., Williams, G.M., 1992. International cancer risk assessment: the impact of biologic mechanisms. Reg. Toxicol. Pharmacol. 15, 41–50.
Williams, G.M., 2001. Mechanisms of chemical carcinogenesis and application to human cancer risk assessment. Toxicology 166, 3–10.
Williams, G.M., Iatropoulos, M.J., Jeffrey, A.M., 2012. Dose-effect relationships for DNA-reactive liver carcinogens, chapter 12. In: Greim, H., Albertini, R.J. (Eds.), The Cellular Response to the Genotoxic Insult: The Question of Threshold for Genotoxic Carcinogens. The Royal Society of Chemistry, London, UK.

Arch Toxicol (2013) 87:2063–2081
DOI 10.1007/s00204-013-1105-6

REVIEW ARTICLE

How the US National Academy of Sciences misled the world community on cancer risk assessment: new findings challenge historical foundations of the linear dose response

Edward J. Calabrese

Received: 24 April 2013 / Accepted: 11 July 2013 / Published online: 4 August 2013

Abstract This paper extends several recent publications indicating that Hermann J. Muller: (1) Made deceptive statements during his Noble Prize Lecture on December 12, 1946, that were intended to promote the acceptance of the linear dose-response model for risk assessment for ionizing radiation and (2) that such actions of Muller were masked by a series of decisions by Muller's long-time colleague and esteemed radiation geneticist Curt Stern, affecting key publications in the mutation literature. Such actions further enhanced acceptance of the linearity dose-response model while preventing Muller's deceptions from being discovered. This paper provides documentation that Muller reinforced such practices within the scientific literature in the early 1950s, by supporting scientifically questionable actions of Stern. Detailed documentation is provided that demonstrates how these actions affected national and international risk assessment policy for ionizing radiation and chemical carcinogens via the recommendations of the National Academy of Sciences Biological Effects of Atomic Radiation committee in 1956, to adopt the linear dose-response model.

Keywords Mutation · Linearity · Dose response · Risk assessment · History of science · Muller

E. J. Calabrese (✉)
Department of Public Health, Environmental Health Sciences, University of Massachusetts, Morrill I, N344, Amherst, MA 01003, USA
e-mail: edwardc@schoolph.umass.edu

Introduction

It was recently discovered that the 1946 Nobel Prize Lecture for Biology and Medicine by Laureate Hermann J. Muller misled the audience on the nature of the dose response in the low-dose zone concerning the effects of ionizing radiation on germ-cell mutagenicity to advance an ideologically motivated risk assessment policy (Calabrese 2011a, b, 2012). Evidence to support this conclusion is found in Muller's own words from letters he sent to Professor Curt Stern of the University of Rochester, an expert in radiation genetics. Stern sent Muller a manuscript by Ernst Caspari and himself on November 6, 1946, for review as Muller was a paid consultant to the project (Calabrese 2011c). This manuscript demonstrated support for a threshold dose response, while challenging the linear dose-response single-hit mutagenicity mechanism model, based on an extensive study of ionizing radiation on mutation in the germ cells of male fruit flies. On November 12, 1946, Muller acknowledged receipt, noting that the findings strongly challenged the linearity dose-response concept and, given their importance, needed to be replicated as soon as possible (Calabrese 2011c). This long-term study used the lowest ionizing radiation dose rate yet reported. Despite this new information, Muller would go on to deliver his Nobel Prize Lecture some 5 weeks later (December 12, 1946), proclaiming that one could no longer consider the possibility of a threshold dose response for germ-cell mutagenicity. The only option, he argued, was to switch to a linearity dose-response model for risk assessment (Muller 1946a).

Muller, of course, made these public claims while knowing that the most extensive and relevant testing supported a threshold interpretation. A letter from Muller to Stern 5 weeks after the Nobel Prize Lecture (January 14, 1947)

confirmed his support for study replication, that he had no technical criticisms of the Caspari study, and supported publication especially in view of the caveats worked into the discussion, hopefully preventing acceptance of a threshold interpretation (Calabrese 2012; Lilly Library 1947a, January 14 letter). In effect, Muller told the Nobel Prize Lecture audience one story while in private correspondence he revealed a profoundly different view. According to his former student, friend, and colleague, Crow (1995), it was well known that Muller would try to win arguments by exaggeration and overstatement. Crow found this behavior exasperating as Muller would often end up hurting his case by unnecessarily misrepresenting facts and circumstances, incorrectly thinking it would help him win his argument. This same behavioral trait was evident at the Nobel Prize Lecture.

Before his Nobel Prize Lecture, Muller sought to raise concern over the public health implications of ionizing radiation and to change the risk assessment process for ionizing radiation from the use of a threshold dose-response model to the far more conservative linear dose response. This goal was essentially shared by the entire radiation geneticist community. Following his Lecture, Muller would now have two goals: Protecting his reputation by ensuring that his misleading comments would not be discovered while still aggressively pushing acceptance of the linearity agenda. Both goals were entangled; being such an important scientist and leader any fall in Muller's status would have a devastating impact on the acceptance of the linearity dose response, especially if it involved an ideological misrepresentation about the linearity concept. Muller achieved both goals due to decisions of Stern that discredited the findings of his colleague and co-author Ernst Caspari, thus saving Muller from criticisms about his Nobel Prize Lecture while supporting the questionable findings of Delta Uphoff, another co-author. Muller's misleading comments and the Stern's apparent data obfuscations would not be revealed for more than 60 years while the linearity acceptance goal by regulatory agencies worldwide was attained. The present paper extends the recent reports of Calabrese (2011a, b, 2012) with newly discovered findings that demonstrate a carefully focused and timed set of inexplicable scientific judgments by Muller concerning the nature of the dose response. These actions reinforced his Nobel Prize Lecture comments and the actions of Stern that enhanced the goal of achieving a switch from threshold to linearity. This paper also demonstrates the profound impact of the Stern/Muller actions on the radiation genetics community based on the scientific publication record and dose–response recommendations/conclusions supporting a linearity dose–response risk assessment model by the highly influential NAS BEAR I Committee, Genetics Panel.

Part 1—Stern's plan to promote linearity

Curt Stern was a long-time supporter of the idea that ionizing radiation affected germ-cell mutation in a linear dose-response manner. He expected that this would be observed in studies he was directing under the aegis of the Manhattan Project using fruit flies. While a linearity dose-response was reported in acute studies with X-rays (Spencer and Stern 1948), the most significant test would take place with the research of Ernst Caspari when gamma radiation would be administered up to a 13,200-fold lower rate than in the Spencer research. In a troubling development, Caspari reported to Stern that his findings did not support a linear interpretation but rather a threshold dose response. Based on letter correspondence between Stern and Caspari, Stern initially refused to accept this interpretation, arguing that the mutation threshold response was most likely due to unusually high control group values (i.e., spontaneous mutations in sperm stored in the spermatheca of the female for 3 weeks) which masked a radiation-induced treatment effect (Calabrese 2011b). Caspari then researched this issue by exploring the literature and obtaining substantial unpublished data on this specific issue from Muller based on research during his appointment at Amherst College (1940–1945). Caspari argued that his control group mutation data were not aberrant but consistent with the literature and Muller's data for aged sperm whether stored in the spermatheca of the female or in the male. As a result of the Caspari analysis, Stern withdrew his objection and accepted the conclusion that the control group spontaneous mutation values were within the normal range. Since Stern could not dismiss the findings of Caspari due to the controls, he then opted for an alternative but bizarre strategy to marginalize the threshold dose-response conclusion. Stern directed the manuscript discussion to explain why these data should not be accepted and utilized until it was determined why Caspari's findings differed from those of Spencer and Stern's acute study which they claimed supported linearity. It was this manuscript of Caspari that was sent to Muller for review just prior to his Noble Prize Lecture.

It is odd that investigators reporting on striking new findings, using the most advanced methods and the lowest dose rate yet studied, would demand the reader not take the data seriously. Stern placed no such restriction upon the Spencer paper, a study with considerable methodological limitations [e.g., inadequate control groups, inappropriate data combining for statistical analysis, lack of adequate X-ray instrumentation calibration, poor temperature control, and dose rates differing by as much as 10-fold (10 and 100 r/min) between treatments, thereby creating two experimental variables within one experiment] (Calabrese 2011b). Furthermore, there were at least two dozen significant methodological differences between the two studies

making them not directly comparable. Stern published the manuscript (Caspari and Stern 1948) with its misdirected discussion, without apparent independent, peer review in the journal for which he was the editor, that is, *Genetics*.

Comment

Based on this temporal sequence, it would appear that the principal driving force to challenge the Caspari findings that supported a threshold interpretation was his advisor and co-author, Curt Stern. It was Muller who indicated that the findings of Caspari needed to be replicated since they were contrary to a linear single-hit dose-response interpretation. Of particular note, however, was that the only changes made to the Caspari manuscript following the review of Muller was to add the name of Muller to the acknowledgments section and to remove the statement from the conclusion that the findings supported a tolerance or threshold interpretation (Calabrese 2011b).

Part 2—the replication studies

Since Ernst Caspari and Warren Spencer were no longer available to continue experimentation, Stern engaged the services of a Master's student, Delta Uphoff, to assess why the Caspari study did not support a linear interpretation. The results of the initial experiment were deemed by Stern as not usable as her control group spontaneous mutation rate was strikingly low, being outside the expected range for aged sperm (~40 % lower than expected); no conclusions could be drawn from the study (Uphoff and Stern 1947). A similar very low control group spontaneous mutation rate response for aged sperm in her second experiment would also make such data uninterruptable. In her third and final experiment, Uphoff reported control values in the normal range for aged sperm but the radiation treatment response was itself aberrant, far exceeding predicted responses assuming low-dose linearity (Calabrese 2011b).

Stern: What to do next

Finding a way to support linearity was the prevailing theme. For example, when Caspari had shared his data with the Head of Genetics at the Brookhaven National Laboratory and future member of the BEAR I Committee/Genetics Panel, Milislav Demerec, he wrote to Caspari asking what can be done to save the single "hit" linearity dose-response paradigm (Calabrese 2011b; American Philosophical Society 1947f, September 25). The "hit theory" for ionizing radiation-induced mutation was first postulated by Timoféeff-Ressovsky et al. (1935), providing a theoretical mechanistic foundation for the LNT dose-response model. Given his goals and ideology, Stern had little choice. Another experiment was not going to be practical as Uphoff would leave for a position with the NIH. In the absence of new data, Stern decided upon a new strategy to "save" the single-hit linearity dose response. In order to achieve this goal, he would have to do two things: (1) Reverse his position on the Uphoff control group data, declare that they are normal, not aberrant, making the Uphoff experiments now interpretable and (2) challenge further the credibility and acceptance of the Caspari study (i.e., beyond the misdirected discussion of the Caspari/Stern paper). Stern took the bold action of asserting that the Uphoff control group data were part of the normal distribution. He offered no explanation or assessment of the literature to justify this conclusion. This would not be difficult as only very few people would have known about his earlier concerns with the Uphoff control group data, since the manuscript (Uphoff and Stern 1947) detailing such concerns was never submitted for publication but was placed in the Atomic Energy Commission (AEC) archives, initially as a classified manuscript. Thus, the written critique of the Uphoff control group data and letter communications on this topic were generally not known or available.

The Uphoff and Stern (1949) paper also raised a number of doubts about the Caspari paper such as whether its non-treatment effect/threshold finding was the result of "errors in sampling." Given standard professional protocol, the "errors in sampling" hypothesis was a surprising and unexpectedly harsh challenge to the work of Caspari, a University of Rochester team member, especially since this criticism had never been raised previously by Stern, Muller, or others in previous detailed evaluations. In fact, there was never any documentation to support this possibility. Further, Stern also raised the specter of the Caspari control being elevated by unnecessarily stating that his control group was higher than each of the controls of the three Uphoff experiments. Stern neglected to state that two of the Uphoff studies had aberrantly low control group values based on the published literature and Muller's data. This decision by Stern would now make the Uphoff experimental data "interpretable," whereas several months before he judged it as "uninterpretable." Also, the third Uphoff experimental control data were indistinguishable statistically from the Caspari control (0.2489 vs. 0.2352 %). Such actions helped to achieve the above-stated goals of enhancing the credibility of the Uphoff data while marginalizing the Caspari findings.

The Uphoff and Stern (1949) paper changed the way the Caspari data (Caspari and Stern 1948) were perceived and accepted by members of the scientific community. Below are quotes from several papers (Higgins 1951; Singleton 1954a, b) and a dissertation (Jolly 2004) that address very

clearly how the Uphoff and Stern (1949) paper marginalized the research of Caspari. Of particular significance is that the judgments drawn by each of these papers were factually and interpretationally incorrect.

Higgins (1951) stated that "Uphoff and Stern (1949)…concluded that low-level radiation does produce mutations in fruit-fly sperm and that the apparent inconsistencies of previous results were due to different experimental techniques and errors in sampling" (page 10, column 1).

Singleton (1954a) stated that "Caspari and Stern (1948) studying chronic gamma radiation found no increase over controls for doses of 2.5 r/day for 21 days. However, it was later documented by Uphoff and Stern (1949) that the controls used by Caspari and Stern had an abnormally high sex linked lethal frequency and that actually there was an effect of the chronic gamma radiation of 2.5 r/day." (page 599)

Jolly (2004) stated (1) that "Stern and Caspari initially detected no significant difference in the mutation rates on the controls and the irradiated flies, though later they corrected for experimental errors and got a statistically significant difference." (pages 78–79) (2) "The results of Stern's initial experiment failed to support the linear hypothesis for genetic injury. Assuming that something must have been wrong with the experiment, he eventually identified experimental errors, which, when corrected for, supported linearity." (pages 80–81).

Caspari's control group data were therefore once again challenged by Stern; the once aberrantly low controls of Uphoff were now seen as being in the normal range. With these changes, the dose response of the collective grouping of the Stern Drosophila experiments would appear linear. This is the conclusion of what Uphoff and Stern published in their one-page technical note in the 1949 *Science* article summarizing the Spencer and Stern (1948) and Caspari and Stern (1948) papers and the three Uphoff experiments. This 1949 paper, as noted above, did not include mention that the previous conclusions (Uphoff and Stern 1947) about the Caspari and the Uphoff control groups that had been reversed by Stern and the role of the Muller data assessment in the decision-making process. Since the Uphoff and Stern (1949) brief technical paper lacked any information on research methods and other relevant data, the authors promised a detailed follow-up publication to correct this critical limitation, a promise never fulfilled. Given the lack of information provided in the *Science* paper and the prestige of this journal, it raises a question about the circumstances surrounding its publication within this context. It should be noted that Hermann J. Muller's first graduate student (i.e., H. Bentley Glass) became an editor at *Science* in 1948, only months prior to the submission of the Uphoff and Stern manuscript. Glass also had a relationship with Stern with whom he had been awarded a National Research Council post-doctoral fellowship at the Kaiser Wilhelm Institute in Berlin (Erk 2009). Since Glass was an expert on *Drosophila* radiation genetics, it is likely that he oversaw the evaluation of the manuscript. One must also question to what extent Muller/Stern may have exploited their relationship with Glass to facilitate the publication of such a limited paper and used the journal to advance an ideological perspective.

Muller's post Nobel Prize dose–response comments about the Caspari and Stern (1948) study

Muller's statement

In his 1950 article entitled "Some present problems in the genetic effects of radiation" in the *Journal of Cellular and Comparative Physiology* Muller (1950a) provided an explicit characterization of the Caspari and Stern (1948) findings. Muller stated on page 10 "A recent paper by Spencer and Stern…….extends the principle (i.e., one-hit principle) down to total doses of 50 r and 25 r." In the next paragraph, he stated: "It is true, in a parallel paper… .Caspari and Stern have reported results somewhat deviating from the above."

Comment

Muller trivialized the significant challenge of the Caspari study to the linearity dose-response paradigm. The key Muller phase concerning the Caspari data is "somewhat deviating". The Spencer and Stern (1948) study involved an acute exposure, that is, all doses of radiation were administered within a few minutes to a few hours. In contrast, the Caspari and Stern (1948) study provided the same total dose as in the Spencer and Stern study but spread over 21 days, at a dose rate up to 13,200-fold lower. The "somewhat deviating" results were such that at the lower dose rate of the Caspari and Stern study, the data supported a threshold interpretation, not the expected linear proportionality response. Muller was quite concerned with the Caspari study as it represented a potentially significant challenge to linearity, repeating this perspective in letters (Lilly Library 1947a, January 14; American Philosophical Society 1946, November 12) to Stern and emphasizing the need to replicate this study, despite the requirement for additional funding and the efforts of multiple scientists and staff for about 1 year. It is also important to note that Muller never mentioned any of the numerous methodological/analysis limitations/flaws of the Spencer and Stern (1948) in any of his publications.

Muller's statement

In footnote 1 on page 10 of the above-cited article, Muller (1950a) stated that "Uphoff and Stern have published a report of further work, with doses as low as 50 r, given an intensity as low as 0.0165 r per minute. The results obtained are entirely in conformity with the one-hit principle. A consideration of these results, together with the early work, leads to the conclusion that the deviation first referred to (the Caspari and Stern 1948 findings) was caused by a value for spontaneous mutation rate that happened to be unusually high."

Comments

Muller claims that the research of Delta Uphoff and Curt Stern is "entirely in conformity with the one-hit principle" (Timoféeff-Ressovsky et al. 1935). What Muller neglected to state was: (1) Uphoff's first experiment displayed an aberrantly low control group response based on Muller's own extensive data involving some 200,000 fruit flies (Muller 1946b). A letter from Curt Stern to Ernst Caspari (undated) (American Philosophical Society Undated, circa July-Aug 1947) addressed the control group issue. It states: "The radiation data continues to be puzzling. Delta's difference between control and exper[imental group] appears to be due mainly to a much lower control group value than yours. However, Muller informs me that his data give an aged control value close to yours. Thus, my first idea that your results could be "explained away" by assuming that your control value happened to be unusually high, seems unlikely. Rather does Delta's control appear too low. Well, we'll have to meet." Muller provided this information to Stern twice in letters dated February 3, 1947, and August 4, 1947 (Lilly Library 1947b, c). It should be noted that the occurrence of increased mutations in aged sperm in the control group as reported by Caspari was not a new concept to Stern. In fact, when Timoféeff-Ressovsky first presented such data in the late 1930s, Stern corresponded with Demerec specifically addressing these findings. These letter exchanges reveal not only Stern's knowledge of the findings, but also of his knowledge that the findings had been subsequently replicated (Lilly Library 1938a, b, c). The report of Rajewski and Timofeeff-Ressovsky (1939) on this topic would most likely have considerable scientific weight as Timoféeff-Ressovsky was on par with Muller for scientific reputation in the area of radiation genetics.

In the Atomic Energy Commission (AEC) manuscript by Uphoff and Stern (1947) concerning her replication of the Caspari study, the low response control group issue was explicitly addressed as follows in their "Discussion" section. "In his extensive studies on the effect of aging on the mutation rate in sperm, H.J. Muller (unpublished) has found a weekly increase of about 0.07 % for sex-linked lethals in various stocks kept at 25 °C. At 18 °C, the temperature used for aging in the laboratory, the weekly increases may be assumed to be slightly less, perhaps 0.05 %. Taking a value of 0.10 %, similar to that of Spencer and Stern's control rate, for sperm before aging, the expected control rate after aging should be approximately 0.25 %. This figure is much closer to the control rate observed by Caspari and Stern than to that found in the present work." In their acknowledgments of this manuscript, Uphoff and Stern stated that "we are very grateful to Dr. H. J. Muller for his permission to quote from his unpublished data." Thus, Muller would have known that his research was used to evaluate the reliability of the Caspari and Uphoff control groups. The control group response of Uphoff and Stern (1947) was sufficiently low such that they stated that the data were uninterpretable (i.e., "a final interpretation of these results cannot be offered."). Uphoff and Stern (1947) explicitly raised the possibility that the low control group values "may reflect a personal bias of the experimenter." The manuscript did not identify whether the bias concern statement was directed to Stern, Uphoff or both, or the type of bias. (2) Uphoff's second experiment also displayed a similarly aberrant low control group response, likewise affecting the possible utility of the data. (3) The third (and final) Uphoff experiment obtained control values in the normal range but an aberrantly high treatment response, even assuming a linearity dose response (see Calabrese 2011a for a detailed evaluation). "Appendix" section provides the temporal letter exchange between Stern and Muller on the key question of control group mutation frequency upon which the acceptance of the Caspari and Uphoff studies are based.

Muller (1950b) discredits the conclusion of Caspari and Stern (1948) by asserting that the control group values were unusually high. (1) Muller failed to state that the "high" control value of Caspari and Stern (1948) was first put forward as a criticism by Stern in the fall of 1946, when Caspari informed Stern that his findings supported a threshold, rather than a linearity interpretation. (2) He also did not report that Caspari successfully rebutted Stern by presenting data on control group responses from published studies in the literature and from unpublished data provided by Muller himself. Muller failed to state that he had published a summary of the mutation rate of sperm stored in the spermatheca for several weeks (Muller 1945). This is the information that he sent to Stern that supported the reliability of the Caspari control group data and marginalized the Uphoff study control group (see "Appendix" section). Later studies by Muller and his student Helen L. Byers at the University of Indiana also supported the Caspari control group mutation frequency (Byers 1954; Byers and Muller 1952). Nonetheless, Muller (1954b) would inexplicably continue his criticism of the Caspari and Stern (1948)

study, repeating the "unusually high control frequency" (page 476) conclusion as a basis to reject its challenge to linearity. The question may be raised as to why Muller would directly contradict himself on such a serious matter and never be exposed to criticism. While any answers to this question must be speculative, Sankaranarayanan and Wassom (2008) unequivocally state that Muller was an "unquestioned authority," suggesting that it would be quite difficult to challenge him or even consider doing so.

It should be noted that in early 1949, Muller became concerned that Robley Evans of MIT was publishing a paper in the journal *Science* on the mutagenic effects of ionizing radiation and the nature of the dose response in the low-dose zone. Muller had reviewed the manuscript prior to publication and was upset that Evans had given credibility to the Caspari and Stern (1948) paper. Muller wrote to Stern (Lilly Library 1949, February 5) requesting that Stern contact Evans and try to convince Evans to withdraw his support for the Caspari and Stern (1948) findings. There is no evidence that Stern did this based on correspondence records. However, it is possible that the subsequent attack of Muller (1950a, b) on the Caspari and Stern (1948) findings was stimulated by this Evans paper (1949) which would need to be "neutralized."

Muller (1954b) also further criticized the Caspari and Stern (1948) paper in a vague manner as being "more doubtful than the others on some other grounds" (page 476), which he never clarified. Such criticism may have referred to the fact that Uphoff and Stern (1947) introduced a modified method of counting sex-linked recessive lethals, one that was different than reported by Caspari and Stern (1948) and also different than Spencer and Stern (1948). Uphoff and Stern (1947) recounted (i.e., adjusted) the Caspari and Stern (1948) data with the new counting method in order for it to be as directly comparable to their study as possible. The results of those adjustments were deemed by Uphoff and Stern to be insignificant in their 1947 paper, resulting in control and treatment responses that were, in fact, even more similar than before the adjustment (i.e., without a treatment effect). The published paper of Caspari and Stern (1948) did not incorporate this adjustment (perhaps resulting in the veiled criticism of Muller 1954a, b), whereas the Uphoff and Stern (1947) manuscript presented the original and adjusted data; only these adjusted data were used for the Caspari and Stern (1948) data as summarized in the 1949 paper in *Science* by Uphoff and Stern. Regardless, the adjustment for differing lethality estimation techniques did not affect the study interpretation. In a letter on February 9, 1949, to Caspari in anticipation of the *Science* publication, Stern (American Philosophical Society 1949, February 9) stated that "It will be shown below (the *Science* manuscript) that the difference in defining a lethal is of no significance in the evaluation of the results."

In his 1950 papers, Muller never addressed any of these critical issues that might affect a decision on the nature of the dose response (Muller 1950a, b). He also failed to state that the Uphoff and Stern (1949) paper was only a one-page summary, has very low control group values, no presentation of research methods and that Uphoff and Stern (1949) promised to publish a detailed paper with all the missing methods and data but had not (and never did). By discrediting the Caspari and Stern (1948) paper and restoring the Uphoff data, Muller was able to protect his scientific reputation, his ethical standing and to give strong support to the linearity single-hit theory dose-response model.

In a second paper in 1950 entitled *Radiation Damage to the Genetic Material* in the *American Scientist*, Muller (1950b) used the findings of Stern and his colleagues to extend "the principle of proportionality of mutation frequency to dose down to doses of 50 r and 25 r and of less than 0.001 r per minute, with a time-intensity relation differing by over 400,000 times from that of our high intensity dose."

Comment

By using the now revitalized data of Uphoff, Muller made the claim of linearity over a 400,000-fold dose range. This was a major conclusion as it gave an assertion of linearity at low dose by a Noble Prize winner who had great authority within the field. Furthermore, Stern (1960) continued to affirm the findings of Uphoff and Stern (1949) in the second edition of his acclaimed genetics textbook, published in English, German, Japanese, Polish, Russian, and Spanish (American Philosophical Society 1973, November) (autobiographical statement), by stating that the dose rate had no impact on the mutation incidence in *Drosophila*, whether administered acutely or given "slowly and continuously, that is, 'chronically,' given over a long period." In order for Stern (1960) to have reached this conclusion, he had to diminish the findings of Caspari and Stern (1948) and accept those of Uphoff and Stern (1949). A further note is that the Muller (1950b) paper contradicted his 1950a paper on the dose rate: The two papers used a different lowest dose rate: 0.001 r/min (Muller 1950b) versus 0.00165 r/min (50 r/30240 min in 21 days) (Muller 1950a)—a 65-fold difference. Muller (1950b) rounded down the 0.00165 r/min rate to 0.001 r/min, increasing the extrapolation range from approximately 250,000- to 400,000-fold. Why Muller rounded the numbers down is not known, nor was it necessary. Secondly, if rounding was to occur it would normally have been rounded up to 0.002 r/min. This action of Muller reveals an effort to exaggerate the linear extrapolation range. Third, Muller (1950b) makes an error in his statement that the linearity was shown with a dose rate "less than 0.001 r per minute" when the actual value was 0.00165 r/min.

Table 1 Hermann J Muller and Curt Stern quotes on low-dose linearity

References	Quote
Muller (1948)	Page 462 "…the frequency of the mutations induced will be proportional to the total dose of radiation received over an unlimited period of time." "There is then absolutely no threshold dose, unlike what is true of many other biological effects of radiation, and even the most minute dose carries a definite chance of producing mutations—a chance exactly proportional to the size of that dose."
Muller (1952)	Page 317 "In making our calculations it is safe, as both the earlier (6–10) and the more recent (11–15) works have agreed, to accept the principle that the frequency of the gene mutations produced is simply (linearly) proportional to the amount of the total accumulated dose received, as expressed in r units. Moreover, as some of these same studies show, this relation holds within wide limits, regardless of how short and concentrated or dilute and protracted the exposure may have been, or whether it was given in one treatment or many." "There are good theoretical grounds for inferring that these principles hold true no matter how small the total dose, or the dose per unit time. Of course, such a sweeping conclusion necessarily involves an extrapolation from actual data. Not until recently has it been possible, because of technical difficulties, to test the mutagenic effectiveness of doses lower than about 13 r per day, totaling 400 r (11–13), and even the most recent work goes down no lower than about 2.5 r per day, totaling 25 r (14, 15)."
Stern (1950)	Page 433 "The proportionality rule has been proven to hold over a wide range. Figure 155 shows that, for Drosophila, the relation is essentially linear over the range from 25 r to several thousand r. It has further been shown that the frequency of induced mutations is independent of the time over which the radiation is applied."
Stern (1960)	Page 491 "It has been established for a variety of experimental organisms that the number of mutations induced by radiation is proportional to the dose. This proportionality has been proven to hold over a wide range of dosages. Figure 202 shows that, for Drosophila, the relation is essentially linear over the range 25–12,500 r (insects, unlike mammals, can survive after exposure to many thousands of roentgens). It would be desirable to extend the data toward dosages lower than 25 r, for instance, to 10 r, 5 r, and still lower. Since, however, the expected differences are small between the rate of mutations in not-artificially irradiated control organisms and that in organisms exposed to low artificial doses, it is difficult to obtain significant results even with large experiments."

Impact of the Stern and Muller deceptions

Effect on the radiation genetics literature/community

In the aftermath of his Nobel Prize Lecture, Muller published his Lecture in the *Journal of Heredity* in 1947 (Muller 1947), assuring its broader distribution. Within 4 months of the Noble Prize Lecture, he gave a lecture to the New York Academy of Medicine during which he affirmed his Nobel Prize Lecture message, stating that there was "absolutely no threshold dose" for mutations and that induced mutational response was proportional to the total dose (Table 1). This presentation was published in the Academy's journal (Muller 1948) soon thereafter. Stern (1950) also cited Spencer and Stern (1948) and Uphoff and Stern (1949) in his acclaimed textbook, emphasizing that the dose response for mutations was linear (Table 1).

These follow-up activities by Stern and Muller had an impact on other leading radiation geneticists influencing them to adopt the linearity dose-response interpretation. Table 2 provides a series of quotations from subsequent publications of leading contemporary radiation geneticists. The quotes are numerous, varied, and a fair representation of what each author stated. These comments strongly support the conclusion that there was a generally consistent view that the nature of the dose response in the low-dose zone for mutations was linear. Most of these quotes directly cite the research of Stern and his colleagues as providing the key evidence supporting linearity, especially that of Spencer and Stern (1948) and Uphoff and Stern (1949). This demonstrates the significance and success of the Stern mediated manipulation of the Caspari and Uphoff studies in affecting mutation dose-response beliefs of key research leaders of the radiation genetics community.

Effect on the BEAR I Committee/Genetics Panel

Crow (1995) noted the following in his historical recounting of the BEAR I Committee Genetics Panel: "the debate over the nature of the dose response for ionizing radiation and mutations had been decided before the convening of the BEAR Committee in November 1955." The accepted view was clear and unified; the answer for the dose response question for mutagenicity was "linearity at low dose."

When reading the transcripts of the BEAR I Committee Genetics Panel, one is struck by the absence of debate and even discussion on the issue of dose response (e.g., linearity vs. threshold). To illustrate the fact that the decision on

Table 2 Radiation genetics quotations about the mutation dose-response following Hermann J Muller's Nobel Prize and Curt Stern's (with Spencer, Caspari and Uphoff) mutagenicity papers

References	Quotes
Catcheside (1950)	Page 592 "The induced mutation is proportional to the total dose over the whole range investigated, down to total doses as small as 25 r. There is good reason to conclude that there is no threshold dose, i.e., no dose so small that it gives no mutational effect. Also, the intensity of the radiation appears to be without effect on the frequency of mutation induced by a given total dose. A dose of 50 r given in a fraction of a minute appears to give no greater effect than the same dose given in the course of a few weeks. There is no threshold, no time factor, and no recovery, the effects being cumulative."
Glucksmann (1950)	Page 42 "The induction of gene mutations is linearly proportional to dose even down to levels of 25 r (Spencer and Stern 1948)."
Lefevre (1950)	Page 341 "It has been amply verified that the number of mutations produced by X-rays is linearly proportional to the total dose applied, even when the total dose received is very small (see Spencer and Stern 1948). Further, the number of mutations produced is independent of the rate of dosage (Uphoff and Stern 1949)."
Sax (1950)	Page 332 "The early work by Muller and by Timoféeff-Ressovsky showed a linear relationship between X-ray dosage and mutation frequency in *Drosophila*. It was also found that the induced mutation rate was independent of radiation intensity. From these observations it was concluded that the X-ray-induced mutations are produced by single 'hits,' and that there is no threshold effect. Spencer and Stern (2) found no increase over the spontaneous mutation rate by irradiating *Drosophila* for 21 days at 2.5 r/day, but later experiments by Uphoff and Stern (3) indicated that low intensities are effective."
Higgins (1951)	Page 9 "As a result of exhaustive experiments on the genetics of the fruit fly, of mice and of many plants, it is held that the number of induced mutations bears a linear relationship to the total amount of radiation absorbed by the sensitive volume of the cell and is independent of either the duration or the intensity of exposure. Consequently, a long exposure to low-level radiation would have the same genetic effect as shorter exposure to a higher level. Experiments of Spencer and Stern (1948) on the fruit fly show that the percentage of sperm containing a sex-linked lethal mutation is increased about .002 per r of radiation exposure and that 50 r exposure is required to double the natural mutation rate." "Spencer and Stern (l.c.) conclude their exhaustive study of the validity of the linear relationship between radiation exposure and mutation frequency with the statement (p. 64): '…for radiation with X-rays, dosages as low as 25 r produce mutations as drastic in their effects and in the same proportion to the dosage as do exposures to high dosages. If an extrapolation is permissible, one may assume that there exists no tolerance dose below which mutations are not induced." "The classical hit theory of induction of mutations, particularly the linear relation between dosage at low levels and mutation rate, has been questioned by Caspari and Stern (1948), who found no significant difference in mutation rates in the sperm of the fruit fly between controls and experimentals exposed to 2.5 r per day for 21 days. Uphoff and Stern (1949), however, after further tests, concluded that low-level radiation does produce mutations in fruit-fly sperm and that the apparent inconsistencies of previous results were due to different experimental techniques and errors in sampling."
Stone (1952)	Page 657 "There is no threshold for genetic mutations…" (cited Muller reference 1950, J Cell Comp Physiol 35(suppl 1):9–70.)
Singleton (1954a)	Page 598 (Discussion) "That a non-linear relationship exists between dose rate of chronic gamma radiation and mutation rate of endosperm characters seems to have been well established by these experiments. This was shown quite conclusively by disproportionately higher mutation rates at the higher dosages, and was definitely indicated by the fact that there seems to be a threshold of dosage required to raise the mutation rate from the spontaneous level to a detectable increase over that level." Page 599 "These data (i.e., data shown in Singleton 1954a study) showing a definite threshold are in contrast to the Drosophila data of Spencer and Stern (1948), where no threshold was indicated even when low doses of radiation were used. In their experiments the effects of acute radiation were studied. Caspari and Stern (1948), studying chronic gamma radiation, found no increase over the controls for doses of 2.5 r/day for 21 days. However, it was later demonstrated by Uphoff and Stern (1949) that the controls used by Caspari and Stern had an abnormally high sex linked lethal frequency and that actually there was an effect of the chronic gamma radiation of 2.5 r/day."
Kelner et al. (1955)	Page 36 "The linear mutation-dose curve indicated for X-ray induced drosophila lethals (Lethals-Dros:X) is perhaps best exemplified by the data of Spencer and Stern (53) for sex linked lethals and may be considered as the classical type of mutation-dose relation. Interpreted within the target theory, the linear relation indicates that a single hit is sufficient to produce a mutation."

Table 2 continued

References	Quotes
Nybom et al. (1956)	Page 81 "In this connection references may be made to the concordant results of Uphoff and Stern (1949) who did not find any threshold in *Drosophila* after low dose rates. A similar result was published by Sax (1950) using chronic irradiation of *Tradescantia* pollen."
Lewis (1957) (This Science article was reprinted in Congressional Testimony)	Page 971 (columns 2 and 3) "Gene mutation has long been known to show a linear relationship with respect to dose of ionizing radiation from studies with *Drosophila*. This linearity has been extended by Spencer and Stern (43) to doses of 50 and 25 roentgens. Gene mutation is also known to be directly proportional to the accumulated dose of radiation, even when the radiation is chronically administered at a relatively low dose rate, as in the studies of Uphoff and Stern (44)."
Norwood (1958)	Page 1929 "Several geneticists[4] have sketched the background which has lead to the concern of this study. Briefly, realization that radiation increases the mutation rate dates back 30 years to Muller's experiments with fruit flies[4e]. Spencer and Stern,[5] using more than 50 million flies, showed that genetic damage was proportional to dosage in the important range of 25 to 50 r. Concern has been heightened by recent findings[4f] that exposure of mice to a given quantity of radiation increases the mutation rate by about 15 times as much as does an equal exposure of Drosophila, which had formerly served as the sole basis for inferring human risks."
Spear (1958)	Page 20 "There is general agreement, however, that mutations can be produced with very low dosage down to a level which approaches natural background (Uphoff and Stern 1949)."
Newcombe (1960)	Page 331 "One basic premise which has not so far been seriously challenged is that the number of gene mutations resulting from irradiation varies in direct proportion to the dose. In other words, there is no threshold level of radiation below which the mutations will not be produced." "In the fruitfly the curve has, by dint of considerable work, been pushed to within 25 roentgens of the origin (Caspari and Stern 1948; Spencer and Stern 1948; Uphoff and Stern 1949) (3, 4, 5)."

LNT had already been settled prior to the creation of the BEAR I Committee, there was no discussion of the scientific foundations of the LNT, including any documenting of its theoretical basis and experimental support, including its strengths and limitations. As noted above, the Genetics Panel placed a high priority on the chronic exposure experiments published under the leadership of Curt Stern. Yet these studies, even ignoring the control group problems of the Uphoff and Stern experiments, had little or no risk assessment relevance. That is, these were sex-linked recessive lethality studies in which the spermatozoa were deposited in the spermatheca of the female. The females were then placed into a type of specialized experimental "hibernation" in which there was a profound alteration of the diet and a lowering of the temperature, changes designed to prevent egg production. The females (with the deposited spermatozoa) were then exposed for 21 days (24 h/day) to gamma irradiation. After the 21 days, the dietary and environmental conditions were changed to permit egg laying so that the testing for sex-linked recessive lethal mutations could take place. In effect, Stern exposed the spermatozoa to ionizing radiation for the equivalent of an entire lifespan, something comparable to a 70–80-year human lifespan. The spermatozoa are known to be highly compromised, having lost much of their normal repair capability. The study represented a worse case exposure scenario, that is, selection of a very susceptible developmental stage linked to a profoundly extended and highly unrealistic exposure period. In effect, the study was a chronic exposure to a cell type that has only a very short developmental stage. The basic concept of the study was not appropriate for a chronic exposure with risk assessment application. The BEAR I Committee incorrectly accepted Stern and Muller's concept of "chronic" for risk assessment purposes as did the entire field and regulatory agencies.

While the BEAR I committee relied upon the findings of the *Drosophila* research directed by Curt Stern, it failed to cite other similarly large-scale *Drosophila* studies (Bonnier and Lüning 1949; Bonnier et al. 1949) in which the lowest total dose was 8 r, below the lowest dose (25 r) of the Spencer and Stern (1948) findings. These papers documented the response of several single genetic loci (e.g., white and forked loci) to which their detailed statistical analysis for mutational studies was applied. The analysis revealed a linear dose response in the dose range of 700–2,800 r, whereas the linearity response was not observed in the low-dose range (8–16 r), where the data were supportive of a threshold response. The authors also suggested that the difference in the shape of the dose response between high and low doses was indicative of differing dose-dependent mechanisms. At the high doses, the linear dose response was consistent with the target theory of Timoféeff-Ressovsky et al. (1935), whereas at lower doses mutational effects could be due to the effects of chemical

mutagens (i.e., hydroxyl radicals from the hydrolysis of water). The dose-dependent mechanism-based hypothesis of Bonnier and colleagues (Bonnier and Lüning 1949; Bonnier et al. 1949) was soon supported with experimental data (Haas et al. 1950; MacKey 1951; Lüning 1954; Barron 1954). According to Barron (1954), "it is dangerous, however, to extrapolate from experimental data with large doses of radiations to what might take place with small doses. In biological systems the effect of ionizing radiations differs qualitatively when the radiation dose is changed. Small doses act by indirect action and produce mainly oxidations. Large doses act by two mechanisms," that is, free radical formation via water hydrolysis and by a direct collision, which is consistent with the target theory.

The Bonnier and Lüning (1949) (Bonnier et al. 1949) papers were also critical of the use of sex-linked recessive lethal experiments for estimating responses in the low-dose zone due to the "impossibility of differentiating between true lethals and semi lethals, and the fact that there are several hundreds of targets per chromosome ready for lethal mutations…" The lack of target specificity would represent an important limitation in the interpretation of dose-response relationships and their potential application to a mechanism-based risk assessment process. Bonnier et al. (1949) also provided a detailed statistical reanalysis of the Spencer and Stern (1948) data challenging the broadly accepted conclusion that the linearity response applied across the entire dose-response range, including the lower dose range. None of these fundamental technical issues were discussed by the BEAR I committee.

Another relevant aspect of the discussion on the nature of the mutation dose response involved the research of Arnold H. Sparrow and W. Ralph Singleton of the Brookhaven National Laboratory. Chairman Warren Weaver introduced their research and its relevance to the BEAR I Committee/Genetics Panel (Weaver W., February 5–6, 1956, see page 110—Transcript) (BEAR I 1956). The discussion of the Sparrow and Singleton data was then led by Committee member Berwind D. Kaufmann, who claimed to have copied several tables from their paper. He stated that Sparrow and Singleton showed that 0.41 r per day yielded a modestly elevated (i.e., less than twice the control values) but statistically significant effect on micronuclei formation. What Kaufmann failed to inform the Committee was that Sparrow and Singleton (1953) specifically stated that a threshold response had been observed at a lower dose. In fact, there was no discussion concerning their threshold dose-response statement by the BEAR I Committee/Genetics Panel. The data in Table 2 (page 35) of the published paper by Sparrow and Singleton (1953) show that 0.084 r per day caused no significant increase in micronuclei. This recounted activity of the BEAR I Committee/Genetics Panel demonstrates that it either ignored or was misled on the published findings of Sparrow and Singleton as the data did not support the pre-determined linear dose-response conclusion. This analysis also suggests that the BEAR I Committee/Genetics Panel was very selective in their choice of what data to consider and that such decisions reveal a prevailing bias supportive of LNT model acceptance.

Since 0.41 r per day of radiation in the Sparrow and Singleton (1953) hypothesis study is more than 1,000 times greater than the naturally occurring intensity, these data do not support the theory that the spontaneously occurring micronuclei are produced by naturally occurring ionizing radiation. The findings of Sparrow and Singleton (1953) were similar to that of Giles (1940) from Harvard who showed that when *Tradescantia* were "subjected to irradiation 1,000 times that due to natural radiation….no increase in aberration was found." Other experiments by Giles indicated that even using ionizing radiation at some 1,800-fold above background no impact on the occurrence of spontaneous mutations occurred.

It is possible to obtain a sense of the personal views of a number of the members of the BEAR I Committee/Genetics Panel on the matter of dose response via two contemporary publication avenues: Testimonies at a 1957 Congressional Hearings (Table 3) and journal publications in the open literature (Table 4) such as a special issue of *Scientific American* on ionizing radiation and several other journals. Based on these collective comments, it follows that the BEAR I Committee/Genetics Panel report and an article in the journal *Science* (Table 5) summarizing the report of the Genetics Panel were replete with statements asserting linearity at low dose.

Placing the new Muller and BEAR I Genetics Panel developments in perspective

The story of Muller's Nobel Prize Lecture is important for its history of science implications, as well as its role in affecting the decision of the US National Academy of Sciences (NAS) to recommend a linearity dose-response policy for assessing risks to the genome from ionizing radiation, replacing the threshold dose-response model. This formal recommendation initiated a series of advisory and regulatory dominoes in essentially all countries to adopt linearity and apply it to somatic effects, that is, cancer risk assessment, for ionizing radiation and later for chemical carcinogens (Calabrese 2009). The linearity decision of the NAS BEAR I Committee/Genetics Panel was strongly championed by Muller, the titular leader of radiation geneticists and with strong ties to all radiation geneticists on the BEAR I Committee/Genetics Panel. In fact, the switch to linearity, which was ushered into the international

Table 3 BEAR I Committee Genetics Panel member quotes at Joint Committee on Atomic Energy—1957

References	Quotes
Muller (1956)	Page 392 "In material of varied kinds, but more especially in *Drosophila*, there is good evidence that over a considerable range of dose (in *Drosophila*, from some 50 r to more than 1,000 r, a more than 20-fold range) the frequency of point mutations (like that of chromosome breaks) is directly proportional to dose."
Crow (1957a)	Page 1013 "4. Evidence from experimental animals, principally Drosophila, indicates that the number of mutations produced is strictly proportional to the amount of radiation received. There are departures from this straight-line relationship at high doses, but these are too high to be likely to be encountered in any ordinary human situation. It is technically impossible to test this relationship for the very lowest doses, but the straight-line relation holds down to the smallest amounts that have been studied." "For these reasons a simple proportionality between the amount of radiation and the number of mutations is fully accepted by geneticists." "The proportionality between dose and mutation production holds irrespective of the intensity or spacing of the dose." Representative Holifield (page 1013) questions Dr. Crow: "This, then, would establish as far as the majority of the geneticists are concerned the principle of linear progression in deleterious effects of radiation regardless of amount?" Dr. Crow answers: "That is correct. A nonthreshold situation, to put this in yesterday's vocabulary." "This means that there is no such thing as a safe dose of radiation to the population. Any amount of radiation, however, small, that reaches the gonads—testes or ovaries—of a person who may later reproduce, involves a risk proportional to that amount."
Glass (1957a)	Page 1030 "The data are most extensive for the fruitfly and the lowest dose that has actually been studied is 25 r." Page 1031 "Because a mutation can be produced by a single ionization in the right place, there is no threshold below which the amount of radiation is too small to produce mutations—that is, every dose produces mutations with a probability equal to its magnitude." "This is to repeat what Dr. Crow said, that there is no safe dose of mutation. This curve continues down without any threshold until it hits the zero point…"
Muller (1957a)	Page 1052 "In respect to the fact that probably there is no threshold, that these effects are proportional to the dose, in this respect these effects of radiation—and also the leukemia—on the exposed individual himself resemble those produced by the radiation in weakening descendants." "You have heard Dr. Glass and Dr. Crow say that geneticists are convinced that there is no threshold for the genetic effects and that others, too, now accept that principle for the genetic effects." "If this is true of these other effects, and it is certainly time we knew whether it was—I think the evidence is convincing that it is—then this important resemblance between the effects on later generations and on the exposed generation is probably not an accidental resemblance. For there is growing reason to infer that this shortening of life and the other long delayed damage done to an exposed individual have their basis in damage done to the genetic material—the chromosomes and their contained genes—of the body's ordinary cells, those of the blood, skin, glands, and so forth, similar to the damage done in his reproductive cells that is passed on to later generations." Page 1056 "Through work on the fruitflies where we have the most exact knowledge to date, unless Dr. Russell has more exact knowledge on mice now, we can get a kind of minimum estimate of the amount of damage to the children by a given amount of irradiation of the parents."
Muller (1957b)	Page 1066 "Since there is much evidence indicating a linear relation between the radiation dose and the frequency of the induced point mutations, even at extremely low doses, and the exactly cumulative nature of these radiation effects, it becomes possible to arrive at probable estimates of the minimum damage done to subsequent generations by any given chronic or acute exposure of parents." Page 1067 "…leukemia and some other malignancies, the induction of which may also be linearly dependent upon radiation dose…"
Joint Committee on Atomic Energy (1957)	Page 12 "…geneticists believe that the direct proportion applied down to zero dose—that is, that there exists no safe "threshold" below which the dose produces no damage, and that damage occurs from any irradiation of the genetic cells, no matter how small the dose."

Table 4 BEAR I Committee Genetics Panel member quotes on low-dose linearity in journals after the BEAR I Committee

References	Quotes
Crow (1957b)	Page 19 (column 2) "2. The number of mutations produced is directly proportional to the dose in roentgens. The linear proportionality over wide dose ranges has been shown in several organisms, especially in *Drosophila*." "Experimental verification in *Drosophila* has been carried to as low as 25 r…" Page 20 (column 1) "The proportionality between dose and mutation production holds irrespective of intensity or spacing,…" Page 20 (column 2) "The conclusions of the previous section imply that there is no such thing as a "safe" dose. Any increase in radiation, however, small, involves a risk proportional to that amount."
Glass (1957b)	Page 956 "Our present evidence indicates that the frequency of these point mutations always increases linearly with the radiation dose (Fig. 1). In Drosophila studies this holds over the range from 25 r to 6,000 r. In some plants, the linear range has been extended down to about 5 r. In mice, the linearity in relation to dose holds over the range from 300 r to 600 r, and there is no sign that it does not hold at lower doses. This linear proportionality to dose, over and above the spontaneous frequency of mutation, implies that (a) as long as dosage is measured in terms of roentgens, that is, in terms of the ionization produced by the radiation, absorbed quanta do not interact to produce effects, but are individually effective; and (b) there is no sign of a threshold dose below which mutations are not produced. Rather, even the lowest doses are proportionally mutagenic, and all doses, however, distributed, are additive or cumulative in effect."
Beadle (1959)	Pages 225 and 226 "…thus there is probably no threshold below which radiation will produce no mutations. Since there is no repair mechanism, once the mutation process is complete, mutations induced at different times will tend to accumulate in a line of descent…"
Hollaender and Stapleton (1959)	"In sum, cell studies have served to elucidate the basic mechanism by which ionizing radiation damages the living organism. They have provided no evidence that there is a true threshold of dosage below which ionizing radiation produces no harmful effects…"

community by the BEAR I Committee Genetics Panel, is the most significant action in regulatory environmental public health history with ever expanding social, political, economic, and public health implications (Hamblin 2007).

The present paper provides the first documentation of how Muller (Muller 1950a, b, 1954a, b) himself used the carefully constructed activities of Stern (described in detail in Calabrese 2011b) to enhance the concept of linearity and to protect his reputation. Muller lent credibility to the technical note of Uphoff and Stern (1949) while further marginalizing the Caspari and Stern study results (Caspari and Stern 1948). The stakes were high on multiple levels and these core individuals knew it. Stern and Muller needed to prevent the acceptance of the Caspari and Stern (1948) study findings in order to sustain the single-hit linearity model. They also needed any criticisms of the Spencer and Stern (1948) and Uphoff and Stern (1949) papers to be muted. They were successful as other leaders of the radiation genetics community simply failed to address the serious limitations of the Spencer and Uphoff findings while incorrectly asserting that the Caspari and Stern (1948) paper suffered from an aberrantly high control value, simply re-stating the demonstrably incorrect, but authoritative conclusion of Muller (1950a).

Despite the fact that Caspari had successfully rebutted the first challenge of Stern concerning the control group spontaneous mutation rate, there is no evidence that he disputed the control group mutation rate reversal decision of Stern barely a year later and of Muller's equally strange affirmation of Stern's position as well (Muller 1950a, b). A January 27, 1949, letter from Caspari to Stern supported the publication of the Uphoff and Stern (1949) paper now adopting part of the mantra of Stern, that is, that there is considerable variability in the mutagenic frequency of sperm prolongedly stored in the spermatheca. This conclusion provided the opportunity to rehabilitate the inexplicitly low control group values of Uphoff. Caspari, however, would not go so far as to also state that his control values were unusually high. At the time of the Uphoff and Stern (1949) article, there were only two papers published in the literature (Rajewski and Timofeeff-Ressovsky 1939; Kaufmann 1947) on aged sperm and mutation and the published abstract of Muller (1946b). Each supported the mutation frequency of Caspari. These findings are consistent with subsequent mutation frequencies in aged sperm stored in the spermatheca of female Drosophila (Byers 1954; Byers and Muller 1952; Rinehart 1969; Graf 1972; Muller et al. 1961). Muller et al. (1961) stated that "The data clearly showed a rise in mutation frequency (averaging some .06 percent of recessive lethals in the X chromosome per week) resulting from storage of the mature spermatozoa in the female" (page 213). Note the striking similarity of how

Table 5 Low-dose linearity quotation in the journal Science from article summarizing the findings of the BEAR I Committee Genetics Panel

References	Quotes
BEAR I (1956)	Page 1159 (column 2) "…the genetic damage done, however, felt and, however, measured, is roughly proportional to the total mutation rate." Page 1160 (column 1) "3) Any radiation dose, however, small, can induce some mutations. There is no minimum amount of radiation dose, that is, which must be exceeded before any harmful mutations occur." Page 1160 (bottom column 1) "The probable number of additional induced mutations occurring in an individual over a period of time is by and large proportional to the total dose of extra radiation received, over that period, by the reproductive organs where the germ cells are formed and stored." Page 1160 (top column 2) "The *total dose* of radiation is what counts, this statement being based on the fact that the genetic damage done by radiation is *cumulative*." Page 1162 (column 2)—how harmful are radiation-induced mutations? "1) Thus the first and unanimous reply to the question posed by the title to this section is simply this: *Any radiation is genetically undesirable*, since any radiation induces harmful mutations. Further, all presently available scientific information leads to the conclusion that *the genetic harm is proportional to the total dose*… This tells us that a radiation dose of 2X must be presumed to be twice as harmful as a radiation dose of X…" Page 1164 (column 1) "…for there is no such figure other than zero." [referring to whether there is an amount of radiation which is genetically harmless (preceding phase)] Page 1164 (column 1) "As geneticists we say: keep the dose as low as you can." Page 1165 (last sentence) "From the point of view of genetics, they are all bad." (referring to the effect of exposures to ionizing radiation)

Uphoff and Stern (1947) characterized Muller's data some 14 years earlier, "a weekly increase of about 0.07 %…" The 0.06 % increase would yield an estimated 0.28 % (i.e., 0.06 % × 3 weeks + 0.10 % background = 0.28 %) mutation incidence after 3 weeks, consistent with the Caspari and Stern (1948) findings, the logic used in Uphoff and Stern (1947) and with the Muller (1946b) statement that "spermatozoa aged several weeks in the female may contain several times as many mutations as they originally had." Furthermore, the reported inter-study variability for mutations of aged sperm and/or stored sperm aged in the spermatheca appears modest with 95 % confidence intervals typically being about ±25–30 % of the mean. The attempt by Stern, therefore to assert that the very low values of Uphoff reflected a highly variable response endpoint was not supported in the contemporary and subsequent literature. Stern never argued his case by a comparative data assessment nor did he address the apparent contradiction with the Muller data and comments which he (i.e., Stern) previously used when he concluded that the Caspari data were credible while those of Uphoff were not. He simply made an authoritative declaration that was accepted without question or comment by the radiation genetics community.

BEAR I Committee/Genetics Panel

The BEAR I Committee/Genetics Panel was comprised of outstanding scientists and national leaders. Despite their significant individual accomplishments in scientific and radiation genetics domains, the committee as a whole lacked extensive experience in conducting low-dose, dose–response studies. Only two of the members had extensive direct experimental dose-response experience (i.e., Demerec and Russell) up to the time of the BEAR I meetings. This experience was essential for evaluating the nature of the dose response in the low-dose zone. Of these two, Demerec had the most extensive and varied experience having dealt with multiple models and agents as well as different types of radiation. His research experience on dose response was spread over a 25-year period starting about 1931. Nonetheless, his dose-response experience with *Drosophila* was limited to only a few high dose studies during the 1930s, a key limitation. Despite his significant and prolonged career at Oak Ridge, Russell was relatively new to the dose-response research area, with about 5–6 years experience at the start of the BEAR I Committee in 1955. In the case of Russell, his developing research findings with mice were still somewhat premature, having little impact on BEAR I Committee/Genetics Panel conclusions. Among the remaining members of the committee, Muller's principal dose-response experience is found in the research of Hanson and Heys (1929), and Oliver (1930, 1931) at the University of Texas and Ray-Chaudhuri (1944) at Edinburgh (completed in 1939), as well as his consultant role with Stern from 1943 to 1946. Limited relevant low dose-response research based on the publication record experience was found for Berwind Kaufmann. Alexander Hollaender, PhD in physical chemistry, had made

important contributions on the effects of UV wavelengths specificity on mutation in bacteria and fungi. He became the director of radiation biology research at Oak Ridge, hiring Russell. Hollaender had no experience with Drosophila research. H. Bently Glass' low-dose experimental research experience was quite limited during BEAR I, becoming far more extensive only after BEAR I. Importantly, very limited to no meaningful dose-response research experience is apparent for the remaining 11 members [George W. Beadle, Charles W. Cotterman, James F. Crow, Gioacchino Failla, Clarence C. Little, James V. Neel, Tracy M. Sonneborn, Alfred H. Sturtevant, Sewall Wright, Warren Weaver (Chair), and Shields Warren] of the BEAR I Committee/Genetics Panel. This situation resulted in the "senior" dose-response experience to reside with Demerec and Muller, two individuals on record to save the "hit" model.

The geneticists on the BEAR I committee were principally basic researchers; their experimental approaches were neither dose response nor risk assessment oriented. Even Muller (1950a, b) claimed that the work of Spencer and Uphoff (with Stern) at low doses would markedly extend his and his students' (e.g., Hanson and Oliver) research conducted at very high doses. Further, in the detailed comments that Muller sent to Stern about the Spencer (Lilly Library 1946, September 13) and Caspari (Lilly Library 1947a, January 14) manuscripts, nearly all dealt with fundamental biological/genetic questions with little direct relevance to risk assessment. Multiple study design issues and other methodological/analysis problems documented in Calabrese (2011b) for the Spencer and Stern (1948) paper were not identified by Muller (Lilly Library 1946, September 13). The members of the BEAR I Committee/Genetics Panel looked to Muller for leadership on matters related to the dose–response. However, Muller displayed critical limitations in assessing such studies based on his written statements. Thus, the methodological and analysis limitations of the Spencer and Stern (1948) paper and the serious flaws of the Uphoff and Stern (1949) paper were missed by the radiation genetics community and the BEAR I Committee/Genetics Panel, a condition that continues (Lipshitz 2005). Of further note is that Muller (1946b) and Kaufmann (1947) published findings on the control group mutation rate of aged *Drosophila* sperm that supported the findings of Caspari and Stern (1948). Kaufmann worked closely with and under the direction of Demerec at Cold Spring Harbor at that time. Furthermore, an October 7, 1947, letter (i.e., 6 weeks before submitting his paper to *Genetics*) from Caspari to Stern (American Philosophical Society 1947g, October 7) stated that "I have discussed the paper (the Caspari/Stern manuscript) with Demerec and Kaufmann. Both did not find very much to suggest……Both Demerec and Kaufmann were impressed by the amount of material which we have. The ageing effect in our experiments is of the same order of magnitude as that found by Timoféeff and Kaufmann." In fact, Caspari and Stern (1948) cited a 1947 paper by Kaufmann as support for control group values of their study. Muller and Kaufmann, both BEAR I committee members, therefore, reported research on mutation incidence of *Drosophila* aged sperm findings consistent with the findings of the Caspari and Stern (1948) paper. Thus, the BEAR I Committee/Genetics Panel should have been informed on the issue of control group validity by Demerec, Kaufmann, and/or Muller as it related to the research of the Caspari and Uphoff studies. However, based on the transcripts of the BEAR I Committee/Genetics Panel, Demerec, Kaufmann and Muller did not provide this information. Knowledge of the mutation rates in aged *Drosophila* sperm should have led to a reconsideration of the Caspari and Stern (1948) paper as well as generated serious questions about the findings and interpretations of the Uphoff and Stern (1949) data. This was a key issue affecting which study would be relied upon by the BEAR I committee. By their actions, the BEAR I committee Genetics Panel came to the erroneous conclusion that the Caspari study was unreliable due to its "unusually high control group value."

The future of ionizing radiation risk assessment was largely determined by the actions of a few, by the failure of the scientific community, especially the radiation genetics community, to probe deeper into the key findings of Stern and his colleagues and journals such as *Science* that published influential but poorly documented findings (Uphoff and Stern 1949). As has been pointed out, the linearity paper of Spencer and Stern (1948) was burdened with numerous methodological limitations that only recently have been documented, as well as statistical analysis limitations that challenged the conclusion of linearity at low dose (Bonnier and Lüning 1949; Bonnier et al. 1949) while the Caspari and Stern (1948) findings supporting a threshold perspective were unfairly marginalized (Calabrese 2011b). Furthermore, the BEAR I Committee/Genetics Panel failed to require Stern to provide the promised detailed accounting for the *Science* article (Uphoff and Stern 1949) upon which they so heavily relied.

According to Muller (1950a, b), by 1950, the radiation genetics community had accepted the linearity risk assessment paradigm (Table 2). Their belief was based largely on the fruit-fly work of Stern and his associates as well as the leadership, prestige, and authority of Muller, as few of the geneticist members of the BEAR I Committee/Genetics Panel had relevant experience with low-dose research. By the time, the National Academy of Sciences BEAR I Committee/Genetics Panel convened, therefore, the decision over the nature of the response in the low-dose zone had been decided by the radiation genetics community as there was no dispute or even debate within the BEAR

I Committee/Genetics Panel over the adoption of linearity to replace the threshold model for germ-cell mutagenicity (Crow 1995). The actions of Stern and Muller had led the way, assuring that the ends (i.e., linearity) justified the means (i.e., unfair/improper scientific evaluation). In fact, it is from this heritage and upon this foundation that regulatory cancer risk assessment theory and practice in the USA and throughout the world was built.

Conclusions

1. This paper provides specific documentation of how Hermann J. Muller supported and extended the like actions of Curt Stern to prevent the scientific community from discovering Muller's Nobel Prize lecture deception and to promote his ideological goal of linearity at low dose for ionizing radiation risk assessment (Table 6).
2. Muller strengthened the questionable actions of Stern in key publications in early 1950s while improperly discrediting the threshold findings of Caspari and supporting the "uninterpretable" data of Uphoff to achieve a linearity interpretation. The bases of these actions are documented in this paper.
3. The paper shows how the actions of Stern and Muller affected numerous publications and the dose–response beliefs of leaders of the radiation genetic community and the NAS BEAR I Committee/Genetics Panel, affecting the adoption of linearity at low dose for ionizing radiation-induced mutation and eventually for carcinogen risk assessment for ionizing radiation and chemical carcinogens.
4. The findings demonstrate that the adoption of the LNT model for risk assessment lacked a proper scientific foundation, yet was accepted by regulatory and public agencies worldwide.

Unresolved issues

1. Why didn't Stern publish the follow-up detailed paper containing the entire methodology for all the relevant data for the Uphoff three experiments?

Table 6 A summary concerning Muller's actions that affected the discrediting of Caspari's findings and acceptance of the Uphoff and Stern conclusions

A five-page detailed letter sent from Muller to Stern dated January 14, 1947, concerning scientific strengths and limitations of the Caspari and Stern manuscript provided no comment on the control group lethality data
Muller was actively researching the area of spontaneous mutations in sex-linked recessive lethality studies using aged sperm stored in the spermatheca of female fruit flies. This was the research method of the Caspari and Stern paper. Muller had been doing extensive research on this topic since the early 1940s. He was a leading authority on the topic
Muller provided his spontaneous control group data to Stern ("Appendix" section) in order to address the concern that Stern expressed about the apparently high control group values of Caspari
Based on the data of Muller, Uphoff and Stern (1947) determined that the average weekly spontaneous mutation rate in *Drosophila* sperm stored in the spermatheca of the female was about 0.07 %, yielding an additional mutation increase in about 0.21 % by 3 weeks, the length of the Caspari sperm storage time. The 0.21 % increase would be added to a background value of about 0.10 %, yielding an estimated control group value of about 0.31 %. The 95 % confidence intervals were about ±0.07 %, with an approximate range of 0.24–0.38 %. The values were obtained when studies were conducted at about 25 °C. At the lower temperature of 18 °C used by Caspari, it was estimated by Stern (and Uphoff) that the rate of increase might be reduced to 0.05 % per week. This would result in an estimated value for the Caspari control of about 0.25 %, nearly identical to his final adjusted value (i.e., 0.2489 %)
Based on these data, Uphoff and Stern (1947) concluded that the Muller data supported the Caspari conclusion that his control data were well within the normal range and not unusual or aberrant. The Muller data lead Uphoff and Stern (1947) to conclude the Uphoff findings were uninterpretable
Continued research in the area of spontaneous mutation in sperm stored in the spermatheca by Muller and his graduate students at the University of Indiana were consistent with this conclusion and quantitative assessment (Byers 1954; Byers and Muller 1952; Graf 1972). These findings were also consistent with that published by other researchers as well (Kaufmann 1947; Rinehart 1969)
Based on this information, the statements of Muller that Caspari's control group data were unusually high are inconsistent with: (1) His own data and that published by other researchers; (2) his previously detailed assessment of the Caspari data; (3) how Uphoff and Stern (1947) evaluated the Muller data, an evaluation that Muller was knowledgeable of, based on an acknowledgment in the Uphoff and Stern (1947) paper, and (4) internal written correspondence between Stern and Caspari
This assessment indicates that Muller's statements that Caspari's control group data were unusually high and adversely affected Caspari's threshold interpretation are contradicted by the body of evidence
While Muller repeatedly challenged the credibility of the Caspari findings by attacking his control group data, he made no statement about the reliability of the extremely low control group data of Uphoff. In fact, he would consistently cite the Uphoff and Stern (1949) paper as being a critical reference to support a linearity perspective
The collective findings on these matters indicate that Muller displayed compromised scientific judgment, having a significant impact on the scientific literature and national and international risk assessment policy that continues to the present

2. Why didn't the radiation geneticist community demand that Stern publish these findings?
3. Why didn't Stern address the scientific basis, if any, of why he reversed his position on the Uphoff control group data?
4. Why didn't Caspari challenge any of the multiple papers that claimed that the Caspari control group data were unusually/abnormally high or that their paper displayed "different techniques" or had "errors in sampling" that accounted for their threshold-like findings?
5. Why did Muller agree to let Uphoff and Stern (1947) acknowledge the use of his aged sperm data that supported the Caspari control groups findings and then repeatedly claim that Caspari's control group values were unusually high, adversely affecting the credibility of this paper?

Acknowledgments The research on the topic of hormesis has been supported by awards from the US Air Force and ExxonMobil Foundation over a number of years. Sponsors had no involvement in study design, collection, analysis, interpretation, writing, and decision to submit.

Conflict of interest The author declares that there is no conflict of interest.

Appendix

Stern–Muller temporal letter exchange concerning the aged-stored sperm control mutation rate (Source: Lilly Library, Stern–Muller correspondence)

Curt Stern wrote a letter to Hermann J. Muller on January 22, 1947 (American Philosophical Society 1947a), informing him that "At the present time it looks as if our new control data (probably the results of the first 3 months of the first Uphoff experiment; note that her first month's reading was an especially low mutation rate of 0.005 %) for aged sperm are considerably below those of Caspari's." He then asked Muller to "send me your figures on rate of sex-linked lethal in sperm aged several weeks, (most desirably, if you have them, data on 3 weeks), in comparison to control data from non-aged sperm?"

On February 3, 1947 (Lilly Library 1947b, February 3), Muller answered by stating that "…. sperm of males which are about a week old and have been copulating freely (as in Caspari's experiment) during that period have only about .07 or .08 % of lethal. Thus, the latter sperm, after 3 weeks, should contain something like .28 % of lethal."

On July 23, 1947 (American Philosophical Society 1947b), Stern writes Muller again stating that "I have mislaid your letter of some months ago (February 3, 1947, letter) in which you gave me some details of your own on the mutation rate under various physiological conditions. May I therefore ask you two questions and will you permit me to use your answers in a report which I am just preparing for the Manhattan Project? Obviously, full credit for it would be given. The questions are: (1) What is the spontaneous mutation rate in sperm derived from Canton-special males of from 3- to 6 days old? (2) What is the weekly increase in mutation rate of sperm from such males stored in females?"

On August 4, 1947 (Lilly Library 1947c), Muller responds "When sperm were stored in females, there was a weekly increase in the mutation frequency of about 0.07 %, on the average." On August 7, 1947 (American Philosophical Society 1947c), Stern cabled Muller asking him the temperature used and on August 8, 1947 (American Philosophical Society 1947d), Muller answered via cable indicating "25 °C." A subsequent undated letter, but most likely prior to September 9, 1947 (American Philosophical Society 1947e), Muller noted "A recalculation of my data gives the figure of 0.08 % instead of 0.07 % as the frequency of lethal accumulating in mature sperm per week." Since Uphoff and Stern (1947) did not include this correction in their report to the AEC it suggests that this undated letter was received after submittal of their report to the AEC.

The control value therefore used by Uphoff and Stern (1947) of 0.07 % for the estimated mutation rate of the sperm stored in the spermatheca was based on the earlier letter correspondence-supplied estimates of Muller (Lilly Library 1947b, c, February 3 and August 4) which Muller later clarified as being slightly in error.

The Caspari and Uphoff studies used Drosophila melanogaster fruit flies, breeding Canton-wild-type (S) males with Muller-5 females. Muller claimed (Lilly Library 1947c, August 4) that he never conducted mutation experiments with aged males of the Canton-wild-type stock. Muller stated that he had tested the aged sperm mutation frequency in "a number of different stocks (of Drosophila males) without finding any difference." The rate of increase on a weekly basis was said to be 0.07 % on average. This value of 0.07 % is believed to be prior to the correction to 0.08 %. This suggests that Muller did not observe significant inter-stock variation in mutation rates of the stored sperm.

Stern seems to have completed his Uphoff and Stern (1947) paper for the Manhattan Project during August, 1947. Stern knew that Uphoff's mean mutation frequency was 0.1682 % (0.1365–0.2097 %). This suggests a weekly mean increase in mutation rate of 0.0227 % (0.0122–0.0366 %), far lower than the 0.07 or 0.08 % mean weekly increase in Muller. When Stern wrote to Muller on September 9, 1947, he stated that for the Canton-special stock "…the weekly increase is considerably less than that found by you and others. It seems to be much more of the order of 0.03–0.05." This September 9,

1947, letter was written probably just after the submission of the Uphoff and Stern (1947) paper to the AEC, and definitely before the submission of the Caspari and Stern (1948) paper for publication by Genetics (i.e., November 25, 1947). Thus, the judgments of Uphoff and Stern that found that Uphoff's data were "uninterpretable" and that supported the reliability of the Caspari control data were made with the information provided by Muller during the summer of 1947. The apparent argument that Stern seems to be suggesting in his September 9, 1947, letter to Muller is that the Canton-wild-type stored sperm in the female may yield uniquely lower control mutation values. The argument is tenuous as the far higher weekly rate was consistently shown by multiple investigators, and with multiple Drosophila stocks, only being low in two Uphoff experiments. In fact, significant inter-strain differences on the frequency of dominant lethal mutations as induced by radiation were not reported in various Drosophila strains, including the Canton-special wild-type strain (Demerec and Fano 1944; Strömnaes 1951). This suggestion by Stern was not included in the Uphoff and Stern (1947) report.

This letter exchange between Stern and Muller fails to provide support for the later statements of Muller that Caspari's control group was unusually high. The Muller data and statements also do not provide support for the conclusion that the low Uphoff control data were in a normal range. None of this information was provided by Stern in his *Science* publication to permit the scientific community to better evaluate the Uphoff and Caspari control group data.

References

American Philosophical Society (1946) Muller letter to Stern. Stern Papers, Muller File—Box 16, 12 Nov

American Philosophical Society (1947a) Stern letter to Muller, 22 Jan 1947

American Philosophical Society (1947b) Stern letter to Muller, 23 July 1947

American Philosophical Society (1947c) Stern letter to Muller (cable), 7 Aug 1947

American Philosophical Society (1947d) Muller letter to Stern, 8 Aug 1947

American Philosophical Society (1947e) Stern letter to Muller, 9 Sept 1947

American Philosophical Society (1947f) Caspari letter to Stern. Stern Papers, 25 Sept 1947

American Philosophical Society (1947g) Caspari letter to Stern. Stern Papers, Box 21, 7 Oct

American Philosophical Society (Undated) Stern letter to Caspari, July-Aug 1947

American Philosophical Society (1949) Stern letter to Caspari. Stern Papers, 9 Feb 1949

American Philosophical Society (1973) A geneticists' journey, pp 1–19, Stern Nov 1973

Barron ESG (1954) The effect of x rays on systems of biological importance. In: Hollaender A (ed) Radiation biology, volume I: high energy radiation, chapter 5. McGraw-Hill Book Company, New York, pp 283–314

Beadle GW (1959) Ionizing radiation and the citizen. Sci Am 201:219–232

BEAR I (1956) Genetic effects of atomic radiation. Science 124:1157–1164

Bonnier G, Lüning HG (1949) Studies of X-ray mutations in the white and forked loci of *Drosophila melanogaster*. I. A Statistical analysis of mutation frequencies. Hereditas 35:116–189

Bonnier G, Lüning HG, Perje AM (1949) Studies of X-ray mutations on the white and forked loci of *Drosophila melanogaster*. II. A study of the formation of Gynandromphs and other kinds of mosaics. Hereditas 35:301–336

Byers L (1954) Thermal effects on the spontaneous mutation rate in mature spermatozoa of *Drosophila melanogaster*. Caryologia Suppl 1(6):694–696

Byers L, Muller HJ (1952) Influence of ageing at two different temperatures on the spontaneous mutation rate in mature spermatozoa of *Drosophila melanogaster*. Genetics 37(5):570–571

Calabrese EJ (2009) The road to linearity: why linearity at low doses became the basis for carcinogen risk assessment. Arch Toxicol 83:203–225

Calabrese EJ (2011a) Toxicology rewrites its history and rethinks its future: giving equal focus to both harmful and beneficial effects. Environ Toxicol Chem 30(12):2658–2673

Calabrese EJ (2011b) Key studies to support cancer risk assessment questioned. Environ Mol Mutagen 52(8):595–606

Calabrese EJ (2011c) Muller's Nobel lecture on dose-response for ionizing radiation: ideology or science? Arch Toxicol 85(12):1495–1498

Calabrese EJ (2012) Muller's Nobel Prize lecture: when ideology prevailed over science. Toxicol Sci 126(1):1–4

Caspari E, Stern C (1948) The influence of chronic irradiation with gamma rays at low dosages on the mutation rate in *Drosophila melanogaster*. Genetics 33:75–95

Catcheside DG (1950) Radiations and genetics. Practitioner 165(990):590–593

Crow JF (1957a) Testimony—Statement of Dr. James F. Crow, Professor of Genetics and Zoology, University of Wisconsin. Hearings before the Special Subcommittee on Radiation of the Joint Committee on Atomic Energy, Congress of the United States. 85th Congress, 1st session, part 1. United States Government Printing Office, Washington DC

Crow JF (1957b) Genetic considerations in establishing maximum radiation doses. Radiology 69(1):18–22

Crow JF (1995) Quarreling geneticists and a diplomat. Genetics 140:421–426

Demerec M, Fano U (1944) Frequency of dominant lethal induced by radiation in sperms of *Drosophila melanogaster*. Genetics 29:348–360

Erk FC (2009) Biographical memoirs. H. Bentley Glass. Proc Am Philos Soc 153(3):327–339

Evans RD (1949) Quantitative inferences concerning the genetic effects of radiation on human beings. Science 109:299–304

Giles N (1940) Spontaneous chromosome aberrations in Tradescantia. Genetics 25:69–87

Glass B (1957a) Testimony—Statement of Dr. Bentley Glass, Professor of Biology, the Johns Hopkins University. Hearings before the Special Subcommittee on Radiation of the Joint Committee on Atomic Energy, Congress of the United States. 85th Congress, 1st session, part 1. United States Government Printing Office, Washington DC

Glass B (1957b) The genetic basis for the limitation of radiation exposure. Am J Roentgen Radium Ther Nucl Med 78(6):955–960

Glucksmann A (1950) Cytological aspects of protection from ionizing radiations. Br J Radiol 23(265):41–45
Graf U (1972) Spontaneous mutations in *Drosophila melanogaster*. Humangen Hum Genet 16(1):27–32
Haas Fl, Clark JB, Wyss O, Stone WS (1950) Mutations and mutagenic agents in bacteria. Am Nat 84(817):261–274
Hamblin JD (2007) A dispassionate and objective effort: negotiating the first study on the biological effects of atomic radiation. J Hist Biol 40(1):147–177
Hanson FB, Heys F (1929) An analysis of the effects of the different rays of radium in producing lethal mutations in Drosophila. Am Nat 63:201–213
Higgins E (1951) Atomic radiation hazards for fish. J Wildl Manag 15(1):1–12
Hollaender A, Stapleton GE (1959) Ionizing radiation and the living cell. Sci Am 201(3):94–100
Joint Committee on Atomic Energy (1957) 85th Congress, 1st session. Summary-analysis of hearings 27–29 May, and 3–7 June, 1957 on the nature of radioactive fallout and its effect on man. United States Government Printing Office, Washington DC, Aug 1957
Jolly JC (2004) Thresholds of uncertainty: radiation and responsibility in the fallout controversy. Ph.D. Dissertation. Oregon State University, Corvallis
Kaufmann BP (1947) Spontaneous mutation rate in Drosophila. Am Nat 81:77–80
Kelner A, Bellamy WD, Stapleton GE, Zelle MR (1955) Symposium on radiation effects on cells and bacteria. Bacteriol Rev 19:22–44
Lefevre G Jr (1950) X-ray induced genetic effects in germinal and somatic tissue of *Drosophila Melanogaster*. Am Nat 84(818):341–365
Lewis EB (1957) Leukemia and ionizing radiation. Science 125(3255):965–972
Lilly Library (1938a) Indiana University, Bloomington, IN. Stern letter to Demerec, 21 Jan 1938
Lilly Library (1938b) Indiana University, Bloomington, IN. Demerec letter to Stern, 16 Mar 1938
Lilly Library (1938c) Indiana University, Bloomington, IN. Stern letter to Demerec, 31 Mar 1938
Lilly Library (1946) Indiana University, Bloomington, IN. Muller letter to Stern, 13 Sept 1946
Lilly Library (1947a) Indiana University, Bloomington, IN. Muller letter to Stern, 14 Jan 1947
Lilly Library (1947b) Indiana University, Bloomington, IN. Muller letter to Stern, 3 Feb 1947
Lilly Library (1947c) Indiana University, Bloomington, IN. Muller letter to Stern, 4 Aug 1947
Lilly Library (1949) Indiana University, Bloomington, IN. Muller letter to Stern, 5 Feb 1949
Lipshitz HD (2005) From fruit flies to fallout: Ed Lewis and his science. Dev Dyn 232:529–546
Lüning KG (1954) Effect of oxygen on irradiated males and females of *Drosophila*. Hereditas 40(3–4):295–312
MacKey J (1951) Neutron and x-ray experiments in barley. Hereditas 37(3):421–464
Muller HJ (1945) Age in relation to the frequency of spontaneous mutations in Drosophila. Yearb Am Philos Soc 150–153
Muller HJ (1946a) The production of mutations. Nobel Lecture, 1946. Nobleprize.org. http:www.nobelprize.org/nobel-prizes/medicine/laureates/1946
Muller HJ (1946b) Age in relation to the frequency of spontaneous mutations in Drosophila. In: Year book 1 Jan 1945, 31 Dec 1945. The American Philosophical Society, Philadelphia, PA, pp 150–153
Muller J (1947) The production of mutations. J Hered 38(9):259–270
Muller HJ (1948) Mutational prophylaxis. Bull N Y Acad Med 24(7):447–469
Muller HJ (1950a) Some present problems in the genetic effects of radiation. J Cell Comp Physiol 35(suppl 2):9–70
Muller HJ (1950b) Radiation damage to the genetic material. Am Sci 38(1):32–59, 126
Muller HJ (1952) Genetic effects of cosmic radiation. Chapter XVII. (Reprinted from Physics and Medicine of the Upper Atmosphere)
Muller HJ (1954a) The nature of the genetic effects produced by radiation. In: Hollaender A (ed) Radiation biology. Volume I: high energy radiation, chapter 7. McGraw-Hill Book Company, New York, pp 351–473
Muller HJ (1954b) The manner of production of mutations by radiation. In: Hollaender A (ed) Radiation biology. Volume I: high energy radiation, chapter 8. McGraw-Hill Book Company, New York, pp 475–626
Muller HJ (1956) How radiation changes the genetic constitution. In: Proceedings of the international conference on the peaceful uses of atomic energy. Volume 11, biological effects of radiation, held in Geneva Aug 1955. United Nations, New York, NY, pp 387–399
Muller HJ (1957a) How radiation changes the genetic constitution: congressional testimony [reprinted from the Bulletin of the Atomic Scientists, Nov 1955]
Muller HJ (1957b) Potential hazards of radiation—congressional testimony. Hearings before the Special Subcommittee on Radiation of the Joint Committee on Atomic Energy, Congress of the United States. 85th Congress, 1st session, part 1. United States Government Printing Office, Washington DC
Muller HJ, Carlson E, Schalet A (1961) Mutation by alteration of the already existing gene. Genetics 46:213–226
Newcombe B (1960) Genetic effects of ionizing radiation. Can J Biochem Physiol 38(3):330–337
Norwood WD (1958) Common sense approach to the problem of genetic hazard due to diagnostic radiology. JAMA 167(16):1928–1935
Nybom N, Gustafsson A, Granhall I, Ehrenberg L (1956) The genetic effects of chronic-gamma irradiation in barley. Hereditas 42(1–2):74–84
Oliver CP (1930) The effect of varying the duration of x-ray treatment upon the frequency of mutation. Science 71:44–46
Oliver CP (1931) An analysis of the effect of varying the duration of x-ray treatment upon the frequency of mutations. Doctor of Philosophy Thesis, University of Texas. Austin, Texas
Rajewski BN, Timofeeff-Ressovsky NW (1939) Höhen-Strahlung und die Mutationsrate von *Drosophila melanogaster*. ZIAV 77:488–500
Ray-Chaudhuri SP (1944) The validity of the Bunsen-Roscoe law in the production of mutations by radiation of extremely low intensity. Proc R Soc Edinb 62:66–72
Rinehart RR (1969) Spontaneous sex-linked recessive lethal frequencies from aged and non-aged spermatozoa of *Drosophila melanogaster*. Mutat Res 7:417–423
Sankaranarayanan K, Wassom JS (2008) Reflections on the impact of advances in the assessment of genetic risks of exposure to ionizing radiation on international radiation protection recommendations between the mid-1950s and the present. Mutat Res 658:1–27
Sax K (1950) The cytological effects of low-intensity radiation. Science 112(2908):332–333
Singleton WR (1954a) The effect of chronic gamma radiation on endosperm mutations in maize. Genetics 39:587–603
Singleton WR (1954b) Radiation effect on living systems. J Hered 45:58–64
Sparrow AH, Singleton WR (1953) The use of radiocobalt as a source of gamma rays and some effects of chronic irradiation on growing plants. Am Nat 87(832):29–48
Spear FG (1958) Some biological aspects of experimental radiology. A historical review. Intern Rev Cytol 7:1–77

Spencer WP, Stern C (1948) Experiments to test the validity of the linear R-dose/mutation at low dosage. Genetics 33:43–74

Stern C (1950) Principles of human genetics. WH Freeman and Company, San Francisco

Stern C (1960) Principles of human genetics, 2nd edn. WH Freeman and Company, San Francisco

Stone RS (1952) The concept of a maximum permissible exposure. Radiology 58(5):639–661

Strömnaes O (1951) X-ray induced lethal mutations in several strains of *Drosophila melanogaster*. Hereditas 37:533–559

Timoféeff-Ressovsky NW, Zimmer KG, Delbruck M (1935) Uber die Natur der Genmutation und der Genstruktur. Nachrichten von der Gesellschaft der Wissenschaften zu Gottingen:Mathematische-Physikalische Klass, Fachgruppe VI, Biologie 1(13):189–245

Uphoff DE, Stern C (1947) Influence of 24-hour gamma-ray irradiation at low dosage on the mutation rate in Drosophila. MDDC-1492, US Atomic Energy Commission, pp 1–6. Hathi Trust Digital Library. Available at http://www.hathitrust.org

Uphoff DE, Stern C (1949) The genetic effects of low intensity in irradiation. Science 109:609–610

Health Physics Journal

THE INTEGRATION OF LNT AND HORMESIS FOR CANCER RISK ASSESEMENT OPTIMIZES PUBLIC HEALTH PROTECTION

--Manuscript Draft--

Manuscript Number:	
Full Title:	THE INTEGRATION OF LNT AND HORMESIS FOR CANCER RISK ASSESEMENT OPTIMIZES PUBLIC HEALTH PROTECTION
Article Type:	Forum Article
Keywords:	hormesis; LNT; risk assessment; cancer
Corresponding Author:	Edward J Calabrese, Ph.D. University of Massachusetts Amherst Amherst, Massachusetts UNITED STATES
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	University of Massachusetts Amherst
Corresponding Author's Secondary Institution:	
First Author:	Edward J Calabrese, Ph.D.
First Author Secondary Information:	
Order of Authors:	Edward J Calabrese, Ph.D.
	Dima Yzaji Shamoun, Ph.D.
	Jaap C. Hanekamp, Ph.D.
Order of Authors Secondary Information:	

Health Physics

SUBMISSION TO *HEALTH PHYSICS*

This form must accompany a new submission and *must be signed by all named authors* of the manuscript.

All corresponding authors submitting manuscripts will receive an acknowledgment of receipt letter that identifies the manuscript by number. If you do not receive this letter within standard mail time (i.e., allow more time for international mail), please contact the Editorial Office.

"This certifies that I have participated in the design of this work, the analysis of data (if applicable), and the writing of the manuscript. Neither this manuscript nor a similar version has been published or is being considered for publication elsewhere."

Manuscript title: The Integration of LNT and Hormesis for Cancer Risk Assessment Optimizes Public Health Protection

Author signatures:

Edward Calabrese

Dimitrios Kapiszi Shamoun

J.C. Hoult

Please provide the following information:

- Manuscript classification (check one):

____ Paper	____ Operational Topics article
____ Review Article	__X__ Forum Article
____ Note	____ Other

- List four key words for subject indexing (list available at http://edmgr.ovid.com/hpj/accounts/keyword_list.doc):
  hormesis, LNT, risk assessment, cancer
- Should your manuscript be accepted for publication, you will be assessed page charges of $70.00 per page. These page charges are essential to the continuing operations of the Journal. As is customary in scientific and technical publishing, authors are expected to pay these charges. Upon payment 50 reprints will be sent to the corresponding author.
- Please print or type the corresponding author's full name, address, telephone number, fax number, and email address, below.

Edward J. Calabrese, Ph.D.
Professor of Toxicology, School of Public Health & Health Sciences, Department of Environmental Health Sciences, Morrill I N344, University of Massachusetts, Amherst MA 01003, Phone: 413-545-3164, Fax: 413-545-4692, E-mail: edwardc@schoolph.umass.edu

Orig. Op.	OPERATOR:	Session	PROOF:	PE's:	AA's:	COMMENTS	ARTNO:

Health Physics Forum Article: Low Dose Symposium

THE INTEGRATION OF LNT AND HORMESIS FOR CANCER RISK ASSESSMENT OPTIMIZES PUBLIC HEALTH PROTECTION

Edward J. Calabrese, Ph.D.* ; Dima Yazji Shamoun, Ph.D. †; Jaap C. Hanekamp, Ph.D. ††

Conflicts of Interest and Source of Funding: Authors declare no conflict of interest. Effort sponsored by the Air Force Office of Scientific Research, Air Force Material Command, USAF, under grant number FA9550-13-1-0047 (to EJC). The U.S. Government is authorized to reproduce and distribute for governmental purposes notwithstanding any copyright notation thereon. Long-term research activities in the area of dose response have been supported also by awards from ExxonMobil Foundation over a number of years. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsement, either expressed or implied, of the Air Force Office of Scientific Research, U.S. Government, or Exxon Mobil Foundation.

*Corresponding Author

ABSTRACT

This paper proposes a new cancer risk assessment strategy and methodology which optimizes population-based responses by yielding the lowest disease/tumor incidence across the entire dose continuum. We argue that the optimization can be achieved by integrating two seemingly conflicting models, i.e., the linear no-threshold (LNT) and hormetic dose-response models. The integration would yield the optimized response at a risk of 10^{-4} with the LNT model. The integrative functionality of the LNT and hormetic dose response models provides an improved estimation of tumor incidence through model uncertainty analysis, and major reductions in cancer incidence via hormetic model estimates. This novel approach to cancer risk assessment offers significant improvements over current risk assessment approaches by revealing a regulatory sweet-spot that maximizes public health benefits while incorporating practical approaches for model validation.

Keywords: hormesis, LNT, risk assessment, cancer

INTRODUCTION

Over the past three decades there has been a resurgence of scientific interest in the hormetic dose response model. The resurgence has been largely driven by the major switch from whole animal to cell culture investigations, which has created the opportunity to efficiently and inexpensively test up to 11 concentrations (plus a control group), replicate it eight times, and evaluate the consistency of responses using a 96 well-plate. This improvement in study design has therefore provided opportunities to explore a broad concentration/dose range in considerable detail, standing in marked contrast to a typical high dose, whole animal bioassay using two to three treatment doses. In fact, from 1990 to 2010, the proportion of *in vitro* studies entered into the hormesis database rose from 47% to 70% of the total entries (Calabrese and Blain 2011).

The *in vitro* experimental transition occurred in parallel with marked advances in the specification of pharmacology/toxicological mechanisms, including the identification of a plethora of receptor and cell signaling pathways that mediate a vast array of hormetic dose responses. For example, Calabrese (2013) reported specific mechanisms for 400 different hormetic dose responses. This assessment indicated that the quantitative features of the hormetic dose response are independent of mechanism. In addition, several large-scale head-to-head comparisons amongst the hormetic, threshold, and the linear no-threshold (LNT) models revealed that the hormetic dose response far outperformed the other two models in making reliable predictions in the low-dose zone (Calabrese and Baldwin 2001, 2003; Calabrese et al. 2006, 2008, 2010).

These developments are significant since they indicate that the hormetic dose response is a central concept that can describe how biological systems respond to low stimulatory doses or to disruptions in homeostasis when the damage induced has been minimal to moderate. Based on

the evidence accumulated over the past 30 years (Calabrese 2008, 2011), a convincing case can be made that hormesis deserves not only formal consideration along with the LNT and threshold models in risk assessment applications, but having repeatedly outcompeted these two models, perhaps an even more prominent position.

A recent historical study (Calabrese 2015) indicates that the evidence used by the U.S. NAS BEAR I Committee, Genetics Panel in 1956 to usher LNT into its dominant position in cancer risk assessment is highly problematic. The Genetics Panel did not base its assessment on appropriate studies as well as falsifying and fabrication of the research record to enhance acceptance of its LNT recommendation. The Genetics Panel also failed to provide a justification for the switch from the threshold to the linear model. However, as a regulatory tool, LNT had two attractive features, namely, ease of application and likelihood to consistently overestimate risk (The U. S. Interagency Staff Group 1986). Therefore, once the NAS recommendation for LNT was accepted, risk estimates became highly sensitive to the LNT model, making the latter a regulatory gold standard without adequate validation.

In the case of the threshold dose response model, it had been adopted by the regulatory and medical communities in the 1930s, yet, during the 20^{th} century, it was never validated for making accurate predictions in the low dose zone (Calabrese 2011). Further, when tested about 60 years later, it performed very poorly (Calabrese and Baldwin 2001, 2003; Calabrese et al. 2006, 2008, 2010).

Having disregarded objective scientific protocol, the dose response judgments of regulatory agencies that have led to the adoption of the LNT and threshold models have therefore been seriously flawed. The hormetic dose response, on the other hand, has not only outperformed the alternative models in direct comparisons, it also best describes responses in cancer bioassays

for numerous high profile agents such as DDT (Sukata et al. 2002), dioxin (Kociba 1978), ionizing radiation (Ootsuyama and Tanooka 1991; Cheda et al. 2006; Mitchell 2006; Ishii-Ohba et al 2007; Lacoste-Collin et al. 2007; Tanooka 2011; Nomura et al. 2013), in the 24,000 mega mouse study with acetyl aminoflurene (Society of Toxicology 1981), and in numerous studies at various early stages in the process of tumor development (Koana et al. 2007, 2012; Elmore et al. 2011a, 2011b).

INTEGRATION OF LNT AND HORMESIS

Despite the above arguments criticizing LNT, the present paper is not proposing replacing the LNT model with the hormetic model. Instead, it presents the LNT and hormetic models under the light of model uncertainty analysis. We propose a unique scientific way to reconcile the two models in a manner that defines the so-called regulatory sweet spot for optimal public health protection; both models are needed to achieve this goal. This paper extends earlier work of Calabrese et al. (2015) by illustrating a practical means to harmonize the LNT and hormetic models for cancer risk assessment. As might be expected under such circumstances, conflicting entities or scientists with differing perspectives may seek at times a common goal (i.e., optimization of public health responses), all the while recognizing that scientific and/or policy differences would remain.

Given their contrasting implications, it may at first glance seem impossible to reconcile the LNT and hormetic dose response models. A type of convergence occurs, however, when the integration of the two models displays a dose where public health gains are optimized as compared to responses at other doses. The optimal dose is at the nadir of the hormetic dose-

response, which occurs at approximately the dose corresponding to 10^{-4} risk using the LNT model (Calabrese et al. 2015). This regulatory sweet spot provides the lowest tumor incidence based on an integration of the risk estimates of both models.

In order to derive such a convergence estimate, the hormetic-based dose corresponding to a 1% response from a specific chronic animal bioassay is estimated, using a standard benchmark dose (BMD) method. This BMD is then divided by 100 (or the product of two standard 10-fold EPA/FDA uncertainty factors for interspecies and interindividual variability) as is typically the case for non-carcinogenic regulatory risk assessment. This process yields a dose that also corresponds to a 10^{-4} risk using the LNT for a substantial number of data sets (Gaylor 1989). The nadir of the normal population response according to the hormetic model coincides with this 10^{-4} risk, whereas the optimal dose for the high-risk group using the hormetic model coincides with the 10^{-5} risk of the LNT (Calabrese and Cook 2005). Since the size of the high-risk segment with unique genetic predisposition features is uncertain, but is assumed to be approximately $\leq$ 1-5% of the population (Calabrese 1978), the overriding population-based optimized benefit would employ the 10^{-4} risk value.

The reconciliation of the hormetic and LNT models in this manner reasonably characterizes the model uncertainty range for cancer risk estimation. The convergence point at 10^{-4} M yields the lowest estimated population-based cancer risk by a considerable margin. If the dose was decreased to the 10^{-5} M value, the LNT would estimate proportionately (i.e. 90%) fewer cancers (i.e., few in absolute terms) but essentially all the considerable "absolute" benefits of the hormetic estimates would be lost based on the most likely range of the hormetic curve (Calabrese and Blain 2005, 2011).

To illustrate, suppose we use a bladder cancer model, with a background effect of 12,500 cases per year in a population of 10^6 people. The 10^{-4} risk estimates 100 new chemically-induced cancers in the same population. At the same time, the hormetic model predicts a 25% decrease in risk, thereby reducing the number of new bladder tumors by 3,150 (Calabrese et al. 2015). Loss of the hormetic benefit to the general population by lowering the dose to a 10^{-5} risk would push the population-response far outside the optimal zone. More specifically, while the LNT model would estimate 90 less affected people, approximately 3,150 new people would be predicted to develop this cancer, thereby creating a large net negative response at the lower dose corresponding to the 10^{-5} LNT risk.

The proposed new cancer risk assessment methodology is superior even if the hormetic dose response was incorrect and the LNT theoretically correct, as risk at 10^{-4} is far below the capacity to be detected even in the most powerful epidemiological studies. Like the current federal cancer risk assessments, the proposed scheme is not verifiable in light of the limitations with experimental/epidemiological methods, but, unlike the current federal cancer risk assessments, it serves to minimize model error. Thus, the proposed scheme can provide an objective basis for regulation.

The present study applies only to chronic animal bioassays which strongly dominate the chemical risk assessment domain. Incorporating the hormetic dose response model offers several significant additional functional components which support its usage, as described here. The hormetic model contains both threshold and hormetic features allowing for the dose response assessment to maintain flexibility if follow up experimental validation tests are undertaken.

Though the hormetic model is complex, it could, if needed, readily default to a threshold model with a long history of use in the risk assessment community and federal regulations.

Additionally, as inferred above, the hormetic dose response can be tested to determine whether it can be confirmed or discredited. A practical example of efficient validation is seen with the research of Sukata et al. (2002) who retested the capacity of DDT to produce liver tumors using liver foci in the F344 male rat. This short-term study involved a large number of doses and it confirmed the cancer-causing effects of DDT at high doses while identifying the hormetic response at low doses. This capacity for validation is not present with the LNT model, where low levels of risk (e.g., less than 1%) cannot be confidently detected and render the LNT model unfalsifiable. The capacity to verify scientific models, even when using short-term predicted biomarker endpoint as in the case of Sukata et al. (2002) is an essential feature of any science. This new proposal brings this important feature back to cancer risk assessment.

FOOTNOTES

*Professor of Toxicology, School of Public Health & Health Sciences, Department of Environmental Health Sciences, Morrill I N344, University of Massachusetts, Amherst MA 01003, Phone: 413-545-3164, Fax: 413-545-4692, E-mail: edwardc@schoolph.umass.edu

†Research Fellow, Mercatus Center, George Mason University, 3434 Washington Blvd., Arlington VA 22201, Phone: 703-220-7308, E-mail:dshamoun@mercatus.gmu.edu

††, University College Roosevelt, Lange Noordstraat 1, NL-4331 CB Middelburg, The Netherlands, Phone: 31-79-346-0304, E-mail: j.hanekamp@ucr.nl

REFERENCES

Calabrese EJ. Pollutants and High Risk Groups. New York: John Wiley & Sons; 1978.

Calabrese EJ. Hormesis: Why it is important to toxicology and toxicologists. Environ Toxicol Chem 27(7):1451-1474; 2008.

Calabrese EJ. Toxicology rewrites its history and rethinks its future: Giving equal focus to both harmful and beneficial effects. Environ Toxicol Chem 20(12):2658-2673; 2011.

Calabrese EJ. Biphasic dose responses in biology, toxicology and medicine: Accounting for their generalizability and quantitative features. Environ Poll 182:452-460; 2013.

Calabrese EJ. On the origins of the linear no-threshold (LNT) dogma by means of untruths, artful dodges and blind faith. Environ Res 142:432-442; 2015.

Calabrese EJ, Baldwin LA. The frequency of U-shaped dose responses in the toxicological literature. Toxicol Sci 62(2):330-338; 2001.

Calabrese EJ, Baldwin LA. The hormetic dose-response model is more common than the threshold model in toxicology. Toxicol Sci 71:246-250; 2003.

Calabrese EJ, Blain R. The occurrence of hormetic dose responses in the toxicological literature, the hormesis database: An overview. Toxicol Appl Pharmacol 202(3):289-301; 2006.

Calabrese EJ, Blain RB. Hormesis and plant biology. Environ Poll 157(1):42-48; 2009.

Calabrese EJ, Blain RB. The hormesis database: The occurrence of hormetic dose responses in the toxicological literature. Reg Toxicol Pharmacol 61(1):73-81; 2011.

Calabrese EJ, Cook RR. Hormesis: How it could affect the risk assessment process. Hum Exper Toxicol 24(5):265-270; 2005.

Calabrese EJ, Staudenmayer JW, Stanek EJ III, Hoffmann GR. Hormesis outperforms threshold model in National Cancer Institute antitumor drug screening database. Toxicol Sci 94(2):368-378; 2006.

Calabrese EJ, Stanek EJ III, Nascarella MA, Hoffmann GR. Hormesis predicts low dose-responses better than threshold models. Intl J Toxicol 2008; 27:369-378; 2008.

Calabrese EJ, Hoffmann GR, Stanek EJ III, Nascarella MA. Hormesis in high-throughput screening of antibacterial compounds in E. coli. Hum Exp Toxicol 29:667-677; 2010.

Calabrese EJ, Shamoun DY, Hanekamp JC. Cancer risk assessment: Optimizing human health through linear dose-response models. Fd Chem Toxicol 81:137-140; 2015.

Cheda A, Wrembel-Wargocka J, Nowosielska EM, Janiak MK. Immune mechanism of the retarded growth of tumor nodules in mice exposed to single low level irradiations with x-rays. Centr Eur J Immunol 31(1-2):44-50; 2006.

Elmore E, Lao XY, Redpath JL. Application of the human hybrid cell neoplastic transformation assay to the study of the effects of low doses and dose-rates of ionizing radiation. In Vitro Cell Dev Biol 47(1):S33-S33; 2011a.

Elmore E, Lao XY, Kapadia RL, Redpath JL. Effect of dose-rate on induction of neoplastic transformation in vitro by low doses of 232 MeV protons 87(6):622-627; 2011b.

Gaylor DW. Quantitative risk analysis for quantal reproductive and developmental effects. Environ Health Perspect 79:243-246; 1989.

Ishii-Ohba H, Kobayashi S, Nishimura M, Shimada Y, Tsuji H, Sado T, Ogiu T. Existence of a threshold-like dose for γ-ray induction of thymic lymphomas and no susceptibility to radiation-induced solid tumors in SCID mice. Mut Res 619:124-133; 2007.

Koana T, Okada MO, Ogura K, Tsujimura H, Sakai K. Reduction of background mutations by low-dose X irradiation of *Drosophila* spermatocytes at a low dose rate. Rad Res 167:217-221; 2007.

Koana T, TakahashiT, Tsujimura H. Reduction of spontaneous somatic mutation frequency by a low-dose X-irradiation of *Drosophila* larvae and possible involvement of DNA single-strand damage repair. Rad Res 177:265-2771; 2012.

Kociba RJ, Keyes DG, Beyer JE, Carreon RM, Wade CE, Dittenber DA, Kalnins RP, Faruson LE, Park CN, Barnard SD, Hummel RA, Humiston CG. Results of a 2-year chronic toxicity and oncogenicity study of 2,3,7,8-tetrachlorodibenzo-para-dioxin in rats. Toxicol Appl Pharmacol 46(2):279-303; 1978.

Lacoste-Collin L, Jozan S, Cances-Lauwers V, Pipy B, Gasset G, Caratero C, Courtade-Saidi M. Effect of continuous irradiation with a very low dose of gamma rays on life span and the immune system in SJL mice prone to B-cell lymphoma. Rad Res 168:725-732; 2007.

Mitchel REJ. Low doses of radiation are protective in vitro and in vivo: evolutionary origins. Dose-Response 4(2):75-90; 2006.

Nomura T, Sakai K, OgataH, Magae J. Prolongation of life span in the accelerated aging *klotho* mouse model, by low-dose-rate continuous γ irradiation. Rad Res 179:717-724; 2013.

Ootsuyama A, Tanooka H. Threshold-like dose of local β irradiation repeated throughout the life span of mice for induction of skin and bone tumors. Rad Res 125:98-101; 1991.

Society of Toxicology. Reexamination of the ED01 study. Adjusting for time on study. Fundam Appl Toxicol 1:67-80; 1981.

Sukata T, Uwagawa S, Ozaki K, Ogawa M, Nishikawa T, Iwai S, Kinoshita A, Wanibunchi H, Imaoka S, Funae Y, Okuno Y, Fukushma S. Detailed low-dose study of 1,1-bis(p-chlorophenyl)-2,2,2-trichloroethane carcinogenesis suggests the possibility of a hormetic effect. International Journal of Cancer 99(1):112-118; 2002.

Tanooka H. Meta-analysis of non-tumour doses for radiation-induced cancer on the basis of dose-rate. Int J Radiat Biol 87(7):645-652; 2011.

The U.S. Interagency Staff Group. Chemical Carcinogens: A review of the science and iIts associated principles. Environ Health Perspect 201-82; 1986.

Human & Experimental Toxicology (2000) 19, 32–40

www.nature.com/het

The marginalization of hormesis

EJ Calabrese*,[1] and LA Baldwin[1]

[1]*Department of Environmental Health Sciences, School of Public Health, University of Massachusetts, Amherst, MA 01003, USA*

Despite the substantial development and publication of highly reproducible toxicological data, the concept of hormetic dose-response relationships was never integrated into the mainstream of toxicological thought. Review of the historical foundations of the interpretation of the bioassay and assessment of competitive theories of dose-response relationships lead to the conclusion that multiple factors contributed to the marginalization of hormesis during the middle and subsequent decades of the 20th century. These factors include: (a) the close-association of hormesis with homeopathy lead to the hostility of modern medicine toward homeopathy thereby creating a guilt by association framework, and the carry-over influence of that hostility in the judgements of medically-based pharmacologists/toxicologists toward hormesis; (b) the emphasis of high dose effects linked with a lack of appreciation of the significance of the implications of low dose stimulatory effects; (c) the lack of an evolutionary-based mechanism(s) to account for hormetic effects; and (d) the lack of appropriate scientific advocates to counter aggressive and intellectually powerful critics of the hormetic perspective.
Human & Experimental Toxicology (2000) **19**, 32–40

Keywords: hormesis; low dose; stimulation; β-curve

Introduction

A recent detailed review of the toxicological literature from the late 19th century to the mid 1930s has revealed the publication of numerous papers indicating that the shape of the dose response curve often displayed a low dose stimulatory response followed by an inhibitory response at higher doses.[1] This was especially the case with respect to the responses of plants, bacteria and fungi. The endpoints measured typically involved growth rate, colony number, germination rate, time to germination, and physiological responses such as carbon dioxide production, sugar production and utilization, waste product generation, ammonification, nitrification and nitrogen-fixation. These low dose stimulatory responses became viewed as reproducible and broadly generalizable. The papers were published in leading scientific journals by investigators with notable reputations from outstanding institutions and universities. In fact, several of the leading researchers in the area of low dose stimulatory research involved Nobel prize winners and/or their students.

By 1905 the fundamental shape of the dose-response curve had been proposed[2] and is remarkably consistent with the modern hormetic β-curve for both the range and magnitude of stimulatory responses (Figure 1).[3] Furthermore, by the mid 1920s a scientific journal, *Zell Stimulations Forschungen* [i.e., Cell Stimulation Research[4]], was published in Germany to assess the biological significance of chemically-induced stimulatory responses typically of a low dose nature.

Despite this substantial development and publication of reproducible scientific data the concept of hormetic dose-response relationships became rapidly marginalized during the mid-decades of the 20th century in the field of toxicology and its related disciplines involving assessment of the effects of pesticides, disinfectants, environmental toxins and radioactivity. Such has been the case of its marginalization that the concept of hormesis rarely, if ever, merits even an historical note in any leading toxicological textbook. Thus, several generations of toxicologists have never been formally exposed to the extensive history of the hormetic response, including how it relates to the broader dose-response continuum, what its potential biological significance may be or how such a phenomenon may be studied. In fact, by ignoring this concept and adhering strongly to the regulatory belief that there are no biologically relevant effects below the no observed adverse effect level (NOAEL), a common and unmistakable conclusion that hormesis does not exist has pervaded the toxicology community.

*Correspondence: EJ Calabrese
Received 15 October 1999; Revised 15 October 1999; Accepted 15 October 1999

Within this framework it is easy to see how fair minded toxicologists of the 1990s could dismiss the concept of hormesis.[3,5] First, hormetic effects are generally of a limited nature, usually being only about 30–60% greater than the values of the unexposed controls at their maximum stimulatory response level. Such modest increases over controls may be readily dismissed as manifestations of background variability. Second, most mammalian toxicity studies involve very few doses. In fact, it appears that only 2% of toxicology studies published in the open literature have utilized six or more doses with only 10% of this two percentage number having three or more doses below the NOAEL (i.e., ~2 of each 1000 studies).[5] Given the low magnitude of stimulatory responses and the limited number of studies with an adequate number of doses especially in the low dose area, the modern mammalian toxicologist is not routinely exposed to studies providing an opportunity to evaluate hormesis as a biological hypothesis.

Therefore, the intention of this paper therefore is to assess the following: (1) how the toxicology community came to establish its concepts of dose-response relationships; (2) how such dose-response relationships were interpreted; (3) why the numerous observations of hormetic dose-response effects became marginalized by the toxicology community; and (4) why the concept of hormesis played no role in the development of the current tenets of toxicology and pharmacology with respect to the bioassay including its design, conduct and interpretation and its relationship to risk assessment principles and procedures.

Interpreting dose-response relationships

Background

The development of conceptual understandings of the dose-response curve has had a long history with its earliest, most definitive and sustained articulation derived from studies on chemical disinfectants [see review[6]]. Pasteur's work with disinfectants first definitively established that preservatives act by means of their toxic effects on microorganisms. These findings were extended by Lister in England to disinfectants and human disease causing microbes. Such scientific and medical milestones were followed by a vast amount of work which began a more formal characterization of the disinfectant properties of numerous agents and the development of the concept of a disinfectant ranking system using comparative potency rankings. However, it was not until 1886 that Robert Koch published the first systematic studies on disinfectants employing pure cultures of bacteria. He assessed the effects on anthrax spores of the then popular disinfectants carbolic acid (i.e., phenol) and sulfur dioxide, plus many other previously unstudied agents including mercuric chloride. Soon these findings were extended to practical applications by establishing a series of emulsified disinfectants thereby making use of soaps as effective agents to improve hygiene.[6] Thus, the waning years of the 19th century brought forth remarkable advances in the understanding of the biology of microorganisms, the development of methods for their evaluation, and their implications for human health and disease. Theories for the intrepretation of dose-response relationships were subsequently proposed and debated.

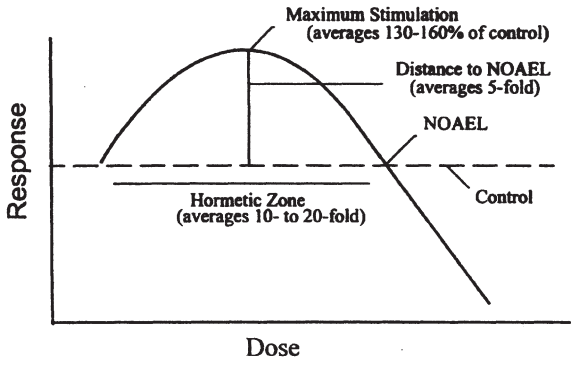


Figure 1 Dose-response curve depicting the characteristics of the chemical hormetic zone [modified from Calabrese and Baldwin[3]]. Abbreviations: NOAEL=no observed adverse effect level; LOAEL=lowest observed adverse effect level; ZEP=zero equivalent point

The unimolecular theory

In 1897 Kronig and Paul published an assessment of disinfectant properties using definite species of micro-organisms and disinfectants of every chemical class in widely varying concentrations.[7] This extremely comprehensive study also employed improved methods for more reliably quantifying bacteria. Such work set the stage for subsequent researchers to develop the means for standardization and comparison of various disinfectants (e.g., carbolic acid coefficient comparisons). The Kronig and Paul paper[7] was also important for the concept of hormesis since it established the importance of assessing biological effects of disinfectants over a broad concentration range and over time, thereby providing a foundation to establish dose-effect and dose-time-effect relationships.

The findings of Kronig and Paul[7] suggested a logarithmic relationship between the numbers of surviving bacteria and survival times. This logarithmic relationship was also reported by Madsen

and Nyman[8] and Chick[6] in the disinfection of anthrax spores with mercuric chloride and by Chick (1908) with phenol. Of significance is that Chick[6] observed that the dose-time-effect curve was quite similar to that expressing the course of a 'unimolecular reaction' (i.e., first order reaction). This follows directly from the Law of Mass Action where the velocity of any reaction is proportional to the active mass of reacting substance present at that moment. Therefore, Chick[6] concluded that the velocity of disinfection at any instant is proportional to the number (or weight) of living bacteria present.

Challenges to the unimolecular theory

Chick extended her original findings[6] of experiments with anthrax spores and *B. paratyphosus* to several additional bacterial models to enhance their generalizability.[9] Her striking claims, however, elicited an enormous debate in the microbiological literature over the next three decades and beyond, involving some of the most luminary scientists of the day [e.g., the Nobel laureate Arrhenius[10]]. The first wave of criticism was offered by Eijkman,[11] Hewlett[12] and Reichel,[13] followed by Loeb and Northrop,[14] Brooks,[15] Peters,[16] Smith,[17,18] Shackell,[19] Shackell *et al*,[20] Buchanan and Fulmer,[21] and Clark.[22]

The duration and intensity of the debate over how to interpret the shape of the dose-response curve reflects how pivotal this issue was in establishing the fundamental tenets of pharmacology, toxicology and even more recently with respect to current practices in risk assessment. In fact, the scope of the debate over the theory of unimolecular reactions was soon applied to the newly emerging field of radiation biology. More specifically, Blau and Altenburger,[23] supporting the view of Chick,[6,7] reported that the destruction of micro-organisms by radiation yielded unimolecular dose-response curves, leading those authors to conclude that the death of cells was produced by one or at most a few quanta of energy. Such an interpretation was subsequently adopted by early radiation response theoreticians.[24–27] Thus, the theoretical constructs of radiation dose-response evolved from the earlier debate sparked over the interpretation of disinfection dose-response issues. But, in fact, as pointed out by Packard[28] and Clark[22] the respective chemical and radiation dose-response interpretational controversies are strikingly similar and essentially hinge on the resolution of the same biological/toxicological points.

By the 1930s the developing consensus had emerged that strict acceptance of the unimolecular theory (i.e., the logarithmic death rate) should be abandoned based on the following arguments originally put forth by Brooks[15] and refined by Buchanan and Fulmer[21]: (a) Therewa a lack of correspondence of the experimental curve with the theoretical curve at the beginning of the experiment. The experimental studies of Chick[6,9] had inadequate data at short intervals at the beginning of the studies. In fact, in numerous instances more than half of the cells had been destroyed by the end of the first time interval reported. Such experimental data precluded an adequate assessment of the distribution of susceptibility in the population of bacteria and is believed to have been a major factor leading to erroneous conclusions supporting the unimolecular response theory. (b) There was a lack of correspondence of actual survivors' curve with the logarithmic curve in the latter part of response curve. When the values of the velocity constant are small there is a tendency for them to decrease rather than remain constant as predicted. (c) There were difficulties in assumption of uniform susceptibility. The overwhelming data for biological studies indicate that not all cells or organisms are equally susceptible; in addition the ratio of susceptible to non-susceptible cells is not a constant (i.e., in equilibrium) as is assumed in the unimolecular theory. Variation in response to toxic substances is widespread rather than the exception. Finally, (d) there was difficulty associated with securing a theoretical basis for interpreting cell death as a unimolecular reaction.

The characteristic curve theory

As a result of the above arguments, the unimolecular interpretation gradually receded and was essentially replaced by what was initially called the 'characteristic curve' which estimates the distribution of individual variation within populations with respect to responses to xenobiotics. Numerous studies of a substantial nature were published which attempted to characterize the extent of variability with animal models in response to toxic substances. For example, in 1924 van Wijngaarden determined the lethal dose of digitalis in nearly 600 cats, which resulted in a symmetrical bell-shaped distribution of variation.[29] Similarly, Morrell and Chapman reported the response of a population of 1331 rats to a series of neoarsphenamine doses, which also showed a symmetrical distribution of variation over a moderate dose range.[30,31]

Perhaps the most influential assessment of the characteristic curve was the work of Trevan[32] which remains the principal cited reference in the major toxicology texts of the 1990s [e.g., *Cassarett and Doull's Toxicology*[33]]. In setting forth to determine the best manner to estimate quantal effects, Trevan

concluded that the preoccupation of previous attempts to estimate the minimal effective (or lethal) doses should be abandoned in favor of estimating the central tendency of the group response.[32] In fact, Trevan established the concept and terminology of the lethal dose 50 (LD50).[32] He also utilized the term 'characteristic' to define the curve expressing the percentage of mortality or other limiting biological effect induced by varying the dose of a drug on animals of a certain species. He went on to provide a statistical foundation for sample size and statistical power. The proposal to estimate a response descriptor that could be accurately measured provided a strong stimulus for further work. While Trevan published a number of follow-up papers, it was the in-depth and extensive writings of Bliss and, to a lesser but still important extent, Gaddum, who applied the fundamental concepts laid out by Trevan to the major biological disciplines. Bliss published statistical methods for interpreting dose-response functions separately for entomology, microbiology, physiology, pharmacology and toxicology.[34–43] His widely cited papers, along with those of Gaddum[44–46] became a powerful standard for dose-response modeling and interpretation that is widely used today.

The influence of these biostatistical leaders of the 1920s and 1930s on the subsequent generations of toxicologists cannot be underestimated. They also influenced the further development of probit analysis by Finney[47–49] with its subsequent widespread application to the field of toxicology. By 1940, such statistical thinking had been integrated into the assessment of toxicology studies including the cancer bioassay as well evidenced in the classic paper of Bryan and Shimkin.[50] This paper focused on assessing the nature of the dose-response curve for chemical carcinogens over a wide range of exposures. Furthermore, some 21 years later probit analysis became the foundation of the Mantel-Bryan method[51] for low dose cancer risk estimation based on extrapolation beyond the observable range of the experimental data.

Therefore, the two predominant theories for interpreting dose-response relationships, the unimolecular and the characteristic curve, dominated the toxicological landscape in the early decades of the 20th century. While the unimolecular theory was losing out to the characteristic curve, there were other influential attempts to reinterpret the characteristic curve as one expressing that of a chemical process rather than a manifestation of interindividual variation. In fact, this concept in different mathematical forms continues to the present time as seen in examples of various hit and stage theories of carcinogenesis. Other biostatistical transformations were also proposed, including the logistic method.[52–54] The use of logits with quantal data used the assumption that the logarithms of the individual doses are distributed in a complex curve slightly different from the normal curve used in the probit method. In contrast to the characteristic curve of Trevan, with its incorporation into the probit concept, Emmens[55] sought to account for the dose-mortality curve by chance alone using the logistic method. He suggested that if the concept of tolerance is rejected, the law of chance supports the use of logits. This concept was developed earlier by Yule[56] in a more defensible statistical framework via a random hit theory dose-response approach which offers a similar dose response curve to that of the probit method.

In the 1930s the chemical process alternative interpretation to the characteristic curve was proposed in the case of the kinetics of drug action. This was based on the assumption that individuals varied symmetrically with respect to the lethal dose of xenobiotic required to produce a given response [see[22] for a review] and that the rate of xenobiotic entrance into the cell depended on diffusion according to a unimolecular formula. Such an hypothesis offered an explanation for the observation that the time-action curve expressing the rate of destruction of a population of small organisms frequently approximates the curve describing a unimolecular reaction.

Role of hormesis in the dose-response debate

So significant has been the need to resolve the above interpretational issues that the widely cited text, *Handbook of Experimental Pharmacology* by AJ Clark devoted approximately 10% of its pages to critiquing the unimolecular theory and another 15% to that of the characteristic curve and its possible integration with kinetic processes.[22]

While there has been much debate over the unimolecular theory, the characteristic curve and chemical process/kinetic-mechanism-based curves, there was almost no discussion concerning the role of hormetic responses during the early decades of the 20th century when the bioassay was being formalized and its interpretation refined. The only explicit discussion of hormetic responses in relationship to generalizable dose-response functions was that of Clark[22] who refers to this phenomenon as the Arndt-Schulz Law and/or diphasic responses [the term hormesis, from the Greek word meaning 'to excite', was not proposed until 1943[59]]. Rahn[57,58] modeled the hormetic response, which he acknowledged as a

widespread and generalizable phenomenon. He attempted to offer an enzymatic explanation for the low dose stimulatory response by developing an analogy between the actions of temperature on chemical metabolism. It was assumed that the toxic compound is a catalyst enhancing enzyme activity as well as enzyme deterioration. The model was built upon the concept that there was a shifting of the optimum enzyme activity with time from higher to lower concentration of the poison. While Rahn[57,58] cited the example of zymase activity enhancement by arsenate as a striking example of this phenomenon, he concluded there were no sufficiently complete data to adequately evaluate this phenomenon in a modeling sense. It appears that this attempt by Rahn had no demonstrable impact on the subsequent interpretation of the bioassay.

The effort made by Clark[22] to assess the significance of hormetic responses was principally judgmental providing extremely limited information with respect to theoretical foundation, historical basis and presentation of relevant data. The Arndt-Schulz Law was characterized as being 'in accordance with homeopathic doctrines and hence has maintained a certain popularity', thereby suggesting that it drew its strength from a homeopathic perspective rather than a biological/toxicological tradition. Clark further claims that the chief objection to the Arndt-Schulz Law is that 'it is obviously untrue in the case of most drugs that have been studied carefully' without providing adequate documentation to support this statement. He also states that 'many of the effects which appear to support this law were found to have simple explanations', again without providing sufficient information to support this broad conclusion. Based on the information presented on the historical foundations of hormesis by Calabrese and Baldwin,[3] it is apparent that the conclusions of Clark[22] were not only unsubstantiated but, in fact, were inconsistent with a large and generally available database that preceded his book. In addition, Clark's concept of hormesis was lacking in toxicological refinement in the sense that he collectively grouped all low dose stimulatory responses together including stimulatory responses of purified enzymes independent of their presence in a *bona fide* biological system. Despite the lack of an adequate review of the underlying foundation of the concept and data relating to the Arndt-Schulz Law, the strident judgments and highly critical characterizations of Clark are believed to have contributed in an important way to facilitate the marginalization of the hormetic concept within the toxicological community. Clark was a pharmacologist of exceptional reputation in the 1930s. He is credited with novel discoveries in the area of receptor mediated mechanisms. In fact, even as late as 1981 his original findings[60–62] were still being referred to as 'landmark' [see[63]]. In his historical foundations of the receptor theory, Parascandola[64] states that interaction between drugs and receptors was not treated quantitatively by pharmacologists until the research of AJ Clark and JH Gaddum in the 1920s and 1930s. Furthermore, in his foreward to the book '*Towards Understanding Receptors*' Robison[65] refers to the 1937 text of Clark[32] as the 'now classic monograph on General Pharmacology, a book that had great influence on a number of individuals.' Under normal circumstances such a negative statement would be challenged. However, none apparently did. Likewise, no alterative biostatistical models were presented to either challenge or refine those of Trevan, Bliss, and Gaddum which did not take the hormetic perspective into account.

The marginalization of hormesis

The marginalization of hormesis is therefore the result of a complex interactive web of circumstances which may be summarized as follows. First, the early identification of hormetic responses with the concept and practice of homeopathy created a 'guilt by association' situation. In fact, this concern was explicitly articulated by Hueppe in his 1896 book.[66] Even 50 years after the original association as stated by Arndt and Schulz [see[67] for review], Clark[22] continued to promote the close association of hormesis with homeopathy.

The strong association between homeopathy and hormesis (i.e., Arndt-Schulz Law) was based on the need of homeopathy to have an underlying biological foundation for this treatment's practices. The capacity for very low quantities of some poisons to act as stimulants to certain life functions as seen in some of the early research activities of Hugo Schulz[68] provided the scientific foundation for this relationship. While the low dose stimulatory actions of hormesis appear to be consistent with the tenets of homeopathy, this relationship is more apparent than real. More specifically, the stimulatory responses in homeopathy are alleged to become heightened as the solutions to be used in treatment became more dilute. In contrast, the stimulatory action seen in hormetic responses are of a very limited dose-range (usually not greater than a factor of ten, with a maximum stimulatory response of about 4–5-fold from the NOAEL. In addition, the most substantial theory of hormetic responses is that it represents an overcompensation to a limited disruption in homeostasis. Thus, if the experiment involves a properly timed sequence a clear decrement in response will

be observed followed by the hormetic overcompensation. Further, as a medical practice homeopathy treats patients with low doses of agents that elicit similar effects of their diseased condition (i.e., the so-called 'law of similars'). Hormesis has a different temporal arrangement such that an hormetic dose administered prior to a more massive dose as in the case of chemical and radiation adaptive responses may provide protection. Thus, there are important conceptual differences between homeopathy and hormesis that should have become progressively apparent over the decades since Arndt and Schulz sought to use hormesis as an underlying feature to this medical practice. In fact, as early as the late 1890s the temporal features of hormetic responses were being established by Townsend,[69–71] later refined by Branham[72] and Smith[73] and developed into a modern toxicological theory by Stebbing in the late 1970s and early 1980s [see[74] for a recent review]. The bias, therefore, against hormesis because of its relationship with the concept of homeopathy reflects (1) how strongly the homeopathic founders sought its association, (2) how intense the tension was between homeopathy and modern medicine, and (3) the lack of toxicological insight into the fundamental differences between the theoretical basis of homeopathy and the experimentally derived toxicological manifestations of hormetic responses.

Second, during the late 19th century up through at least the first half of the current century, the principal practical interests in dose-response curves of have focused on the upper end of the curve. This has been the case for disinfection biology, for the eradication of pests such as insects, weeds, etc., as well as the need to characterize hazardous situations especially for occupational exposures to workers. Even Winslow, who played such a critical role in establishing hormetic responses in bacteria demonstrably involved himself in the debate over the assertions of Chick and the unimolecular theory of dose response relationships[75] even though Falk,[76] a colleague of Winslow, emphasized the high significance of the low dose stimulatory responses of Hotchkiss,[77] Winslow's PhD student. Winslow never articulated its (i.e., the hormetic response) implications for the generalized dose-response function. Unfortunately, Hotchkiss did not follow up on her findings but redirected her research to other bacteriological questions. Despite the lack of follow through by Winslow and Hotchkiss it should be noted that a number of subsequent leading microbiological textbooks [e.g.,[78–80]] and significant research papers [e.g.,[81]] highlighted these findings and emphasized the hormetic nature of disinfectant dose-response relationships. Nonetheless, these multiple leading microbiological publications had little apparent impact on other related fields in relationship to establishing hormetic effects in the dose-response continuum.

Even though such references as noted above emphasized the regular occurrence of low dose stimulatory responses, the widely cited paper by Marshall and Hrenoff[81] explicitly stated that the stimulatory response 'is frequently of no practical value'. The inclusion of the stimulatory range in the dose-response continuum at least as far as disinfectants was involved, was principally for completeness rather than relevance. Such interpretations would have contributed to a further lowering of the concept of hormesis as a research priority, thereby perhaps inadvertently further trivializing its status.

Third, the lack of a sound evolutionary-based mechanism providing a credible conceptual framework for both evaluating and interpreting new experimental data was a serious limitation affecting the incorporation of the hormetic concept into the mainstream of biological/toxicological theory. This lack of a coherent framework is evident in the limited and confused presentation on this topic in 1937 by Clark, a highly respected pharmacologist.[22] If the evolutionary based mechanism of Stebbing[74] that hormesis may represent an overcorrection of a disruption in homeostasis had been available, it is likely that hormesis would have been seen more as a manifestation of an evolutionary based adaptive strategy than as part of the rationale for the practice of homeopathy.

Fourth, hormetic responses needed not only a leading toxicological advocate to develop its evolutionary basis but also a biostatistical/dose-response modeling one. For example, when Winslow was at Yale University (New Haven, CT, USA), the outstanding biostatistician Bliss was at the Connecticut Agricultural Research Station also in New Haven, Connecticut. If Winslow and Bliss had collaborated on this issue, it is likely that the course of low dose modeling of biological effects would have changed. However, Bliss did not arrive in New Haven until approximately 1943, some 20 years after the ground breaking work of Hotchkiss[77] and toward the end of Winslow's career. It should be mentioned that Bliss[38] was clearly aware of the 1937 book by Clark[22] which was so critical of the Arndt-Schulz Law. In addition, in 1935 Bliss acknowledged the assistance of Clark in his paper on estimating the dosage-mortality curve.[34] Furthermore, in 1936 Clark provided the opening address at a discussion of the chemical and physical foundations of pharmacological actions where one of the leading presenters was Gaddum [see[82]]. In addition, Clark and Gaddum were professors at the same university (University of Edinburgh) and in the same academic

department! Gaddum[44] also acknowledged the assistance of Trevan in reviewing his manuscript for publication. Such a close interaction of Clark with the leading biostatisticians of that era and the acknowledged respect for his professional advice would seem to have precluded the necessary biostatistical interest in the concept of hormesis.

The only notable figure in the biostatistical area, who potentially could have readily explored the statistical features of hormetic responses would have been the American Frank Wilcoxon, famous for the Wilcoxon Signed Rank Test, who explicitly reported hormetic responses in several fungal species exposed to low levels of hydrogen sulphide.[83] Nonetheless, McCallan and Wilcoxon[83] went on to emphasize the upper end of the dose-response curve with no further consideration of the significance of these observations. Note that Clark[22] discussed the report of McCallan and Wilcoxon[83] in considerable detail with no mention of the acknowledged and obvious low dose stimulatory responses.

Discussion

This paper described the historical unfoldings of the concept of dose-response relationships (i.e., dose-effect, dose-time-effect, etc.) and the intellectual struggle over how to interpret their biological significance. Resolution of such complex issues is rarely completed overnight as the debate over the unimolecular theory was very active for over three decades. This is often the nature of scientific debate that depends upon the conduct of new experiments, the subsequent publication of the data and the further refinement of the debate question.

The unimolecular theory had long staying power because of its theoretical linkage with a well-defined chemical process, its formulation in an established statistical model, and its strong advocates including leading researchers (e.g., Chick, Martin) at the world famous Lister Institute, the nobel laureate Svante Arrhenius and the highly regarded Cornell University Professor, Otto Rahn. The appeal to authority, even in scientific circles, can be substantial as Clark[22] noted that 'it is obvious that a physico-chemical theory (i.e., unimolecular theory) regarding the mode of action of drugs, which has received the support of Arrhenius must be considered carefully.'

While this constellation of factors can help to explain the long and continuing debate over the unimolecular theory of dose-response relationships, a more complex interaction of factors was likely at work which served to undermine the successful integration of the concept of hormesis into the mainstream of toxicological thought. Even though hormesis has been an unintended victim of the continuing struggle between traditional medicine and homeopathy, even though society has seen the upper end of the response curve as rightfully more significant than low dose responses, even though many decades of reproducible hormetic responses lacked coherent underlying mechanisms and even though mammalian toxicology studies have generally been relegated to the use of such small numbers of doses that hormetic effects could not be legitimately assessed, there still existed numerous cumulative *bona fide* examples in the literature leading up to the 1930s that should have accorded hormesis the status of a mainstream and an adequately supported biologically-based hypothesis. While any one of the above factors should not have been sufficient to marginalize this concept, their independent, yet collective, force is believed to have achieved that result.

Nonetheless, as has been shown, hormetic responses continued to be reported by numerous investigators over the following decades in various fields of biology. The biological basis therefore of hormesis as a credible hypothesis has been immeasurably strengthened over the past 50 years. Equally important, various toxicological explanations have begun to be offered which provide not only a sound theoretical foundation for hormesis, the hypothesis, but also a means of experimentally evaluating these proposed mechanisms [see[74,84]]. In addition, the issue of low dose cancer risks based on biostatistical models has clearly elevated the significance of hormesis in the risk assessment debate. Likewise, advances in molecular biology are also revealing that hormetic mechanisms may have important biomedical implications with respect to affecting adaptive responses of patients. These collective and relatively recent advances provide a foundation upon which hormesis as a biological hypothesis may now receive a more objective and comprehensive assessment.

Acknowledgements

This article was previously published in *Toxicologic Pathology* 1999; **27**:187–194 and has been reproduced in *Human and Experimental Toxicology* with the approval of *Toxicologic Pathology* and the Society of Toxicologic Pathologists. This work has been sponsored in part by an award to the University of Massachusetts (Edward J Calabrese, PhD, Principal Investigator) by the Texas Institute for Advancement of Chemical Technology, Inc. (TIACT). The text is based upon Appendix 8 of *Hormesis: The Stimulatory Action of Sub-Inhibitory Amounts of Toxins*, a report prepared for TIACT.

References

1 Calabrese EJ, Baldwin LA. Chemical hormesis: Its historical foundations as a biological hypothesis. *Toxicologic Pathology* 1999; **27:** 195–216.
2 True RH, Oglevee CS. The effect of the presence of insoluble substances on the toxic action of poisons. *Botanical Gazette* 1905; **39:** 1–21.
3 Calabrese EJ, Baldwin LA. A quantitatively-based methodology for the evaluation of chemical hormesis. *Human and Ecological Risk Assessment* 1997; **3:** 545–554.
4 *Zell-Stimulations-Forschungen.* **Vol. 1–3.** M Popoff and W Gleisberg (eds.). Paul Parey: Berlin, 1924–1930.
5 Calabrese EJ, Baldwin LA. The dose determines the stimulation (and poison): Development of a chemical hormesis database. *International Journal of Toxicology* 1997; **16:** 545–559.
6 Chick H. An investigation of the laws of disinfection. *Journal of Hygiene* 1908; **8:** 92–158.
7 Kronig B, Paul T. Die chemischen Grudlagen der Lehre von der Ciftwirkung und Disinfection. *Zeitschrift fur Hygiene und Infectionskrankenheiten* 1897; **25:** 1–112.
8 Madsen T, Nyman M. Zur Theories der Desinfektion. I. *Zeitschrift fur Hygiene und Infectionskrankenheiten* 1907; **57:** 388–404.
9 Chick H. The process of disinfection by chemical agents and hot water. *Journal of Hygiene* 1910; **10:** 237–286.
10 Arrhenius S. Quantitative Laws in Biological Chemistry. London, 1915.
11 Eijkman C. Die Ueberlebungskurve bei Abstotung der Bakterien durch Hitze. *Biochemische Zeitschrift* 1908; **11:** 12–20.
12 Hewlett RT. The Milroy Lectures of disinfection and disinfectants. *The Lancet* 1909; March **13:** 20–27.
13 Reichel H. Zur theorie der desinfektion. I, II, and III. *Biochemische Zeitschrift* 1909; **22:** 149–177, 177–199, 201–232.
14 Loeb J, Northrop JH. On the influence of food and temperature upon the duration of life. *Journal of Biological Chemistry* 1917; **32:** 103–121.
15 Brooks SC. A theory of the mechanism of disinfection, hemolysis and similar processes. *Journal of General Physiology* 1918; **1:** 61–80.
16 Peters RA. Variations in the resistance of protozoon organisms to toxic agents. *Journal of Physiology* 1920; **54:** 260–266.
17 Smith JH. The killing of Botrytis spores by phenol. *Annals of Applied Biology* 1921; **8:** 27–50.
18 Smith JH. The killing of Botrytis cinerea by heat, with a note on the determination of temperature coefficients. *Annals of Applied Biology* 1923; **10:** 335–347.
19 Shackell LF. The relation of dosage to effect. II. *Journal of Pharmacology and Experimental Therapeutics* 1925; **25:** 275–288.
20 Shackell LF, Williamson W, Deitchman MM, Katzman GM, Kleinman BS. The relation of dosage to effect. *Journal of Pharmacology and Experimental Therapeutics* 1924/1925; **23/24:** 53–65.
21 Buchanan RE, Fulmer EI. Physiology and Biochemistry of Bacteria. Vol. Effects of Environment Upon Microorganisms. The Williams and Wilkins Co.: Baltimore, MD, 1930.
22 Clark AJ. Handbook of Experimental Pharmacology. Verlig Von Julius Springer: Berlin, 1937.
23 Blau M, Altenburger K. Uber einige Wirkungen von Strahlen. II. *Zeitschrift fur Physik* 1922; **12:** 315–329.
24 Crowther JA. The action of x-rays on *Colpidium colpoda*. *Proceedings of the Royal Society of London B* 1926; **100:** 396.
25 Holweek F, Lacassagne A. Sur le mecanisme de l'action cytogaustique des radiations. *Comptes Rendus Societe Biologie de Paris* 1930; **103:** 766–768.
26 Wyckoff RWG. The killing of certain bacteria by x-rays. *Journal of Experimental Medicine* 1930; **52:** 435–447.
27 Pugsley AT, Oddie TH, Eddy CE. The action of X-rays on certain bacteria. *Proceedings of the Royal Society of London B* 1935; **118:** 276–298.
28 Parkard C. The biological effects of short radiations. *Quarterly Review of Biology* 1931; **6:** 253–280.
29 van Wijngaarden L. *Archiv fur Experimentelle Pathologie und Pharmakologie* 1924; **113:** 40 [cited in 22].
30 Morrell CA, Chapman CW. On the determination of the toxicity of neoarsphenamine I. The increase in toxicity on exposure to air. *Journal of Pharmacology and Experimental Therapeutics* 1933; **48:** 375–390.
31 Morrell CA, Chapman CW. On the determination of the toxicity of neoarsphenamine II. The determination of the characteristic curve for rats. *Journal of Pharmacology and Experimental Therapeutics* 1933; **48:** 391–409.
32 Trevan JW. The error of determination of toxicity. *Proceedings of the Royal Society of London B* 1927; **101:** 483–514.
33 Klaassen C (editor). Casserett and Doull's Toxicology: The Basic Science of Poisons, Fifth Edition, McGraw-Hill: New York, 1996.
34 Bliss CI. Estimating the dosage-mortality curve. *Journal of Economic Entomology* 1935; **25:** 646–647.
35 Bliss CI. The calculation of the dosage-mortality curve. *Annals of Applied Biology* 1935; **22:** 134–167.
36 Bliss CI. The comparison of dosage-mortality data. *Annals of Applied Biology* 1935; **22:** 307–333.
37 Bliss CI. The toxicity of poisons applied jointly. *Annals of Applied Biology* 1939; **26:** 585–615.
38 Bliss CI. The relation between exposure time, concentration and toxicity in experiments on insecticides. *Annals of the Entomological Society of America* 1940; **33:** 721–766.
39 Bliss CI. Biometry in the service of biological assay. *Industrial Engineering and Chemistry* 1941; **13:** 84–88.
40 Bliss CI. The calculation of microbial assays. *Bacteriological Reviews* 1956; **20:** 243–258.
41 Bliss CI. Some principles of bioassay. *American Scientist* 1957; **45:** 449–466.
42 Bliss CI, Cattell M. Biological Assay. *Annual Review of Physiology* 1943; **5:** 479–539.
43 Bliss CI, Packard C. Stability of the standard dosage-effect curve for radiation. *American Journal of Roentgenology and Radium Therapy* 1941; **46:** 400–404.

44 Gaddum JH. The action of adrenalin and ergotamine on the uterus of the rabbit. *Journal of Physiology* 1926; **61:** 141–150.
45 Gaddum JH. Lognormal distributions. *Nature* 1945; **156:** 463–466.
46 Gaddum JH. Bioassays and mathematics. *Pharmacological Reviews* 1953; **5:** 87–134.
47 Finney DJ. The application of the probit method to toxicity test data adjusted for mortality in the controls. *Annals of Applied Biology* 1943/1944; **30/32:** 68–74.
48 Finney DJ. The choice of a response metameter in bio-assay. *Biometrics* 1949; **5:** 261–272.
49 Finney DJ. Probit Analysis, A Statistical Treatment of the Sigmoid Response Curve. 2nd. Edition, Cambridge Univ. Press: London, 1952.
50 Bryan WR, Shimkin MB. Quantitative analysis of dose-response data obtained with carcinogenic hydrocarbons. *Journal of the National Cancer Institute* 1940; **1:** 807–833.
51 Mantel N, Bryan WR. 'Safety' testing of carcinogenic agents. *Journal of the National Cancer Institute* 1961; **27:** 455–470.
52 Reed LJ, Berkson J. The application of the logistic function to experimental data. *Journal of Physical Chemistry* 1929; **33:** 760–779.
53 Berkson J. Application of the logistic function to bio-assay. *Journal of the American Statistical Association* 1944; **39:** 357–365.
54 Berkson J. Why I prefer logics to probate. *Biometrics* 1951; **7:** 329–339.
55 Emmens CW. The dose/response relation for certain principles of the pituitary gland, and of the serum and urine of pregnancy. *Journal of Endocrinology* 1940/1941; **2:** 194–225.
56 Yule GU. On the distribution of deaths with age when the causes of death act cumulatively, and similar frequency distributions. *Journal of the Royal Statistical Society* 1910; **73:** 26–38.
57 Rahn O. Der Einfluss der Temperature und der Gifte aluf Enzymwirkung, Garung und Wachstrum. *Biochemische Zeitschrift* 1915; **72:** 351.
58 Rahn O. Physiology of Bacteria. P. Blakiston's Son and Co.: Philadelphia, 1932.
59 Southam CM, Ehrlich J. Effects of extracts of western red-cedar heartwood on certain wood-decaying fungi in culture. *Phytopathology* 1943; **33:** 517–524.
60 Clark AJ. The reaction between acetyl choline and muscle cells. *Journal of Physiology* 1926; **61:** 530–546.
61 Clark AJ. The antagonism of acetyl choline by atropine. *Journal of Physiology* 1926; **61:** 547–556.
62 Clark AJ. Mode of Action of Drugs on Cells. Arnold: London, 1933.
63 Hollenberg MD, NexΦ E. Receptor binding assays. In: Jacobs S and Cuatrecasas P (eds). *Membrane Receptors. Series B*, Vol. 11. Chapman and Hall: London, 1981, pp. 1–22.
64 Parascandola J. Origins of the receptor theory. In: Lamble JW (ed). *Toward Understanding Receptors.* Elsevier/North-Holland: New York, 1981, pp. 1–7.
65 Robison GA. Forward. In: Lamble JW (ed). *Towards Understanding Receptors.* Elsevier/North-Holland: New York, 1981, pp. v–x.
66 Hueppe F. Principles of Bacteriology. Translated from the German by E.O. Jordon. The Open Court Publishing Company: Chicago, 1896.
67 Boyd LJ. A Study of the Stimile in Medicine. Boericke and Tafel: Philadelphia, 1936.
68 Schulz H. Uber Hefegifte. *Pfluger's Archiv fur die gesemmte Physiologie* 1888; **42:** 517–541.
69 Townsend CO. The effects of ether upon the germination of seeds and spores. *Botanical Gazette* 1899; **27:** 458–466.
70 Townsend CO. The effect of hydrocyanic acid gas upon the germination of seeds. *Scientific American Supplement* 1899; **48:** 20010–20011.
71 Townsend CO. The effect of hydrocyanic acid gas upon grains and other seeds. *Botanical Gazette* 1901; **31:** 241–264.
72 Branham SE. The effects of certain chemical compounds upon the course of gas production by Baker's yeast. *Journal of Bacteriology* 1929; **18:** 247–284.
73 Smith EC. Effects of ultra-violet radiation and temperature on Fusarium. II. Stimulation. *Bulletin of the Torrey Botanical Club* 1935; **62:** 151–164.
74 Stebbing ARD. A theory for growth hormesis. *BELLE Newsletter* 1997; **6:** 1–11.
75 Watkins JH, Winslow C-EA. Factors determining the rate of mortality of bacteria exposed to alkalinity and heat. *Journal of Bacteriology* 1932; **24:** 243–265.
76 Falk IS. The role of certain ions in bacterial physiology. A Review. (Studies of salt action. VII). *Abstract of Bacteriology* 1923; **7:** 33–50; 87–105; 133–147.
77 Hotchkiss M. Studies on salt action. VI. The stimulating and inhibitive effect of certain cations upon bacteria growth. *Journal of Bacteriology* 1923; **8:** 141–162.
78 Salle AJ. Influence of Environment Upon Bacteria. In: *Fundamental Principles of Bacteriology with Laboratory Exercises.* McGraw-Hill Book Company, Inc.: New York, 1939.
79 Clifton CE. Death of bacteria. In: *Introduction to Bacterial Physiology.* McGraw-Hill Book Company, Inc.: New York, 1957.
80 Lamanna C, Mallette MF. Basic Bacteriology: Its Biological and Chemical Background. 3rd Ed. The Williams & Wilkins Co.: Baltimore, 1965.
81 Marshall MS, Hrenoff AK. Bacteriostasis. *Journal of Infectious Diseases* 1937; **61:** 42–54.
82 Clark AJ. Discussion on the chemical and physical basis of pharmacological action, 12 November, 1936. *Proceedings of the Royal Society of London B* 1936/1937; **121:** 580–609.
83 McCallan SEA, Wilcoxon F. The fungicidal action of sulphur. II. The production of hydrogen sulphide by sulphured leaves and spores and its toxicity to spores. *Contributions to the Boyce Thompson Institute* 1931; **3:** 13–38.
84 Hart RW, Frame LT. Toxicological defense mechanisms and how they may affect the nature of dose-response relationships. *BELLE Newsletter* 1996; **5:** 1–16.

This article has been reproduced in *Human & Experimental Toxicology* with the approval of *Toxicologic Pathology* and the Society of Toxicologic Pathologiests

Environmental Research 142 (2015) 432–442

Contents lists available at ScienceDirect

Environmental Research

journal homepage: www.elsevier.com/locate/envres

On the origins of the linear no-threshold (LNT) dogma by means of untruths, artful dodges and blind faith

Edward J. Calabrese*

Department of Environmental Health Sciences, School of Public Health and Health Sciences, University of Massachusetts, Amherst, MA 01003, USA

ARTICLE INFO

Article history:
Received 2 June 2015
Received in revised form
16 July 2015
Accepted 17 July 2015

Keywords:
Risk assessment
Dose–response
Linear dose response
Cancer
Mutation
LNT
Ionizing radiation

ABSTRACT

This paper is an historical assessment of how prominent radiation geneticists in the United States during the 1940s and 1950s successfully worked to build acceptance for the linear no-threshold (LNT) dose–response model in risk assessment, significantly impacting environmental, occupational and medical exposure standards and practices to the present time. Detailed documentation indicates that actions taken in support of this policy revolution were ideologically driven and deliberately and deceptively misleading; that scientific records were artfully misrepresented; and that people and organizations in positions of public trust failed to perform the duties expected of them. Key activities are described and the roles of specific individuals are documented. These actions culminated in a 1956 report by a Genetics Panel of the U.S. National Academy of Sciences (NAS) on Biological Effects of Atomic Radiation (BEAR). In this report the Genetics Panel recommended that a linear dose response model be adopted for the purpose of risk assessment, a recommendation that was rapidly and widely promulgated. The paper argues that current international cancer risk assessment policies are based on fraudulent actions of the U. S. NAS BEAR I Committee, Genetics Panel and on the uncritical, unquestioning and blind-faith acceptance by regulatory agencies and the scientific community.

1. Introduction

In the course of recent assessments of the historical and scientific foundations of dose responses models, it was learned that the linear dose response model was deliberately promoted to advance ideological agendas of some of the world's most prestigious radiation geneticists (Calabrese, 2008; Calabrese, 2013a, 2015a, 2015b). These individuals intentionally misled/deceived the scientific and world communities at the highest possible levels, including in a 1946 Nobel Prize Lecture (Calabrese, 2011a; Calabrese, 2012), in their scientific publications (Calabrese, 2011b; Calabrese, 2013b; Caspari and Stern, 1948; Muller, 1950a, 1954; Uphoff and Stern, 1949), in their role as members of the U.S. NAS (Calabrese, 2013a; Calabrese, 2015b, 2015a) and in publications of the NAS [BEAR Committee, Genetics Panel – (Anonymous, 1956a; National Academy of Sciences NAS)/National Research Council NRC, 1956). Collectively, these deceptive actions became highly significant when they facilitated an unchallenged and blind-faith adoption of the Linear Dose Response (LDR) model for cancer risk assessment of ionizing radiation and later of chemical carcinogens (Calabrese, 2011b, 2013b, 2009a). The adoption of the LDR model affected the magnitude of financial resources involved in regulatory actions, toxic tort decisions and medical practices; it also affected risk communication messages to the general public, educational practices, governmental research funding priorities, as well as decisions related to lifestyle and child rearing.

The impact of these deceptions has been substantial and, to this day, they significantly affect and dominate regulatory policies and risk assessment practices. Since these disturbing findings were published as a series of separate papers in diverse scientific journals, (e.g. mutation, radiation and toxicology journals) (Calabrese, 2015b, 2015a, 2011c, 2012, 2011b, 2013b, 2009a, 2014a, Calabrese, 014b), it has become necessary to develop an integrated and holistic version of this complex story. In addition, newly unearthed materials on key individuals have been discovered and incorporated herein to clarify previous historical frameworks. Finally, critical feedback recently received from reviewers, editors and others in the research community has proven invaluable in tempering the perspective and improving the content and context of this assessment.

This paper follows an historical timeline, starting with the professional/scientific relationship between Hermann Muller and Curt Stern and their subsequent collaborations on ionizing radiation during the Manhattan Project. The many, and, at times, bizarre ways in which Stern tried to prevent acceptance of the threshold model supportive findings of Ernst Caspari, a member of

* Fax: +1 413 545 4692.
E-mail address: edwardc@schoolph.umass.edu

http://dx.doi.org/10.1016/j.envres.2015.07.011
0013-9351/

the Manhattan Project team, in order to promote the LNT model, are detailed. Muller's Nobel Prize Lecture with emphasis on his assessment of the nature of the dose response in the low dose range, especially in light of the Caspari findings, is critiqued, leading to an assessment of how he and Stern acted to cover up Muller's Nobel Prize Lecture deceit via obfuscation of the Manhattan Project findings and the strikingly false subsequent statements of Muller in the scientific literature. The paper then assessed how the leadership of Muller and Stern profoundly affected beliefs on dose response within the genetics community during the 1950s, especially seen through the actions of the NAS BEAR I Genetics Panel in 1956 which assured the acceptance of the LNT by falsifying and fabricating the research record, thereby constituting scientific misconduct at the highest possible level.

2. The Curt Stern–Hermann J. Muller connections

Previously, this author had extensively researched the history of the non-linear (hormetic) dose–response model, its scientific foundations and its failure to thrive and out-compete the linear no-threshold (LNT) dose–response model during the first half of the 20th century (Calabrese, 2011b, 2005, 2009b; Calabrese and Baldwin, 2000a, 2000b, 2000c, 2000d, 2000e). As a continuation of this research activity, efforts have been exerted to assess in detail the historical and scientific origins that have resulted in the validation and acceptance of today's LNT model. During this investigation, it became evident that the role of Hermann J. Muller was essential to the adoption of the LNT model and needed greater clarification.

During this assessment of Muller, interest grew in the research activities of the Manhattan Project at the University of Rochester, especially those under the direction of Curt Stern who employed the fruit fly to investigate the nature of the dose response in the low dose range. Stern was of particular interest because he had a long personal and professional relationship with Muller that would markedly impact the LNT deception story. Stern had helped to organize the Fifth International Genetics Congress in Berlin during the fall of 1927 (Carlson, 1981). It was at this meeting that Muller first presented his landmark findings on X-ray-induced mutations in fruit flies (Muller, 1927, 1928), research that would eventually lead to his Nobel Prize in 1946 (Muller, 1946a). Later, Muller and Stern would have a conflict over Muller's deliberate failure to acknowledge a prior discovery by Stern that provided proof for the linear arrangement of genes, an issue that was then a very significant question in biology. Stern would challenge Muller on this point directly via a carefully documented letter dated August 8, 1929 [American Philosophical Society (APS) (American Philosophical Society, 1929a). Stern informed Muller that his earlier publication in Biologischen Zentrablatt (September, 1926) addressed the “theory of the linear arrangement and have specifically stated it in the title of the paper”. Stern concluded his letter to Muller with the statement that his manuscript "had been written before your [Muller's] first papers about them appeared.” Nearly six weeks later, in a letter dated October 3, 1929, Muller would respond “I am very sorry to have omitted mention of your work in my discussion of translocation and not to have given you credit for having made the first cytological demonstration of a genetically demonstrated translocation and pointed out its significance for the theory of linear arrangement”. He then indicated that he had enclosed a “carbon copy of a note I am sending in on the subject to the American Naturalist, which I hope you will consider as rectifying this mistake” (American Philosophical Society, 1929b). While Stern caught Muller in a significant professional indiscretion, he let Muller “control” the narrative by not objecting to Muller's version of the correction. Nonetheless, this arrangement proved to be acceptable to Stern as seen in an October 23, 1929, letter from Stern to Muller, restoring a positive tone to their relationship (American Philosophical Society, 1929c). One could speculate what might have happened to the LNT story if Muller and Stern had not reconciled, possibly preventing Muller's involvement in the Manhattan Project as described below.

3. The Manhattan Project: Curt Stern and LNT

After Stern[1] initiated research on the Manhattan Project in 1943, he contacted Muller, then a professor of biology at Amherst College (1940–1945), to serve as a consultant to the project. Under normal circumstances this might have been routine, but Muller had a questionable past, abandoning the US to live and research in the Soviet Union from about 1934–1938 (Carlson, 1981). Stern nonetheless obtained approval by the U.S. government for Muller's participation in the radiation genetics project. Muller's involvement proved to be extensive, involving detailed technical written communications with Stern and other team members, visits to the University of Rochester, and a donation of his Muller-5 strain of Drosophila (Calabrese, 2011c).

The Manhattan Project of Stern was designed to expand the study of high dose ionizing radiation on genomic mutations to include the area of chronic, lifetime exposures at relatively low doses and very low dose rates. The first experiment under Stern's direction was an acute (i.e., short duration) exposure study over a broad dose range. It was conducted by Warren Spencer, a professor at the College of Wooster with a PhD from Ohio State University in the area of Drosophila biology. Previous research by several of Muller's students (Hanson and Heys, 1929, 1932; Oliver, 1930, 1931), at very high doses and over a limited dose range, provided support for the hypothesis that the nature of the dose response for X-ray-induced mutation was linear.

In the Spencer study, the effects of X-rays were assessed on sex-linked recessive lethality using Drosophila with acute/short term (2–40 min) exposures and a dose-rate ranging from 10 to 96 r/h. This resulted in a range of cumulative doses from 4000 r down to 25 r (i.e., lowest cumulative dose yet tested). Following a data collection period from December, 1944 to June, 1955, Spencer reported that X-rays induced gonadal mutations in a manner that were linear across the dose response continuum, just as Stern and Muller had predicted (Calabrese, 2011c).

Ernst Caspari, a Ph.D. in insect behavior, directed the next study. From October 1945, to August 1946, Caspari assessed the effects of gamma rays on Drosophila sex-linked recessive lethality. In Caspari's study the females were first mated, placed on an egg laying suppression diet, and then exposed to the gamma radiation (2.5 r/day) for 21 days with sperm stored in the female's spermatheca. In the Caspari study, there was an aging component to the sperm that was not in the Spencer study. The dose rate used in Caspari's study was much lower (13,200 times lower) than that used in Spencer's acute study at the same cumulative dose (Calabrese, 2011c).

The data from the chronic exposure study of Caspari supported a threshold dose–response model. Stern initially rejected the

[1] In the case of the University of Rochester mammalian radiation geneticist Donald Charles, despite the use of over 400,000 mice, his research was largely unproductive, with no methodologically-based technical publications during the time of the Manhattan Project which ended in 1946 (see Charles (1950) for a brief descriptive paper). An additional summary paper (Charles et al., 1961) was published [after Charles's death (Anonymous, 1955a) that tried to salvage the research effort with no obvious success. The failure of Charles to deliver a scientifically significant product for the Manhattan Project, given the level of resources directed to it, represented a substantial failing.

interpretation of Caspari as seen in written correspondence (American Philosophical Society, 1947a). Stern thought the findings were aberrant due to an unexpectedly high mutation rate of the "controls" that obscured a linear dose response, yielding only the appearance of a threshold response. Despite this rejection by his mentor, Caspari dug into the published literature and found convincing support for his rather than Stern's interpretation (Kaufmann, 1947; Muller, 1945, 1946b; Rajewsky and Timofeef-Ressovsky, 1939). To his credit, Stern accepted the data-based argument of Caspari.

Caspari's data were unexpected and somewhat troubling to him because they challenged the linear paradigm of the radiation genetics community. Therefore, Caspari decided to send his findings to another leading researcher, Milisav Demerec, head of genetics at Cold Spring Harbor, for review and comment. Caspari was looking for a way around this problem (i.e., alternative interpretation) and hoping that the influential Demerec might offer a solution. Reflecting the bias of the radiation genetics research community at this time, Demerec wrote back to Caspari, acknowledging the problematic nature of the data, and rather than himself providing the hoped for insight, asked Caspari what could be done to "save the hit theory" (American Philosophical Society, 1947b). There was little question that the Caspari data had created a problem and, in fact, it would be referred to by Stern as a "problem" in future correspondence [Letter of Stern to Noviski – (American Philosophical Society, 1948). Demerec would later become a member of the BEAR I Committee, Genetics Panel that recommended the acceptance of the linear dose–response model.

While Stern seemed to accept Caspari's findings that supported the validity of his control data, he nonetheless challenged the authenticity of the data in other ways. The manuscript that Stern and Caspari developed in the late summer/early fall of 1946 contained a six-page discussion, mostly arguing that Caspari's (rather than Spencer's) findings should not be accepted until it could be shown why his threshold-supporting data differed from the earlier linear dose–response findings of Spencer. This position, in and of itself, was problematic in that the two papers had several dozen important methodological differences (e.g., temperature of 18 °C vs. 24 °C, egg-laying suppression vs. enhancement diets, irradiation by X-rays vs. gamma rays, young vs aged sperm, male vs female exposures and numerous other differences – [see Table 2, (Calabrese, 2011c), making it virtually impossible (if not impractical) to resolve the differences.

Even though the Caspari study adopted technical and methodological improvements over the Spencer study and had avoided serious operational errors of the Spencer study (e.g., Spencer's failure to control temperature, his combining of treatment groups with the same cumulative exposure but with dose rates that differed by up to 2.5 fold, his failure to match control and treatment groups over the same time periods, and his inconsistent calibration of the X-ray machine, etc.) and errors in the modeling of low dose responses (see detailed criticisms – (Bonnier and Lüning, 1949; Bonnier et al., 1949), it was strangely the Spencer study with its linear dose response that became the gold standard and not the Caspari study.

Discussion in the Caspari paper, as noted above, made it clear that the findings in support of a threshold should not be accepted until the differences between the two papers could be resolved. As untenable as this position was, Stern's actions were even more inexplicable as he would not place a similar constraint upon the flawed Spencer paper that supported linearity. It is bizarre, if not unheard of, for investigators to ask the scientific community not to accept the validity of their findings until it could be reliably determined why their findings differed from a study of considerably lesser quality and reliability. Moreover, not placing at least the same constraints on the weaker study, for which Stern was also a co-author, calls into question the investigator's non-biased and objective approach to research. As a very accomplished scientist, Stern should have known that resolving differences between these two studies was not realistically possible.

Stern's unusual behavior makes sense when viewed as an attempt to blunt any challenge to the linear dose–response model (i.e., by demanding that the data of Caspari not be accepted). Stern ensured the success of this strategy by sending the Spencer and Caspari manuscripts to his own journal, *Genetics*, and by fully controlling their publication, including the Caspari discussion. There is no evidence that he submitted either of the papers for an independent peer review as the papers were submitted to the journal on November 25, 1947, and published less than five weeks later in January 1948 (Caspari and Stern, 1948; Spencer and Stern, 1948).

At this point it was not clear whether Muller had seen the Caspari data prior to his Nobel Prize Lecture on December 12, 1946. During the Lecture he disavowed any possibility that a threshold dose response could occur in the induction of mutations by ionizing radiation. He demanded a switch to the linear dose–response model, stating, "there is no escape from the conclusion that there is no threshold" (Muller, 1946a). Not knowing whether Muller had seen Caspari's data in support of a threshold model prior to his Nobel Prize Lecture, several science historians with considerable knowledge of Muller and that era were then contacted. Yet, none of these attempts answered the question. Fortunately, substantial correspondence between Muller and Stern, Caspari, Spencer and others was obtained from archival libraries. The archived records revealed that Stern wrote to Muller on September 24, 1946, to request his services in reviewing the Caspari manuscript in preparation for journal submission. A follow-up letter from Muller on September 27, 1946, accepted this invitation and on November 6, 1946, Stern sent the manuscript to Muller at the University of Indiana. On November 12, 1946, Muller acknowledged receipt of both the letter and the manuscript. He also indicated that he had briefly read the manuscript and recognized that the findings supported a threshold dose response, seriously challenging the linear model. Muller strongly encouraged Stern to find the means to undertake a replication study and indicated that he would try to provide a detailed evaluation prior to his Nobel Prize trip to Europe in early December. Clearly, this November 12th letter acknowledged that Muller had seen Caspari's data, understood the challenge to the linearity model, was not dismissive of the findings and acknowledged Caspari's competence and the need to repeat the findings (see Table 1 for the series of Stern/Muller correspondence statements).

Muller's evaluation of the Caspari manuscript occurred five weeks after his Nobel Prize Lecture in the form of a detailed letter to Stern dated January 14, 1947 (American Philosophical Society, 1947c). Based on this analysis, Muller had not changed his opinion. He unequivocally stated that he could not find any meaningful criticism of Caspari's work (i.e., "I have so little to suggest in regard to the manuscript.") and he restated the need to replicate the findings (i.e., "Unfortunately, therefore a replication seems to be imperative."). Thus, the statements written in private by Muller to Stern were those of a scientist, while his unequivocal public rejection of the threshold model at the Nobel Prize Lecture was deceptive and not without ideological underpinnings. Knowing that uncertainty existed in the low dose zone and that further study was needed, Muller could have acted more forthrightly by pronouncing his conditional rather than categorical support of the LNT model in Stockholm. Even four months later he remained steadfast and continued to advocate his unqualified support for the linear dose–response model. In a presentation to the New York Academy of Medicine in 1947, he stated that "there is then absolutely no threshold dose…and even the most minute dose carries a

Table 1
Letter correspondence demonstrating that Muller had seen and considered Caspari's threshold supportive findings prior to his Nobel Prize lecture on December 12, 1946 (American Philosophical Society, 1946/1947; Calabrese, 2011c).

September 24, 1946 – Stern to Muller:
"Dr. Caspari's report on his work is now being typed and I wonder whether we could bother you with sending you a copy for your new comments."

September 27, 1946 – Muller to Stern:
"Also, I'd be glad to see Caspari's paper too."

November 6, 1946 – Stern to Muller:
"Caspari's manuscript has finally been typed and we would appreciate very much your critical reading of it."

November 12, 1946 – Muller to Stern:
"I have just arrived from an absence of over 2 weeks and find the Caspari manuscript here waiting for me. Unfortunately, it catches me again when I am in a tremendous pressure of work, trying to make up both the trip just passed and for another one to come in a few weeks. However, I see that it is very important and shall do all I can to go through it in a reasonable time, surely before I leave again early in December. I hope that Caspari can wait that long if necessary. In the meantime I wonder whether you are having any steps taken to have the question tested again, with variations in technique. It is of such paramount importance, and the results seem so diametrically opposed to those which you and the others have obtained, that I should think funds would be fourth coming for a test of the matter. It is not, of course, that I doubt Caspari's reliability at all, but only that I naturally share the same doubts which he himself expressed. Of course, I am only judging by the summary and a quick glance through the paper, and have not had the opportunity to read the details."

definite chance of producing a change exactly proportional to the size of the dose" (Muller, 1948).

Muller's statement in a letter to Stern (American Philosophical Society, 1947c) about having "so little to suggest in regard to the [i.e., Caspari] manuscript" may not have been quite truthful, as Muller himself was most likely responsible for the only two changes introduced to the paper prior to its submission to the journal Genetics. With the exception of these two changes, the published study in Genetics was identical in every way to that paper which was sent to both Muller for his pre-submission review and to the Atomic Energy Commission (AEC) in 1947. In the journal version, the first and most significant change was the deletion of a key sentence in the Conclusion of the 1947 AEC version (Caspari and Stern, 1947). The deleted sentence is as follows: "From the practical viewpoint, the results presented open up the possibility that a tolerance dose for radiation may be found, as far as the production of mutation is concerned" (page 15). This statement indicated support for the threshold dose–response model. The second change was significant in that it added the name of Hermann J. Muller to the Acknowledgments of the published paper. It seems more than just coincidence that the only two changes imparted to the journal version consisted of (1) the deletion of a concluding statement in support of a threshold dose–response model and (2) the simultaneous addition of Muller's name to the acknowledgment section. There should be little doubt that removing the threshold conclusion statement was of profound benefit to Muller as it would help him sustain the ideological dominance of his favored LNT model. Muller clearly had the means, motive and opportunity to mitigate the threat imposed by Caspari's paper on the LNT model. So, was Muller responsible for deleting the key concluding sentence in support of a threshold model? Well, we may never know for sure, but strong circumstantial evidence seems to point in that direction.

In the aftermath of the Nobel Lecture, Stern followed Muller's suggestion to repeat the findings of Caspari. However, his two experienced doctoral researchers, Spencer and Caspari, had left for the College of Wooster and Wesleyan University in Middleton, Connecticut, respectively. Consequently, Stern tapped a new Master's student, Delta Uphoff, a recent graduate of Russell Sage College of Albany, New York, to replicate the Caspari research (Calabrese, 2011c). Data from her first experiment piqued Stern because her control values for mutation rates were about 40% below those found in the literature, including Caspari's study. Stern expressed his concern to Muller and also asked Muller to share his largely unpublished data with him on variation among controls for the mutation rates of aging sperm in the fruit fly. In a series of letters between Muller and Stern, Muller confirmed that the findings of Uphoff were not reliable and that the unpublished (and published) data were supportive of the Caspari control results. Muller's data led to an acknowledgment in the discussion section of the Uphoff and Stern manuscript (Uphoff and Stern, 1947) that the control group data were not interpretable and that the low control group value was most likely due to investigator bias. Thus, in a rather unprecedented move, Stern was quick to place blame on the inexperienced Uphoff. This manuscript, which importantly acknowledged the assistance of Muller, was sent to the Manhattan Project/AEC where it became classified and publicly unavailable. Thus, the acknowledgment by Stern of Uphoff's unreliable control data, together with the letter exchanges between Muller and Stern regarding the reliability of Caspari's control data, clearly indicated that Muller had strong confidence in the Caspari and not the Uphoff control data (Calabrese, 2011b).

Stern then had Uphoff undertake a follow up replication study. She again reported a similar unacceptably low control group response. As in the first case, the findings were again not interpretable. Finally, in a third experiment that was undertaken, another problem arose. This time it was not the control group, which seemed to respond as expected, but the treatment group whose response far exceeded that predicted by a linear dose–response model. At this point, Uphoff had finished her degree and eventually joined the National Institutes of Health (NIH) as a staff researcher. However, the damage was done to the Stern initiative regarding the Manhattan Project/AEC. Each attempt to replicate the Caspari findings had significant problems. Could anything be salvaged?

In January of 1949, Stern decided to submit a technical note to the journal *Science*, integrating the five major experiments conducted under his direction for the Manhattan Project/AEC. These involved the studies of Spencer and Caspari and the three Uphoff replications. In this *Science* paper, Stern attempted to rescue the first two Uphoff experiments that he already knew had aberrant control groups (Uphoff and Stern, 1947) and, according to multiple letter exchanges (Table 2), Muller also knew. Stern also chose to ignore certain data that were not in support of the linear model (Caspari and Stern, 1947) and, again attacked the Caspari study as aberrant even though nothing had changed except for the occurrence of even more data supporting the reliability of Caspari's

Table 2
Stern–Muller temporal letter exchange concerning the aged-stored sperm control mutation rate [see (Calabrese, 2015a) – supplement for a more complete letter exchange].

Curt Stern wrote a letter to Hermann J. Muller on January 22, 1947 (American Philosophical Society, 1947d) informing him that "At the present time it looks as if our new control data [probably the results of the first three months of the first Uphoff experiment; note that her first month's reading was an especially low mutation rate of 0.005%] for aged sperm are considerably below those of Caspari's." He then asked Muller to "send me your figures on rate of sex-linked lethal in sperm aged several weeks, (most desirably, if you have them, data on three weeks), in comparison to control data from non-aged sperm?"

On February 3, 1947 (Lilly Library, 1947) Muller answered by stating that "…. sperm of males which are about a week old and have been copulating freely [as in Caspari's experiment] during that period have only about 0.07 or 0.08 per cent of lethal. Thus the latter sperm, after three weeks, should contain something like 0.28 per cent of lethal."

control group. These multiple flip–flops by Stern were befuddling and surely required explanation, yet none were provided. The inferior Spencer study continued to receive strong support from both Stern and Muller even though, as noted above, it had very significant problems, none of which was noted by Muller in his letters to Stern regarding the research of Spencer, September 13, 1946 (American Philosophical Society, 1946) and Caspari on January 14, 1947 (American Philosophical Society, 1947c).

The *Science* paper of Uphoff and Stern (1949) was beneficial both to the LNT model and to Muller himself as its chief advocate. Stern was successful in artfully molding the interpretations of experimental data to fit the LNT mantra. He achieved this goal while the scientific community remained unaware that he and Uphoff (with Muller's support) had acknowledged just a year earlier that their own findings were not interpretable. Now, in the absence of any new data, these same findings were not only acceptable but also argued in support of the LNT model. And Caspari, who had successfully challenged Stern earlier, now remained silent as his findings in support of a threshold model were being undercut in favor of Muller's LNT model. As for Muller, he must have surely felt relief as he was spared the trouble of having to defend his highly deceptive comments at the Nobel Prize Lecture. Since the *Science* paper (Uphoff and Stern, 1949) was only a short one-page note, consisting mostly of a single table, Stern and Uphoff promised the science community a more detailed follow-up paper that would provide important methodological information and other relevant data. However, Stern and Uphoff never did publish the promised follow-up study and there exists no evidence that their colleagues in radiation genetics ever requested them to do so.

The strategy of Muller and Stern to deceive and obfuscate on the nature of the dose response in the low dose zone was successful. This is evidenced by the fact that the Spencer and Stern paper (Spencer and Stern, 1948) and the *Science* technical note by Uphoff and Stern became the highly influential and commonly cited papers. These "flawed" papers provided the scientific foundations upon which the linear dose response model was justified to the science community and, nearly a decade later, to the U.S. Congress at hearings (Congressional Hearings of 1957) partially inspired by the NAS report of the BEAR Genetics Panel (Calabrese, 2013a; Crow, 1957; Glass, 1957; Joint Committee on Atomic Energy, 1957; Muller, 1957). On the other hand, the technically superior and more relevant paper by Caspari in support of a threshold interpretation received virtually no attention; it was, in essence, unfairly but successfully marginalized. Various leaders in the field repeated false limitations of the Caspari study (Higgins, 1951; Jolly, 2004; Singleton, 1954) that were inspired by the deceptive comments of Stern and Muller e.g., (Muller, 1950b, 1954; Uphoff and Stern, 1949). For example Singleton (1954) echoed that Caspari's study could not be accepted because it had an aberrantly high control group. Ironically, this was Stern's original challenge that already had been so effectively rebutted by Caspari and Muller's own data (see Table 2 for letter exchange between Stern and Muller).

After the *Science* paper, Muller published several papers that repeatedly criticized Caspari's study as being too unreliable because of its high control group data. For example, in his 1950 article entitled "Some present problems in the genetic effects of radiation" in the *Journal of Cellular and Comparative Physiology,* Muller (1950a) provided an explicit characterization of the findings produced by Caspari and Stern (1948). Muller states on page 10 "A recent paper by Spencer and Stern….extends the principle (i.e., one-hit principle) down to total doses of 50 r and 25 r". In the next paragraph, he states: "It is true, in a parallel paper….Caspari and Stern have reported results somewhat deviating from the above." In footnote 1 on page 10 of the article cited above, Muller adds "Uphoff and Stern have published a report of further work, with doses as low as 50 r, given an intensity as low as 0.0165 r per minute. The results obtained are entirely in conformity with the one-hit principle. A consideration of these results, together with the early work, leads to the conclusion that the deviation first referred to (the Caspari and Stern (1948) findings) was caused by a value for spontaneous mutation rate that happened to be unusually high." Although this repeatedly false criticism by Muller was indeed highly disconcerting, other geneticists seemed too willing and ready to accept it, more or less on 'blind faith" and without proper review and verification. If they had chosen to follow the data originating from Muller himself (Muller, 1945) and his own graduate students (Byers, 1954; Byers and Muller, 1952) as well as others (Graf, 1972; Rinehart, 1969) then perhaps the findings of Caspari, and not of Uphoff, would have received public attention and support. Thus, Muller continued to perpetuate a false view that was discredited by his own statements/data. Shamefully, there is no evidence that anyone challenged Muller on these contradictions. Furthermore, Muller claimed that the research of Delta Uphoff and Curt Stern was "entirely in conformity with the one-hit principle" (Timoféeff-Ressovsky et al., 1935). What Muller neglected to state was that Uphoff's first two experiments displayed an aberrantly low control group responses based on Muller's own extensive data involving some 200,000 fruit flies (Muller, 1946). A letter from Curt Stern to Ernst Caspari (fall 1947) (American Philosophical Society, 1947a) addressed the control group issue. It states: "The radiation data continues to be puzzling. Delta's difference between control and exper[imental group] appears to be due mainly to a much lower control group value than yours. However, Muller informs me that this data give an aged control value close to yours. Thus, my first idea that your results could be "explained away" by assuming that your control value happened to be unusually high, seems unlikely. Rather does Delta's control appear too low". Muller's false and self-contradictory statements about Caspari's findings may be understood within the context of his ideological focus on establishing the LNT model for risk assessment and in the preservation of his legacy – a legacy that would have been severely tarnished if the deceptive remarks he made during his Nobel Prize Lecture had been discovered.

A further example of Muller's duplicity in promoting the LNT concept was his inaccurate characterization of the dose-rate used in the Uphoff experiments (Uphoff and Stern, 1949), which was 0.00165 r/min, i.e., 50 r in 30,240 min or in 21 days) (Uphoff and Stern, 1949). In his paper entitled "Radiation Damage to the Genetic Material" in the *American Scientist*, Muller (1950b) indicated that their research extended "the principle of proportionality of mutation to doses down to doses of 50 r and 25 r and of less than 0.001 r/min with a time–intensity relation differing by over 400,000 times from that of our high intensity dose." By using the incorrect dose-rate of <0.001 r/min (instead of 0.00165 r/min) Muller (1950b) extended the linear extrapolation over 400,000-fold, some 150,000-fold greater than what the correct dose-rate would have predicted. Just as in the case of validating the Uphoff control groups (discussed above), no one challenged Muller on this point. It is doubtful that Muller's actions was a simple editorial-typo as it involved two discrete changes, removing a 65 and adding a $<$ sign. Furthermore, Muller (1950b) had correctly cited the value as 0.00165 r/min in a previous paper.

4. The NAS BEAR I Committee Genetics Panel

The actions of Muller and Stern (cited above) were critical in persuading the radiation genetics community to adopt the LNT perspective, which was reinforced at multiple levels. By the early 1950s, according to Crow (1995), LNT had become the dominant

view of this group, despite having little support elsewhere. This timing is important as it set the stage for the actions of the NAS Genetics Panel on the Biological Effects of Atomic Radiation, which issued its landmark report on June 12, 1956, and published its technical report in the journal *Science* (Anonymous, 1956a) later that month.

Since the nature of the dose response in the low dose range was a critical issue, it would be important to know how the Genetics Panel debated this issue, what the nature of the debate was, what votes were taken on the general dose response issues, and who were the leading participants in the discussions. The Genetics Panel formally met on November 20 and 21, 1955, at Princeton University and on February 5 and 6, 1956, in Chicago. Transcripts were obtained for both of these meetings. The Panel had a follow up meeting March 1, 1956, with partial attendance and only a meeting summary (i.e., no transcript was taken). Intermeeting communications among Panel members were encouraged via the exchange of working documents and draft materials. These communications were typically preserved in the historical record, and it was generally possible to obtain copies of papers and correspondences of the Panel members on BEAR I from their respective institutional libraries. Although that which was archived varied according to each person, an effort was made to obtain complete sets of information on all Panel members. As a result, copious files on Panel members were obtained, enabling the reconstruction of Panel activity to a high degree.

The transcripts of the Genetics Panel indicate that the members debated neither the nature of the dose response at low doses, the expectations of a linear or a threshold dose response nor any other dosimetric issue. Dr. Tracy Sonneborn from the University of Indiana, a Panel member and colleague of Hermann Muller, wrote a general guiding statement of principles for the Panel to follow; see (Calabrese, 2015a) – Supplementary material. The basic framework consisted of four principles, i.e., that all doses of ionizing radiation were (1) harmful, (2) irreversible, (3) cumulative, and (4) displayed a linear dose–response relationship. No member of the Panel challenged these perspectives. In fact, at the Princeton meeting of the Genetics Panel, Professor Alfred H. Sturtevant from California Tech asserted his disdain for the medical profession that still adhered to an anachronistic belief in the threshold dose response model. Sturtevant stated that he had "no doubt about the correctness of the linear dose response" and that any effort to further document support for it would only be for the "propaganda value" needed to educate and convince the non-geneticists; see (National Academy of Sciences (NAS), 1955) – Transcription, November 21, 1955.

The Panel's single-minded uniformity of belief regarding the nature of the low-dose response was profoundly significant as it tended not only to limit discussion and preclude debate but also to ensure adoption of their preconceived notions. Due to this lack of discussion and absence of debate, the Panel was challenged to identify other activities that could productively fill its meeting times. The Panel Chair, Dr. Warren Weaver of the Rockefeller Foundation, forged ahead and challenged the 13 geneticists on the 17-member Panel to provide estimates of genetic damage to the U.S. adult population given a specific exposure to the gonads. The purpose of this exercise was to see how closely individual estimates of damage might converge among a blended mix of high level expert geneticists who had collective experiences studying an array of diverse populations, including fruit flies, bacteria, paramecia, yeast, human populations and clinical patients, among others. Weaver argued that a greater convergence (i.e., agreement) among individual damage estimates would tend to yield a greater confidence by society in the Panel's scientific conclusions and recommendations. Although one geneticist resigned from the Panel due to overriding academic commitments, the remaining 12 considered the challenge and the need to independently complete the assignment within about one month following the meeting of February 5–6, 1956. Of the 12 geneticists three (Tracy Sonneborn, Clarence C. Little and James V. Neel) eventually decided that there was too much uncertainty for the question to be quantitatively addressed with any degree of accuracy or reliability and that any population-based estimates would simply be misleading. For example, Neel stated that the scientific foundations needed to make such estimates of genetic damage were so uncertain that providing them would be a violation of his obligation to society as a scientist; see the April 6, 1956 letter from Neel to Weaver, cited in Jolly (2004). After the refusal of these three Panel members to participate in the exercise and provide estimates, the nine remaining geneticists may have had similar misgivings, at least to some extent, but nonetheless provided quantitative estimates of genetic damage within the prescribed time; see (Calabrese, 2015a) – Supplementary material.

When the Panel finally published its paper in *Science*, it indicated erroneously that six (instead of nine) geneticists took up the challenge and provided such estimates (i.e., "Six of the geneticists on this committee considered the ....problem."). This apparent discrepancy triggered a more extensive assessment of communications among panel members and related information regarding the estimates of damage. Chairman Weaver gave James Crow the task of organizing the submitted material and integrating tables listing the damage estimates of each participating geneticist. As a result of this process, it quickly emerged that there was considerable disagreement among Panel members concerning the identity and appropriate use of methods and assumptions in conducting the assignment. Thus, as one can imagine, confusion about the assignment and the lack of a clear protocol yielded estimates of extreme variability. Panel members were highly uncertain of their own estimates, which often radically disagreed with the estimates of fellow Panel members. In spite of the fact that each geneticist employed the linear dose–response assumption, the results of this exercise led to anything but a convergence. A close reading of all the contributions reveals that some of the "experts" had little idea how to approach the problem. This can be highlighted in the case of James Crow, the last surviving member of the Panel, who died in 2012. For example, on March 29, 1956, Crow stated (Crow, 1956): "I shall use as a minimum estimate a direct extrapolation from Drosophila and as a maximum some calculation from the sex-ratio in the Japanese cities. An estimate from mouse data turns out to be just about half way between these, so I shall use it as the most probably estimate." The non-sequiturs inherent in such biological reasoning demonstrate how poorly some of the leading experts addressed this issue. As the other geneticists expressed similar levels of uncertainty and disagreement, it is not surprising then that the Panel would share their documentation with neither external reviewers nor the interested public.

A major problem arose as a result of the extreme variability among the individual estimates. That is, the uncertainty of these estimates would erode public confidence in the Panel's pronouncements. Crow perceived the problem and memorialized his concern in a letter to Chairman Weaver of March 29, 1956: "The limits presented on our estimates of genetic damage are so wide that the reader will, I believe, not have any confidence in them at all." Thus, Crow believed that if the Panel shared its uncertainty with the public then the likelihood of winning their acceptance of any scientific and policy guidance would be seriously threatened. Crow then made a unilateral decision to exclude the estimates of three of the geneticists (i.e., Kaufmann, Wright and Demerec), the three with the lowest estimated damage values; see (Calabrese, 2015a) – Supplementary material for a detailed assessment for each of these three excluded values. Although Crow's decision

markedly reduced the amount of variation within the group, this initial "adjustment" was simply not enough to solve the variability problem. Crow then strongly urged the Panel not to share the six remaining and highly variable assessments with the scientific community and public. The Panel eventually voted on Crow's recommendation, and the majority decided in favor of it, thus essentially eliminating anyone from the interested public or the science community from critically examining the data or the process by which these estimates were derived. While a copy of the voting tally was obtained, specific information on votes of individual members was discovered for four members. Based on their preserved correspondence, (Calabrese, 2015a) – Supplementary material, Crow, Glass, Muller and Sonneborn all voted not to share the data.

The aforementioned analysis reveals that the Genetics Panel deliberately falsified the research record in the *Science* article by reporting that only six geneticists provided estimates of radiation induced genetic damage. This was patently false as nine geneticists provided detailed estimates within the prescribed period of time. There was no expectation and no established protocol for the exclusion of estimates as each geneticist on the Panel was considered an independent world-class expert in his own area of genetics. The person who excluded the three estimates was Crow, who lacked the authority to do so. In fact, the exercise on estimating risk of genetic damage was designed to develop a gage of expert agreement or lack thereof. Removing the three estimates was a deliberate act to obscure and mitigate the magnitude of disagreement and uncertainty that existed among the experts. Furthermore, the report did not even acknowledge that three other Panelists refused to participate in the exercise because too much uncertainty precluded the possibility of making any reliable estimates, (Calabrese, 2015a) – Supplementary material. Finally, the *Science* article contained an inaccurate estimate of response variability in the range of plus or minus ten-fold on either side of the mean. More specifically, the *Science* paper states, "These six geneticists concluded, moreover, that the uncertainty in their estimation of the most probable value was about a factor of 10. That is to say, their minimum estimates were about 1/10, and their maximum estimates about 10 times the most probable estimate". This 100-fold uncertainty markedly misrepresented the range of uncertainty of the six remaining Panel geneticists for estimating the next generation, which had a mean uncertainty value of 756 (312.5 median). See Table 1 of identified individual values in Calabrese (2015a) – Supplementary material.

The Genetics Panel of the NAS, as a group, therefore deliberately sought to misrepresent the research record in their landmark *Science* publication on three distinct aspects. These included: the incorrect statement that only six geneticists provided genetic damage estimates when nine did; the failure to report that three other geneticists refused to provide any estimates at all because of the high level of uncertainty of this exercise; and, finally, the uncertainty range for the six geneticists was given as 100 fold when the mean value was actually 756 fold. These actions of fabrication and falsification by the Genetics Panel were undertaken to ensure that governmental agencies, legislative bodies and the general public would be more likely to accept the Panel's LNT-derived policy recommendations for assessing the risk of ionizing radiation.

5. BEAR I Genetics Panel report – fallout

Following its acts of falsification and fabrication of the research record, the Genetics Panel continued to show its arrogance in the aftermath of the BEAR I Panel and at the start of BEAR II (fall, 1956). In this case, several leading biologists had requested that the Genetics Panel provide documentation that would explain/support its decision to recommend the adoption of the linear dose–response model for risk assessment purposes, (Calabrese, 2015b) – Supplementary material and Glass (1956). The biologists noted that the BEAR I Panel had proclaimed the correctness of the LNT model, but it failed to provide any written scientific basis for its decision. Since providing documentation to support major decisions is the main mission of any NAS Committee, the BEAR I Genetics Panel, by this standard, clearly failed to perform its mission. However, in a decision that may be difficult to understand, the Panel actually refused to do so, deciding instead to redirect its efforts to identifying research areas for future funding. Furthermore, it is highly unusual, if not astonishing, that the Panel actually informed the President of the NAS, Detlev Bronk that it had decided not to provide documentation to support the LNT recommendation. In fact, no documentation in support of the LNT decision ever existed at the time of the BEAR I Genetics Panel report on June 12, 1956, and now it would have to be written well after the fact – a serious problem in and of itself. Secondly, the Panel members openly noted that they preferred to spend their time identifying research priorities for funding opportunities, some of which would be of interest to their own research laboratories. No evidence has been found to suggest that President Bronk ever objected to the Panel's no documentation decision, which was shared with him in a letter from George Beadle, Chair of the BEAR II, Genetics Panel (Beadle, 1957) on September 11, 1957. Thus, the President of the NAS was complicit in the decision not to require the BEAR Genetics Panel to document its support of the LNT model.

The BEAR I and II Panels consisted of essentially the same individuals except for two changes. The Chair (i.e., Warren Weaver) stepped down so he could award grants from the Rockefeller Foundation to Panel members without an obvious conflict of interest, and one new person (TG Dobzhansky) who had been invited for BEAR I, but was unavailable at the time.

The BEAR I, Genetics Panel released their report amongst a flurry of media attention with front page stories in the *New York Times* (Leviero, 1956) and *Washington Post* (Haseltine, 1956). Other leading venues, including *US New and World Report* (Anonymous, 1956b), *The Saturday Review* (Muller, 1956), *Time Magazine* (Anonymous, 1956c, 1956a), *Science* journal (Anonymous, 1956c), *The Lancet* (Anonymous, 1956f, 1956g) and others, also had articles on the BEAR I Genetics Panel report. The *New York Times* called it the most extensive study ever conducted by such a leading group of experts. Yet, in retrospect the evidence shows that the effort failed in critical ways, especially in not even debating the key question concerning the nature of the low dose zone in the dose–response paradigm. The Panel proclaimed the validity of the linear model at the start and never felt the need to justify this fundamental decision, even following a subsequent challenge by leading biologists. Such inappropriate actions of the Panel continued, as it even deemed it necessary to fabricate and falsify the record in their key *Science* publication to ensure that their views would be accepted. All this was clearly expressed in newly unearthed records of the Panel's correspondence. The dishonesty of the Panel was nothing new as it was simply carrying on a tradition seeded a decade earlier by Hermann J. Muller at his Nobel Prize Lecture.

The explicit deceptions of some Panel members continued even some 35 years after the fact. For example, Panel member and geneticist Bentley Glass (Glass, 1991), in a book review about the Rockefeller Foundation, retold the BEAR I, Genetics Panel story reported in the 1956 *Science* article concerning how the Panel obtained its estimates of genetic damage in the U.S. population. Glass wrote that Chairman Weaver sought to overcome vast disagreements among Panelists by instructing them to return to their hotel rooms and work out their damage calculations individually.

The following day, Glass reports, the disagreements were profoundly diminished and a strong consensus emerged. The story by Glass may well be how he remembered the event but his memory is strongly contradicted by the factual record. The fabrications of Glass started with his "authoritative" quote from Weaver that inspired the geneticists to return to their rooms. The quote does not exist in the meeting transcripts. The story of Weaver sending Panelists to their hotel rooms to work on their estimates and of their returning the next day in triumphal consensus likewise never occurred. In fact, Weaver charged them to return to their respective homes and gave them about a month to work on the estimates. Thus, once again, based on the transcripts and substantial subsequent written communications, Glass bears false witness. Glass's most significant fabrication is that the Panelists actually reached a strong quantitative agreement. The consensus story was not real but faked by Weaver and the Panel as discussed above and detailed elsewhere, (Calabrese, 2015a) – Supplementary material.

The highly regarded Glass, among whose honors included being a President of the AAAS and Phi Beta Kappa, amongst numerous other honors, repeated, therefore, the long established false narrative, reinforcing the LNT mantra well into the modern era of risk assessment and doing so with great appeal to his authority. This is therefore the story of not only how the U.S. and world governments came to adopt the linear dose response for risk assessment but also how its origins were forged by deception, artful dodges and blind faith to become established, preserved, protected and reinforced by those very people (e.g. Genetics Panelists) and organizations (e.g. NAS) that society is supposed to trust.

6. The Rockefeller Foundation and the LNT

In 1954, the Board of Trustees of the Rockefeller Foundation (RF) developed the proposition that it was necessary for the United States (U.S.) to undertake a major assessment of ionizing radiation on humans and the environment. One of their Board members was Dr. Detlev Bronk, who was also serving at that time as the President of the Rockefeller Institute for Medical Research (which would become Rockefeller University in 1965) and President of the U.S. National Academy of Sciences (NAS). Prior to this time, Dr. Bronk had also been the President of Johns Hopkins University and the President of the American Association for the Advancement of Science (AAAS) in 1952. Bronk took the proposal of the RF Board of Trustees to the NAS and received permission to undertake this project as an official NAS activity (Hamblin, 2007). This new project was called the NAS Biological Effects of Atomic Radiation (BEAR) Committee. The project involved six independent technical panels for different areas of concern (e.g., genetics, pathology, oceanography and fisheries, agriculture, meteorology, and waste disposal and dispersal). The panels were created by Dr. Bronk and administratively overseen by the RF.

All six BEAR Committee expert panels were chaired by renowned experts in their respective fields except for the Genetics Panel, which was chaired by Warren Weaver, a mathematician and long-time administrator at the RF (Rees, 1987). Interestingly, Bronk selected Weaver to chair the Genetics Panel and, as such, this selection represented a striking deviation in panel construction and leadership. Although multiple individuals with considerable relevant scientific expertise and strong leadership skills were already on the Genetics Panel, none of them would be selected as Chair. Overlooked in the selection process were: George Beadle, the future President of the University of Chicago (and 1958 Nobel Prize winner); Alexander Hollender, the highly regarded scientific administrator at Oak Ridge; Clarence C. Little, the past President of the Universities of Maine and Michigan; and Milislav Demerec, Head of Genetics at Cold Spring Harbor.

In the selection of panel members, one suspects that Bronk and Weaver may have intended to “stack the deck” with radiation geneticists who supported the LNT. For example, Ralph Singleton was a radiation geneticist at the Brookhaven National Laboratory who at the time, questioned the linearity hypothesis and reported a non-linear relationship between mutation rate and dose rate, with disproportional increases at higher doses (Singleton, 1954; Richter and Singleton, 1955; Sparrow and Singleton, 1953). In an April 17, 1955 article in the *New York Times,* (Anonymous, 1955b) Singleton challenged the linearity concept for genetic damage stating “there probably is a safe level of radiation, below which no genetic changes occur.” Singleton’s expertise and the timing and topic of his publications would seem to have easily qualified him for membership on the Genetics Panel, assuming of course that the key objective was to form a panel representing diverse viewpoints to encourage discussion and thoughtful consideration. As it turns out, Singleton was not appointed to the Genetics Panel but to the Agriculture Panel of BEAR I.

The BEAR Panels were the creation of the RF, fully funded by the RF, administered by RF staff and directed by a member of the RF Board of Trustees, who was also President of the NAS. Not only did Dr. Bronk help to conceptualize the project, but he was also part of the organization that funded the project and led the organization that received the funding and oversaw the project, including guiding the selection of panel chairs and their members.

For a long time, the RF was a major funding organization for radiation geneticists, including members of the Genetics Panel. The funding of such members extended over three decades, much of which was during the employment of Weaver and also under his direction. As noted in Wynchank (2011) and prior to the creation of the Genetics Panel, the RF had funded nearly four million dollars to the University of Indiana for research in the area of radiation genetics alone. Such funding supported the research activities of Professors Sonneborn and Muller, both members of the BEAR Genetics Panel.

Weaver was clearly aware of the importance of RF funding to radiation geneticists and showed no reluctance in connecting the Panel’s success to opportunities of lavish funding for its members. Weaver specifically stated at the February 5, 1956 meeting of the Genetics Panel that he would "try to get a very substantial amount of free support for genetics if at the end of this thing we have a real case for it. I am not talking about a few thousand dollars, gentlemen. I am talking about a substantial amount of flexible and free support to geneticists", (National Academy of Sciences (NAS)/National Research Council (NRC), 1956) – NAS transcripts, February 5, page 35. As part of his interaction with the Genetics Panel, he prefaced his funding remarks with the statement that "There may be some very practical results – and here is the dangerous remark – don't misunderstand me. We are just all conspirators here together." The remarks of Weaver were blunt and remarkably focused linking the project outcome to the funding interests of the geneticists on the Panel. Such a blatant coupling of funds and outcome were highly manipulative.

Could such an inducement, as grant support, really be persuasive enough to affect the performance, judgment or integrity of esteemed scientists on an NAS Panel? In his 2007 dissertation (Seltzer, 2007), Seltzer sheds some light on this question. He concluded that members of the Genetics Panel saw themselves as funding advocates for radiation genetics (p. 285 footnote 208). Furthermore, it was hoped that the Genetics Panel, which would continue into the foreseeable future, would affect the directions and priorities of funded research in genetics. Seltzer (2007) also further showed that such expectations were in fact evidenced in correspondence between members of the Genetics Panel, i.e.,

Beadle, Dobzhansky, Muller and Demerec. In a letter to Beadle, Demerec (American Philosophical Society, 1957a) offered a funding plan that could be achieved by "setting aside a fund (let us say, one hundred million dollars), to be administered by some competent organization (such as the National Academy of Sciences) and used during a period of 20 or 25 years to fund already functioning research centers so as to attract and train first rate scientists". Dobzhansky (American Philosophical Society, 1957b) responded to this proposal by stating that he would "needless to say, be all in favor (of) $100,000,000 for research in general genetics…. but I would find it hard to keep a straight face arguing that they (general genetics) must be studied to evaluate the genetic effects of radiation on human populations". This evoked from Demerec (American Philosophical Society, 1957c) the statement that "I, myself, have a hard time keeping a straight face when the talk is about genetic deaths and the tremendous dangers of irradiation. I know that a number of very prominent geneticists, and people whose opinions you value highly, agree with me". Finally, Dobzhansky (American Philosophical Society, 1957d) responded by saying "Let us be honest with ourselves – we are both interested in genetics research, and for the sake of it, we are willing to stretch a point when necessary. But let us not stretch it to the breaking point! Overstatements are sometimes dangerous since they result in their opposites when they approach the levels of absurdity. Now, the business of genetic effects of atomic energy has produced a public scare, and a consequent interest in and recognition of (the) importance of genetics. This is to the good, since it will make some people read up on genetics who would not have done so otherwise, and it can lead to the powers-that-be giving money for genetic research which they would not give otherwise" (American Philosophical Society, 1957d).

These shared comments by key members of the Genetics Panel provide previously unknown insights into motivations of the leading radiation geneticists of that era and the group that legitimized LNT for use by society. According to Seltzer (2007), these letters made two points: (1) that the geneticists were quite focused on the viability of their discipline and (2) that they were cognizant of and acted upon opportunities to manipulate the current situation (e.g., to stretch a point) for the purpose of increasing the likelihood of greater funding. It seems as though the persuasiveness of grant funding is more powerful than one could have imagined, even for esteemed scientists.

When viewed from a grander perspective, the RF displayed an undue and unheard of influence over the course of cancer risk assessment within the United States and throughout the world. The RF directed and funded the entire process that resulted in the adoption of the LNT, all hidden within the prestige of the U.S. NAS due to the multiplicity of roles played by Bronk. Weaver used his long-honed knowledge and skills concerning the vulnerability of academics for external grant funding and lured Panel members with funding possibilities on the basis that their area would be seen as important to society. Such manipulations raise serious ethical issues. In fact they paved the way for the very activities that occurred within the Genetics Panel, that is, misrepresenting the research record to enhance its policy recommendations. To ensure a "proper" narrative, Weaver the mathematician, and not one of the geneticists, drafted the final report of the Genetics Panel (Glass, 1991). At an organizational level, the RF manifested hegemony over the BEAR Genetics Panel, warping and corrupting a risk assessment process that had lasting, social and economic public policy consequences. At an individual level, Bronk's failure to require the panel to document the scientific basis for the LNT recommendation and the Panel members' self-serving decision to identify funding opportunities instead of writing the report, together represent unscrupulous behaviors that enabled them to establish the legitimacy of the LNT model without having to defend their position and, at the same time, optimizing their future funding options.

7. Conclusions

- The recommendation by the U.S. NAS in 1956 to adopt the LNT model was rapidly accepted by governments worldwide and provided the basis for estimating cancer risks from ionizing radiation and chemical carcinogens over the past six decades.
- The recommendations of the U.S. NAS BEAR I Committee, Genetics Panel were ideologically-driven with no written scientific basis provided by the Panel.The Genetics Panel explicitly refused to provide a written documentation when formally challenged to explain their recommendations. Moreover, the President of the NAS became complicit in the Panel's questionable and irregular actions by taking no corrective action, even after receiving notification by letter of the Panel's refusal to provide such a report.
- Studies under the direction of Curt Stern at the University of Rochester/University of California-Berkley using Drosophila provided the scientific basis for the LNT of the BEAR I Genetics Panel. Detailed re-analyses of these studies has revealed serious flaws in the acute study by Warren Spencer and in key follow up chronic exposure experiments by Delta Uphoff. Curt Stern intentionally concealed critical limitations of the Uphoff findings which had Stern and Uphoff characterize these findings as "uninterpretable". Stern, in cooperation with Hermann Muller, deliberately misrepresented and marginalized the findings of Ernst Caspari which supported a threshold model.
- The NAS Genetics Panel committed scientific misconduct by falsifying, fabricating and then publishing in the journal Science its doctored estimates of human genetic risk to radiation exposures. The Panel's deceits were designed to prevent the scientific community and the general public from knowing the profound uncertainties entailed in its genetic risk estimates, thereby insuring the ready acceptance of its policy recommendations.
- Current cancer risk assessment policy and practices are based on fraud and deception by key leaders of the radiation geneticist community and by the U.S. NAS, BEAR I, Genetics Panel. Their deceptions were uncritically adopted by regulatory agencies and the scientific community worldwide and provide the foundation of cancer risk assessment and risk communication messages. The implications of such fraudulent actions are profound and likely to affect: human health risk assessment, adoption and use of new technologies, cost benefit assessments at multiple societal levels, toxic tort actions/decisions, and in the education of the public on vast areas of environmental health and medical treatment practices.

Funding sources

Research activities in the area of dose response have been funded by the United States Air Force and ExxonMobil Foundation over a number of years. However, such funding support has not been used for the present manuscript.

References

American Philosophical Society, 1929a. Stern letter to Muller, August 8, 1929.
American Philosophical Society, 1929b. Muller letter to Stern, October 3, 1929.
American Philosophical Society, 1929c. Stern letter to Muller, October 23, 1929.

American Philosophical Society, 1946. Muller Letter to Spencer and Stern, September 13, 1946.
American Philosophical Society, 1946/1947. Curt Stern Papers, Hermann J. Muller File, Philadelphia. ⟨http://amphilsoc.org⟩.
American Philosophical Society, 1947a. Stern Letter to Caspari. Stern Papers, Caspari File, Fall 1947, undated.
American Philosophical Society, 1947b. Caspari Letter to Stern. Stern Papers, Caspari File, September 25, 1947.
American Philosophical Society, 1947c. Muller Letter to Stern. Stern Papers, Muller File, January 14, 1947.
American Philosophical Society, 1947d. Stern letter to Muller, January 22, 1947.
American Philosophical Society, 1948. Stern Letter to Edward Noviski. Stern Papers, Noviski File, March 19, 1948.
American Philosophical Society, 1957a. Demerec Letter to Beadle. Milislav Demerec Papers, August 1, 1957.
American Philosophical Society, 1957b. Dobzhansky Letter to Demerec. Milislav Demerec Papers, August 3, 1957.
American Philosophical Society, 1957c. Demerec Letter to Dobzhansky. Milislav Demerec Papers, August 9, 1957.
American Philosophical Society, 1957d. Dobzhansky Letter to Demerec. Milislav Demerec Papers, August 13, 1957.
Anonymous, 1955a. Dr. Donald R. Charles mourned by Educators. Rochester Democrat and Chronicle. Wednesday, November 30, 1955.
Anonymous, 1955b. Experts Explode Fall-out Myths, New York Times, 42.
Anonymous, 1956a Genetics Panel and W. Weaver, Chair). National Academy of Sciences (NAS), Biological Effects of Atomic Radiation (BEAR). Genetic effects of atomic radiation **123–124**, 1157–1164.
Anonymous, 1956b. What You Should Know About Danger from X-rays. US News and World Report, June 29, 1956, pp. 44–48.
Anonymous, 1956c. X-ray Danger. Time Magazine – Medicine, October 10, 1956, p. 67.
Anonymous, 1956d. Atomic Radiation: The rs are Coming. Time Magazine-Science, June 25, 1956, 64–65.
Anonymous, 1956e Biological effects of atomic radiation Science – *News Sci.* **123**, 1110–1111.
Anonymous, 1956f. Biological effects of atomic radiation. Lancet **167**, 1007.
Anonymous, 1956g. Radiation hazards.Lancet **167**, 999–1000.
Beadle, G.W., 1957. Letter to Detlev Bronk. American Philosophical Society, Philadelphia, PA.
Bonnier, G., Lüning, H.G., 1949. Studies of x-ray mutations on the white and forked loci of Drosophila melanogaster. I. A statistical analysis of mutation frequencies. Hereditas **35**, 116–189.
Bonnier, G., Lüning, H.G, Perje, A.M., 1949. Studies of x-ray mutations on the white and forked loci of Drosophila melanogaster. II. A study of the formation of Gynandromphs and other kinds of mosaics. Hereditas **35**, 301–336.
Byers, L., 1954. Thermal effects on the spontaneous mutation rate in mature spermatozoa of *Drosophila melanogaster*. Caryologia **6** (Suppl), 694–696.
Byers, L., Muller, H.J., 1952. Influence of ageing at two different temperatures on the spontaneous mutation rate in mature spermatozoa of *Drosophila melanogaster*. Genetics **37** (5), 570–571.
Calabrese, E.J., 2005. Historical blunders: how toxicology got the dose-response relationship half right. *Cell. Mol. Biol.* **51**, 643–654.
Calabrese, E.J., 2008. Why it is important to toxicology and toxicologists. Environ. Toxicol. Chem **27**, 1451–1474.
Calabrese, E.J., 2009a. The road to linearity: why linearity at low doses became the basis for carcinogen risk assessment. Arch. Toxicol. **83**, 203–225.
Calabrese, E.J., 2009b. Getting the dose response wrong. Why hormesis became marginalized and the threshold model accepted. Arch. Toxicol. **83**, 227–247.
Calabrese, E.J., 2011a. Muller's Nobel lecture on dose–response for ionizing radiation: ideology or science? Arch. Toxicol **85** (12), 1495–1498.
Calabrese, E.J., 2011b. Toxicology rewrites its history and rethinks its future: giving equal focus to both harmful and beneficial effects. Environ. Toxicol. Chem **30** (12), 2658–2673.
Calabrese, E.J., 2011c. Key studies used to support cancer risk assessment questioned. Environ. Mol. Mutagen **52** (2011), 595–606.
Calabrese, E.J., 2012. Muller's nobel prize lecture: when ideology prevailed over science. Toxicol. Sci. **126** (1), 1–**4**.
Calabrese, E.J., 2013a. How the US National Academy of Sciences misled the world community on cancer risk assessment: new findings challenge historical foundations of the linear dose response. Arch. Toxicol. **87**, 2063–2081.
Calabrese, E.J., 2013b. Origin of the linearity no threshold (LNT) dose–response concept. Arch. Toxicol. **87**, 1621–1633.
Calabrese, E.J., 2014a. The Genetics Panel of the NAS BEAR I Committee (1956): epistolary evidence suggests self-interest may have prompted an exaggeration of radiation risks that led to the adoption of the LNT cancer risk assessment model. Arch. Toxicol. **88** (9), 1631–1634.
Calabrese, E.J., 2014b. Response to letter of R.J. Cicerone and K. Crowley regarding "How the US National Academy of Sciences misled the world community on cancer risk assessment: new findings challenge historical foundations of the linear dose response. Arch. Toxicol. **88** (1), 173–177.
Calabrese, E.J., 2015a. Cancer risk assessment foundation unraveling: new historical evidence reveals that the US National Academy of Sciences (US NAS), Biological Effects of Atomic Radiation (BEAR) Committee Genetics Panel falsified the research record to promote acceptance of the LNT. Arch. Toxicol. **89** (4), 649–650, and supplementary material.
Calabrese, E.J., 2015b. An abuse of risk assessment: how regulatory agencies improperly adopted LNT for cancer risk assessment. Arch. Toxicol. **89** (4), 547–648, and supplementary material.
Calabrese, E.J., Baldwin, L.A., 2000a. The marginalization of hormesis. Hum. Exp. Toxicol. **19**, 32–40.
Calabrese, E.J., Baldwin, L.A., 2000b. Radiation hormesis: its historical foundations as a biological hypothesis. Hum. Exp. Toxicol. **19**, 41–75.
Calabrese, E.J., Baldwin, L.A., 2000c. Radiation hormesis: the demise of a legitimate hypothesis. Hum. Exp. Toxicol. **19**, 76–84.
Calabrese, E.J., Baldwin, L.A., 2000d. Tales of two similar hypotheses: the rise and fall of chemical and radiation hormesis. Hum. Exp. Toxicol. **19**, 85–97.
Calabrese, E.J., Baldwin, L.A., 2000e. Chemical hormesis: its historical foundations as a biological hypothesis. Hum. Exp. Toxicol. **19**, 2–31.
Carlson, E.A., 1981. Genes, Radiation and Society: The Life and Work of H.J. Muller. Cornell University Press, Ithaca, NY.
Caspari, E., Stern, C.,1947. The influence of chronic irradiation with gamma-rays at low dosages on the mutation rate in *Drosophila melanogaster*. MDDC-1200. U.S. Atomic Energy Commission, Hathi Trust Digital Library, pp. 1–18. Available at: ⟨http://www.hathitrust.org⟩.
Caspari, E., Stern, C., 1948. The influence of chronic irradiation with gamma-rays at low dosages on the mutation rat in *Drosophila Melanogaster*. Genetics **33**, 75–95.
Charles, D.R., 1950. Radiation-induced mutations in mammals. Radiology **55** (4), 579–581.
Charles, D.R., 1961. Tihen J.A., Otis E.M. and Grobman A., Genetic effects of chronic X-irradiation exposure in mice. Genetics **46**, 5–8.
Crow, J.F., 1956. Letter to Weaver. Lilly Library (Sonneborn manuscripts), Manuscript Department 1956. Indiana University, Bloomington, IN.
Crow, J.F., 1957. Testimony – Statement of Dr. James F. Crow, Professor of Genetics and Zoology, University of Wisconsin. Hearings before the Special Subcommittee on Radiation of the Joint Committee on Atomic Energy, In: Congress of the United States, 85th Congress, 1st Session, Part 1. United States Government Printing Office, Washington DC.
Crow, J.F., 1995. Quarreling geneticists and a diplomat. Genetics **140**, 421–426.
Glass, B., 1956. Memo to the Committee on Genetic Effects of Atomic Radiation. American Philosophical Society, Philadelphia, PA.
Glass, B., 1957. Testimony – Statement of Dr. Bentley Glass, Professor of Biology, the Johns Hopkins University. United States Government Printing Office, Washington, DC.
Glass, B., 1991. The Rockefeller Foundation: Warren Weaver and the launching of molecular biology. Q. Rev. Biol. **66** (3), 303–308.
Graf, U., 1972. Spontaneous mutations in *Drosophila melanogaster*. Humangen Hum. Genet **16** (1), 27–32.
Hamblin, J.D., 2007. A dispassionate and objective effort: negotiating the first study on the biological effects of atomic radiation. J. Hist. Biol **40** (1), 147–177.
Hanson, F.B., Heys, F., 1929. An analysis of the effect of the different rays of radium in producing lethal mutations in Drosophila. Am. Nat. **63**, 201–213.
Hanson, F.B., Heys, F., 1932. Radium and lethal mutations in Drosophila further evidence of the proportionality rule from a study of the effects of equivalent doses differently applied. Am. Nat. **66**, 335–345.
Haseltine, N., 1956. Nation's top scientists call for atomic radiation control; fear shorter life expectancy and mentally deficient babies. The Washington Post, 14, June 13.
Higgins, E., 1951. Atomic radiation hazards for fish. J. Wildl. Manag **15** (1), 1–12.
Joint Committee on Atomic Energy, 1957. 85th Congress, 1st Session. Summary-analysis of Hearings 27–29 May, and 3–-7 June, 1957 on the Nature of Radioactive Fallout and its Effect on Man. United States Government Printing Office, Washington, DC.
Jolly, J.C., 2004. Thresholds of Uncertainty: Radiation and Responsibility in the Fallout Controversy (Dissertation). Oregon State University, Oregon, p. 591.
Kaufmann, B.P., 1947. Spontaneous mutations rate in Drosophila. Am. Nat. **81**, 77–80.
Leviero, A., 1956. Scientists term radiation a peril to future of man: even small dose can prove harmful to descendant of victim, report states. The New York Times, 1, June 13, 1956.
Lilly Library, 1947. Muller Letter to Stern. Indiana University, Bloomington, IN, February 3, 1947.
Muller, H.J., 1927. Artificial transmutation of the gene. Science **66**, 84–87.
Muller, H.J., 1928. The problem of genic modification. Verhandlungen des V. Internationalen Kongresses fur Vererbungswissenschaft (Berlin, 1927). Z. Induct. Abstamm. Vererb. Suppl. Band **1**, 234–260.
Muller, H.J., 1945. Age in relation to the frequency of spontaneous mutations in Drosophila. Yearb. Am. Philos. Soc, 150–153.
Muller, H.J., 1946a. The production of mutations. Nobel Lecture. Nobleprize.org. ⟨http:www.nobelprize.org/nobel-prizes/medicine/laureates/1946⟩.
Muller, H.J., 1946b. Age in relation to the frequency of spontaneous mutations in Drosophila. American Philosophical Society, Philadelphia, PA, pp. 150–153.
Muller, H.J., 1948. Mutational prophylaxis. Bull. N. Y. Acad. Med. **24** (7), 447–469.
Muller, H.J., 1950a. Some present problems in the genetic effects of radiation. J. Cell. Comp. Physiol. **35** (Suppl. 2), 9–70.
Muller, H.J., 1950b. Radiation damage to the genetic material. Am. Sci **38** (1), 32–59.
Muller, H.J., 1954. The nature of the genetic effects produced by radiation. In: Hollaender, A. (Ed.), 1954. McGraw-Hill Book Company, New York, pp. 351–473.
Muller, H.J., 1956. Race Poisoning by Radiation. The Saturday Review 37–39, 9–11, June 9, 1956.
Muller, H.J., 1957. Potential Hazards of Radiation– Congressional Testimony. Hearings before the Special Subcommittee on Radiation of the Joint Committee on

Atomic Energy. In: Congress of the United States. 85th Congress, 1st Session, Part 1. United States Government Printing Office, Washington, DC.
National Academy of Sciences (NAS). 1955. Transcript of the Genetics Panel. National Academy of Sciences Committee to Study the Biological effects of Atomic Energy. First Meeting, Princeton NJ, November 20–21, 1955.
National Academy of Sciences (NAS)/National Research Council (NRC). 1956. The Biological Effects of Atomic Radiation (BEAR): A Report to the Public, NAS/NRC; Washington DC.
Oliver, C.P., 1931. An Analysis of the Effect of Varying the Duration of X-ray Treatment upon the Frequency of Mutations (Doctor of Philosophy thesis). University of Texas, Austin.
Rajewsky, B.N., Timofeef-Ressovsky, N.W., 1939. Holen-stahlungund die Mutationsrate von *Drosophila melanogaster*. Zeitscher. Indukt. Abstam. Vererb. **77**, 488–500.
Rees, M., 1987. Warren Weaver 1894–1978 – A Biographical Memoir, 1987. National Academy of Sciences, Washington, DC.
Richter, A., Singleton, W.R., 1955. The effect of chronic gamma radiation on the production of somatic mutations in carnations. Proc. Natl. Acad. Sci **41** (5), 295–300.
Rinehart, R.R., 1969. Spontaneous sex-linked recessive lethal frequencies from aged and non-aged spermatozoa of *Drosophila melanogaster*. Mutat. Res. **7**, 417–423.
Seltzer, M.W., 2007. The Technological Infrastructure of Science (Dissertation). Virginia Polytechnic Institute and State University, Blacksburg, VA.
Singleton, W.R., 1954. The effect of chronic gamma radiation on endosperm mutations in maize. Genetics **39**, 587–603.
Sparrow, A.H., Singleton, W.R., 1953. The use of radiocobalt as a source of gamma rays and some effects of chronic irradiation on growing plants. Am. Nat. **87** (832), 29–48.
Spencer, W.P., Stern, C., 1948. Experiments to test the validity of the linear R-dose/mutation at low dosage. Genetics **33**, 43–47.
Timoféeff-Ressovsky, N.W., Zimmer, K.G., Delbruck, M., 1935. Uber die Natur der Genmutation und der Genstruktur. Nachrichten von der Gesellschaft der Wissenschaften zu Gottingen: Mathematische-Physikalische Kalass, Fachgruppe VI. Biologie **1** (13), 189–245.
Uphoff, D.E., Stern, C., 1947. Influence of 24-hour gamma-ray irradiation at low dosage on the mutation rate in Drosophila. MDDC-1492, U.S. Atomic Energy Commission, Hathi Trust Digital Library, pp. 1–6. Available at:⟨http://www.hathitrust.org⟩.
Uphoff, D.E., Stern, C., 1949. The genetic effects of low intensity irradiation. Science **109** (2842), 609–610.
Wynchank, S., 2011. The Rockefeller Foundation and its Support of Radiobiology up to the 1970s. Rockefeller Archive Center, Research Reports, Online. ⟨http://www.rockarch.org/publications/resrep/rronlinealpha.php⟩.

Human & Experimental Toxicology (2000) 19, 76–84

www.nature.com/het

Radiation hormesis: the demise of a legitimate hypothesis

EJ Calabrese*,[1] and LA Baldwin[1]

[1]*Department of Environmental Health Sciences, School of Public Health, University of Massachusetts, Amherst, MA 01003, USA*

This paper examines the underlying factors that contributed to the marginalization of radiation hormesis in the early and middle decades of the 20th century. The most critical factor affecting the demise of radiation hormesis was a lack of agreement over how to define the concept of hormesis and quantitatively describe its dose-response features. If radiation hormesis had been defined as a modest overcompensation to a disruption in homeostasis as would have been consistent with the prevailing notion in the area of chemical hormesis, this would have provided the theoretical and practical means to blunt subsequent legitimate criticism of this hypothesis. A second critical factor undermining the radiation hormesis hypothesis was the generally total lack of recognition by radiation scientists of the concept of chemical hormesis which was markedly more advanced, substantiated and generalized than in the radiation domain. The third factor was that major scientific criticism of low dose stimulatory responses was galvanized at the time that the National Research Council (NRC) was organizing a national research agenda on radiation and the hormetic hypothesis was generally excluded from the future planned research opportunities. Furthermore, the criticisms of the leading scientists of the 1930s which undermined the concept of radiation hormesis were limited in scope and highly flawed and then perpetuated over the decades by other 'prestigious' experts who appeared to simply accept the earlier reports. This setting was then linked to a growing fear of radiation as a cause of birth defects, mutation and cancer, factors all reinforced by later concerns over the atomic bomb. Strongly supportive findings on hormetic effects in the 1940s by Soviet scientists were either generally not available to US scientists or disregarded as part of the Cold War mindset without adequate analysis. Finally, a massive, but poorly designed, US Department of Agriculture experiment in the late 1940s to assess the capacity for low dose plant stimulation by radionuclides failed to support the hormetic hypothesis thereby markedly lessening enthusiasm for research and funding in this area. Thus, the combination of a failed understanding of the hormetic hypothesis and its linkage with a strong chemical hormesis database, flawed analyses by prestigious scientists at the critical stage of scientific research development, reinforced by a Cold War mentality led to marginalization of an hypothesis (i.e., radiation hormesis) that had substantial scientific foundations and generalizability.
Human & Experimental Toxicology (2000) **19,** 76–84

Keywords: hormesis; low dose; stimulation; radiation; β-curve

Introduction

There is little question that the radiation hormesis hypothesis had considerable support in the peer-reviewed, experimental and clinical literature during the first 50 years after the discoveries of X-rays and radionuclides. As was presented in the previous paper,[1] this support was well founded based on the quality of the studies, reproducibility, generalizability of the findings, and the remarkable similarity of the hormetic dose-response relationship between chemical and radiation effects. However, there is also little question that the radiation hormesis hypothesis not only never achieved the status of a central core dose-response hypothesis within the field of radiobiology and health, but was relegated to a very tenuous hypothesis status that was never taken very seriously as is evidenced by its omission from all leading radiation health and toxicological texts, its lack of inclusion within symposia at leading scientific society conferences and lack of consideration by regulatory agencies. This paper set forth to examine why the concept of radiation hormesis which had a strong and generalizable scientific foundation up the 1940s became a marginalized hypothesis within the US and western countries.

*Correspondence: EJ Calabrese, Department of Environmental Health Sciences, School of Public Health, N344 Morrill Science Center, University of Massachusetts, Amherst, MA 01003, USA
Received 15 October 1999; Accepted 15 October 1999

Factors affecting the demise of radiation hormesis

The underlying factors for the demise of radiation hormesis are, as expected, complex, multiple factor-

ial, and dynamic entities that differentially affected the hormetic hypothesis over the first half of the 20th century. Despite this complex web of interacting factors affecting the acceptance of hormesis as a legitimate scientific hypothesis, it is both important and possible to prioritize the influential factors affecting the rejection of this hypothesis and to clarify to some degree the interaction of these factors. While it is tempting to look outside of the limitations of the radiation hormesis hypothesis to a grander conspiracy theory undermining radiation hormesis, it is best to consider the hypothesis itself and how its limitations may have contributed to its own demise before considering external, although potentially important, factors.

Experimental design challenges

The overwhelming data on hormetic responses in well-designed studies indicate that the maximum stimulatory response is quite modest, being only about 30–60% greater than the unexposed control.[2,3] In addition, the maximum stimulatory response is relatively close to the toxic threshold [e.g., the no observed adverse effect level (NOAEL), zero equivalent point (ZEP)], that is, a factor of only 4–5-fold (Figure 1). This quantitative relationship, even if assumed to be real, placed great constraints on the hypothesis because it created the need for more powerful study designs, especially for an adequate number of properly spaced doses below the NOAEL; furthermore, concern over distinguishing normal variability from an apparent modest stimulatory response affected factors such as endpoint selection, sample size, and statistical power.

These experimental dimensions of hormesis made it more difficult to establish evidence to support this hypothesis and increased the level of effort by requiring more treatment groups per experiment and more subjects per treatment group. Given those constraints, it was certainly easier to conduct experiments at higher doses and define the upper end aspects of the dose-response curve. Such high dose experimentation was less controversial, more reproducible, required less resources, and was more certain of being published. This set the stage for failure for the hormetic hypothesis since hormetically oriented research offered more professional risk along with few obvious benefits and limited economic applications. Thus, the burden of proof fell on an unorganized (e.g., no scientific society) and a limited number of scientists to establish the data that were to evaluate radiation hormesis as a biological hypothesis. Radiation hormesis was to become an easy target for legitimate methodological critiques that demanded objective answers that were based on proper study design and statistical power.

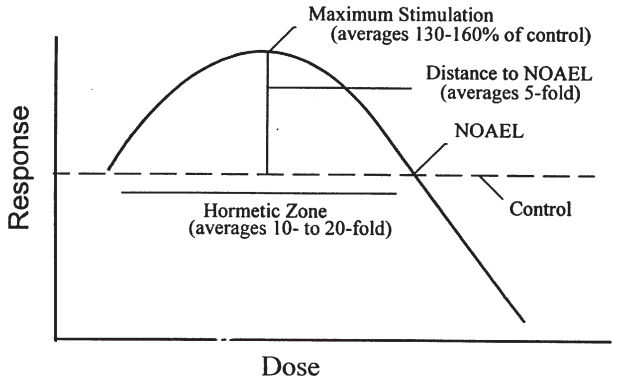


Figure 1 Dose-response curve depicting characteristics of the chemical hormetic zone (modified from Calabrese and Baldwin[3]). Abbreviations: NOAEL=no observed adverse effect level; LOAEL=lowest observed adverse effect level; ZEP=zero equivalent point

Lack of awareness of research on chemical hormesis

In addition to its inherent limitations as a dose-response theory, the opportunity to provide support for this perspective by citing the substantial and earlier supportive work on chemical hormesis was essentially totally ignored by those publishing in the area of radiation hormesis. This lack of linkage between radiation and chemical hormesis denied those radiation scientists interested in the hormesis concept the opportunity to become aware that similar low dose stimulatory phenomena had been observed by numerous highly regarded scientists, over nearly three decades of previous research. If such information had been considered by the critics of radiation hormesis it is likely that their perspectives could have been altered.

By 1910, the concept of chemical hormesis was well established in the areas of plant and algal biology, fungal responses, and bacterial growth.[1] In fact, the basic hormetic curve (i.e., β-curve) was clearly published as early as 1905 by True and Oglevee in the journal Botanical Gazette.[4] By 1920, low dose stimulation had been extended to insect responses to toxic substances and by the 1930s hormetic responses by bacteria to low doses of toxic substances were highlighted in leading microbiological texts along with adequate documentation.[5] In fact, the concept that this low dose stimulation represented an overcompensation to a disruption in homeostasis was first proposed by Townsend[6] in 1897 and then supported by Branham[7] and Colley.[8] This was also an important methodological concept since it required a proper temporal component to such experiments thereby adding further resource and time demands on study protocols.

Despite the substantial supportive information on low dose stimulatory responses to highly diverse chemical agents, none of these papers were ever cited in the research that comprises the database on the historical foundations of radiation hormesis.[9] The only direct linkage between the chemical and radiation hormesis areas is believed to be that of FL Stevens who published evidence of low dose chemical stimulation of fungal growth in 1898[10] and then 30 years later a series of highly influential papers on low dose UV radiation as a stimulatory influence on fungi.[11–15] However, the later work of Stevens, having moved from the University of Chicago to the University of Illinois, never cited his earlier stimulatory work on chemical agents.

Scientific criticisms of radiation hormesis

One can observe the type of framework being established by the 1920s and 1930s in which criticisms of low dose stimulatory responses emphasized poor study design features, inadequate sample size, and inconsistent reproducibility.[16,17] This view became the dominant technical perspective in the mid-1930s following deafening criticism on radiation hormesis' strongest area (i.e., X-ray induced plant growth stimulation) by Edna Johnson, Professor at the University of Colorado in her capacity of invited author in the highly prestigious volume of the NRC on the toxicological effects of radiation.[17] In many ways, such criticism was reinforced by Professor Elizabeth Smith from the University of Wisconsin writing in the same prestigious publication who critically assessed the effects of radiation on fungal growth.[18] However, in the case of Smith,[18] she recognized that stimulation of mycelium growth was a verifiable phenomenon except that it only occurred AFTER the UV-induced initial damage with stimulation representing an overcompensation response.

Defining the concept of hormesis

Such a recognition of the stimulation not being a direct one, but only in response to damage, was viewed by some as a direct refutation of the hormesis hypothesis. For example, while Manfried Fraenkel argued that small doses can stimulate by a direct biopositive action of the X-rays,[19] Holzknecht and Pordes denied the possibility of a direct stimulatory response without simultaneous damage.[19] This confusion over whether the stimulatory response of the Arndt-Schulz Law was a direct one or only in response to damage became an important issue that was still highly visible several decades later.[19,20] Given the predominant role of Holzknecht in the early development of the field of medical applications of X-rays (i.e., he studied with Roentgen for 3 years; he established the first method of measuring X-rays; he created the International Society of University Professors of Medical Radiology; he was the first European professor of medical roentgenology),[21] such disputes remained active and further confounded the issue of hormesis since it was not clear, even to the established experts and advocates, exactly what constituted an hormesis stimulation (i.e., direct or indirect).

The lack of understanding of hormesis continued to be a critical factor in its rejection as the field of radiation health rapidly matured into the 1940s. For example, the prestigious Harvard professor and first Director of the Division of Biology and Medicine at the US Atomic Energy Commission,[22] Shields Warren, continued to promote the concept of Holzknecht and Pordes by stating that 'the assumption that small doses of X-ray or radium radiation are stimulatory (the Arndt-Schulz 'law') is invalid. The slight evidences of proliferative activity offered as evidence by the proponents of this hypothesis are in fact only reparative responses to the injury that has been done'![23] Recognition of reparative overcompensation due to radiation-induced damage was proposed in 1920 by Hektoen,[24] head of Pathology at the University of Chicago, with respect to antibody production, and by Pohle,[25] Koga,[26] Teneff and Stoppani,[27] and Schurer[28] for enhancement of reticuloendothelial activity. The key element in this assessment is the incorporation of an adequate temporal component in the study design. For example, in the case of Schurer[28] phagocytosis was inhibited during the initial 4 h after exposure to X-rays; however, by 8 h after treatment this condition had yielded to one of enhanced phagocytic activity. These findings indicating an overcompensation response to an initial toxic insult have been supported in later reports of Bloom and Jacobson,[29] Dunlop,[30] and Taliaferro and Taliaferro.[31] These radiation-induced reparative responses were also comparable to the responses reported by Smith[20] for UV-induced fungal mycelium growth. That is, that enhanced growth was observed only after damage and that it was necessary to include a repeat measures design to properly describe this phenomenon.

Thus the rejection of the Arndt-Schulz Law by prominent individuals such as Warren over the observation that the stimulatory response was merely a response to damage rather than a direct stimulatory effect was perhaps the critical judgmental factor in marginalizing the hormesis concept. In fact, these dismissing individuals neglected to hypothesize that the process that they were marginalizing was a basic feature of the toxicologic dose-response curve observed in plant and animal

models without regard to whether the damage was induced by chemicals or radiation. The fact that the 'stimulation' (i.e., overcompensation) was modest, consistently distanced (i.e., 3–5-fold) from the traditional NOAEL, and with a modest overall range of about one order of magnitude supported the fact that this response was likely due to a limited induction of damage. Rather than offering a refinement of an hypothesis (i.e., the Arndt-Schulz Law) to incorporate an appropriate temporal experimental feature in the study design and to recognize the possible or likely role of an overcompensation reparative response to account for the quantitative aspects of the low dose stimulatory response, the rather astonishing collective conclusion was to reject the Arndt-Schulz Law and the hormesis concept on the simple equivalent of a yes or no vote.

It is ironic that over 50 years later that definition of hormesis that is most prominently articulated is that of an overcompensation response following a disruption in homeostasis.[32] This is the very concept that was recognized as being most consistent with the available data in the 1930s and 1940s and yet dismissed because it was not a 'direct' stimulation. It thus appears that Warren and others have derived that proper scientific concept, but they marginalized its role to the point of irrelevancy, as is seen in the following paragraph on how the concept became ignored in the conceptual development of the dose-response relationship.

Of particular importance to the field of radiation hormesis was the fact that the concept of dose-response, as developed by leading biostatisticians, occurred in the mid-1930s through the 1940s. These authors totally ignored the concept of hormetic dose-response relationships and developed mathematical models more catered to fit high dose data sets. Of further note is that the well-known biostatistician Bliss who developed biostatistical models of radiation effects data worked closely with the world-renowned pharmacologist AJ Clark, an ardent and articulate opponent of the Arndt-Schulz Law, especially influential in European circles (see Clark[33]). The collaboration of the laboratory-bench scientists with the biostatistician as partners[34,35] in the articulation of the nature of the dose-response relationship was a powerful and dominating combination that would long suppress challenges to the so-called dominating toxicologic paradigm of linear or threshold dose-response relationships.

Hormesis and economic implications and charlatans

While much confusion ruled the debate over whether hormesis or low dose stimulation following radiation exposure occurs, the fledgling field of radiation hormesis was further hampered by both legitimate and charlatan-like desires to exploit the concept of low dose stimulation for a range of applications including the enhancement of agricultural production[36–38] to providing a rejuvenating quality to human life.[39] For the most part, these attempts at commercialization of the hormesis concept never really were sufficiently convincing to establish a long-term successful commercial presence. Furthermore, the concept of low dose consumption of radioactivity became embodied in what was called 'mild radiation therapy' to separate it from the more destructive treatment in the case of tumor destruction therapies.

According to Macklis,[39] the mild radiation therapy approach had its foundation in the American homeopathic and physical medicine movements of the late 19th century. Mild radiation therapy was more associated with endocrinology than oncology and was based on the premise that low doses of radiation could serve as a powerful metabolic catalyst.[39,40] The principal belief of the mild radiation therapists was that the beneficial effects were mediated by the alpha particles of the radium nucleus. This was linked with the use of hot springs throughout Germany, Italy, and France which had been touted to cure numerous illnesses. Once radon was found in 1903 in the Gastein Springs by the famous German chemist Justus von Liebig, alpha particle emitting isotopes became a great rage, becoming used as natural elixirs which were believed to provide direct energy transfusions to depleted organs.[39]

As Macklis[39] noted, the discovery of the therapeutic uses of radon marked the start of an important era of radioactive patent medicines. Since consumption of mineral water from hot springs having high background radon levels had a long history without known adverse health effects, it was then assumed that consumption of long-term use of small quantities of radon would also likely be without harm in commercial products. Supportive of this assumption was the study of the German physiologist, George Wendt, who claimed in 1929 that moribund vitamin-depressed rats could be rejuvenated following radium exposure. In fact, the radium was prescribed for nearly three decades with numerous commercial products on the market claiming to treat just about every human ailment imaginable. However, according to Macklis,[39] the reign of the radioactive elixirs and alpha particle emitting liniments came to an abrupt halt on March 31, 1932 when the well-known millionaire industrialist Eben M Byers died a highly disfiguring death from radium induced bone cancer which received first page coverage in the New York Times, 'Eben M.

Byers Dies of Radium Poisoning!' Byers had consumed the radium containing product Radithor on a routine basis for several years.

Such publicity of the death of the well-known Byers and the somewhat earlier recognition of osteosarcoma in female radium dial painters marked a turning point leading to the demise of mild radium therapy. The mounting criticism of scientists such as Johnson[17] and the negative publicity such as the Byers tragedy and the lack of successful commercial applications went a long way to undermine the scientific and medical belief in the stimulatory effects of low level radiation effects.

At the same time, the use of low levels of X-rays had been employed to treat many human diseases with an apparent record of good success (see reviews by Desjardins[41–45]). This application of X-ray treatment was usually a single low dose treatment that was quite distinct from that used at higher doses for tumor destruction therapies. Typically, a single dose of 50–100 r was all that was used to successfully treat a large variety of human diseases such as furuncles (boils), carbuncles, pneumonia, sinusitis, gas-gangrene, and others. However, the use of even successful low dose X-ray therapy was severely challenged during the early decades of the 20th century by attractive alternative new therapies such as vaccines, antiseptics, and antibiotics. This was especially true from the 1930s onward as sulfa drugs, penicillin, and streptomycin and their derivatives became more available.

Opposing scientific leadership

The availability of new magic chemical bullet treatments, the concern over toxicity at high doses, and the knowledge as of 1927 by Muller that X-rays could cause mutations, all contributed to a very precautionary era of radiation use and exposure. These developments occurred very closely in time and with reinforcement of the limitations of the low dose stimulation theory of radiation. In addition, there appears to have been no powerful intellectual counterforce to defend the radiation hormesis perspective and at sometime in this temporal window of crisis (1930s–1940s), radiation hormesis became rejected by science, medicine, and society and therefore became marginalized. It is interesting to note that the most likely individual to step forward and become a visible advocate for radiation hormesis was Benjamin M Duggar, a professor at the University of Wisconsin. Professor Duggar had studied under the internationally renowned German botanist Pfeifer in the late 1890s at the University of Lipzig and became interested in various types of adaptive responses and low dose stimulatory effects. Upon his return to the US in 1896 he proceeded to finish his PhD at Cornell University and published a very significant paper on low dose chemical stimulation on fungi.[46] Duggar eventually moved to Wisconsin and became the mentor of the well known Alexander Hollender, co-founder of the Environmental Mutagen Society (EMS), and the source of acknowledged guidance for University of Wisconsin Professor Elizabeth Smith in her research in UV stimulation of fungal growth. Duggar also was the editor of the NRC publication[47] in 1936 when Johnson[47] and Smith[18] authored their highly influential articles. Duggar was to later move on to American Cyanimide, assisting in the search for new antibiotics after the remarkable work of Dubos and Waksman.[48]

Nonetheless, Duggar had a long career of leadership in the area of low dose stimulation, had achieved an influential position, had the respect of leading experts and the NRC, and yet he did not accept the challenge at the critical juncture to advocate hormesis. In addition, it should be noted that the criticisms of Johnson[17] could have been addressed in a very direct manner by Shull, who along with Mitchell,[49] published a widely cited study on the low dose stimulatory effects of X-rays on multiple plant species. As noted by Calabrese and Baldwin,[9] Professor Shull helped guide Edna Johnson's dissertation in the mid-1920s when she was a student at the University of Chicago. His latter widely-cited work[49] directly contradicted the conclusions of Johnson and others who emphasized high dose radiation experiments while he established that the nature of the biological response was principally a function of dose with high doses of X-rays causing inhibition and low doses stimulation. However, neither Johnson or Shull ever directly confronted each other on this issue in print, and even more oddly, essentially only tangentially cited each other's work.

Other factors

After the 1930s the field of low dose stimulatory research became subsumed within the unfoldings of World War II and the development and concerns of the atomic bomb. The course of research was also affected by the development of new gamma products from fission reactions making radium studies in low dose studies almost passé with cesium becoming prominent.

At the same time the issue of low dose stimulation became a progressively more central theme among eastern block biomedical and agricultural researchers, especially among a large number of Soviet scientists. In the 1930s–1950s Soviet scien-

tists published a remarkable series of papers on low dose stimulatory responses to X-rays and later X-rays and gamma rays. This research was viewed with suspicion by many US and western scientists because of both political influence on many aspects of Soviet science as well as frequent inadequacies in the reporting of research methods by Soviet scientists. However, most US scientists had little access to and knowledge of these findings since their papers were usually published in the Russian language in Soviet journals having poor circulation in the west. In fact, one of the more significant works of the Soviet scientists of 1946 addressing low dose stimulation was not translated into English until 1960,[50] reflecting the time lag between east-west scientists, even in areas of considerable importance.

Further undermining of scientific support of the low dose radiation stimulatory hypothesis were the results of the then famous US Department of Agriculture (USDA) 22 center, three radionuclide study in 1948 to assess the stimulatory hypothesis in 20 plant species.[51] The data, which were generally not supportive of the stimulatory hypothesis, are believed to have had a major impact on how the concept of hormesis was to be considered not only by the scientific community but also by potential US and international funding agencies in the west. It was unfortunate that this remarkably large study undercut the hormetic hypothesis since it only utilized from one to three doses per experiment without providing any information on how doses were selected for any of the species studied. Thus, this was an example of a disproportionate influence that had major long-term impact since it was such a massive study, conducted under the leadership of the US government at a time of major political uneasiness with respect to radionuclides and their usefulness.

The relegation of hormesis to a marginalized status was significant because it achieved this position just at the same time that the US government was organizing a national radiation research program under the influence of the NRC. Thus, leading researchers were not encouraged to pursue the hormesis tract nor would low dose stimulatory hypotheses be granted any reasonable priority. The continuation and reinforcement of such practices were clearly seen with the program activities of the Radiation Research Society throughout the 1950s and 1960s in which this topic never surfaced at national meetings.[22] This organization was key in bringing together the major leaders of the radiation health research community, including the likes of Alexander Hollander, Gino Failla, and others, and especially guided US research in this critical area. Within this highly influential group there was no leading advocate for radiation hormesis.

It should be noted that the world-renowned radiation physicist Gioacchino (Gino) Failla published a paper in 1922 which reported that low doses of radium enhanced the growth of mice.[52] These authors offered the following comments about the effect of radiation dose on the growth of mice: '(a) Sufficiently small doses of radiation accelerate the growth of suckling white mice. (b) A larger dose of the proper value will have no influence on the body growth of mice. (c) A still larger dose, up to a certain limit, will retard growth, but the animals will eventually attain normal size. (d) Still larger doses cause premature death. Similar results have been obtained before in experiments on seeds and plants, also on lower forms of animal life exposed to X-rays. From these it is commonly assumed that the action of radiation on the living cell follows the same general law which governs the action of all anesthetics, as well as chemical, mechanical, and electrical stimulants; that is, if some form of energy is gradually brought to bear on the cells, at first they may be stimulated to greater activity, then their normal function may be arrested, and finally they may be destroyed.'

One might have thought that with such support for the hormetic perspective Failla could have been the scientific leader that radiation hormesis needed at this critical junction. Failla, who obtained his doctoral degree at the Sorbonne in 1923 under Madame Curie, became one of the most noted leaders in the field of radiation and health physics in the US. He was the recipient of numerous prestigious achievement and career awards and the co-founder and second president (1953–1954) of the Radiation Reseach Society (see Failla obituary by Marinelli[53]). Following his death, the Society created the annual Failla Lecture which is published in his honor.

The question is why did Failla not become a leading supporter of radiation hormesis since he was a strong and effective leader of so many other important aspects of the field? First, the above cited and highly supportive paper on hormesis[52] was published in 1922, one year prior to Failla's completing his dissertation. Consequently, he had relatively young professional status at that time. Second, he was principally a radiation physicist and devoted his activities to that area. Later he did co-publish a paper with Henshaw in 1931 on the effects of X-rays and gamma rays on wheat.[54] This extremely well designed and conducted study was conducted using high doses of radiation and induced inhibitory growth. Most of his other

research was directed to physical phenomena and not only low dose response experimentation. Much of his public service activities were devoted to worker protection and establishing safe exposure standards. Thus, even though Professor Failla had a knowledge of the concept of hormesis, published supportive original data on this topic, and was aware of other supportive findings in the literature, he pursued other interests more germane to his training in radiation physics leaving hormesis research behind. While it is unclear how he considered the hormesis hypothesis in his later years, his early positive encounter with it never materialized into Failla being either an advocate or critic of hormesis.

Arnold H Sparrow of Brookhaven National Laboratories and later president of the Radiation Research Society reported on the capacity of gamma radiation to stimulate plant growth.[55,56] In fact, Sparrow was influential in securing the translation of the above mentioned 1946 Russian study into English in 1960. Also, the highly regarded Professor Karl Sax of Harvard University published two limited but important and supportive literature reviews on the stimulatory effects of X-rays and gamma rays on plants in 1955 and 1963, respectively.[57,58] In fact, it is noteworthy that Sax's graduate student Sheldon Wolf in the mid-1980s was a co-discoverer of the concept of adaptive response with radiation. However, the involvement of Sax with the issue of hormesis was limited to the modest reviews and was not of a transforming nature to the field.

Discussion

Why did the radiation hormesis hypothesis become marginalized in the scientific community in the first half of the 20th century? While the reasons were numerous, it definitely could and should have been avoided. As the previous assessment of Calabrese and Baldwin[9] has shown, the data were available to have secured a firm place for the radiation hormesis as a legitimate hypothesis. Yet a combination of factors acting collectively led to its undermining (Table 1). It appears that much of the 'blame' can be placed primarily on the lack of critical reviews of the available literature on low dose stimulation by chemical agents and radiation by the scientific community, little apparent communication between those researching the biological effects of chemicals and radiation at low levels, a heavy reliance on the judgment of a few scientists of solid reputation (e.g., Johnson, Warren) to analyze the main body of radiation hormesis evidence, and lack of scientific leadership to step forward to challenge 'authoritatively' erroneous and perpetuated conclusions[59,60] by other recognized experts. Furthermore, the criticism of radiation hormesis by leaders such as Johnson[17] and Warren[23] which addressed agricultural and medical perspectives, respectively, occurred precisely during the time period US federal agencies were enhancing research on the biological effects of radiation. Such timing of events relegated the hormesis hypothesis to a position out of the mainstream of power and influence.

These central factors were reinforced by the progressive recognition within the scientific community, governmental agencies, the general public and the media of the adverse effects of high and perhaps much lower doses of radiation and the failure of exaggerated commercial and health claims of low dose exposures.

It is hoped and expected that a scientific hypothesis will rise or fall on its own merits. We have found that the outcome of this process for radiation hormesis was complicated by lack of available knowledge, as well as scientific, medical, societal, and political factors operating within a dynamic temporal context. While the concept of hormesis is now being revived as a biological hypothesis, the thought that an hypothesis with

Table 1 Summary of the factors involved in the demise of the radiation hormesis hypothesis

Factors
1. Hormetic responses are modest and can be hard to reproduce without an adequate study design
2. Researchers in the radiation area did not link hormetic findings to the more substantial and earlier chemical hormesis database
3. Confusion existed over what hormesis was even among supporters
4. Prestigious scientists offered flawed criticism that was perpetuated throughout the literature, and negatively influenced funding programs
5. Low dose stimulation failed to be a commercial success in various areas such as agriculture reinforcing the above criticism
6. Biostatistical modeling ignored hormetic responses linking only with the alternative traditional dose-response paradigm
7. No leading/respected scientist supportive of hormesis countered opposition
8. Radiation research funding emphasized high dose effects, ignored low dose effects
9. Supportive evidence in foreign literature was not generally available to US scientists
10. As a result of WWI and WWII US science became dominant; there was a strong bias to exclude hormesis
11. Soviet support of hormesis was largely disregarded in the Cold War
12. Major US test of hormesis in plants by USDA in 1948 failed to support hormetic claims; this poorly designed study had a long-term dominant influence on governmental programs

substantial supportive data could be so quickly marginalized without either notable scientific refutation nor with at least a modest but visible debate within the scientific community is a sobering thought.

Acknowledgements

This work has been sponsored in part by a grant to the University of Masschusetts (EJ Calabrese, Principal Investigator) by the US Nuclear Regulatory Commission.

References

1 Calabrese EJ, Baldwin LA. Chemical hormesis: Its historical foundations as a biological hypothesis. *Toxicologic Pathology* 1999; **27:** 195–216.
2 Calabrese EJ, Baldwin LA. The dose determines the stimulation (and poison): development of a chemical hormesis database. *International Journal of Toxicology* 1997; **16:** 545–559.
3 Calabrese EJ, Baldwin LA. A quantitatively-based methodology for the evaluation of chemical hormesis. *Human and Ecological Risk Assessment* 1997; **3:** 545–554.
4 True RH, Oglevee CS. The effect of the presence of insoluble substances on the toxic action of poisons. *Botanical Gazette* 1905; **39:** 1–21.
5 Salle AJ. (1939) Fundamental Principles of Bacteriology. McGraw-Hill Book Co. Inc.: New York, pp. 166–197.
6 Townsend CO. The correlation of growth under the influence of injuries. *Annals of Botany* 1897; **11:** 509–532.
7 Branham SE. The effects of certain chemical compounds upon the course of gas production by baker's yeast. *Journal of Bacteriology* 1929; **18:** 247.
8 Colley AW. Stimulation phenomenon in the growth of bacteria as determined by nephelometry. *American Journal of Botany* 1931; **18:** 266–287.
9 Calabrese EJ, Baldwin LA. Radiation hormesis: Its historical foundations as a biological hypothesis. *Human and Experimental Toxicology* 1999; in press.
10 Stevens FL. 1898. The effect of aqueous solutions upon the germination of fungus spores. *Botanical Gazette* 1898; **26:** 377–406.
11 Stevens FL. Effects of ultra-violet radiation on various fungi. *Botanical Gazette* 1928; **86:** 210–225.
12 Stevens FL. The sexual stage of fungi induced by ultra-violet rays. *Science* 1928; **67:** 514–515.
13 Stevens FL. The effects of ultra-violet irradiation on various Ascomycetes, Sphaeropsidales and Hyphomycetes. *Centralblatt fur Bakteriologie und Parasitenkunde* (Abt. 2) 1930; **82:** 161.
14 Stevens FL. Relation of nutrients to erithecial production under ultra-violet irradiation. *Philippine Agriculturist* 1930; **19:** 265–272.
15 Stevens FL. The response to ultra-violet irradiations shown by various races of Glomerella cingulata. *American Journal of Botany* 1930; **17:** 810–881.
16 Ancel S. Action de faibles doses de rayons X sur des graines seches. *CR Soc Biol* 1924; **91:** 1435–1436.
17 Johnson E. (1936) Effects of X-rays upon green plants. In: Duggar BM (ed). *Biological Effects of Radiation*, Vol. II. McGraw-Hill Book Co., Inc.: New York. pp. 961–985.
18 Smith EC. (1936) The effects of radiation on fungi. In: Duggar BM (ed). *Biological Effects of Radiation*, Vol. II. McGraw-Hill Book Co., Inc.: New York. pp. 889–918.
19 Gordon MB. The stimulative effect of roentgen rays upon the glands of internal secretion. *Endocrinology* 1930; **14:** 411–437.
20 Smith EC. Effects of ultra-violet radiation and temperature on Fusarium. II. Stimulation. *Bulletin of the Torrey Botanical Club* 1935; **62:** 151–164.
21 Josephs I. Professor Doctor Guido L. Ed. Holzknecht. *Radiology* 1931; **17:** 1316–1318.
22 Patt HM, Brues AM, Casarett AP, Hollaender A, Magee JL, Nygaard OF, Wyckoff HO. Radiation Research Society: The first quarter century: 1952-1977. A Report by the History Committee. *Radiation Research* 1977; **70:** 3–54.
23 Warren S. The histopathology of radiation lesions. *Physiological Reviews* 1945; **25:** 225–238.
24 Hektoen L. Further observations on the effects of roentgenization and splenectomy on antibody production. *Journal of Infectious Diseases* 1920; **27:** 23–30.
25 Pohle EA. Effect of roentgen rays on the reticuloendothelial system. *American Journal of Roentgenology* 1929; **22:** 439–447.
26 Koga Y. Uber die Wechselbeziehungen zwischen den Veranderungen des Farbstoffspeicherungsvermogens des Retikuloendothelial-Systems, der Hamobakterizidie und des Mineralstoffgehaltes der Gewebe bei bestrahlten Kaninchen. *Strahlentherapie* 1933; **47:** 201–232.
27 Tenneff S, Stoppani F. L'influenza delle irradiazioni sulle linfoghiandole e sulla circolazione linfatica. *Radiol Med* 1935; **22:** 768–787.
28 Schurer F. *Wien. Klin. Wchnschr.* 1928; **41:** 1581 (as cited in Dunlap[30]).
29 Bloom W, Jacobson LO. Some hematologic effects of irradiation. *Blood* 1935; **3:** 586–592.
30 Dunlap CE. Effects of radiation on the blood and the hemopoetic tissues, including the spleen, the thymus and the lymph nodes. *Archives of Pathology* 1942; **34:** 562–608.
31 Taliaferro WH, Taliaferro LG. Effect of X-rays on immunity: A review. *Journal of Immunology* 1951; **66:** 181–212.
32 Stebbing ARD. A theory for growth hormesis. *BELLE Newsletter* 1997; **6:** 1–11.
33 Clark AJ. Handbook of Experimental Pharmacology. Verlig Von Julius Springer: Berlin, 1937.
34 Bliss CI, Packard C. Stability of the standard dosage-effect curve for radiation. *American Journal of Roentgenology and Radium Therapy* 1941; **46:** 400–404.

35 Reed LJ. (1936) Statistical treatment of biological problems in irradiation. In: Duggar BM (ed). *Biological Effects of Radiation*, Vol. I. McGraw-Hill Book Co., Inc.: New York. pp. 227–251.
36 Russell EJ. The effect of radium on the growth of plants. *Nature* 1915; **96:** 147–148.
37 Hopkins CG, Sachs WH. Radium fertilizer in field tests. *Science* 1915; **41:** 732–735.
38 Gager CS. (1936) The effects of radium on plants. In: Duggar BM (ed). *Biological Effects of Radiation*, Vol. II. McGraw-Hill Book Co., Inc.: New York. pp. 987–1013.
39 Macklis RM. Radithor and the era of mild radium therapy. *Journal of the American Medical Association* 1990; **264:** 614–618.
40 Borland V. Mild radium therapy. *British Journal of Phys Medicine* 1932; **6:** 226–228.
41 Desjardins AU. Radiotherapy for inflammatory conditions. *Journal of the American Medical Association* 1931; **98:** 401–408.
42 Desjardins AU. The action of roentgen rays or radium on inflammatory processes. *Radiology* 1937; **29:** 436–445.
43 Desjardins AU. Dosage and method of roentgen therapy for inflammatory conditions. *Radiology* 1939; **32:** 699–707.
44 Desjardins AU. Radiotherapy for inflammatory conditions. *New England Journal of Medicine* 1939; **221:** 801–809.
45 Desjardins AU. The action of roentgen rays on inflammatory conditions. *Radiology* 1942; **38:** 274–280.
46 Duggar BM. Physiological studies with reference to the germination of certain fungous spores. *Botanical Gazette* 1901; **31:** 38–66.
47 Duggar BM (1936). *Biological Effects of Radiation,* Vol. II. McGraw-Hill Book Co., Inc.: New York.
48 Duggar BM. Aureomycin: A product of the continuing search for new antibiotics. *Annals of the New York Academy of Sciences* 1948; **51:** 177–181.
49 Shull C, Mitchell J. Stimulative effects of X rays on plant growth. *Plant Physiology* 1933; **8:** 287–296.
50 Breslavets LB. (1946). Plants and X-rays. Translation by A. Elbl. A.H. Sparrow (ed.) The American Institute of Biological Sciences, Washington.
51 Alexander LT. Radioactive materials as plant stimulants–field results. *Agronomy Journal* 1950; **42:** 252–255.
52 Sugiura K, Failla G. Some effects of radium radiations on white mice. *Journal of General Physiology* 1922; **4:** 423–436.
53 Marinelli LD. Obituary: Gioacchino Failla. *Radiation Research* 1962; **16:** 617–625.
54 Failla G, Henshaw PS. The relative biological effectiveness of X-rays and gamma rays. *Radiology* 1931; **17:** 1–43.
55 Sparrow AH. Stimulation and inhibition of plant growth by ionizing radiation. *Radiation Research* 1954; **1:** 562.
56 Sparrow AH, Gunckel JE. (1956). The effects of plants of chronic exposure to gamma radiation from radiocobalt. In: *Proceedings of the International Conference on the Peaceful Uses of Atomic Energy*, Vol. 12, pp. 52–59.
57 Sax K. The effect of ionizing radiation on plant growth. *American Journal of Botany* 1955; **42:** 360–364.
58 Sax K. The stimulation of plant growth by ionizing radiation. *Radiation Botany* 1963; **3:** 179–186.
59 Packard C. Roentgen radiations in biologial research. *Radiology* 1945; **45:** 522–533.
60 Kimball RF. (1955). The effects of radiation on protozoa and the eggs of invertebrates other than insects. In: Hollander A (ed). *Radiation Biology* Vol. III: Ultraviolet and Related Radiations. McGraw-Hill Book Co., Inc.: New York. pp. 285–331.

Human & Experimental Toxicology (2000) 19, 85–97

www.nature.com/het

Tales of two similar hypotheses: the rise and fall of chemical and radiation hormesis

EJ Calabrese*,[1] and LA Baldwin[1]

[1]*Department of Environmental Health Sciences, School of Public Health, N344 Morrill Science Center, University of Massachusetts, Amherst, MA 01003, USA*

This paper compares the historical developments of chemical and radiation hormesis from their respective inceptions in the late 1880's for chemical hormesis and early 1900's for radiation hormesis to the mid 1930's to 1940 during which both hypotheses rose to some prominence but then became marginalized within the scientific community. This analysis documents that there were marked differences in their respective temporal developments, and the direction and maturity of research. In general, the formulation of the chemical hormesis hypothesis displayed an earlier, more-extensive and more sophisticated development than the radiation hormesis hypothesis. It was able to attract prestigious researchers with international reputations from leading institutions, to be the subject of numerous dissertations, to have its findings published in leading journals, and to have its concepts incorporated into leading microbiological texts. While both areas became the object of criticism from leading scientists, the intensity of the challenge was greatest for chemical hormesis due to its more visible association with the medical practice of homeopathy. Despite the presence of legitimate and flawed criticism, the most significant limitations of both chemical and radiation hormesis and their respective ultimate undoing were due to their: (1) lack of development of a coherent dose-response theory using data of low dose stimulation from both the chemical and radiation domains; (2) difficulty in replication of low dose stimulatory responses without an adequate study design especially with respect to an appropriate number and properly spaced doses below the toxic threshold; (3) modest degree of stimulation even under optimal conditions which was difficult to distinguish from normal variation; and (4) lack of appreciation of the practical and/or commercial applications of the concepts of low dose stimulation.
Human & Experimental Toxicology (2000) **19,** 85–97

Keywords: hormesis; low dose; stimulation; β-curve; radiation

Introduction

We have recently assessed the early history of chemical[1] and radiation[2] hormesis and factors contributing to their respective marginalization within the scientific community.[3,4] As a result of these assessments it became clear that chemical and radiation hormesis displayed an historical development quite distinct from each other with respect to temporal development, scientific maturity and sophistication, animal and plant models studied, type and motivation of scientific opposition, and consideration for commercial application. It should be noted that the term hormesis was proposed in 1943 by Southam and Erhlich[5] who observed that chemical extracts of cedar stimulated fungal growth at low doses, but inhibited at higher doses. The authors were apparently unaware of the fact that this phenomenon was previously characterized as either the Arndt-Schulz Law or Hueppe's Rule. This paper will use the term hormesis to describe the low dose stimulatory phenomenon but is cognizant that this term was created after the early historical development and scientific marginalization of both chemical and radiation low dose stimulation (e.g. Arndt-Schulz Law) hypotheses.

The present paper will extend these previous efforts by directly comparing both chemical and radiation hormesis with respect to their scientific development and research directions, the quality of their developmental maturity, and generalizability of their two hypotheses, as well as their respective underlying weaknesses which lead to the demise of each hypothesis.

Chemical hormesis develops earlier than radiation hormesis

The concept of low dose chemical stimulation had its origins in the later decades of the 19th century. The predominant direction in these early years concerned the effects of various chemical agents on plant and fungal growth (Table 1). In fact, prior to

*Correspondence: EJ Calabrese
Received 15 October 1999; accepted 15 October 1999

Table 1 Early historical references for low dose stimulatory response by chemicals

Year	*Person*	*Comment*
Plants		
1874	Bottger[6]	KOH enhanced germination of coffee
1881	Gustavson[7]	Argued that A1, Zn and other substances may act as stimulants or accelerators
1883-1885	Koenig[8]	Reported a stimulatory effect of iron sulphate upon the growth of plants in culture experiments
	Griffiths[9,10]	
1893	DeToni and Mach[11]	Alkaloid enhanced tobacco growth
1897	Townsend[12]	Response to injury was stimulation; low dose of ether stimulated growth; hydrocyanic acid stimulated
1898	Coupin[13]	Low doses stimulate growth
1899	Jacobi[14]	Effects of stimulatory chemicals on respiratory proceses: KCl, NaCl, iodine, oxalic acid, antipyrine and others
1900	Lovinson[15]	Germination enhanced by CS_2
1900	Johnsson[16]	Ether enhanced flowering plants; CO_2 production enhanced
1899, 1903	Morkowin[17,18]	Respiration is enhanced by ether, alcohol, and various alkaloids
1900-1903	Dandano[19]	Dilute solution of Zn, Cu solutions stimulate leaf tissue
		KI in high dilutions enhanced growth of peas and oats, but not vanadium or KCN
1902-1903,	Suzuki[20,21]	
1902-1903a		Low CO_2 stimulates plant growth; high CO_2 inhibits plant growth
1902	Chapin[22]	Mn at low doses stimulates many plant species; toxicity at higher doses
1902-1903	Aso[23]	Support Aso findings
1902	Lowe and Sawa[24]	Injection of Cu salt into plants such as vines enhanced growth; similar results with Fe
1902, 1903	Chuard and Porchet[25,26]	or Cd; for small amounts of Co enhanced maturation at first similar to Bordeaux mixture; at higher doses Cu was toxic
		Boron at low doses enhanced growth of peas and spinach
1903	Nakamura[27]	Assessed possible stimulatory actions of copper in several species of plants; copper
1904	Kanda[28]	did not stimulate in solution; said to stimulate in soil; Zn at low doses was stimulatory
		Hormetic-like response by sodium fluoride in peas and barley (4 doses+control)
1906-1908	Aso[29]	38 compounds tested in wheat in hydroponics; 10 of 38 compounds were stimulatory,
1908	Schreiner and Reed[30]	strong study design
		Boron at low doses enhanced wheat growth in water and sand culture; stimulation
1910	Agulhon[31]	also seen with oats, radish, lipine and peas
		Abstract – no data showing claims; 71 experiments; $MnSO_4$, $CuSO_4$ stimulated growth
1911	Montemartini[32]	of various plants at very dilute concentrations
		Stimulation of vetch and wheat by several inorganic salts; toxicity at higher doses
1913	Lipman and Wilson[33]	(used up to 9 doses)
Plants and dissolved salts/electrolytic production		
1896	Kahlenberg and True[34]	72 experiments, no stimulation
1896	Heald[35]	Follow-up to Kahlenberg; no stimulation
1899	Copeland and Kahlenberg[36]	In effort to assess the oligodynamic theory of Nageli these authors unexpectedly reported a low dose stimulation, this is the first indication of stimulation using highly dilute solutions
1903	True and Gies[37]	Multiple agents stimulated plant growth (up to 5 doses)
1904	Cameron and Breazeale[38]	Extremely comprehensive; numerous experimentation; large number of plants/ experiment; claimed stimulation at low doses; did not show data; could have been an important study if data presented
1905	True and Oglevee[39]	Stimulatory response; dose-response similar to current hormetic concept
1907	Jensen[40]	9 agents stimulated wheat at low doses; used soil and hydroponics; many doses employed
Plants and use of copper as fungicide		
1893	Rumm[41]	Stimulation reported for grapes by Bordeaux mixture
1895	Berlese and Sostegni[42]	
1896	Evans[43]	
1899	Frank and Kruger[44]	Bordeaux mixture enhanced potato growth
Fungi		
1869	Raulin[45]	Sulphates of Zn, Fe stimulate dry weight; worked in Louis Pasteur's laboratory
1887	Schulz[46]	Numerous chemicals stimulate yeast metabolism at low levels; developed into Arndt-Schulz Law
1895	Pfeffer[47]	Development of *Aspergillus* and *Penicillum* in glycerol solutions was stimulated by Zn, Mn, and Co
1895	DeSeynes[48]	$CuSO_4$ stimulated penicillin growth
1896	Stevens[49]	Chemical agents stimulate fungal growth
1897,1899	Richards[50,51]	Very significant publications; use of Zn, Co, Fe and Ni sulphates, NaCl, LiCl, Na_2SiO_2, cocaine, morphine, and amigdaline stimulated three species of fungi (*Aspergillus niger, Penicillum glacom, Botulytis)*; introduced concept of economic coefficient
1899,1899a	Clark[52,53]	Three papers; last paper is particularly relevant for stimulatory responses
1900	Ono[54]	Sulphates of Zn, Ni, Fe, and Co, NaF, HNO_3, K_2As_3 and HCl stimulated fungal growth; findings supported Richards
1901	Kosinki[55]	Chemical agents enhanced fungal growth

(Continued)

Table 1 *(Continued)*

Year	*Person*	*Comment*
Fungi (continued)		
1901	Richter[56]	Zn ion stimulated growth of *Aspergillus*
1901	Yasuda[57]	Alkaloids stimulated fungal growth
1901	Hattori[58]	$CuSO_4$ stimulated fungal growth
1901	Duggar[59]	Fungal growth was stimulated by ethanol, ether, and camphor
1902	Clark[60]	Bordeaux mixture stimulated fungal growth
1902	Pulst[61]	Cd and Ti stimulated fungal growth
1902	Orlawski[62]	Small doses of arsenic (1/1000–1/100%) stimulated growth of *Aspergillus niger;* larger doses (1/8%) retarded growth; larger doses (1/6%) killed colonies
1903	Coupin[63]	Investigated some of Raulin's earlier work under more antiseptic conditions; he denied the earlier stimulation by Raulin
1904	Watterson[64]	Followed up on Richard's work; extended economic coefficient concept
1905	Latham[65]	Enhanced fungal economic coefficient by low doses of chloroform (6 doses+control)
1906	Fitch[66]	Low levels of copper sulphate (N/256) stimulated *Penicillum glacum* Zn stimulated vegative growth in yeast
1907	Javillier[67]	Similar to 1905 study but with Zn sulphate (3 doses+control) enhanced nitrogen fixation
1909	Latham[68]	Measured fermentation with arsenate and arsenite (13 doses+control); strong stimulation
1910-1911	Harden and Young[69]	Extension of earlier paper
1911	Harden[70]	
Algae		
1900	Ono[54]	Low doses of toxic substances enhanced algal growth
1905	Livingston[71]	Reached conclusions similar to Ono
Insects		
1915	Woodworth[72]	Scale insect eggs fumigated with sublethal dose of HCN; this resulted in earlier hatching
Bacteria		
1894	Effront[73]	Growth of bacterial colonies is enhanced by low doses of toxic substances
1896	Hueppe[74]	Development of broad generalization of low dose stimulation for bacteria
1904	Lindet[75]	Low doses of toxic substances enhance bacterial growth
1909	Hune[76]	Formaldehyde enhanced bacterial growth; many doses and temporal factors assess
1912	Lipman[77]	Low levels of Na_2CO_3 enhanced ammonia formation; 14 doses; replication
1913	Greaves[78]	Sodium arsenate stimulates ammonia formation (26 doses+control); arsenic trioxide stimulated ammonia formation (36 doses+control); 3 other compounds stimulated ammonia formation
1914	Delepine and Greenwood[79]	Low levels of many metals enhanced bacterial growth
1917	Brown and Hitchcock[80]	Nitrification is stimulated in normal soil by small amounts of NaCl, Na_2SO_4, and $MgSO_4$; toxicity at higher doses; similar findings with alkali soil
1917	Winslow and Falk[81]	Sodium enhanced bacterial reproduction; $CaCl_2$, NaCl and KCl enhanced bacterial respiration
1919	Gustafson[82]	Stimulated bacterial respiration by NaCl and $CaCl_2$
1921	Brooks[83]	Enhanced reproduction in bacteria by low levels of $La(NO_3)_3$ (6 doses+control)
1922	Winslow and Hotchkiss[84]	Preliminary report from Hotchkiss' dissertation on stimulation of bacterial reproduction by multiple chemical agents; 15 of 23 salts stimulated growth at low levels
1923	Hotchkiss[85]	Dissertation and paper

1900, the general belief had emerged in the realm of chemical toxicology that low doses as a general rule had the capacity to stimulate, while higher doses would inhibit the activity. This so-called truism became referred to as either the Arndt-Schulz Law or Hueppe's Rule as a result of Hugo Schulz's research on chemical stimulation of yeast metabolism[46,86,87] and Ferdinand Hueppe's research on chemical stimulation of bacterial growth.[74]

The concept of radiation hormesis followed that of chemical hormesis and of course had to await the discovery of X-rays in 1895 by Roentgen and uranium by Becquerel in 1896 and radium by the Curies in 1898. However, the biological effects of ultraviolet (u.v.) radiation were actively researched much earlier with the report of Downs and Blount[88] being credited with the discovery of the killing action of sunlight on bacteria and other microorganisms, thereby drawing attention to the importance of chemically active spectral radiation. By the 1890's it was clear that u.v. light was an important factor, if not the principal factor, in the reported lethal effects of u.v. By the early years of the 20th century, the region of lethal actions had been considerably refined to progressively more specific u.v. wavelengths[89] (Table 2).

Thus, it was quite clear that evaluation of the concept of low dose chemical stimulation had a

Table 2 Early historical references for low dose stimularory response by radiation

Year	*Person*	*Comment*
Plants and radium		
1908	Gager[90]	Reported on up to 91 experiments; most of poor quality; some limited supportive evidence of low dose stimulation
1912-1915	Ewart[91]; Ross[92]; Sutton[93]; Hopkins and Sachs[94] Ramsey[95]	All explored the commercial application of radium soil treatment for agricultural purposes; generally not very enthusiastic; Hopkins and Sachs – 19 of 36 experiments showed stimulation (4 doses+control)
1912	Dormer[96]	Seed germination enhanced by radiation
1913	Stoklasa[97]	Growth of cucumber, mint and tobacco seedings enhanced by radium
1915	Agulhon and Roberts[98]	Seed germination enhanced by radium
1922, 1930, 1932	Stoklasa[99]; Stoklasa *et al*[100] Stoklasa and Penkava[101]	Enhanced seed germination in multuple species by radium
1932	Montet[102–104]	Seed germination enhanced by radium
Fungi and radium		
1908	Gager[90]	Yeast fermentation enhanced noticeable by radium source
1923	Kotzareff and Chodat[105]	Fermentation enhanced by radium
Mice and radium		
1922	Sugiura and Failla[106]	Growth in mice enhanced at low levels by radium
Fungi and ultraviolet radiation		
1907	Purvis and Warwick[107]	One of earliest reports on stimulation of spore products with small amounts of radiation; *Saccharomyces* culture exposed to 10–20 hours at 30cm from lamp; edges of cultures grew very fast
1924	Chavarria and Clark[108]	Short exposures always produced a stimulation of growth rates of *Montonella* cultures; longer exposures were lethal (up to 6 doses)
1928	Nadson and Philoppov[109]	Observed that yeast colonies had much greater growth around the edges of an irradiated zone; growth in the middle was retarded; stimulation attributed to small amounts of scattered radiation
1928	Hinrichs[110]	UV stimulated fermentation by yeast; longer duration exposure was inhibitory
1930	Hutchinson and Ashton[111]	Reported growth stimulation following radiation of *Colletitrichum phomoides*
1928, 1930, 1931	Stevens[112–115]	UV initiated the development of reproductive structures in great numbers; these were caused by exposure of less than 1 min at 20 cm from lamp; numerous fungal species (9 doses)
1928	Porter and Bockstahler[125]	Acceleration of numbers of spores forming on a black border to the sterile zone
1930	Ramsey and Bailey[117]	Marked stimulation of fungal growth with short duration exposure
1932	Bailey[118]	Marked stimulation of fungal growth with short duration exposure
1932	Dillon-Weston[119]	Marked stimulation of fungal growth with short duration exposure
1935	Smith[120]	Stimulation of mycelium growth after retardation; there was direct stimulation of spore formation
1937	Sperti *et al*[121]	Yeast cells stimulated to growth after death/injury of other cells
Insects and X-rays		
1912	Hastings *et al*[122]	Hatching of silkworm eggs was accelerated by X-rays (dose not given)
1917	Davey[123]	Low doses of X-rays increased longevity of confused flour beetle
1919	Davey[124]	Confirmed and extended 1917 study
Rats and X-rays		
1919, 1921	Russ *et al*[125,126]	Body weight enhanced in rats exposed to low doses of X-rays; immunity towards tumor grafts enhanced in rats exposed to low doses of X-rays

strong headstart over its radiation counterpart. This differential advance in favor of chemical hormesis can be seen not only in the sheer volume of published research activity, but also in what was studied (Figure 1). In general, the most active early areas of low dose chemical stimulation research was that of plant growth, followed by fungal growth and metabolism, with bacterial metabolism a distant third (Table 1, Figure 1). By the end of the 19th century, there had been considerable activity in these two dominant areas principally driven by hopes of industrial and/or agricultural applications. In the case of plant research, there was the obvious interest in the enhancement of agricultural productivity, while in the case of fungal growth and metabolism there was great interest in finding ways to enhance, refine, and apply the process of fermentation.

In the case of plant growth a separate line of research developed in the late 1890's by Kahlenberg and his associates at the University of Wisconsin [34–36] who were more concerned with the assessment of the biological effects of highly dilute solutions based principally on the application of newly developed concepts of physical chemistry permitting the use of molar solutions rather than percentages as had been the case.

In the case of chemical stimulation of plant growth, the first decade of the 20th century witnessed the transformation of relatively naïve experimental de-

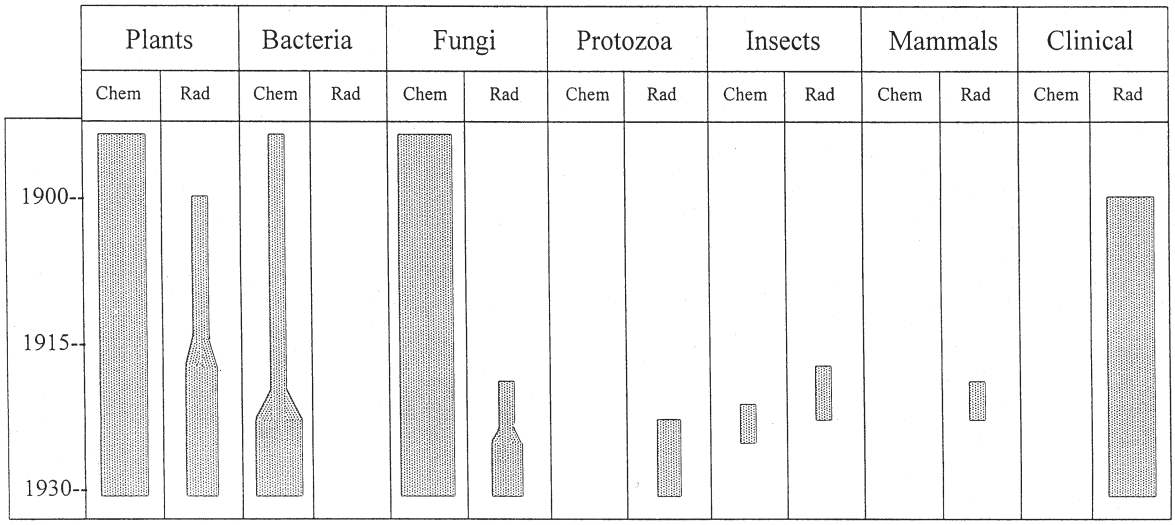


Figure 1 Historical differences between chemical (Chem) and radiation (Rad) hormesis for direction of research during the early decades of the 20th century. Relative widths represent qualitative differences in the amount of research at any particular time point

signs into an impressive formulation of concepts that established a strong foundation for low dose chemical research. Most notably during the first decade of the 20th century were the contributions of True and Gies,[37] Cameron and Breazeale,[38] True and Oglevee,[39] Jensen,[40] and Schreiner and Reed.[30] These investigators incorporated the concept of large numbers of doses, doses below toxic thresholds, and high sample sizes along with the concurrent measurement of multiple endpoints. In addition, these authors began to display these data on highly illustrative graphs enhancing presentation of the data and conceptual understanding.

During the late 1890's initial hypotheses were formulated concerning the underlying mechanism of the low dose response and the role of such mechanisms in the organism's metabolism. For example, Townsend[12] first proposed that low dose chemical stimulation was an overcompensation for chemically induced injury, an observation that was subsequently supported and generalized to other models during these early years of development.[127,128] In fact, this initial concept of Townsend[12] remains the dominant conceptual explanation of the low dose stimulatory phenomenon.[129] In addition, Richards[50,51] proposed the concept of enhancement of metabolic economic efficiency as a result of exposure to low levels of chemical stress. While these conceptual frameworks were being developed and debated, True and Oglevee[39] developed the first graphing of the so-called hormetic β-curve with respect to dose stimulatory range, maximum stimulatory response, and relationship of the maximum stimulatory response to the toxic threshold which were remarkably similar to more modern representations.[130,131] All of these developments occurred well before the genuine onset of low dose stimulatory research in the radiation domain.

The course of low dose radiation research displayed both similar and different research tracts than that for the chemical hormesis research. The most significant area of similarity was the interest in the effects of X-rays and radium on plant growth. The quality of the study designs with respect to low dose radiation stimulation of plant growth did not reflect the rapidly developing maturity of the chemical hormesis researchers. For example, by 1900 the area of low dose chemical stimulation was well on its way to having adopted modern study design criteria, whereas the early researchers in the area of radiation hormesis often displayed noticeable deficits in study design. This was clearly illustrated in the publication of Gager[90] who often employed in his over 90 experiments inadequate sample size, inadequate reporting of experimental methods and overinterpretation of preliminary findings. However, this differential maturity seen in the design, conduct, and interpretation of early low dose stimulation studies appeared to be related to the fact that the first wave of chemical researchers in the US, such as True, Jensen, and Stevens, were part of the broader and well established plant/agricultural research community with considerable laboratory and field research experience. In addition, there was a strong tendency to publish in the most well established journals of that era, such as the Botanical Gazette, which further insured a more advanced professional product.

Part of the differential quality of the earlier investigations on radiation effects on plant growth

was the difficulty in establishing a quantifiable radiation dose metric. When this became established by the early 1920's it encouraged researchers in the plant area to collaborate more effectively with persons trained in radiation dosimetry. For example, the 1931 study by Failla and Henshaw[132] represents an excellent collaboration between radiation dosimetry expertise (Failla) and plant biology (Henshaw). However, this challenge of bringing together such initially divergent expertise resulted in the differential rate of maturity between the chemical and radiation domains.

In addition to the above mentioned factors affecting the limitations in study design, it was also influenced by both the object of exposure and technological developments. That is, the generally lower sample size among early plant radiation researchers was influenced by the fact that exposure was usually administered to seeds rather than the seedling or later developing plant and the fact that the X-ray Coolidge tube was of limited size and could only accommodate a fixed number of seeds.[133] This also affected studies with larger seeds more differentially. While it would have appeared that such a technological limitation should not have been an issue, it nonetheless affected the sample size of numerous earlier plant studies. The area of chemical plant research did not have such constraints.

Of particular note is that the first American researchers assessing potential low dose stimulatory effects of X-rays on plant growth were at the University of Chicago in the late 1920's,[134,135] some 30 years after the discovery of X-rays! The reasons for this late application to plant growth of low doses of X-rays in the US are unknown, given its widespread use in the plant domain in other countries including various European countries and Japan and the earlier assessment and claims of radium induced plant growth stimulation in the US.[2]

Perhaps the most aggressive area of research in low dose radiation was in the area of medical applications for the treatment of various diseases. While X-rays were being used to treat tumors within a year of their discovery, it took about 10 years for the concept to emerge that certain 'low' doses of X-rays (i.e., about 10–50% of the human erythema dose, 60–300 R) to treat a wide spectrum of human inflammatory diseases.[136–140] This concept of employing radiation at relatively low doses for therapeutic purposes never developed in an analogous fashion in the area of chemical stimulation except in so far as the Arndt-Schulz Law became a theoretical framework to support the medical practice of homeopathy. However, in the case of radiation the low dose X-ray treatment was part of the traditional medical establishment with publications and advances in this area in such journals as the Journal of the American Medical Association, the New England Journal of Medicine, Radiology, and others. Thus, in contrast to the relationship of chemical hormesis to the fringe medical practice of homeopathy low dose radiation therapy was part of the traditional medical establishment.

Other notable developments included the later and active research of u.v. on fungi in the 1920's and 1930's (Table 2). In fact, it was during this research that the first conceptual development of hormetic mechanisms was presented by radiation hormesis researchers. That is, Smith[120] reported that u.v.-induced mycelium growth occurred only after an initial injury. This concept was quite similar to that reported almost 40 years before by Townsend.[12]

Similarities and differences in the objects of study in chemical and radiation hormesis

A marked difference between the chemical and radiation low dose stimulatory response was the near total absence of such observations with radiation on bacteria, but the striking productivity of this area in the chemical domain in the 1920's and 1930's (Tables 1 and 2) particularly at Yale University where a long series of PhD students under the highly respected Professor Winslow clearly established the reproducible nature of the hormetic response. Of particular note was the research of Hotchkiss[85] who assessed the effects of twenty-three chemical agents on bacterial growth, with fifteen demonstrating low dose stimulatory effects. The work of Hotchkiss[85] was remarkable for its strong study design features, large number of doses especially below the toxic threshold, and consistent nature of the low dose stimulatory response. In fact, these and the related findings of Winslow's other students became incorporated into leading microbiological texts of the mid 20th century[141–143] along with incorporation of standard assays in laboratory exercises for introductory college students.[144]

A general area that was pursued by those involved in radiation but not chemical hormetic research was the area of cell division. This is seen in research concerning cell division in paramecia, the chick embryo, and various cell types. This research became more substantial in the 1920's with the generally consistent conclusion that low doses of X-rays can stimulate cell division in a variety of models.

The role of low dose stimulatory responses generally did not address the issue of longevity during the early years of the 20th century. However, two excellent papers were provided by Davey of

General Electric in which he unexpectedly reported a low dose of X-rays enhanced longevity in the confused flour beetle.[123] These findings were confirmed and extended in a follow-up paper by the same author.[124] No comparable paper was presented on the chemical side. However, even more surprising is that the striking findings of Davey[123,124] were not followed up for some 40 years until Cork[145] confirmed the life extending response with a gamma ray source using the same animal model.

In summary, the development of research in the area of chemical hormesis occurred earlier, was more extensive, and considerably more mature with respect to the quality of study design and conceptual understanding of a mechanistic framework. However, the two areas are similar in that both were influenced in their early development by commercial applications.

With respect to commercial applications the most visible and high stakes activities were concentrated within the area of radiation hormesis. In these cases a number of attempts explored the use of radionuclides to enhance plant growth. By 1923 a patent was issued on a process to cause radiation induced stimulation of plant growth.[147] A number of commercial businesses were created for this purpose but with little tangible and no long lasting success.[146,148] Factors affecting the lack of apparent commercial success were complex, involving technological, biological, social, and economic factors. From the biological perspective, the 'fact' of stimulation was assumed to have been proven before reasonably convincing data had been established. In addition, there was little appreciation at that time (i.e., 1915–1930) concerning the nature of the low dose stimulation dose-response relationship, including the recognition that the average maximum increase would only be 30–60% above controls and that each plant species and perhaps each set of experimental conditions could display a different optimal dose. Such complexity especially in a new developing area clearly provided the basis for failure for commercial success. In addition, the cost of radium for potential use as fertilizer was quite high, being about $100 000/gram in 1915.[94] In order to double the background levels of radium emanation (radon) in the soil, Ramsey[95] estimated one must use 75 milligrams/acre (i.e., $7500/acre).

The area of low dose clinical treatment with X-rays had a long term series of successes that were adequately documented in highly prestigious journals.[136–140] However, X-ray treatment, like other treatments, competed as a therapeutic option with other available treatments and/or procedures. In the case of X-ray treatment of inflammatory diseases, it eventually lost out to advances in chemotherapeutics which was reinforced with a growing concern over potentially harmful effects of X-rays even at low doses.

Why radiation and chemical hormesis were rejected

The rejecting of chemical and radiation hormesis hypotheses has some general overlapping features but a number of distinctive aspects as well. First, while this discussion has divided the debate into chemical and radiation hormesis, it is not clear that either one of these areas were identified as a stand alone 'field'. In the case of chemical hormesis it was uncommon for plant chemical hormesis researchers to cite those in other chemical areas such as fungi and bacteria. Furthermore, the chemical and radiation hormesis researchers generally never cited each other. In addition, there was a dearth of review papers on the topic of low dose stimulatory responses. This lead to only a very limited summarization of relevant papers in either the introduction or discussion of focused research reports. While this intellectual truncation is appropriate for narrow research papers, the general lack of critical broad reviews of the literature limited the capacity to develop more integrated assessments of the broad scientific literature. While the publication of critical and integrated reviews is common today, in the early decades of the 20th century it was not common. In fact, it is ironic that the first major review of the literature on radiation hormesis was of a highly critical nature (Johnson[149] see below). This lack of broad integration coupled with the absence of a central dose-response concept resulted in a poorly developed general understanding of hormetic dose-response relationships. Thus, low dose stimulatory findings were quite truncated into very narrow model (e.g., plant, bacteria, etc.) specific responses with little attempt to develop a general focus on dose-response relationships. Such a lack of an integrated focus on the hormetic dose-response was a fundamental underlying factor that contributed to the inability of these hypotheses to better establish themselves.

Radiation hormesis, with a more limited database to support its premise than chemical hormesis, became the object of a highly successful attack by Edna Johnson in the 1930's.[149,150] Of particular significance was that the Johnson criticism targeted the strongest experimental basis of radiation hormesis, that is, the effects of X-rays on plant growth. As noted above, the review of Johnson[149] was not only one of the first major reviews of the literature on the effects of X-rays on plant growth it also received additional prestige for being part of a major National Research Council (NRC) assessment of the biological

effects of radiation. This combination proved to have considerably greater impact than criticism tucked away in a discussion section of a focused research paper. With the X-ray/plant component of the radiation hormesis perspective placed on weakened grounds by the review of Johnson,[149] there was little countervailing opposition to offset the criticism or limit its impact. More specifically, there was no corresponding supportive evidence with bacteria, and only limited supportive evidence with fungi (Table 2). The only other potential widespread strength supporting the concept of low dose stimulation was in the area of medical treatment and this itself was on weak grounds due to overzealous claims for beneficial radiation exposure (e.g., mild radium therapy)[151] and growing fears of low dose mutational and cancerous effects of X-rays.[152]

The demise of radiation hormesis is understandable especially in light of its limited database, difficult findings to reproduce, harsh criticism from leading scientists coupled with a growing fear of radiation and an emphasis on defining safety standards that required defining frankly toxic effect levels, lowest adverse effect levels and toxic thresholds. These became the predominant questions in the mid 1930's, not whether low doses cause a marginal and hard to reproduce and even harder to interpret stimulatory response. Once radiation hormesis was pushed aside and not considered credible, funding became generally unavailable and it became further marginalized.

The chemical hormesis area should have survived as a central hypothesis not only as a result of its better general database but also because it had direct linkage with numerous well known scientists or their students such as Louis Pasteur [Raulin's work[45] in Pasteur's laboratory], Robert Koch [Ferdinand Hueppe[74] was a protégé of Koch], Wilhelm Ostwald [Kahlenberg (see Kahlenberg and True;[34] Copeland and Kahlenberg[36]) received his PhD with Ostwald in Germany before returning to the University of Wisconsin], Charles Richet[153, 154] (the Nobel laureate for discovering anaphylaxis) who demonstrated low dose stimulatory effects in fermentation systems, and a strong series of US academics at prestigious institutions (i.e., Duggar at the University of Wisconsin, Townsend and Richards at Columbia University, Stevens at the University of Chicago and later at Stanford University, Winslow at Yale University) and True with the US Agricultural Research Service after moving from the University of Wisconsin. No comparable grouping of outstanding researchers with such powerful lineages and/or institutions and strong publication records were present with radiation hormesis. Yet despite the greater historical foundations, stronger data, acceptance as a central concept in bacteriology and long listings of prestigious scientists supporting it, all factors less developed in radiation hormesis, both chemical and radiation hormesis met the same fate of marginalization about the same time.

Despite the above outstanding research and academic pedigree of hormesis researchers of the early decades of the 20th century, the area of low dose chemical stimulation was to become the object of intense criticism by the next generation of dominant features in the field of pharmacology and toxicology. This criticism was to have its origin in the fact that this area of research was too closely allied with the controversial medical practice of homeopathy.[1] The area of chemical hormesis had become used as an explanatory factor by advocates of the medical practice of homeopathy. In fact, Hugo Schulz, the microbiologist who first reported that low doses of numerous chemicals stimulated yeast metabolism, joined with (the homeopathic physician) Rudolph Arndt and together promoted the broad generalizability of the low dose stimulatory curve into a prime explanatory framework of how homeopathic drugs worked. This close association of a scientific hypothesis with a politicized medical practice was criticized as early as 1896 by Hueppe.[74] Nonetheless, the association of hormesis to homeopathy remains even to the present.[155] However, in 1937 the prestigious pharmacologist AJ Clark, of the University of Edinburgh, published his classic text, 'Handbook of Pharmacology', in which he devoted 15% to a refutation of the Arndt-Schulz Law.[156] Clark, the discoverer of the first molecular receptor (i.e., the acetylcholine receptor), was a towering scientific feature by himself, but he also had an unusually strong collaboration with several of the most powerful and respected biostatisticians of that era.

At this time, the fundamental nature of the dose-response was powerfully articulated and was greatly affected by the very biostatisticians (e.g., Bliss, Trevan) who worked with Clark. Lacking any comparable countervailing intellectual force at the time, the concept of hormesis, especially chemical hormesis, became a cultural victim of guilt by association with homeopathy. This marginalization was encouraged by traditional medical philosophy because of the long standing antipathy with homeopathy. Since pharmacology and toxicology developed most extensively within traditional medical schools, it was only natural to have physician-trained pharmacologists/toxicologists lump hormesis with homeopathy and the marginalization was complete.

Table 3 Comparison of factors leading to the demise of chemical and radiation hormesis

Limitations more unique to either chemical or radiation hormesis
Chemical hormesis
Close association with homeopathy brought criticism from those trained in traditional medicine
No practical application recognized and/or advocated
Inadequate human data supporting hypothesis
Radiation hormesis
Weaker database than chemical hormesis
Failure in highly visible applications such as with agricultural productivity
Technical criticism of strongest aspect of supportive data (e.g., plant research criticism of Johnson for NRC report)
Low dose clinical human applications were out-competed by advances in chemotherapy
Fear of exposure even at low doses; this was reinforced by failure of mild radium therapy, fear of the atomic bomb, and later fear of radioactive fallout
Limitations shared by both hypothesis
Lack of quantitative understanding of low dose stimulatory response relationships
Magnitude of stimulatory response is modest even at optimum stimulation (approximately 50% greater than control)
Need for strong study designs with numerous, properly spaced doses below the toxic threshold placed greater time and resource constraints on low dose stimulatory investigations; this has many implications including the encouragement of investigators to emphasize high concentrations for the few doses they would test; this has been embodied in hazard assessment testing and is also more easily reproducible leading to greater ease of acceptance for publication
Lack of intergration of low dose stimulatory response studies in which chemical and radiation hormesis researchers recognized contributions of both fields as providing supportive data
Small pool of scientists researching in the low dose area due to above limitations; this further reinforces bias to continue funding the research of the majority who perform high dose studies
'Timing' was not right to appreciate its significance; society did not have to contend with low dose risk extrapolation; it was more important to establish effects of high doses for worker and community protection
Strong linkage of funded studies with the goals of government which were to define the upper end of the dose-response curve for worker protection and standard setting purposes; this further de-emphasized interest and recognition of low dose effects
Lack of organized presence of low dose researchers (i.e., no society outlet) to consider criticism of low dose effects by established high dose researchers

Discussion

This paper has argued that the concepts of chemical and radiation hormesis had remarkably independent histories with respect to temporal development, direction of research, selection of experimental model, quantity and quality of supportive data, acceptance by the broader scientific community and commercial applications. While it may be the case that the entire fields of chemical and radiation hormesis are perceived as simply one concept, the actual unfolding of these two areas of research has been quite distinct. The separate developments of chemical and radiation hormesis which are seen at the start of the 20th century has been maintained to the present time. Thus, even today there is very little cross communication between those working in the areas of radiation and chemical hormesis.

Even in their respective demises there was also considerable uniqueness (Table 3). That is, the area of chemical hormesis was plagued by its long standing and close association with the medical practice of homeopathy which set the stage for a guilt by association response from traditional medicine which was strongly influencing textbook development, professional society activities and funding programs. In contrast, radiation hormesis was plagued by the high dose applications of radiation which dominated medical practices and made many ill, and the overzealous claims of radium profiteers with the highly visible and tragic death of the millionaire industrialist and playboy Eben M Byers, which resulted in the end of the era of mild radium therapy.[151]

Both areas of hormesis were also plagued by different but highly visible critics. In the case of chemical hormesis the attack was profoundly more intense, intentional, and systematic. As noted above, the allocation of 15% of what has been referred to as a major and classic text for the repudiation of the Arndt-Schulz Law by AJ Clark was the type of challenge that radiation hormesis did not experience. Radiation hormesis certainly had its critics, such as Johnson,[149] but they were more generally limited and focused within the narrower context of a particular research paper. The use of low dose radiation to treat human inflammatory diseases became widely integrated within modern medical practice from the early 1900's through the 1940's. While these practices were not generally referred to

as being related to the Arndt-Schulz Law, there is little question that its advocates such as Desjardin, chief of radiology at the Mayo Clinic, clearly articulated the view that low doses were beneficial for the patients' conditions while higher doses were progressively less effective and even higher doses harmful. The low dose X-ray therapy which typically utilized only a single exposure was simply outcompeted by novel therapies of the mid 1940's such as the progression of antibiotics which brought rapid cures without the residual fears of adverse effects of radiation treatment.

Despite dissimilarities in both chemical and radiation hormesis, the most serious challenges were ones they held in common (Table 3). That is the basic reality that hormesis affected a modest stimulatory response over a limited range of doses. It also requires very stringent study design criteria and endpoint selection in order to properly assess it. These factors affected both its reproducibility and its commercial applications and in the end these most fundamental factors are the principal determining factors for their common demise.

Acknowledgements

This work was sponsored in part by a grant to the University of Massachusetts (Edward J Calabrese, Principal Investigator) by the US Nuclear Regulatory Commission.

References

1 Calabrese EJ, Baldwin LA. Chemical hormesis: Its historical foundations as a biological hypothesis. *Toxicologic Pathology* 1999; **27:** 195–214.
2 Calabrese EJ, Baldwin LA. Radiation hormesis: Its historical foundations as a biological hypothesis. *Human and Experimental Toxicology* 1999; in press.
3 Calabrese EJ, Baldwin LA. The marginalization of hormesis. *Toxicologic Pathology* 1999; **27:** 187–194.
4 Calabrese EJ, Baldwin LA. 1999c. Radiation hormesis: The demise of a legitimate hypothesis. *Human and Experimental Toxicology* 1999; in press.
5 Southam CM, Erhlich J. Effects of extracts of western red-cedar heartwood on certain wood-decaying fungi in culture. *Phytopathology* 1943; **33:** 517–524.
6 Bottger R. Beforderung-des Keimprocesses. *Botanischer Jahresbericht* 1874; **I:** 833.
7 Gustavson G. Chemische Rolle der Mineralsalze in der organischen Natur. *Mitt. Landw. Forstwirthsch. Acad. Petrowskoe-Ravsumowskow* 1881; **3:** 1–15.
8 Koenig J. *Landwirtwchraftliche Jahrbucher* 1883; **12:** 837.
9 Griffiths. *Journal of the Chemical Society* 1884; p. 71.
10 Griffiths. *Journal of the Chemical Society* 1885; p. 46.
11 DeToni BG, Mach P. Sopra l'influenza esercitata della nicotina, etc. *Boll. R. Ist. Bot. Univ. Parmiense* 1893.
12 Townsend C.O. The correlation of growth under the influence of injuries. *Annals of Botany* 1897; **11:** 509–532.
13 Coupin H. Sur la toxicite de sodium et de l'eau de mer a l'egard des vegetaux. *Rev. Gen. Bot.* 1898; **10:** 177–190.
14 Jacobi B. Ueber den Einfluss verschiedener Substanzen auf die Athmung und Assimilation suberser Pflanzen. *Flora* 1899; **86:** 289–327.
15 Lovinson O. Ueber Keimungs- und Wachsthumsversuche an Erbsen in Losungen von fettsauren Salzen unter Ausschluss von Mineralsauren. *Botanisches Zentralblatt* 1900; **83:** I, 33, 65, 97, 129, 185, 210.
16 Johnsson WI. Das Aether-Verfahren beim Fruhtreiben. *Jena.* 1900.
17 Morkowin N. Recherches sur l'influence des anesthesiques sur la respiration des plantes. Revue Gen. de Bot. 1899. **11:** 289–303.
18 Morkowin N. Ueber den Einfluss der Reizwirkunzen auf die intramoleculare Athmung der Pflanzen. *Berlin Deutsches Botanische Gesellschaft* 1903; **21:**72.
19 Dandeno JB. 1900–1903. As cited in Brenchley WE. Inorganic Plant Poisons and Stimulants. Cambridge University Press: London, 1914.
20 Suzuki S. On the poisonous action of potassium ferrocyanide on plants. *Bulletin of the College of Agriculture of Japan* 1902–1903; **5:** 203–205.
21 Suzuki S. On the action of vanadium compounds on plants. *Bulletin of the College of Agriculture of Japan* 1902–1903; **5:** 513–515.
22 Chapin P. Einfluss der Kohlensaure auf das Wachsthum. *Flora* 1902; **91:** 348–379.
23 Aso K. On the physiological influence of manganese compounds on plants. *Bulletin of the College of Agriculture, Tokyo Imperial University* 1902–1903; **5:** 177–185.
24 Loew O, Sawa S. On the action of manganese compounds on plants. *Bulletin of the College of Agriculture, Tokyo* 1902; **5:** 161–172.
25 Chuard E, Porchet F. L'action des sels de cuivre sur les vegetaux. *Archives des Sciences Phys. Nat.* 1902; **14:** 502–505.
26 Chuard E, Porchet F. L'action des sels de cuivre sur les vegetaux. *Archives des Sciences Phys. Nat.* 1903; **15:** 91.
27 Nakamura S. Comparative experiments of the effect of X-rays. *Proceedings of the Koto-Kwai* 1903; No. 111. Referenced in the *Journal of the College of Agriculture, University of Tokyo*, 1923; **8,** No. 2.
28 Kanda M. Studien uber die Reizwirkung einiger Metallsalze auf das Wachsthum hoherer Pflanzen. *Journal of the College of Science of the Imperial University of Tokyo* 1904; **19:** 1–39.
29 Aso K. On a stimulating action of calcium fluoride on phaenograms. *Bulletin of the College of Agriculture, Tokyo* 1906–1908; **7:** 90.
30 Schreiner O, Reed HS. The toxic action of certain organic plant constituents. *Botanical Gazette* 1908; **45:** 73–102, 271.

31 Agulhon H. Recherches sur la presence et la role du bore chez les vegetaux. Thesis, Paris, 1910.
32 Montemartini L. The stimulatory effect of sulphate of manganese and sulphate of copper on plants. Inter. Inst. Agr. [Rome], *Bul. Bur. Agr. Intell. and Plant Diseases,* Year 2, 1911; **No. 11/12,** p. 2467 (abstract).
33 Lipman CB, Wilson FH. Toxic inorganic salts and acids as affecting plant growth. *Botanical Gazette* 1913; **55:** 409–420.
34 Kahlenberg L, True RH. On the toxic action of dissolved salts and their electrolytic dissociation. *Botanical Gazette* 1896; **22:** 81–124.
35 Heald FD. On the toxic effect of dilute solutions of acids and salts upon plants. *Botanical Gazette* 1896; **22:** 125–153.
36 Copeland EB, Kahlenberg L. The influence of the presence of pure metals upon plants. *Wisconsin Academy of Sciences, Arts, and Letters* 1899; **12:** 454–474.
37 True RH, Gies WJ. On the physiological action of some of the heavy metals in mixed solutions. *Bulletin of the Torrey Botanical Club* 1903; **30:** 390–402.
38 Cameron FK, Breazeale JF. The toxic action of acids and salts on seedlings. *Journal of Physical Chemistry* 1904; **8:** 1–13.
39 True RH, Oglevee CS. The effect of the presence of insoluble substances on the toxic action of poisons. *Botanical Gazette* 1905; **39:** 1–21.
40 Jensen GH. Toxic limits and stimulation effects of some salts and poisons on wheat. *Botanical Gazette* 1907; **43:** 11–44.
41 Rumm C. Ueber die Wirkung der Kapferpraparate bei Bekampfung der sogenannten Blattfallkrenkheit der Weinrebe. *Botanische Zeitung* 1893; **51:** 163–164.
42 Berlese, Sostegni. Recherches sur l'action des sels de cuivre sur'la vegetatuib de la vigne et sur le sol. *Botanischer Jahresbericht* 1985 **23:** 292–294.
43 Evans WH. Copper sulphate and germination. U.S. Department of Agriculture Division of Vegetable Physiology and Pathology, Bulletin No. 10, 1896.
44 Frank B, Kruger F. Ueber den Reiz, welchen die Behandlung mit Kupfer auf die Kartoffulpflanze hervorbringt. *Berlin Deutsches Botanische Gesellschaft* 1899; **12:** 8–11.
45 Raulin J. Etudies chimiques sur la vegetation. *Ann. Sc. Nat. Bot. V.* 1869; **11:** 93–299.
46 Schulz H. Zur Lehre von der Arzneiwirdung. *Virchows Archiv fur Pathologische Anatomie und Physiologie* 1887; **108:** 423–445.
47 Pfeffer W. Ueber Election organischer Nahrstoffe. *Jahrbuch fur Wissenschaftliche Botanische.* 1895; **28:** 205–268.
48 DeSeynes J. Resultats de la culture du *Penicillium cupricum* Trabut. *Bull. Soc. Bot. France* 1895; **42:** 451–455.
49 Stevens FL. The effect of aqueous solutions upon the germination of fungus spores. *Botanical Gazette* 1898; **26:** 377–406.
50 Richards HM. Die Beeinflussung des Wachsthums einiger Pilze durch chemische Reinze. *Jahrbuch fur Wissenschaftliche Botanische* 1897; **30:** 665–688.
51 Richards HM. The effect of chemical irritation on the economic coefficient of sugar. *Bulletin of the Torrey Botanical Club* 1899; **26:** 463–479.
52 Clark JF. Electrolytic dissociation and toxic effect. *Journal of Physical Chemistry* 1899; **3:** 263–316.
53 Clark JF. On the toxic effect of deleterious agents on the germination and development of certain filamentous fungi. *Botanical Gazette* 1899; **28:** 289–327, 378–404.
54 Ono N. Notes on the stimulating effect of certain substances upon the growth of algae and fungi. *Botanical Magazine, Tokyo* 1900; **14:** 75–78.
55 Kosinki I. Die Athmung bei Hungerzustanden und unter Einwirkung von mechanischen und chemischen Reizmitteln bei *Aspergillus niger*. *Jahrbuch fur Wissenschaftliche Botanische* 1901; **37:** 137–204.
56 Richter A. Zur Frage der chemischen Reizmittel. *Centralblatt fur Bakteriologie* 1901; **7:** 417–429.
57 Yasuda A. On the effect of alkaloids upon some moulds. *Botanical Magazine, Tokyo* 1901; **15:** 79–83.
58 Hattori H. Studien uber die Einwirkung des Kupfersulfate auf einige Pflanzen. *Journal of the College of Sciences, Imperial University of Tokyo* 1901; **15:** pt. 3.
59 Duggar BM. Physiological studies with reference to the germination of certain fungous spores. *Botanical Gazette* 1901; **31:** 38–66.
60 Clark JF. On the toxic properties of some copper compounds with special reference to bordeaux mixture. *Botanical Gazette* 1902; **33:** 26–48.
61 Pulst C. Die Widerstandsfahigkeit einiger Schimmelpilze gegen Metallgifte. *Jahrbuch fur Wissenschaftliche Botanische* 1902; **37:** 205–263.
62 Orlawski SF. Ueber die Wirkung des Arsen auf das Wachstum und die chemische Zusammensetzung von *Aspergillus niger*. Petersburg, Dissertation, 1902.
63 Coupin H. Sur la nutrition du sterigmatocystis nigra. *Comptes Rendus* 1903; **136:** 392–394.
64 Watterson A. The effect of chemical irritation on the respiration of fungi. *Bulletin of the Torrey Botanical Club* 1904; **31:** 291–303.
65 Latham ME. The stimulation of *Sterigmatocystis* by chloroform. *Bulletin of the Torrey Botanical Club* 1905; **32:** 337–351.
66 Fitch R. The action of insoluble substances in modifying the effect of deleterious agents upon the fungi. *Annales de Chimie et de Physique* 1906; **4:** 313–322.
67 Javillier M. Sur l'influence favorable del petites doses de zinc sur la vegetation de *Sterigmatocystis nigra* V. Tgh. *Comptes Rendus* 1907; **145:** 1212–1215.
68 Latham ME. Nitrogen assimilation of Sterigmatocystis nigra and the effect of chemical stimulation. *Bulletin of the Torrey Botanical Club* 1909; **36:** 235–244.
69 Harden A, Young WJ. The alcoholic ferment of yeast-juice. Part VI. The influence of arsenates and arsenites on the fermentation of sugars by yeast-juice. *Proceedings of the Royal Society, London (Series B)* 1910–1911; **83:** 451–475.
70 Harden A. Alcoholic Fermentation. Longmans, Green and Co.: New York, 1911.
71 Livingston BE. Chemical stimulation of a green alga. *Bulletin of the Torrey Botanical Club* 1905; **32:** 1.
72 Woodworth CW. The toxicity of insecticides. *Science* 1915; **41:** 367–369.
73 Effront. *Comptes Rendus* 1894; **119:** 169.

74 Hueppe F. Principles of Bacteriology. Translated from the German by E.O. Jordan. The Open Court Publishing Company: Chicago, 1896.
75 Lindet. *Comptes Rendus* 1904; **138:** 508.
76 Hune. Die begunstigende Reizwirkung kleinster Mengen von Bakteriengiften auf die Bakterienvermehrung. *Centralblatt fur Bakteriologie* I. Abt. Orig. 1909; **48:** 135–140.
77 Lipman CB. Toxic effects of 'Alkali Salts' in soils on bacteria. I. Ammonification. *Centralblatt fur Bakteriologie* 1912; **32:** 5864.
78 Greaves J.E. The influence of arsenic upon the biological transformation of nitrogen in soils. *Biochemical Bulletin* 1913; **3:** 2–16.
79 Delepine AS, Greenwood A. The action of some metals upon certain water and other bacteria. *Royal Sanitation Institute Journal* 1914; **35:** 317–327.
80 Brown PE, Hitchcock EB. The effects of alkali salts on nitrification. *Soil Science* 1917; **4:** 207–229.
81 Winslow C-EA, Falk IS. Studies on salt action. I. Effect of calcium and sodium salts upon the viability of the colon bacillus in water. *Proceedings of the Society of Experimental Biology and Medicine* 1917; **15:** 67–69.
82 Gustafson FG. Comparative studies on respiration: IX. The effects of antagonistic salts on the respiration of *Aspergillus niger*. *Journal of General Physiology* 1919; **2:** 17–24.
83 Brooks MM. Antagonistic action of lanthanum as related to respiration. *Journal of General Physiology* 1921; **3:** 337–342.
84 Winslow C-EA, Hotchkiss M. Studies on salt action. V. The influence of various salts upon bacteria growth. *Proceedings of the Society of Experimental Biology and Medicine* 1922; **19:** 314–315.
85 Hotchkiss M. Studies on salt action. VI. The stimulating and inhibitive effect of certain cations upon bacteria growth. *Journal of Bacteriology* 1923; **8:** 141–162.
86 Schulz H. Uber Hefegifte. *Pflugers Archiv fur die gesamte Physiologie* 1888; **42:** 517–541.
87 Grote LR. Die Medizin der Gegenwart in selbstdarstellungen: Hugo Schulz. Verlag von Felix Meiner, Leipzig, pp. 217–250, 1923 (NIH Library Translation by Ted Crump).
88 Downs A, Blount TP. Researches on the effect of light upon bacteria and other organisms. *Proceedings of the Royal Society (London)* 1877; **26:** 488–500.
89 Duggar BM. Effects of radiation on bacteria. In: Biological Effects of Radiation, Vol. II BM Duggar, (ed) McGraw-Hill Book Co., Inc.: New York, pp. 1119–1149, 1936.
90 Gager CS. Effects of the rays of radium on plants. *Memoires of the New York Botanical Gardens* 1908; **4:** viii+278.
91 Ewart AJ. *Journal of the Department of Agriculture, Victoria* 1912; **10:** 417–421.
92 Ross GG. X-ray treatment of carbuncle of face. *Annals of Surgery* 1914; **66:** 99–100.
93 Sutton HF. Radium and plant growth. *Gardeners' Chronicle 3rd Series* 1915; **58:** 102.
94 Hopkins CG, Sachs WH. Radium fertilizer in field tests. *Science* 1915; **41:** 732–735.
95 Ramsey RR. Radium fertilizer. *Science* 1915; **42:** 219.
96 Doumer E. Radiumenanation und die Keimung der Samen. Beiblatter, VI. *Internat. Kongress Allgem. und Arztl. Elektrol. und Radiol. Prog,* 1912.
97 Stoklasa J. Bedeutung der Rakioaktivitat in der Physiologie. *Chemische Zeitschrift* 1913; **37:** 1176.
98 Agulhon H, Robert T. The action of radium and radioactivity on germination in higher plants. *Annales de L'Institut Pasteur* 1915; **29:** 261–273.
99 Stoklasa J. Influence du selenium et du radium sur la germination des grains. *Comptes Rendus Academie Science [Paris]* 1922; **174:** 1075–1077.
100 Stoklasa J, Penkava J, Strupl M, Tjukov D, Vrbensky V. Die Dynamik der physiologischen Wirkung der Radioaktivitat auf das Protoplasma. *Beitrage zur Biologie des Pflanzen* 1930; **18:** 185–224.
101 Stoklasa J, Penkava J. Biologie des Radiums und Uraniums. Paul Parey: Berlin, 1932.
102 Montet D. Action des faibles radioactivities sur la germination des graines. *Comptes Rendus* 1932; **194:** 304–306.
103 Montet D. L'action de la radioactivite dur la germination des bulbes. *Comptes Rendus* 1932; **194:** 1093–1095.
104 Montet D. De l'influence des faibles radioactivities sur la germination. *Comptes Rendus* 1932; **109:** 678–680.
105 Kotzareff A, Chodat F. De l'action exercee par l'emanation du radium sur les levures. *Compte Rendu Societe de Physique et d'Histoire Naturelle [Geneve]* 1923; **40:** 36–39.
106 Sugiura K, Failla G. 1922. Some effects of radium radiations on white mice. *Journal of General Physiology* 1922; **4:** 423–436.
107 Purvis JE, Warwick GR. The influence of spectral colors on the sporulation of Saccharomyces. *Proceedings of the Cambridge Philosophical Society* 1907; **14:** 30–40.
108 Chavarria AP, Clark JH. The reaction of pathogenic fungi to ultra-violet light and the role played by pigment in this reaction. *American Journal of Hygiene* 1924; **4:** 639–649.
109 Nadson GA, Philoppov GS. Action excitante des rayons ultra-violets sur le developpement des levures et des moisissures. *Comptes Rendus* 1928; **98:** 366–368.
110 Hinrichs MA. Ultraviolet radiation and division in *Paramecium caudatum*. *Physiological Zoology* 1928; **1:** 394–415.
111 Hutchinson AH, Ashton MR. The effects of radiant energy on growth and sporulation in *Colletotrichum phomoides*. *Canadian Journal of Research* 1930; **3:** 187–198.
112 Stevens FL. Effects of ultra-violet radiation on various fungi. *Botanical Gazette* 1928; **86:** 210–225.
113 Stevens FL. The sexual stage of fungi induced by ultra-violet rays. *Science* 1928; **67:** 514–515.
114 Stevens FL. The effects of ultra-violet irradiation on various Ascomycetes, Sphaeropsidales and Hyphomycetes. *Centralblatt fur Bakteriologie* (Abt. 2) 1930; **82:** 161.
115 Stevens FL. The ascigerous stage of *Colletotrichum lagenarium* induced by ultra-violet irradiation. *Mycologia* 1931; **23:** 134–139.
116 Porter CL, Bockstahler HW. Concerning the reaction of certain fungi to various wave lengths of light. *Proceedings of the Indiana Academy of Sciences* 1928; **38:** 133–135.
117 Ramsey GB, Bailey AA. Effect of ultra-violet radiation on sporulation in *Macrosporium* and *Fusarium*. *Botanical Gazette* 1930; **89:** 113–136.

118 Bailey AA. Effects of ultraviolet radiation upon representative species of Fusarium. *Botanical Gazette* 1932; **94:** 225–271.
119 Dillon-Weston WAR. III. Studies on the reaction of disease organisms to light. The reaction of disease organisms to certain wave-lengths in the visible and invisible spectrum. *Science of Agronomy* 1932; **12:** 352–356.
120 Smith EC. Effects of ultra-violet radiation and temperature on *Fusarium*. II. Stimulation. *Bulletin of the Torrey Botanical Club* 1935; **62:** 151–164.
121 Sperti GS, Loofbourow JR, Dwyer CM. Proliferation-promoting factors from ultraviolet injured cell. *Ins. Divi. Thomae* 1937; **1:** 163–191.
122 Hastings, Beckton, Wedd. Arch. Middlesex Hosp., 11th Cancer Report, 1912.
123 Davey WP. The effect of X-rays on the length of life of *Tribolium confusum*. *Journal of Experimental Zoology* 1917; **22:** 573–592.
124 Davey WP. Prolongation of life of *Tribolium confusum* apparently due to small doses of X-rays. *Journal of Experimental Zoology* 1919; **28:** 447–458.
125 Russ S, Chambers H, Lond MD, Scott GM, Mottram JC. Experimental studies with small doses of X rays. *Lancet* 1919; **1:** 692–695.
126 Russ S, Chambers H, Scott GM. On the local and generalized action of radium and X-rays upon tumour growth. *Archives of Radiation and Electrotherapy* 1921; **26:** 129–137.
127 Branham SE. The effects of certain chemical compounds upon the course of gas production by baker's yeast. *Journal of Bacteriology* 1929; **18:** 247.
128 Colley MW. Stimulation phenomena in the growth of bacteria as determined by nephelometry. *American Journal of Botany* 1931; **18:** 266–287.
129 Stebbing ARD. A theory for growth hormesis. *BELLE Newsletter* 1997; **6:** 1–11.
130 Calabrese EJ, Baldwin LA. The dose determines the stimulation (and poison): development of a chemical hormesis database. *International Journal of Toxicology* 1997; **16:** 545–559.
131 Calabrese EJ, Baldwin LA. A quantitatively-based methodology for the evaluation of chemical hormesis. *Human and Ecological Risk Assessment* 1997; **3:** 545–554.
132 Failla G, Henshaw PS. The relative biological effectiveness of X-rays and gamma rays. *Radiology* 1931; **17:** 1–43.
133 Breslavets LB. Plants and X-rays. Translation by A. Elbl. AH Sparrow (ed.) The American Institute of Biological Sciences, Washington, 1946.
134 Johnson E. Effect of X-rays on growth, development and oxidizing enzymes of *Helianthus annuus*. *Botanical Gazette* 1926; **82:** 373–402.
135 Johnson E. Growth and germination of sunflowers as influenced by X-rays. *American Journal of Botany* 1928; **15:** 65–76.
136 Desjardins AU. Radiotherapy for inflammatory conditions. *Journal of the American Medical Association* 1931; **98:** 401–408.
137 Desjardins AU. The action of roentgen rays or radium on inflammatory processes. *Radiology* 1937; **29:** 436–445.
138 Desjardins AU. Dosage and method of roentgen therapy for inflammatory conditions. *Radiology* 1939; **32:** 699–707.
139 Desjardins AU. Radiotherapy for inflammatory conditions. *New England Journal of Medicine* 1939; **221:** 801–809.
140 Desjardins AU. The action of roentgen rays on inflammatory conditions. *Radiology* 1942; **38:** 274–280.
141 Buchanan RE, Fulmer EI. Physiology and Biochemistry of Bacteria, Vol. II. Effects of Environment Upon Microorganisms. The Williams and Wilkins Co.: Baltimore, MD, 1930.
142 Clifton CE. Introduction to Bacterial Physiology. McGraw-Hill Book Company, Inc.: New York, pp. 317–338, 1957.
143 Lamanna C, Mallette MF. Basic Bacteriology Its Biological and Chemical Background. Third Edition. The Wilkins and Wilkins Company: Baltimore, MD, 1965.
144 Salle AJ. Fundamental Principles of Bacteriology. McGraw-Hill Book Co., Inc.: New York, pp. 166–197, 1939.
145 Cork JM. Gamma radiation and longevity of the flour beetle. *Radiation Research* 1957; **7:** 551–557.
146 Gager CS. Present status of the problem of the effect of radium rays on plant life. *Memoires of the New York Botanical Gardens* 1915; **6:** 153–160.
147 Adams RR. Radioactive Spray Material. *The Official Gazette of the U.S. Patent Office* 1923; **312:** 320.
148 Gager CS. The effects of the rays of radium on plants. In: Biological Effects of Radiation, Vol. II, (BM Duggar, editor) McGraw-Hill Book Co., Inc.: New York, pp. 987–1013, 1936.
149 Johnson E. Effects of X-rays upon green plants. In: Biological Effects of Radiation, Vol. II, (BM Duggar, editor) McGraw-Hill Book Co., Inc.: New York, pp. 961–985, 1936.
150 Johnson E. On the alleged stimulating action of X-rays upon plants. *American Journal of Botany* 1931; **18:** 603–614.
151 Macklis RM. Radithor and the era of mild radium therapy. *Journal of the American Medical Association* 1990; **264:** 614–618.
152 Weart SR. Nuclear Fear: A History of Images. Harvard University Press: Cambridge, MA, 1988.
153 Richet C. De l'action de doses minuscules de substances sur la fermentation lactique. *Archives Internationales de Physiologie* 1905; **3:** 203–217.
154 Richet C. De l'action de doses minuscules de substance sur la fermentation lactique-troisieme memoire–Periodes d'acceleration et ralentissement. *Archives Internationales de Physiologie* 1906–1907; **4:** 18–50.
155 Bellavite P, Andrioli G, Lussignoli S, Ortoloni R, Conforte A. A scientific reappraisal of the 'principle of similarity.' *Medical Hypotheses* 1997; **49:** 203–212.
156 Clark AJ. Handbook of Experimental Pharmacology. Verlig Von Julius Springer: Berlin, 1937.

HET

Human and Experimental Toxicology
29(8) 667–677

DOI: 10.1177/0960327109358917
het.sagepub.com
SAGE

Hormesis in high-throughput screening of antibacterial compounds in *E coli*

Edward J Calabrese[1], George R Hoffmann[2], Edward J Stanek[3] and Marc A Nascarella[4]

Abstract

This article assesses the response below a toxicological threshold for 1888 antibacterial agents in *Escherichia coli*, using 11 concentrations with twofold concentration spacing in a high-throughput study. The data set had important strengths such as low variability in the control (2%–3% SD), a repeat measure of all wells, and a built-in replication. Bacterial growth at concentrations below the toxic threshold is significantly greater than that in the controls, consistent with a hormetic concentration response. These findings, along with analyses of published literature and complementary evaluations of concentration-response model predictions of low-concentration effects in yeast, indicate a lack of support for the broadly and historically accepted threshold model for responses to concentrations below the toxic threshold.

Keywords

hormesis, biphasic, dose response, U-shaped, adaptive response, antibiotics

Introduction

Since the later part of the 19th century, the threshold concentration/dose response has been widely accepted as the dominant and most fundamental concentration/dose-response model across a broad range of biologically based disciplines.[1] However, the capacity of the threshold concentration/dose-response model to predict accurately low-concentration/dose responses, that is, responses below the threshold, has been challenged over the past decade. The challenge is based on numerous published reports demonstrating that biphasic concentration/dose responses (i.e., hormetic concentration/dose responses) are common in the literature regardless of biological model, endpoint measured, and chemical class/physical stressor.[1-5] In addition, several large-scale direct comparisons of the threshold and the hormetic models using a priori entry and evaluative criteria have been published, indicating that the threshold model poorly predicts responses in the below-threshold zone, while the reverse is the case for hormesis.[2,3,6,7]

The present study evaluates the capacity of the threshold and hormesis concentration response models to predict below-threshold responses in a high-throughput bioassay in which 1888 prospective antibacterial compounds were tested for effects on growth in *Escherichia coli*. The design of the concentration responses provides a rich opportunity to evaluate effects above and below the toxic threshold, in that there were 11 concentrations spaced at twofold intervals and built-in replication procedures. We analyze these data to identify the pattern of responses at concentrations below the estimated threshold.

Materials and methods

The study data came from McMaster University, HTS Laboratory, Hamilton, Ontario, Canada. *E coli* strain

[1] Department of Public Health, Environmental Health Sciences, University of Massachusetts, Amherst, MA, USA
[2] Department of Biology, College of the Holy Cross, Worcester, MA, USA
[3] Department of Public Health, Biostatistics and Epidemiology, Arnold House, University of Massachusetts, Amherst, MA, USA
[4] Gradient, 20 University Road, Cambridge, MA, USA

Corresponding author:
Edward J. Calabrese, Department of Public Health, Environmental Health Sciences, Morrill I, University of Massachusetts, Amherst, MA 01003, USA.
Email: edwardc@schoolph.umass.edu

MC1061[8], containing plasmid pET26b(+), was used to screen for toxicity in 50,000 small molecules that had been acquired from Maybridge plc (Cornwall, UK). This strain has chromosomal and plasmid-borne genes that confer resistance to streptomycin and kanamycin, respectively, and it exhibits enhanced permeability to chemicals relative to that of wild-type *E coli*. The initial screen of 50,000 compounds involved testing at a single concentration (50 μM) to assess cell killing. Hits, which were defined in the McMaster study as those compounds that reduced bacterial growth to 75% as compared to solvent-treated controls, occurred with 1925 compounds. The OD_{600} at time zero was not measured for about 2% of the 1925 compounds, reducing the number of chemicals that were initially assessed in the present article to 1888.

Single colonies were picked from culture plates of *E coli* MC1061 pET26(+) and grown overnight in Luria-Bertani broth (3 mL) containing 50 μg/mL kanamycin (LB-kan), then diluted 100-fold into LB-kan, and grown at 37°C with shaking until OD_{600} reached ca. 0.5. This subculture was diluted 10^3-fold into screening broth (LB-kan containing 50 μg/mL streptomycin and 1 mM isopropyl-β-D-thiogalactopyranoside) and used for concentration-response determination. Concentration-response evaluations were carried out in duplicate in separate 96-well plates with 190 μL screening broth and 10 μL test compound dissolved in DMSO. Eleven concentrations (0.244 μM; 0.488 μM; 0.976 μM; 1.95 μM; 3.91 μM; 7.81 μM; 15.6 μM; 31.2 μM; 62.5 μM; 125 μM; and 250 μM) of each compound were evaluated. The solvent DMSO had a final concentration of 5% v/v. Solvent and positive controls (four of each) in each 96-well plate contained DMSO or ampicillin (100 μg/mL) and chloramphenicol (25 μg/mL) instead of tested compounds, respectively. All transfers were performed using a Biomex FX liquid handler (Beckman-Doulter Inc., Fullerton, CA). Upon mixing the bacteria with the tested compounds, the OD_{600} of each plate was measured with a SpectraMax Plus384 plate reader (Molecular Devices Corp., Sunnyvale, CA, USA; OD_{600}zero). Plates were then incubated at 37°C, 85% humidity, for ca. 20 hours, after which OD_{600} was measured again (OD_{600}growth). Evaluation of growth kinetics indicated that the bacteria were in log growth phase at 20 hours when the measurements were taken. The responses reported are the fraction of growth of *E coli* exposed to the compound relative to the growth of the solvent controls.

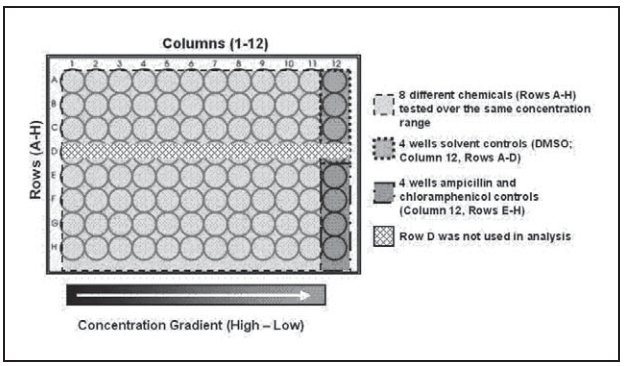


Figure 1. Layout of 96-well plates. Each chemical was tested at the same 11 concentrations on a different 96-well plate. The highest concentration (250 μM) was always placed in column 1 on the left edge of the well plate, with lower concentrations placed in the next well progressing to the right side. The column on the far right was used for controls. See the text for details.

Replication procedure

Each chemical was tested twice at the same 11 concentrations on a different 96-well plate on the same day. The study was conducted on 11 days between 20 January 2005 and 3 February 2005. The highest concentration (250 μM) was always placed in column 1 on the left edge of the well plate. Each progressively lower concentration was placed in the next well progressing to the right side. Eight different chemicals were tested over the same concentration range on one plate, leaving the eight wells adjacent to the low treatment group on the right edge (column 12) of the plate for controls (Figure 1). Of the eight control wells, four were allocated to solvent controls (i.e., DMSO) and four to controls using ampicillin and chloramphenicol. Since *E coli* MC1061 pET26(+) is sensitive to these agents, the concentration response was deemed invalid if growth occurred in their presence.

Optical density (OD) measurements were taken at time zero and after 20 hours of incubation. The time zero readings were subtracted from those made at 20 hours, correcting for possible absorbance by tested chemicals at 600 nm. The response data for each concentration consisted of a ratio of the OD_{600} of the well for the treatment, divided by the mean of the OD readings of three solvent controls in rows A–C. The control response in row D was not used due to an apparent calibration error in the OD reader for that row. Controls and treatments for row D were found to be systematically low by approximately 20% in the time 0- and time 20-hour readings, thereby supporting the decision to eliminate row-D chemicals and controls

from the analysis. As a result of this calibration error, all chemicals (i.e., 236) and controls tested in row D were dropped from the analysis.

Evaluation strategy

Individual concentration-response relationships were used to evaluate whether there is nonrandom biological activity, measured by cell proliferation, below the toxicological threshold as estimated by a benchmark dose (BMD) approach. In the 11-concentration protocol, the concentration-response study should ideally have at least one concentration in the toxic domain (i.e. above the BMD), a concentration with a response that approximates the control response, and several lower concentrations that would be evaluated for biological activity below this threshold-like value. The McMaster University data are useful for this purpose, in that most experiments show toxicity but also have multiple concentrations below an estimated toxicological threshold.

For a concentration-response to be included in the analysis, we required evidence of toxicity at high concentrations (i.e. response $\leq 80\%$ of the controls), along with one or more concentrations below the threshold concentration, subsequently defined as the benchmark dose BMD(5). In addition, concentration responses had to pass a screening for response outliers in the below-BMD(5) range based on the standard deviation of responses below the BMD(5).

Once eligible concentration responses were identified with one or more concentration below the BMD(5), we describe the response pattern in two ways. First, we calculated the mean response for concentrations below the BMD(5) for a concentration response and summarized these average responses over concentration responses. Our summary stratified the concentration responses according to the number of concentrations below the BMD(5).

The second analysis accounted for replicate concentration responses for a chemical, and plate effects using a mixed model analysis using SAS PROC Mixed. In the model, random effects were included for plates, chemicals, and replications. Fixed effects corresponded to the number of concentrations below the BMD(5), as well as concentration effects and their interaction. This analysis accounts separately for variation due to plates, chemicals, and replicates, and it more closely parallels the physical experimental structure. The p values for tests of hypotheses that the average response differs from 100% of control are based on the mixed model analyses.

Since the controls were located in column 12, an edge, the most direct comparison of response with controls is for chemicals that were also assigned to an edge (rows A and H; see Figure 1). We report results for these chemicals first, followed by results for chemicals assigned to 'internal' rows (i.e. rows B, C, E, F and G) versus the controls. Emphasizing edge rows in our analysis has the advantage of eliminating the possible influence of edge effects, that is, differences in results ascribable to positioning on edges or in interior wells. Edge effects have been reported to occur in high-throughput screening, but they are not consistent with respect to whether responses on edges are higher or lower than those of interior wells.[9,10] Because of the uncertainty, we thought it prudent to emphasize edge rows, thereby maximizing the equivalency between treatment groups and controls.

Threshold estimation strategies

A Benchmark Dose (BMD) methodology was adopted to estimate the toxicological threshold in a manner identical to that reported earlier.[7] The BMD concept, which was first proposed by Crump,[11] is a widely accepted general approach for estimating threshold responses in toxicology. In contrast to a linear or nonlinear regression approach to calculating a BMD,[11] the procedure that we used, as described below, is 'local' in the sense that the BMD is calculated using only the responses at concentrations that are adjacent to the BMD.

The toxicological threshold was estimated using a BMD(5) as the threshold estimate. This value represents the concentration at which the response is estimated to have decreased 5% below control values. This BMD was selected since it closely reflected the variability (SD) of the *E coli* controls on 96-well plates. Since our goal is to classify toxicity, a lower bound of a confidence interval for that concentration was not calculated as may be performed in a risk assessment process. The BMD approximates the control response but probably includes a low degree of toxicity. That is, it corresponds to a concentration that is slightly higher than the toxicological threshold. This suggests that a concentration immediately below and very close to the BMD may itself be within the toxic zone (i.e. its concentration may be higher than the actual toxicological threshold). This would become less likely with increasing distance between

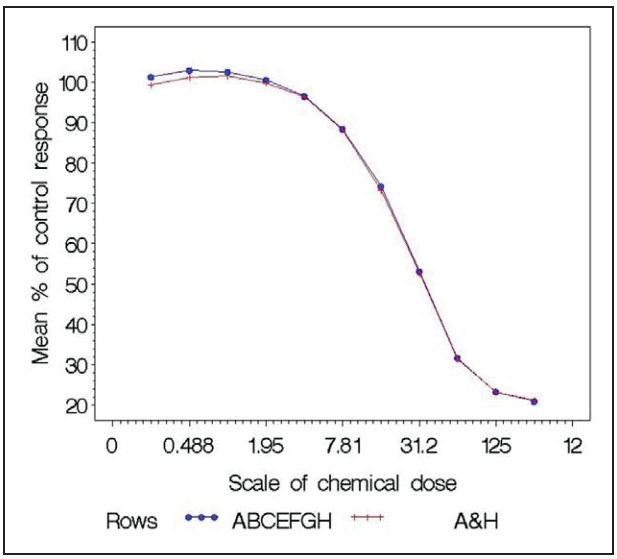


Figure 2. Mean responses (% of control growth) by dose in the *Escherichia coli* database: The curves compare the dose-response relationship for edge rows (rows A and H) with that for all evaluated rows (rows A–C and E–H).

the BMD and the concentration below the BMD. For example, for agents with a BMD near 0.488 μM, the 0.244 μM concentration would be close to the toxicity threshold. In contrast, for agents with a BMD approaching 31 μM, the 0.244 μM concentration would be approximately two orders of magnitude below the toxicity threshold.

BMD values were estimated for concentration responses where the response was less than 80% of control at a higher concentration, which we call a 'toxic concentration.' Using concentrations lower than the toxic concentration, the BMD(5) was determined by the following procedure:

1. The largest concentration with a response below 95.0% is identified. Let this concentration be C_{below}, and let the associated response be R_{below}.
2. If the response at the next smallest concentration is at least 95.0%, then let this concentration be C_{above}, and let the associated response be R_{above}. The BMD(5) is determined to be in the concentration range between C_{above} and C_{below}.
3. If the average response at the next lowest concentration below C_{below} is less than 95.0%, then let this concentration be C_{below} with response R_{below}, and return to step two.

The concentrations defining the BMD(5) range were not included in analyses of responses below the BMD(5).

Results

Figure 2 shows concentration-response relationships for the data satisfying the entry criteria. Two curves are presented, one for all analyzed rows (A–C and E–H) and the other for edge rows A and H. The concentration-response relationships are similar, with slightly smaller responses at low concentrations in the edge rows. Analysis of the edge rows (i.e. A and H) and internal rows (i.e. B, C, E, F, and G) follows in separate sections.

Edge row chemical assessment

We first report results for chemicals where the concentration response occurred on the edge of a plate (row A or H). A total of 471 plates were eligible for analysis. Two plates were analyzed for each of 235 sets of eight chemicals, and one set of chemicals occurred on a single plate. This resulted in a total of 472 chemicals whose concentration responses were on an edge of a plate, of which replicate concentration responses were available for 470 chemicals. The total number of concentration responses for these chemicals was 942. Among the 942 concentration responses, 36 did not satisfy the toxicity ($\leq$80% of control at high concentration) requirement. Of the remaining 906 concentration responses, 51 were highly toxic, with no concentrations below the BMD(5) range. A total of 855 concentration responses had one or more concentration below the BMD(5), including replicate concentration responses on 411 chemicals and a single concentration response on 33 chemicals.

We examined concentration responses for possible outliers by calculating the standard deviation of response among concentrations below the BMD(5) for the 827 concentration responses with more than one concentration below the BMD(5) and reviewing the entire response profile for the 42 concentration responses where the standard deviation was greater than 8.85 (corresponding to the 95th percentile). The median standard deviation was 2.39, while the 75th percentile standard deviation was 3.83. Among the 42 concentration responses with high standard deviations, 40 concentration responses were for chemicals where a replicate concentration response had a standard deviation below the 95th percentile. Two concentration responses were for the same chemical, with the high standard deviation apparently due to precipitation at high concentrations. When replicate concentration responses were available for a chemical, profiles of both concentration responses were

Table 1. Edge row analysis[a]

	Number of concentrations below BMD(5)								
	1	2	3	4	5	6	7	8	9
Median	97.68	99.03	99.72	101.09	102.00	101.31	101.96	100.78	100.33
Mean	97.59	98.70	99.64	101.08	101.90	101.60	103.50	101.46	100.33
Standard deviation.	3.85	3.32	3.77	3.45	3.65	3.89	5.02	4.62	3.34
Maximum	104.58	108.87	107.62	114.42	113.88	113.23	117.85	106.38	102.69
# Concentration responses	28	78	135	188	224	142	25	3	2

[a] Summary of median/mean percentage of control response by number of concentrations below the BMD(5) averaged over concentration responses for 825 concentration responses on 441 chemicals.

reviewed to identify outlier responses for individual wells. Based on this review, we decided to eliminate concentration responses with standard deviations greater than 10 since there was evidence of one or more outlier response. This excluded 30 concentration responses (3.5% of the 855). The remaining 825 concentration responses included replicate concentration responses on 384 chemicals and a single concentration response on 57 chemicals. This corresponds to 87.5% of the concentration responses conducted in an edge row, and it includes one or more concentration response on 93.4% of the chemicals assigned to a plate edge.

Assessment of edge-row replication consistency

Of the edge chemicals, 384 had replicates with at least one concentration response that satisfied the toxicity requirement, had a local BMD(5) and at least one concentration below the BMD(5), while 57 chemicals had only one replicate meeting these criteria. We compared the BMD(5) range for the 384 chemicals with replicates to assess consistency. Chemicals were judged as to how close each chemical's BMD(5) was to its replicate's BMD(5) based on whether they were within the same or different twofold concentration increment. Of the 384 chemicals, 198 (52%) had BMD(5) values that were within the same twofold concentration range. An additional 146 chemicals (38%) were within the next twofold concentration increment, while 97% were within two, twofold concentration increments. We inspected the 11 replicated concentration responses where the BMD(5) range differed by more than two adjacent concentrations and concluded that the differences were a result of normal variability that was a consequence of relatively low toxicity. For this reason, all 825 concentration responses were included in subsequent analyses.

Edge rows: concentration response analysis

We summarize the average response below the BMD(5) for concentration responses (Table 1) and chemicals (Table 2, Figure 3) depending on the number of responses below the BMD(5). The data indicate a general inverted U-shaped concentration response, which becomes progressively more prominent when there are 3–8 concentrations below the BMD(5) category. The results in Table 1 provide a simple summary using a concentration response as a unit. This analysis does not account for the fact that some of the concentration responses are replicates on the same chemical, but it allows each concentration response to be uniquely assigned to a column corresponding to the number of concentrations below the BMD(5). The results in Table 2 are based on a mixed model analysis that accounts for variation between plates, chemicals, and replications estimated as 6.57, 4.34, and 11.71, respectively. The estimated means in Table 2 show the pattern of response for concentrations below the BMD(5) depending on the number of concentrations below the BMD(5), with means that are statistically significantly different from the control 100% so indicated. As seen in Table 2, the first concentration below the BMD(5) tends to have an average response that is slightly less than that of lower doses, suggesting that the BMD(5) was likely to be slightly below the true threshold (i.e. displaying slight toxicity). At progressively lower concentrations, there is evidence of an increase in the proliferation response and then a decrease toward control values, conforming to the inverted U-shape. Figure 3 summarizes the estimates from Table 2 when there were more than 50 concentration responses with the given number of concentrations below the BMD(5).

Table 2. Edge row analysis[a]

	Concentrations (μM)								
# of Conc. Below BMD (5)	0.244	0.488	0.976	1.95	3.91	7.81	15.6	31.2	62.5
1	98.41								
2	98.75	98.76[b]							
3	98.90	99.61	100.12						
4	100.07	101.45	101.79	100.79[b]					
5	99.78	101.78	102.96	102.44[b]	101.48[b]				
6	99.64	101.58	103.18	102.62[b]	103.13[b]	101.78[b]			
7	100.73	101.59	102.82	103.68[b]	104.00[b]	103.86[b]	101.35		
8	95.64	101.63	108.09	104.57	107.17[b]	103.72	102.13	98.33	
9	98.54	105.27	101.33	103.38	104.86	103.96	97.74	96.54	94.73

[a] Estimated mean percentage of control response by number of concentrations below the BMD(5) based on mixed model analysis with random plate, chemical, and replications for 467 plates, 441 chemicals, and 825 concentration responses.
[b] p Value < .05.

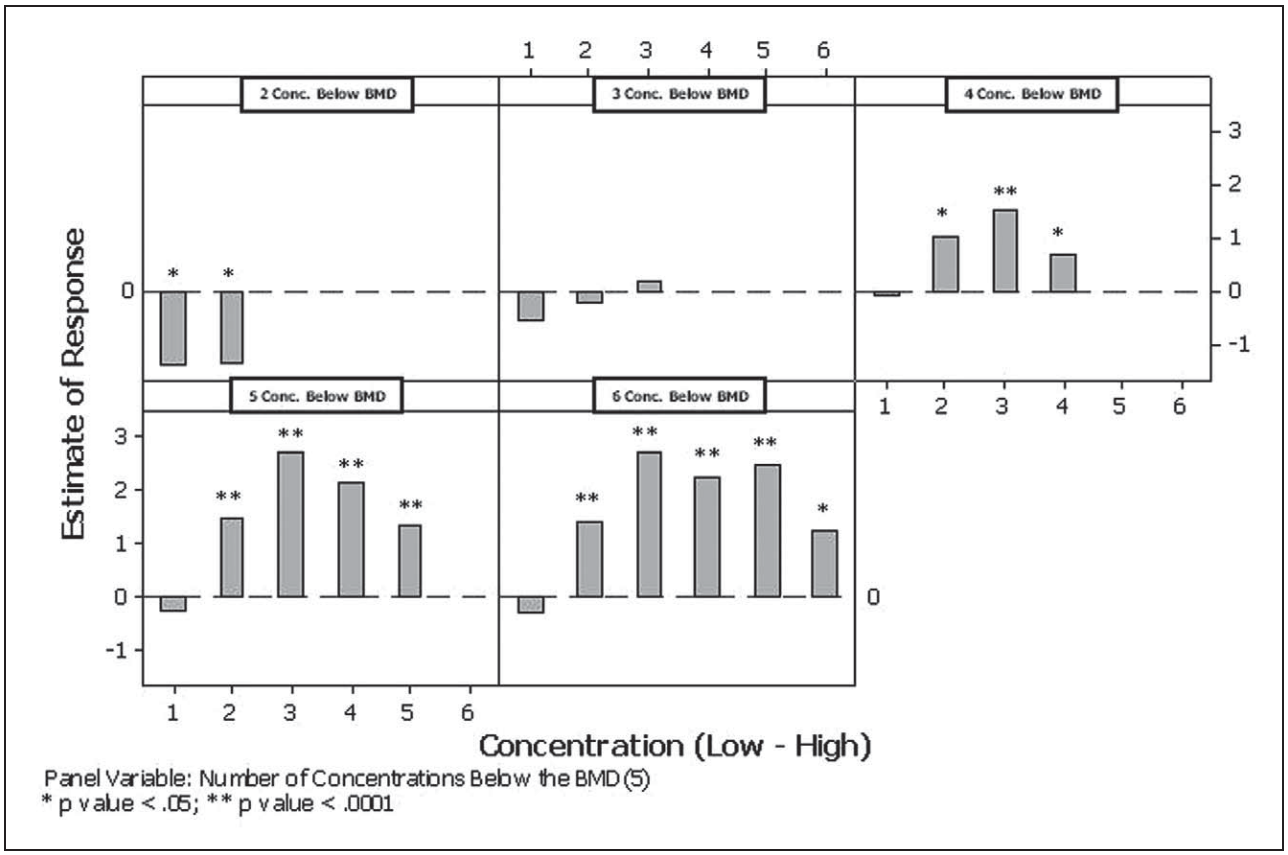


Figure 3. Edge row mixed model analysis. Estimates of the average response [y-axis; expressed as percentage deviation from the control (100%)] over the concentration of antibacterial agent (x-axis; see Table 2) and number of concentrations below the BMD(5) (panel variable). The figure is based on chemicals in an edge row (A and H) of a 96-well plate (with controls placed in column 1, rows A–C). Data are only shown (in panels) when there were >50 concentration responses in the respective category (see number of concentration responses in Table 1).

Table 3. Internal row analysis[a]

	Number of concentrations below BMD(5)								
	1	2	3	4	5	6	7	8	9
Median	98.11	99.68	103.08	103.94	105.16	106.27	106.72	104.05	97.52
Mean	98.56	100.27	103.10	103.43	104.83	105.84	106.61	106.52	96.91
Std. Dev.	9.21	5.99	5.64	5.74	5.43	6.06	5.04	5.33	5.77
Maximum	146.40	126.00	121.32	116.25	123.66	122.35	119.53	115.78	105.31
# Conc. Responses	72	140	245	407	652	465	104	13	7

[a] Summary of median/mean percentage of control response by number of concentrations below the BMD(5) averaged over concentration responses for 2105 concentration responses on 1135 chemicals.

Internal row chemical assessment

There were 1180 chemicals assayed on 471 plates in internal rows (rows B, C, E, F, and G), of which replicate assays were available for 1125 chemicals. The total number of concentration responses for these chemicals was 2355. Among the 2355 concentration responses, 60 did not satisfy the toxicity ($\leq$80% of control at high concentration) requirement. Of the remaining 2295 concentration responses, 110 were highly toxic, with no concentrations below the BMD(5) range. A total of 2185 concentration responses had one or more concentration below the BMD(5), including replicate concentration responses on 1031 chemicals, a single concentration response on 111 chemicals, and four concentration responses on three chemicals.

We examined concentration responses for possible outliers by calculating the standard deviation of response among concentrations below the BMD(5) for the 2185 concentration responses with more than one concentration below the BMD(5) and reviewing the entire response profile for the 106 concentration responses where the standard deviation was greater than 11.65 (corresponding to the 95th percentile). The median standard deviation was 3.30, while the 75th percentile standard deviation was 5.36. Among the 106 concentration responses with high standard deviations, 94 chemicals had a single concentration response with a high standard deviation, while six chemicals had two concentration responses with standard deviations above the 95th percentile. When replicate concentration responses were available for a chemical, profiles of both concentration responses were reviewed to identify outlier responses for individual wells. Based on this review, we decided to eliminate concentration responses with standard deviations greater than 13 since there was evidence of one or more outlier response. This excluded 80 concentration responses (3.7% of the 2185). The remaining 2105 concentration responses included replicate concentration responses on 961 chemicals, a single concentration response on 171 chemicals, and three chemicals with four concentration responses each. This corresponds to 89.4% of the concentration responses conducted in internal rows, and it includes one or more concentration response on 96.2% of the chemicals assigned to these internal rows.

Assessment of replication consistency for internal rows

Of chemicals in internal rows, 961 had two replicates with at least one concentration response that satisfied the toxicity requirement, had a local BMD(5), and had at least one concentration below the BMD(5), while 171 chemicals had only one replicate meeting these criteria and three chemicals had four replicates. We compared the BMD(5) range for the 961 chemicals with replicates to assess consistency. Chemicals were judged as to how close each chemical's BMD(5) was to its replicate's BMD(5) based on whether they were within the same or different twofold concentration increment. Of the 961 chemicals, 433 (45%) had BMD(5) values that were within the same twofold concentration range. An additional 367 chemicals (38%) were within the next twofold concentration increment, while 94% were within two of the twofold concentration increments. We inspected the 28 replicated concentration responses where the BMD(5) range differed by more than three adjacent concentrations and concluded that the differences were a result of normal variability that was a consequence of relatively low toxicity. For this reason, all 2105 concentration responses were included in subsequent analyses.

Internal rows: concentration response analysis

We summarize the average response below the BMD(5) for concentration responses (Table 3) and

Table 4. Internal row analysis[a]

	Concentrations (μM)								
# of conc. below BMD(5)	0.244	0.488	0.976	1.95	3.91	7.81	15.6	31.2	62.5
1	101.50[b]								
2	102.47[b]	102.94[b]							
3	101.99[b]	104.31[b]	103.36[b]						
4	102.65[b]	104.54[b]	103.96[b]	102.39					
5	102.62[b]	105.31[b]	105.44[b]	104.38	102.56				
6	103.21[b]	105.51[b]	106.25[b]	105.92	105.09	102.99			
7	102.68[b]	106.16[b]	107.28[b]	106.91	106.48	105.15	103.08[b]		
8	103.38[b]	106.70[b]	106.87[b]	108.14	107.54	106.68	102.86[b]	96.41[b]	
9	101.74	97.26	103.13	100.64	103.60	104.31	97.93	86.05[b]	91.03[b]

[a] Estimated mean % of control response by number of concentrations below the BMD(5) based on mixed model analysis with random plate, chemical, and replications for 471 plates, 1135 chemicals, and 2105 concentration responses.
[b] p Value < .05.

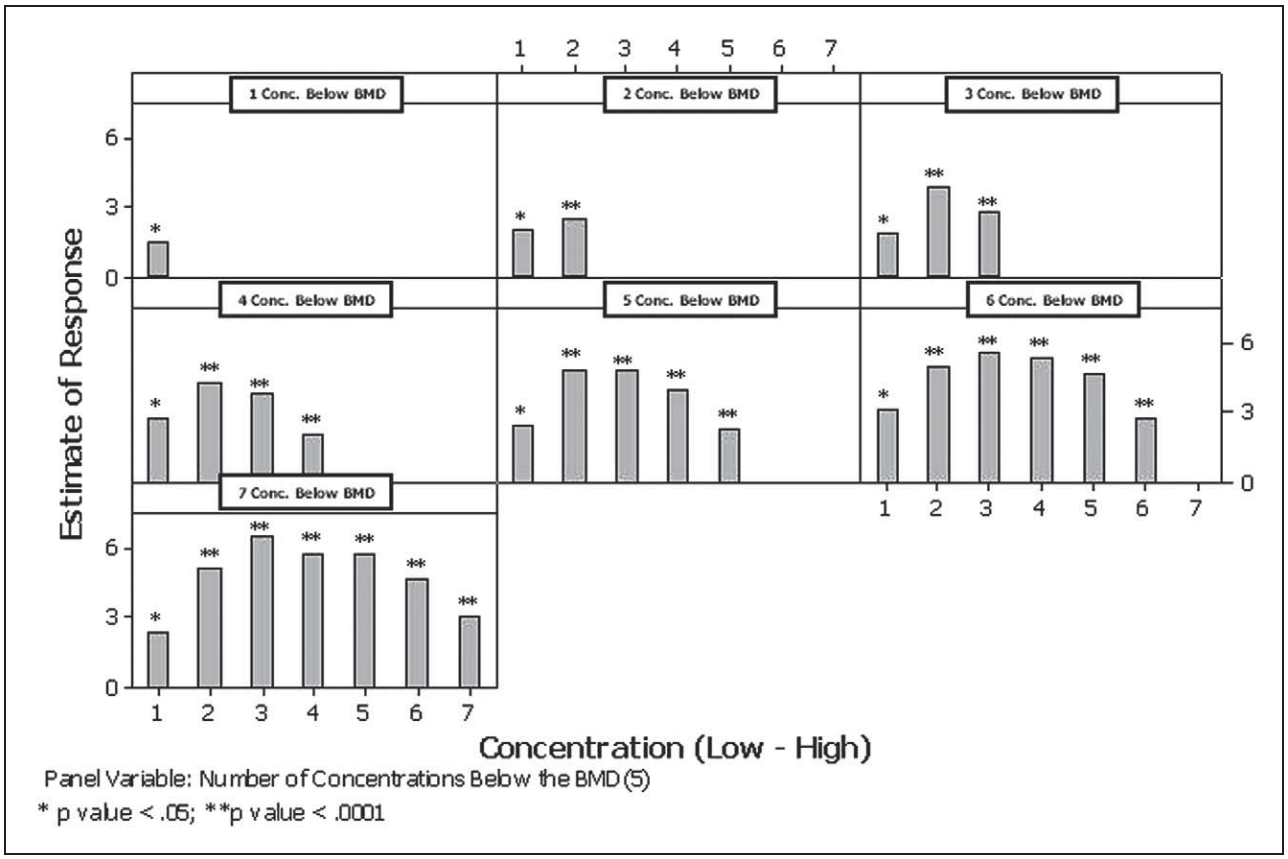


Figure 4. Internal row mixed model analysis. Estimates of the average response [*y*-axis; expressed as percentage deviation from the control (100%)] over the concentration of antibacterial agent (*x*-axis; see Table 4) and number of concentrations below the BMD(5) (panel variable). The figure is based on chemicals in interior rows (B, C, E, F, and G) of a 96-well plate (with controls placed in column 1, rows A–C). Data are only shown (in panels) when there were >50 concentration responses in the respective category (see number of concentration responses in Table 3).

chemicals (Table 4, Figure 4) depending on the number of responses below the BMD(5). The data indicate a general inverted U-shaped concentration response that becomes progressively more prominent when there are 3–8 concentrations below the BMD(5) category. The results in Table 3 provide a simple

summary using a concentration response as a unit. This analysis does not account for the fact that some of the concentration responses are replicates on the same chemical, but it allows each concentration response to be uniquely assigned to a column corresponding to the number of concentrations below the BMD(5). The results in Table 4 are based on a mixed model analysis that accounts for variation between plates, chemicals, and replications estimated as 26.08, 4.92, and 21.06, respectively. The estimated means in Table 4 shows the pattern of response for concentrations below the BMD(5) depending on the number of concentrations below the BMD(5), with means that are statistically significantly different from 100% so indicated. The responses in Table 4 conform to an inverted-U curve, as described above for Table 2. Figure 4 summarizes the estimates from Table 4 when there were more than 50 concentration responses with the given number of concentrations below the BMD(5).

Discussion

These results extend the findings of Calabrese and Baldwin[2,3] and Calabrese et al.[6,7] that the hormetic concentration response is commonly observed in studies with rigorously defined a priori entry and evaluative criteria. The average stimulatory response was highly statistically significant but modest, usually being about 1%–4% above the controls in the analysis of edge rows (Tables 1 and 2; Figure 3). These values were somewhat lower than the stimulatory profile seen with yeast,[7] but both data sets support the hormesis model.

While low-concentration stimulation was a common response for edges and internal rows, the internal rows of replication 2 showed a greater difference between treatment wells and controls than did similar rows with the same chemicals in replication 1. Such differences between replications 1 and 2 were not seen for the edge comparisons. It is possible that an unknown methodological factor may have caused the larger difference in replication 2. Because of the positional correspondence of the edges to the controls and the greater consistency of the stimulatory response between replicates, we place greater confidence in the edge rows than the interior rows. We note, however, that both support the occurrence of hormesis.

The findings support the conclusion that the threshold concentration-response model does not accurately predict bacterial growth responses to chemicals at concentrations below the toxicological threshold, and they are consistent with another large-scale study that was based on 57,000 concentration response relationships for 2200 chemicals in 13 yeast strains.[7] That study used BMD criteria (2.5, 5.0, 7.5, 1.0, and 12.5) and a NOEL criterion. Previous publications applying statistical evaluative criteria to published findings in the toxicological and pharmacological literature also support the hormesis model.[12,13] Analyses of published literature[1-4,14] show that hormesis is a commonly observed concentration response, broadly applicable across biological models, end points measured, and chemical classes. The finding of hormesis in the *E coli* data set is consistent with thousands of examples of hormetic concentration responses in published literature and a large hormesis database.[15,16]

The McMaster bacterial database is notable for the large number of chemicals tested, the use of 11 concentrations with twofold concentration spacing, low variability in the control (2%–3% SD), a repeated measure of all wells at time zero and 20 hours, and built-in replication. The large number of closely spaced concentrations represents an advantage over the yeast study,[7] which has only five concentrations per replication. The low degree of variability of the control responses is an important factor, especially in light of the use of a study design that included all concentrations and controls on the same 96-well plate. The low variability minimizes the possibility of positive bias that might otherwise occur. In the yeast data set, this possible source of bias was eliminated because each concentration of a concentration response was on a separate plate with separate plate–specific controls.[7] The consistency of the yeast and bacterial findings add to the weight of evidence for hormesis.

Since column 12 controls would have one control well with a ‘double’ edge, something not present in chemical-treatment wells at low concentrations, we tested this well against the two ‘single’-edged control wells. There was no significant difference between the ‘double’ edged control and the two ‘single’ edged controls.

While our analysis was designed to compare the below-threshold predictive capacity of the threshold and hormetic concentration response models, these findings also may have implications for drug development, especially in the case of antibacterial agents. While antibacterials are usually administered with the intention of killing or inhibiting bacteria, the data indicate that a substantial proportion of these

compounds may stimulate *E coli* proliferation in the low-concentration zone. That antibacterials can stimulate bacterial growth under some conditions has been known since the mid decades of the 20th century.[17-23] Although there is growing recent awareness of the implications of this phenomenon,[24,25] its clinical implications and relevance for drug development deserve greater consideration.

In conclusion, our findings document the inability of the threshold-concentration response model to accurately predict responses in the low-concentration zone in a large-scale high-throughput study in *E coli*. They also support the predictions of the hormesis concentration-response model. The large database with nearly 2000 agents, favorable study-design features, and general consistency of replicated responses add weight to the findings. Likewise, the consistency of the *E coli* findings with those of a complementary analysis in yeast[7] challenges the continued use of the threshold model in study design and in predicting effects in the low-concentration zone. The findings suggest that the hormetic concentration-response model provides more reliable estimates of responses in the low-concentration zone than does the threshold model.

Acknowledgements

We are grateful to Eric Brown, McMaster University, for making the *Escherichia coli* database available for our use. We also thank Eric Brown, Jan Blanchard, and John Staudenmayer for valuable discussions.

Funding

This research was sponsored by the Air Force Office of Scientific Research, Air Force Material Command, United States Air Force (USAF).

References

1. Calabrese EJ. Historical blunders: how toxicology got the dose-response relationship half right. *Cell Mol Biol* 2005a; 51: 643–654.
2. Calabrese EJ, Baldwin LA. The hormetic dose response model is more common than the threshold model in toxicology. *Toxicol Sci* 2003a; 71: 246–250.
3. Calabrese EJ, Baldwin LA. The frequency of U-shaped dose-responses in the toxicological literature. *Toxicol Sci* 2001a; 62: 330–338.
4. Calabrese EJ. Paradigm lost, paradigm found: the re-emergence of hormesis as a fundamental dose response model in the toxicological sciences. *Environ Pollut* 2005b; 138: 379–411.
5. Calabrese EJ. Hormesis: from marginalization to mainstream. A case for hormesis as the default dose-response model in risk assessment. *Toxicol Appl Pharmacol* 2004; 197: 125–136.
6. Calabrese EJ, Stanek EJ, Nascarella MA, Hoffmann GR. Hormesis predicts low-dose responses better than threshold models. *Int J Toxicol* 2008; 27: 369–378.
7. Calabrese EJ, Staudenmayer JW, Stanek EJ, Hoffmann GR. Hormesis outperforms threshold model in NCI anti-tumor drug screening data. *Toxicol Sci* 2006; 94: 368–378.
8. Casadaban MJ, Cohen SN. Analysis of gene-control signals by DNA-fusion and cloning in *Escherichia-coli. Mol Biol* 1980; 138: 179–207.
9. Faessel HM, Lavasseur LM, Slocum HK, Greco WR. Parabolic growth patterns in 96-well plate cell growth experiments. *Vitro Anim Cell Dev Biol* 1999; 35: 270–278.
10. Brinkmann M, Lütkemeyer D, Gudermann F, Lehmann J. New technologies for automated cell counting based on optical image analysis 'The Cellscreen.' *Cytotechnology* 2002; 38: 119–127.
11. Crump KS. A new method for determining allowable daily intakes. *Fund Appl Toxicol* 1984; 4: 854–871.
12. Calabrese EJ, Baldwin LA. Special issue: hormesis: environmental and biomedical perspectives. *Crit Rev Toxicol* 2003b; 33: 213–424.
13. Calabrese EJ, Baldwin LA. Special issue: scientific foundations of hormesis. *Crit Rev Toxicol* 2001b; 31: 351–695.
14. Calabrese EJ. Hormesis: why it is important to toxicology and toxicologists. *Environ Toxicol Chem* 2008; 27: 1451–1474.
15. Calabrese EJ, Blain R. Hormesis and plant biology. *Environ. Pollut* 2008; 157: 42–48.
16. Calabrese EJ, Blain R. The occurrence of hormetic dose responses in the toxicological literature, the hormesis database: an overview. *Toxicol Appl Pharmacol* 2005; 202: 289–301.
17. Smith MI. An investigation into the chemotherapy of the acridine dyes in experimental tuberculosis. *J Pharmacol Exper Ther* 1923; 20: 419–434.
18. Marshall MS, Hrenoff AK. Bacteriostasis. *J Infect Dis* 1937; 61: 42–54.
19. Salle AJ. *Fundamental principles of bacteriology*. New York: McGraw-Hill, 1939.
20. Welch H, Price CW, Randall WA. Increase in fatality rate of E-typhosa for white mice by streptomycin. *Am Pharmacol Assoc* 1946; 35: 155–158.
21. Randall WA, Price CW, Welch H. Demonstration of hormesis (increase in fatality rate) by penicillin. *Am J Public Health* 1947; 37: 421–425.

22. Clifton CE. *Introduction to bacterial physiology*. New York: McGraw-Hill, 1957.
23. Lamanna C, Mallette MF. *Basic bacteriology its biological and chemical background*. 3rd ed. Baltimore, MD: The Williams & Wilkins Company, 1965.
24. Davies J, Spiegelman GB, Yim G. The world of subinhibitory antibiotic concentrations. *Curr Opin Microbiol* 2006; 9: 445–453.
25. Linares JF, Gustafsson I, Baquero F, Martinez JL. Antibiotics as intermicrobial signaling agents instead of weapons. *Proc Nat Acad Sci USA* 2006; 103: 19484–19489.

Human & Experimental Toxicology (2000) 19, 2–31

www.nature.com/het

Chemical hormesis: its historical foundations as a biological hypothesis

EJ Calabrese*,[1] and LA Baldwin[1]

[1]*Department of Environmental Health Sciences, School of Public Health, University of Massachusetts, Amherst, MA 01003, USA*

Despite the long history of hormesis-related experimental research no systematic effort to describe its early history has been undertaken. The present paper attempts to reconstruct and assess the early history of such research and to evaluate how advances in related scientific fields affected the course of hormesis-related research. The purpose of this paper is not only to satisfy this gap in current knowledge, but also to provide a foundation for the assessment of how the concept of hormetic dose-response relationships may have affected the nature of the bioassay especially with respect to hazard assessment practices within a modern risk assessment framework.
Human & Experimental Toxicology (2000) **19**, 2–31

Keywords: hormesis; low dose; stimulation; β-curve

Introduction

The belief that low doses of poisons or toxic substances could stimulate biological systems has a long history with scattered reports in the early published literature. Several of the earlier research papers that preceded the work of Schulz[1,2] are: (1) Reveill[3] found that sodium hypochlorite in a solution of 0.1% stimulated germination and growth, but was harmful to plants when applied in higher concentrations; (2) Raulin[4] reported that the sulfates of zinc and iron markedly stimulated the growth of the fungus *Aspergillus*; (3) Storp[5] obtained a stimulating effect upon the germination of seeds by immersing them in a 0.01% solution of sodium chloride. The individuals usually viewed as the founders of modern hormesis (i.e., the concept of a low-dose stimulation/high-dose inhibition) are Rudolph Arndt and Hugo Schulz. Schulz was one of the first to report the phenomenon of low dose stimulatory responses by poisonous substances based on experimental work at the University of Greifswald assessing the effects of numerous chemicals on yeast fermentation as measured by the liberation of CO_2.[1,2] However, it was not the outstanding scientific achievement of his research that brought recognition (and notoriety) to Schulz but the claim that the findings were universally generalizable to all organisms and toxic agents with particular emphasis on its medical applications for treating patients within the homeopathic tradition. This limited, though experimentally based perspective of Schulz, soon became merged with similar views of the homeopathic physician Arndt which thereby lead to the Arndt-Schulz Law.

The Arndt-Schulz Law, as will be shown, became the object of a number of scientific investigations to not only replicate some of the initial experimental data upon which it was based, but to also refine the extent to which the concept was generalizable. It was in this latter domain that some aspects of the original assertion were most successfully challenged, especially by its supporters.[6,7] Nonetheless, major support of the biological principle of Schulz was provided in 1896 by Ferdinand Hueppe,[8] a very distinguished bacteriologist who had studied under the world-renowned and Nobel laureate Robert Koch, in his major bacteriological text book. Hueppe, in fact, claimed to have independently discovered the same principle in his bacteriological studies but clearly provided credit to Schulz for its first reporting. However, Hueppe had such a strong scientific reputation as a result of his numerous original contributions and substantial textbook writings that the concept of low dose stimulation by poisons soon became known as Hueppe's Rule rather than (or in addition to) the Arndt-Schulz Law. In fact, the term 'Hueppe's Rule' was adopted earlier in the broader international literature than that of the 'Arndt-Schulz Law' which seems to have received more provincial recognition. For example, Suzuki,[9,10] researching in Japan, cited Hueppe's 'biological law' to indicate that 'poisonous compounds can in very high dilution often produce a

*Correspondence: EJ Calabrese
Received 15 October 1999; accepted 15 October 1999

stimulating action'. A similar recognition occurred for Hueppe in the 1903 and 1905 publications in the US of Copeland[11] and Latham,[12] respectively. None of these papers noted the earlier work of Schulz or the Arndt-Schulz Law.

It is interesting to consider how the more modern toxicological texts have dealt with this issue of early recognition. For example, in his well known text, *Toxicology of Pesticides*, Hayes[13] states that Schulz[2] may have been the first to report the occurrence of stimulation from very low concentrations of toxic substances based on research with mercuric chloride, iodine, bromine, arsenious acid, chromic acid, formic acid and salicylic acid on yeasts. However, as Hayes noted, only a few years later, the bacteriologist Hueppe[8] confirmed these findings emphasizing their ability to be generalized with a 'rule that has come to bear his name'. In stating this basic toxicological principle, Hueppe also acknowledged certain limitations or exceptions thereby suggesting something less than the universality of the Arndt-Schulz proclamation. As suggested above, the concept of the low dose stimulation as advocated by Arndt and Schulz came to be seen as a scientific pillar of the controversial practice of homeopathy. In fact, Hueppe[8] emphasized that the work of Schulz on low dose stimulation stood on its own as a solid scientific foundation and should not be rejected because some homeopathic practitioners had adopted its premise to support their practices.

Despite the long history of hormesis-related experimental research there has been no systematic attempt to describe its early history including how and why it developed and to evaluate how advances in related scientific fields affected the course of hormesis-related research. The present paper attempts to reconstruct and assess the early history of such research not only with respect to satisfying this gap in current knowledge but to also provide a foundation for the assessment of how the concept of hormetic dose-response relationships may have affected the nature of the bioassay especially with respect to hazard assessment practices within a modern risk assessment framework.

The approach selected here for the historical assessment of hormesis considered how this concept evolved in separate yet broad classes of biological models from fields such as plant biology, bacteriology and mycology including that of yeast biology. When there was cross-fertilization between these scientific disciplines, such information has been incorporated into the appropriate section.

This paper addresses the period of the initial unfolding of research in this area (i.e., 1880s) to the 1930s, the time when significant conceptualization and an application of bioassay methods were being established in the various scientific disciplines. This temporal foundation will provide a framework to assess how the concept and interpretation of the modern bioassay considered the knowledge of hormetic responses within the context of the dose-response continuum.

Plants and hormesis

Introduction

The relationship of plant research to hormesis originates from a wide range of field and laboratory investigations. However, common to all cases is that the possibility of enhanced plant growth by low doses of toxic agents was an unexpected occurrence which engendered considerable subsequent comment and in many cases serious follow-up investigations. The principal areas where such observations emerged included the following: (1) the estimation of toxicity thresholds similar to what is currently defined within the context of a hazard assessment; (2) the evaluation of how plants respond to physical and chemical stressor agents of a limited nature; (3) the ability to differentiate essential nutrient functions from the capacity of non-nutritive agents to enhance growth and other metabolic functions; and (4) the claims that fungicidal and insecticidal treatments had a direct stimulatory impact on plant growth separate from their pesticidal actions.

The subsequent section will explore the historical foundations of the relationship of plant biology to the concept of hormesis. Each of the above four independent origins of the concept of low dose stimulatory plant responses will be initially discussed as a distinct phenomenon and then in relationship to overall development of the concept of hormesis.

Plant toxicology

The area of plant research with the most significant impact on the development of hormesis as a serious biological hypothesis was that of plant toxicology as seen within the context of the emerging field of physical chemistry and its subspeciality of electrochemistry. In the 1890s this field was being transformed by exciting developments with respect to enhanced understandings of the concept of molecular dissociation of inorganic acids and salts and their application to assessing the biological effects of highly dilute solutions. This area was particularly significant on two levels.

First, it originated directly from the theoretical work of Jacobus van't Hoff, Svante Arrhenius and Wilhelm Ostwald, Nobel prize winners in chemistry in 1901, 1905 and 1909, respectively. Second, their efforts had developed a highly credible theory of solution based on analogy between gases and dilute solutions that permitted investigators for the first time to utilize molar rather than percentage based solutions. The number of molecules in their solutions was emphasized which standardized comparisons in a more appropriate manner.[14]

The initial plant experiments in this area of the biological effects of highly dilute solutions were undertaken by Louis Kahlenberg at the University of Wisconsin; Kahlenberg had received his PhD (*summa cum laude*) at the University of Leipzig in Germany under the direction of Ostwald.[14] The 1896 publication by Kahlenberg and True was in the *Journal of Physcial Chemistry*, the first journal on this subject outside Germany. The publication which started in 1896 was initiated at Cornell University under the leadership of William D Bancroft, a professor of physical chemistry. Bancroft, like Kahlenberg, studied under Ostwald at Leipsig from 1890–1892 then returning to faculty positions at Harvard (his Alma Mater) and then to Cornell University. Ostwald's students believed that problems in other scientific disciplines might be more successfully clarified and advanced when approached from a physico-chemical perspective. This was particularly the case for areas such as physiology, toxicology, and medicine. In fact, Kahlenberg devoted several years to assessing the applications of electrolytic dissociation to solutions with respect to toxicological effects of dissolved salts on bacteria, seedlings, and plants.[14]

Kahlenberg and True[15,16] set forth to assess the role of various cations and anions and their respective concentrations on plant toxicity outcomes. The goal of their research was to determine the concentration (i.e., toxicology threshold) at which each solution allowed the plant root (i.e., radicle) to live. The plant *Lupinus albus* L. was selected as the biological model because of its straight, clear radicle, the ease of obtaining uniform specimens and its high sensitivity to changes in solution characteristics (e.g., turgor pressure). The basis for judging whether the plant was alive was to obtain measurements of the radicle growth rate. The sample size per concentration for these initial studies was meager by today's standards usually involving only two seedlings. In fact, at times only one seedling was employed in a concentration. Control groups were evaluated but these were not conducted for each experiment. When the authors were uncertain of their findings the experiment was repeated. The experiments typically utilized three concentrations: (1/3 200, 1/6 400, and 1/12 900 grams of hydrogen ions per liter).

Kahlenberg and True[16] reported the findings of 74 different experiments with a wide range of inorganic and organic acids. In general, toxicity was caused principally from the actions of the hydrogen ions while the anions displayed little or no toxicity. In the cases of copper, iron, nickel, cobalt, cadmium, mercury, silver, and cyanide salts the concentrations were much broader with up to seven doses employed with dilutions between 50 000–100 000th of a gram molecular weight.

Despite the obvious and important design limitations (e.g., the lack of concurrent controls for each experiment and the small sample size) of the initial work of Kahlenberg and True,[16] their efforts were ground breaking since they seized upon the biological implications of a major advance in theoretical and applied physical chemistry. To some extent this limitation is offset by the wide range of doses often employed, and the retesting of questionable experimental findings. Statistical methods were not generally available since the chi-square test, the *t*-test of Student and analysis of variance were not developed until 1900, 1908, and 1918, respectively.[17] Despite the noted weakness in these studies, the findings of Kahlenberg and True[15] were considered to have provided reliable and reproducible findings. They sought to determine and did gain insight into the potential toxicity of cations, anions and undissociated molecules. Likewise, they began the development of dose ranging studies in hydroponics for a large number of agents. Their work set the stage for other investigators from their institution and elsewhere to standardize doses based on molar concentrations rather than on a percentage basis as had been typical at that time. It should be noted that in their initial work with approximately 70 agents there was no case where a stimulatory response was noted. The lack of a stimulatory response in these numerous experiments was not, in retrospect, unexpected, given that the intent of the authors was to determine the concentration at which the solutions first permitted the radicle to survive.

As a result of his paper Professor Kahlenberg encouraged a colleague at the University of Wisconsin, FD Heald, to extend his original findings[16] to three additional plant species (*Pisum sativum, Zea mais, Cucurbita pepo*) to assess the ability to generalize their findings. To that end Heald assessed the effects of 18 agents in a hydroponic (300 cc) setting over three to six concentrations depending on the experiment.[18]

The methodology followed closely that of Kahlenberg and True[16] and True[19] with respect to sample size (n=two plants/concentration) and endpoint (i.e., measurement of root growth). Two measurements of root growth were taken, one at 24 and the other at 48 h. In contrast to the experiments of Kahlenberg and True, no control results were reported. Again, no report of stimulatory responses were noted in this study.

Stimulatory effects first reported

A major conceptual advance occurred several years later in 1899 with Kahlenberg now working with Edwin Copeland.[20] The two researchers reported stimulatory effects with copper, cobalt, boron, lead, and tungsten with the plant *Lupinus albus*. In his previous work Kahlenberg had focused on answering the question of what chemical concentration would kill the root of the *L. albus* plant. However, in this 1899 report with Copeland[20] the attention was focused on the striking report of Nageli[21] which was edited by Cramer and published several years after Nageli's death.

The inspiration to assess Nageli's work is important to consider since it provides a fundamental reorientation from that of the assessment of very high levels of exposure to very low concentrations. Copeland and Kahlenberg pointed out that growing plants hydroponically was often problematic since the distilled water used was thought to be in itself toxic to the plant.[20]

The significance of Nageli's research is that he conclusively established that water itself was not toxic but that various metals in solutions in infinitesimal quantities may be highly toxic. Nageli asserted that the most toxic metal tested was copper, which was fatal to *Spirogyra* at the estimated concentration of 1 p.p.b., a value unheard of in 1880 and many decades later. So low was the toxic concentration that this response led Nageli to hypothesize that the adverse effect was of a different nature from 'ordinary poisoning'. He called this new phenomenon an 'oligodynamic effect'. Nageli[21] combined the Greek words 'oligos' meaning small and 'dynamis' meaning power to form a word meaning effective in small quantities. The oligodynamic phenomenon was hypothesized to involve metallic copper particles passing into solution and killing plants by a different mechanism from toxic copper salts, perhaps by physical rather than chemical means.

The principal point of the Copeland and Kahlenberg work was that most metals in solution will become chemically altered by contact with oxygen or other agents in water leading to the formation of oxides, hydroxides, or carbonates which are more soluble in water.[20] Such metals are therefore chemically transformed into various salts with resulting enhanced toxicity. Thus, Copeland and Kahlenberg concluded that the toxicity of copper (as seen by Nageli) was due to the salts of the metals and not the metals themselves, as had been interpreted by Nageli.

While Copeland and Kahlenberg used four plant models (*Zea mais, Lupinus albus, Avena sativa*, and *Soja hispida*), the most extensive and accurate work was with *L. albus* and experiments with this model were emphasized.[20] The toxic potential was tested in over 30 metals which were delivered as a rolled foil with a piece cut out and placed at the bottom of the beaker. Presumably some of the metal would dissolve from the foil and enter solution. As in earlier studies, only root growth was measured with two plants for each agent. A concurrent control was employed. Growth, which was followed for up to 11 days, was enhanced by 14 agents (i.e., gold, platinum, palladium, silver, aluminum, tin, bismuth, sulfur, carbon, chromium, indium, selenium, and rhodium). Those resulting in the largest increase in growth were Bi and Se which grew 2–3-fold greater than the controls. The remaining 12 agents had growth stimulation between 12–86% greater than controls.

While the study was not designed to assess stimulatory responses, such responses were convincing to Copeland and Kahlenberg who then cited a variety of other authors [e.g., Raulin,[4] Richards,[22] Guenther[23]] who reported on low dose stimulatory responses. They then concluded their paper with the statement 'the subject of chemical stimulants is a most inviting one for further study.'

As a result of these findings Copeland was encouraged (in the spring of 1899) to extend this work. In pilot studies with Zn and Cl in water solutions he again showed an acceleration of growth but with stimulatory concentrations not a great deal more dilute than those that were distinctly toxic.[11] Such an observation is remarkably consistent with the overwhelming majority of subsequent hormetic observations in which the toxic threshold is typically within a factor of five of the highest stimulatory response.[24]

In the summer of 1899 Jacobi published a paper on the influence of various chemical stimulants (including KNO_3, KCl, NaCl, chinin, antipyrin, iodine, oxalic acid, and $CuSO_4$) over a limited concentration range on the respiration of several water plants (i.e., *Elodea* and *Myriophyllum*).[25] The positive findings of Jacobi lead Copeland, having moved from the University of Wisconsin to Stanford University, to consider respiration (under experimental conditions where CO_2 has been removed)

rather than growth as an improved index of *Elodea* metabolic processes.[11] While the findings of Copeland convincingly establish that numerous toxic substances enhance CO_2 liberation over broad concentration ranges, the findings generally suffer from poor presentation which affects their utility and interpretation. Despite the limitations in the data presentation Copeland offered unusual insight that has potentially highly significant implications to the theory of hormesis and the interpretation of experimental data.

Copeland observed that the concentrations producing either the stimulation or inhibition differed by less than the inter-individual variation of the seedlings used.[11] Thus, on some occasions some plants were injured while others were stimulated by the same concentration. Such observations have important implications for study design, sample size and interpretation of data. While Copeland[11] focused on the effects of chemical stimuli on growth secondary to that of respiration, effects on reproductive performance are often inversely associated with growth. Klebs had earlier reported that the demands of reproduction upon a number of physiological processes are stricter than that of growth.[26] In this respect, the optimal temperature for reproduction is typically below that of growth. Thus, a chemical stimulus enhancing growth may be predicted to inhibit reproduction. This suggestion is in striking agreement with much of the data presented in the fungal section of this chapter where low dose effects which typically stimulate mycelium growth reduce fruiting responses.

While the initial breakthrough in stimulatory responses was that of Copeland and Kahlenberg[20] the key transitional paper was by Rodney H True, having moved from Kahlenberg's group at Wisconsin to the US Department of Agriculture, and Gies who suggested that low levels of heavy metals may stimulate plant growth based on their work at Woods Hole.[27] The significance of this report is that it compared the primary radicle growth of *Lupinus albus* seedlings with 16 compounds over four or five concentrations (over a dosage range of 8- or 16-fold) with a concurrent control, thereby, providing the first dose-response continuum for dose stimulatory/toxic responses. The number of plants for each concentration was equal to or greater than four. Growth of the radicle was measured at 48 h. A stimulatory response was reported for the dilute solutions for $CuCl_2$, $CuCd_2(CO_2)_2$, $AgNO_3$, $HgCl_2$, $ZnSO_4$, $Ca(NO_3)_2$, $CaCl_2$, $MgCl_2$, but not for $CuSO_4$, NaCl, Na_2SO_4, KCl, KNO_3 and urea (Table 1). Despite the significance of the True and Gies findings,[27] it was frequently not possible to assess the range of stimulatory responses in their study since six of the nine stimulatory agents displayed only one (i.e., the lowest) concentration that was in the stimulatory range.

Even though this experiment could have been strengthened by the presence of a larger number of concentrations, increased sample size and statistical analysis, it represented a significant advance since it not only confirmed the concept of low dose induced metal salt stimulatory responses in *L. albus* but did so within the context of a dose response relationship. This provided a solid foundation for the subsequent papers with stronger study design of True and Oglevee,[28] Jensen,[30] and Schriener and Reed[31] that played an important role in establishing the concept that dilute solutions may produce stimulatory effects in plants. The paper by True and Oglevee considered the effect of the presence of insoluble substances on the toxic action of poisons.[28] Earlier research had shown that when antagonistic salts were added to toxic solutions the toxicity was reduced and growth was enhanced.[27] By extending this concept to various insoluble substances (i.e., clean sea glass, powdered Bohemian glass, filter paper, powdered anthracite coal), True and Oglevee sought to alter the toxicity of numerous inorganic (i.e., copper sulfate, silver nitrate, mercuric chloride, HCl, NaOH) and organic (i.e., the chemical disinfectants thymol, resincinol) toxicants on *L. albus* root growth.[28] The number of plants/treatment was at least four with a 48-h period of observation. The typical experiment involved the determination of the maximum concentration permitting survival. This was then followed by the introduction of the specific insoluble substances into the toxic solutions. In general, the inclusion of the insoluble substances not only reduced toxicity but typically stimulated root growth well beyond that of the controls. Based on these findings True and Oglevee[28] developed what is believed to be the first modern schematic representation of what now is called the hormetic β-curve (Figure 1). Additional

Table 1 Average stimulatory response of the primary radicle of *Lupinus albus* seedlings in the maximum stimularory concentration of various chemicals (data from Copeland and Kahlenberg[20])

Chemical	*Radicle growth (percentage greater than control)*
$CuCl_2$	14.3
$Cu(Cd_2Cl_2)_2$	25.0
$HgNO_3$	6.0
$HgCl_2$	6.6
$ZnSO_4$	16.0
$Ca(NO_3)_2$	210.0
$CaCl_2$	28.0
$CaSO_4$	289.0
$MgCl_2$	25.0

support for True and Oglevee's findings was reported by Fitch[29] who conducted similar experiments with fungi. He concluded that insoluble substances in a solution act as agents of dilution or absorption, thereby reducing the concentration of available toxin to the fungus.

Several years later Jensen at Stanford University presented the results of an extensive study to assess the toxic and stimulatory effects of nine salts and poisons (e.g., copper and zinc sulfates, lead, silver, and iron nitrates, phenol, ethanol) on the growth of wheat in artificial soil (i.e., ground quartz) and hydroponics.[30] This study employed the up to 16 doses plus a control group with five plants/ concentration. The experiment lasted for 15–30 days depending on the specific experiment. The findings are remarkable since all agents displayed the typical hormetic dose-response curve for all four measures of growth (i.e., growth of stems, transpiration, fresh and dry weights). Direct comparisons with solutions and soil were made for each substance. All agents induced stimulatory growth in the quartz soil and all but two (i.e., $ZnSO_4$ and $CuSO_4$) induced stimulation in solution. In addition, the relationship between the dose causing the highest stimulatory response and the NOAEL were determined as well as the dosage range of the stimulatory responses (Table 2).

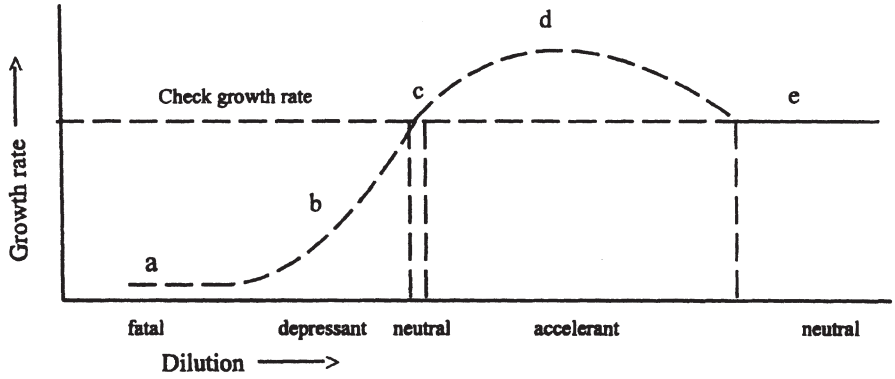


Figure 1 Hormetic β-curve representing the different physiological phases resulting from the progressive dilution of a fatal concentration: a=fatal concentration; b=survival with depressed activity compared to control; c=growth rate approximates control; d=acceleration phase marked by activity beyond the control; e=growth rate approximates control (redrawn from True and Gies[27])

Table 2 The maximum stimulatory response for four endpoints (growth, transpiration, fresh weight and dry weight), the fold difference from the maximum response to the NOAEL and the stimulatory dose range for the multiple chemical agents tested by Jensen[30]

	Cu Soil	*Pb Soil*	*Pb Sol.*	*$AgNO_3$ Soil*	*$AgNO_3$ Sol.*	*$ZnSO_4$ Soil*	*$ZnSO_4$ Sol.*	*$Fe_2(NO_3)_6$ Soil*	*$Fe_2(NO_3)_6$ Sol.*
Growth[a]	9.1	39.5	18.8	9.4	20.6	14.2	15.5	11.4	8.8
Transpiration[a]	25.0	79.1	7.6	10.6	No Data	9.7	16.6	19.7	25.0
Fresh weight[a]	32.5	86.7	8.2	10.5	10.9	4.7	32.2	28.8	43.9
Dry weight[a]	40	53.6	11.8	6.9	11.0	6.4	16.9	35.2	43.5
Max-NOAEL[b] (fold)	2	2.6	2	2.5	2.0	1.4	3	1.4	3.3
Stim. Range[c] (fold)	1	100	1	3.75	2.0	3.0	10	15	9

	Fe_2Cl_6 Soil	Fe_2Cl_6 Sol.	$NiNO_3$ Soil	$NiNO_3$ Sol.	Phenol Soil	Phenol Sol.	ETOH Soil	ETOH Sol.
Growth[a]	No Data	13.9	12.6	No Data	11.7	15.5	6.5	29.6
Transpiration[a]	No Data	10.1	8.9	No Data	22.7	No Data	3.7	No Data
Fresh weight[a]	7	26.2	21.2	17.2	27.7	No Data	10.7	26.5
Dry weight[a]	18	10.1	21.4	20.1	32.2	3.7	19.8	21.7
Max-NOAEL[b] (fold)	3	7.5	5	2	1.8	5.7	1.25	5
Stim. Range[c] (fold)	3	7.5	4	3	10	1.5	1	7

[a]Values represent the per cent increase above the unexposed controls for the treatment group showing the highest stimulatory response (i.e., a comparison of the mean response of the highest stimulatory group and the mean response of the controls). [b]Values represent the dosage difference between that dosage with the maximum stimulation and the dosage displaying the estimated NOAEL. [c]Values represent the estimated dosage range of the observed stimularory response

In 1908 Schreiner and Reed[31] at the US Department of Agriculture, Bureau of Soils (Washington, DC, USA) performed a similar experiment on wheat as Jensen.[30] However, the experiment employed 38 organic agents all in solution. The number of plants per concentration was increased to 20, the number of concentrations per chemical was five and duration of the experiments was 8–10 days. Three endpoints were used to estimate growth (i.e., transpiration, green weight and turgidity). Of the 38 compounds tested ten (i.e., alanin, cholin, picolin, neutralized piperidin, pyrocatechin, resorcin, hydrochinone, phloroglucin, arbutin, and vanillic acid) displayed distinct stimulatory responses. While the actual data were not presented so that a representation of the dose-response curves was not possible, a ratio of the LOAEL and the dose causing the maximum stimulatory response for nine of the ten stimulatory responses was shown to be $\leqslant$25-fold. If one assumes that the LOAEL/NOAEL ratio approximates 5–10-fold, then the findings of Schreiner and Reed[31] would be very consistent with those of Jensen.[30]

Stimulatory response reported by Japanese scientists

While research in the US was proceeding to establish a foundation to assess chemical hormesis as a biological hypothesis,[27,28,30–33] a comparable effort was being undertaken in Japan by Suzuki and Aso.[9,10,34,35] In this research low quantities of potassium iodide (KI) were shown to moderately stimulate the growth in soil of peas,[36] oats and radish.[34] These experiments were limited in scope utilizing only one dose in the pea experiment, three doses in the oat experiment and two doses in the radish study. In general, each treatment had five plants and each experiment had a concurrent control.

The stimulatory responses were quite similar for each model (i.e., 12–50% depending on the endpoint measured). Similar experiments with oat (three doses) and radish (two doses) treated with sodium fluoride (NaF) likewise displayed stimulatory responses with the radish being markedly enhanced (i.e., 2–3-fold) in the treatment group. Follow-up experiments were conducted with peas using four doses of calcium fluoride in a water solution.[35] The growth of the peas after 18 days of calcium fluoride (CaF) treatment displayed weak evidence of a stimulatory response based on the nature of the response and the variability between treatments. A soil experiment with four doses with three plants/concentration displayed a 20–35% increase in growth at the upper two doses. Later harvesting of fruits from these plants and fresh weight of the fruits were stimulatory at the two lowest doses (i.e., the doses that were not stimulatory at the earlier measurement time). A third experiment with barley revealed that CaF treatment (four doses) enhanced the growth in length and fresh weight.

In general, the experiments of Suzuki offer some support that dilute concentrations of KCl and CaF can enhance plant growth. The studies with KCl are very limited for peas since only one dose was used and only five plants/concentration were employed. The KI experiment was stronger due to the use of three doses and the internal consistency of the responses of the three doses for each endpoint measured. Similar consistency was seen with radish but only two doses were used. The fluoride experiments were generally consistent with these stimulatory responses with the exception of the very weak stimulatory response for peas. In general, these experiments are supportive of that reported by the above US references despite their generally limited sample size and modest number of doses and dose ranges.

Plant responses to limited physical/chemical stresses

Another area of interest in the concept of plant responses to injury was in Germany at the Botanical Institute in Leipzig. There, under the direction of Professor Pfeffer, CO Townsend undertook various experiments from 1896–1897 to assess to what extent an injury inflicted upon one part of the plant will affect the growth of the injured and uninjured parts of the plant.[37] One note regarding Professor Pfeffer. In addition to being the mentor of the American scientists Townsend and Duggar, two important figures in the development of early experimental findings relating to hormesis, the German botanist WFP Pfeffer provided key insight on osmotic pressure which van't Hoff used, his noted prize research on solution theory. More specifically, Pfeffer established that osmotic pressure was proportional to concentration, that is, inversely proportional to the volume of the solution in which a specific amount of solute is contained.[14] This observation lead van't Hoff to conclude that osmotic pressure was not the result of specific chemical interactions with the semipermeable membrane but a physical process analogous to the kinetic theory of gas pressure and ultimately explainable by the second law of thermodynamics.

Townsend was interested in the first effects of injury upon the plant growth curve. While this may seem initially unrelated to the concept of chemical/radiation hormesis, it represents perhaps the initial systems-based mechanistic concept of hormesis as

an overcompensation to a disruption in homeostasis, conceptualized and elaborated on nearly 80 years later by Stebbing.[38] Prior to his experimentation no systematic investigation had been undertaken to evaluate plant injury with respect to the effects of a single irritation of short duration or the time needed for an effect (i.e., response) to become manifest. Of special interest to the present analysis is that Townsend initiated the stress via a series of different physical injuries or via the administration of a chemical stressor agent, ether.[37]

In the case of ether, high doses caused a marked retardation of growth compared to concurrent controls over a 7-day period of observation. However, at the lowest doses there was a distinct acceleration of growth (n=12/concentration). The growth period was directed towards various time intervals (i.e., 24, 48, 72, and 192 h). While the low dose-induced growth rate was reduced by 25% after the initial 24 h, the low dose treated plants displayed a marked stimulation such that by 72 h the growth was twofold greater than controls. By 192 h the growth differential had notably diminished to only 12.3% greater in the treated plants. These responses, as noted above, are strikingly similar to the later observations of Stebbing based on investigations with marine micro-organisms.[38] The findings with the ether-induced injury were also in accord with the physical injury-induced growth responses. That is, according to Townsend, 'if the injury is slight, sign of an acceleration in the rate of growth will be apparent from 6–24 h, and will continue for approximately 1 to several days. If the injury is severe, the acceleration of growth will be preceded by a period of retardation of growth of longer duration depending upon the severity of the injury and upon the condition of the plant injured'.[37] The growth resulting from the influence of a single irritation was enhanced by up to 70%. This observation is remarkably consistent with the recent reports of Calabrese and Baldwin[24,39] who developed a quantitative description of hormetic responses based on a review of several hundred examples of hormesis in the chemical toxicology literature.

Following his investigations in Germany during 1896–1897, Townsend moved to Barnard College, New York, and while there undertook a series of studies to assess the influence of ether on seed and fungal spore germination (see section on fungi for a detailed assessment on the development of the concept of hormesis in this biological model).[40] In this experimentation seeds of four plant species (*Zea mais, Cucurbita pepo, Avena sativa*, and *Phaseolus vulgaris*) were soaked in pure water for 24 h at room temperature and then transferred to an air-tight chamber which contained respectively 1, 2.5, 5, and 10 cc of ether dissolved in 100 cc of water. In each of the four species of seeds tested, exposure at the lowest dose (1 cc) produced a marked shortening of the number of hours required for germination (i.e., reduced from 48 to 36 h for *Zea mais* and *Avena sativa*, reduced from 40 to 24 h for *Phaseolus vulgaris*, and reduced from 36 to 24 h for *Cucurbita pepo*). At the higher concentrations the number of hours needed for germination was notably increased (i.e., 84–96 h). With respect to fungal spores (i.e., species *Mucor, Penicillium*), exposure was applied over five doses (0.1, 1.0, 2.5, 5 and 10 cc in an air-tight chamber). According to Townsend, the 0.1 cc group appeared to have a slightly accelerated rate of germination.[40] However, he noted that this was difficult to determine since the normal period of germination in these fungal species does not extend beyond 14 h under the experimental conditions used. The 1.0 cc concentration-treated spores displayed an initial retardation in growth but then a marked adaptation and acceleration of growth thereby overcoming the toxic effects.

Pesticide induced stimulation of plant growth

Hydrocyanic acid In 1886 the US Department of Agriculture introduced the use of hydrocyanic acid for fumigating orange trees. Its use as a fumigant soon spread to greenhouses, nurseries, stock and orchard trees of all types. Such widespread use lead Townsend to assess whether or not hydrocyanic acid gas may adversely affect grains or other seeds.[41] In his study, seeds were constantly under the influence of the gas from 15 days to 1 year. The results indicted that the seeds germinated more readily after they had been constantly exposed for a 15–60 day period while with 153 days exposure the time to germination had notably increased. In addition to enhanced germination, the hydrocyanic gas treatment also stimulated growth but for a limited time only. Townsend indicated that the growth was enhanced to almost twice as great as the controls.[41–43] However, the growth acceleration was of a short duration, rarely lasting more than 1 week after which the treated plants resumed the growth rate of the controls.

Since the initial publications of Townsend, several reports have indicated that hydrocyanic acid gas fumigation enhances the growth and fruit yield of tomato.[44,45] The Clayton report was of particular interest since it used equal to or greater than eight doses depending on the experiment plus a control group.[45] It is of additional significance that considerable anecdotal comments of fruit growers

of the western part of the US as noted by Moore and Willaman indicate that these farmers were concerned that whether their trees were infected with insects or not they would not achieve a maximum yield unless they had received the cyanide fumigation.[44]

Copper salts Of practical significance with respect to pesticidal treatment of plants was the earlier interest in the actions of various copper salts as seed fungicides and how their application may affect seed germination and plant growth.[46] Within the conduct of such investigations concerning the refinement of how dose affects plant toxicity reports began to emerge that the highly toxic copper salt, copper sulfate, may act as a plant stimulant in highly dilute solutions.

This early observation of Rumm[47] encouraged further debate and research on the topic of copper fungicides on plant growth separate from its action as a fungicide. For example, Rumm reported that grapes on vines treated with Bordeaux mixture (i.e., the copper based fungicide) matured 2 weeks earlier than grapes from untreated vines.[47] Rumm considered this effect as separate or independent from the action of the Bordeaux mixture on mold growth. In addition to the claims of enhanced maturation, there was evidence that the leaves of treated plants were notably greener (see Fairchild[48] for an extensive early review of this issue). However, the principal interest in this area remained the inactivation of destructive fungal spores. In fact, numerous chemicals were assessed for their capacity to alter fungal spore growth and reproduction. As the results of these experiments emerged, it became more obvious that highly dilute solutions of some toxic agents were capable of stimulating fungal growth and reproductive processes (see fungal section on the history of hormesis).

Essentiality versus *non-nutritive stimulation*
The mid 1890s brought forth considerable interest in the importance of inorganic or mineral salts in plant nutrition. The initial findings of this plant related research lead to a type of classification of elements into three groups (i.e., nutritive, indifferent and toxic). Examples of the nutritive or essential elements were calcium, phosphorus, nitrogen and potassium while arsenic, copper and boron were viewed at that time as fundamentally toxic.[49]

Within this context emerged the issue of nutrient essentiality and optimal doses as well as whether non-essential substances were also able to cause stimulatory growth at low doses. This was not only an issue for plant related research but also for other types of organisms as well. In fact, this debate in the area of plant biology was foreshadowed by a highly contentious debate in the area of fungal growth, development and reproduction on how to distinguish essentiality responses from what was called at that time stimulatory responses resulting from 'chemical irritation'.[4,22,50] Debate on essentiality *versus* non-nutritive chemical stimulation became linked with issues of when an element is a contaminant or a nutrient.

In 1913 Lipman and Wilson directed their attention to the impact of smelter fumes and smelter waste deposition on crop growth in the vicinity of smelters.[51] This interest was also coupled with enhanced scientific interest in the physiological effects of metallic compounds of copper, lead, zinc and other agents on plant growth[52–54] (see the extensive review by Brenchley[49]). The general belief at that time, as widely promoted by influential publications such as the classic work 'How Plants Grow' by Johnson,[55] was that copper was extremely toxic to plants even in very minute doses.

This conclusion was consistent with the work of Heald[18] and Harter.[56] In fact, Heald noted that ~2.5 p.p.m. of copper in water was lethal to garden peas while maize seedlings were killed in the presence of ~1.25 p.p.m. copper in water.[18] Similar work also revealed copper in soil to be toxic to wheat and rye at low p.p.m. concentrations.[53,54]

Among the waste fumes from smelter operations was H_2SO_4, which was thought to solubilize minerals such as Cu and Zn from soil. In this fashion, the smelter operations could affect plant growth in numerous ways both as a direct toxic exposure and indirectly. In follow-up experimental work Lipman and Wilson assessed the effects of $CuSO_4$, $ZnSO_4$, $MnSO_4$ and H_2SO_4 on the growth of *Vicia sativa* (vetch) and the Little Club variety of wheat.[51] In their experiments 6–9 treatment groups were employed along with concurrent controls with four plants/concentration. In general, it was shown that the $CuSO_4$, $ZnSO_4$ and $MnSO_4$ enhanced the growth of the vetch while only $MnSO_4$ enhanced the growth of the wheat. The dose-response relationships conformed very closely to that reported earlier by True and Oglevee[28] and Jensen[30] with respect to the range of stimulatory responses (10–20-fold) and the maximum stimulatory response (50–60% greater than controls).

In 1914 Brenchley published a highly influential book[49] that offered a substantial review of the literature on the capacity of several inorganic agents (i.e., copper, zinc, arsenic, boron, and manganese) to affect plant growth along with integrating these literature findings with new data from the well-

regarded Rothemsted research station where she was located. Considerable emphasis was placed on the concept of inorganic or mineral substances in plant nutrition and the capacity of such agents to induce stimulatory growth or toxicity depending on the experimental conditions including dose.

As set forth in her review, Brenchley stated that the earliest observations on the effects of metallic agents on plant growth emphasized their capacity to induce toxic responses.[49] However, it was soon recognized that under certain conditions these agents appeared to cause beneficial rather than harmful effects. Such so-called poisons were then seen to act as 'stimulants' if they were administered to the plant in a sufficiently dilute solution. Such stimulatory responses were reported for various plants with a range of toxic substances and the hypothesis was brought forth that any poison may act as a stimulant if given in sufficient low doses, as Hugo Schulz had proposed following his earlier work with yeast in 1888.[2]

The concept that dose determines the stimulation and the poison as articulated by Brenchley[49,57] was linked to practical applications. For example, in 1913, Lindsay reported experiments utilizing the poisonous action of copper to clear ponds of algale growths.[58] A quantity of copper sulfate sufficient to make a solution of $1/50 \times 10^6$ was needed to kill the *Spirogyra*. In contrast, Allison *et al.* reported that totally unproductive soil could be made to produce high crop yields of lettuce, radish, turnips, rye, tomato, etc. by the addition of 30 pounds of copper sulfate per acre.[59] Such findings and numerous other examples lead Brenchley to speculate on the possibility that copper may be an essential nutrient for many plant species.[60] However, she emphasized that an element may stimulate growth without being essential in the sense that in its absence important aspects of growth are inhibited. Thus, the issue of determining essentiality is a complex one and how that relates to hormesis is an important issue needing resolution.

While the above examples relate to the issue of copper as a growth stimulant and toxic agent as a function of dose, species, and experimental conditions (e.g., soil *versus* hydroponics), Brenchley also discussed this within the context of other inorganic agents as well.[49,57,60] It is clear that most substances she discussed stimulate at low doses while being inhibitory at higher doses.

Summary

When an agent stimulates plant growth it is difficult to determine the relationship of the stimulatory response to the underlying mechanism. Is the mechanism one that relates to a *bona fide* essentiality function, the substitution of a function of a genuine essential element that may be present in low doses, or is the stimulatory response an overcompensation to injury? By the mid 1930s the plant research community clearly recognized the concept that the dose determines the stimulation or poison and initiated the process of determining the underlying causes of the stimulatory response.

The early history of hormetic effects in plant biology represents an impressive evolution of independent observations of low dose stimulatory responses within the context of progressively enhanced rigor in the study design requirements. The early studies of Copeland and Kahlenberg represented a clear break from past observations with the intriguing conclusion that chemical stimulants offered a 'most inviting one' for future investigation.[20] In fact, it was the continued follow-through of this suggestion by associates of Professor Kahlenberg, such as Edwin Copeland and Rodney True, that quickly transformed the strikingly insightful above noted quote to the now widely recognized β-curve of the hormetic response.

Not to be overlooked in affecting the acceptance of low dose stimulation by toxic substances in plants was the detailed, objective and highly supportive reviews of this topic by various technical publications by the US Department of Agriculture.[56,61]

The combination of these developments with the observations of Townsend, Suzuki and Aso in Japan and the recognition of comparable developments with other biological models made the hypothesis that toxic substances may be stimulating in dilute concentrations a progressively supportive one during the early decades of the 20th century. Such was the position of this hypothesis as the plant research community was unknowingly positioning itself for the striking development in the 1940s to the present of the role of plant auxins and other growth regulators on plant growth.

Bacteria

Introduction

The concept that low doses of toxic substances may be stimulatory to bacteria has been credited to Ferdinand Hueppe.[8] This seems to be based on the explicit statements in his 1896 text indicating that he independently discovered this phenomenon in bacteria which had been shown in yeast by Schulz[2] during the previous decade. As result of his strong

reputation in the field of bacteriology and promotion of the concept that low doses of toxic substances stimulate biological processes, this phenomenon became known as Hueppe's Rule (as well as the Arndt-Schulz Law). Despite his leadership on this concept Hueppe neither presented data in his 1896 book which documented his claim to subsequent co-discovery nor provided references to primary literature which would document his claim.

Even though the original basis of the Hueppe claim was not apparently reported, a number of researchers in the field of bacteriology have unequivocally substantiated this concept with work following along several parallel lines. These include effects on colony growth, bacterial size, and physiological processes such as ammonification, nitrification, nitrogen-fixation and respiration from inorganic salts, heavy metals, various disinfectants and other organic compounds.

Stimulation of colony growth by low doses

It was not until over a decade after the announcement by Hueppe[8] supporting the conclusions of Schulz,[1,2] that substantial research came forth extending these findings in Europe and the United States. In Germany, Hune assessed the effects numerous bactericides (e.g., fluoride, $CuSO_4$), sublimate (i.e., a sulfur compound), thymol, alcohol, ether, formaldehyde, capsicum (as derived from the dried ripe seed of the species *C. frutescens*, used as a gastric and intestinal stimulant) and other agents on the colony growth of four types of bacteria (*E. coli*, typhus, dysentery and cholera).[62] He observed that all such agents displayed the capacity to stimulate the multiplication of bacteria at low doses. Of particular interest is that Hune also incorporated an exposure duration and concentration interaction for the effects of formaldehyde on *E. coli* plate count and generation time.[62] Such findings were confirmed and extended in a series of subsequent studies over the next 30 years by numerous investigators.[6,63–70] It is worthwhile to see reviews of the above information in leading bacteriological texts (e.g.,[71–74]) of the above information; in all cases they reiterated the consistency of the low dose stimulatory findings and the ability to generalize the results across a broad range of chemical classes including the non-essential elements such as lead. The original papers are remarkable with respect to the wide range of agents evaluated, the consistency with respect to the stimulatory responses and the quantitative description of the dose-response relationships including the stimulatory range, the maximum stimulatory response and the relationship of the maximum stimulatory dose to the NOAEL (or the toxic threshold dose).

It should be emphasized that with the exceptions of Hune[62] and Hofman[6] the above cited references were principally designed to clarify various mechanistic aspects of the rapidly evolving field of chemical disinfection with no reference to the Arndt-Schulz Law or Hueppe's rule. As noted in the detailed review by Falk[75] on the role of ions in bacterial physiology and disinfectant theory, Winslow and Hotchkiss[76] principally set forth to assess the capacity of different salts to affect growth of gram-negative bacteria in an effort to extend in part the massive effort of Eisenberg[77] on the physical, chemical or physio-chemical actions of ions specifically in disinfection and the nature of the differential responses between gram-positive and -negative organisms. While the work of Hotchkiss[64] established an ionic potency gradient for chlorides, Falk[75] noted that this work contained 'a striking original contribution', demonstrating a definite stimulatory action exerted by a variety of numerous well known toxic agents including lead and mercury at concentrations below the inhibitive level.

Hotchkiss, working as a pre-doctoral student under the direction of Professor Winslow at Yale University, published the results of an extensive assessment of 23 cations with the same anion (Cl^-) on the growth of *E. coli* [see Winslow and Hotchkiss[76] for a preliminary report of these findings]. The author was interested in providing a more comprehensive survey of the stimulatory, as well as toxic effects of a broad range of chemicals. The types of chemicals used included a broad survey of minerals (e.g., calcium, sodium) and toxic (lead, mercury, tin, cadmium, titanium, nickel) salts. Since that study involved a large number of measurements, turbidity was used to estimate the number of bacteria per cubic centimeter since this was quick and reliable. The average value for each treatment (five tubes/treatment) was determined after 3 days of incubation. The dose range and the number of doses varied according to the salt. Of the 23 salts, 20 had $\geqslant 3$ doses and 17 had $\geqslant 6$ doses. Of the eight salts that showed no stimulation, six had $\leqslant 4$ doses, while of the 15 with a stimulatory response all had $\geqslant 5$ doses. In addition, of the 15 chemicals with a stimulatory response seven had three or more doses below the NOAEL.

These findings indicate that the observations of the stimulatory response at low levels were influenced to a large extent by the study design. Experiments with a large number of doses, especially those with multiple treatments below the NOAEL, displayed stimulatory responses. The magnitude of the enhancement response varied considerably among the metallic salts. The max-

imum stimulatory responses ranged from 17–540%, with the median value 50%. The dose range of the stimulatory zone was between 2–100-fold with all except one experiment within an approximate 2–50-fold range.

Of considerable interest were the observations that, in addition to potassium, sodium, NH3, lithium, magnesium, calcium, and barium salts, the toxic salts of titanium, tin, nickel, lead, cesium, and mercury were also stimulatory. The stimulatory responsiveness of the toxic salts was generally comparable to the mineral salts with respect to maximum extent of stimulation. Perhaps the most important observations are the differences in molar concentrations for stimulation of bacterial growth. Table 3 indicates that the mineral salt stimulatory concentrations were much higher (i.e., typically >1000-fold higher) than the toxic metals. Nonetheless, once the hormetic threshold was achieved the range, shape and magnitude of responses were similar. The hormetic findings of Hotchkiss[64] for the various mineral and toxic (e.g., lead, mercury) salts were confirmed and extended by subsequent PhD students of Winslow's[66–68] using similar (peptone) and different (tartrate, asparagine) growth media as well as different measurement techniques for estimating bacterial counts.

By the late 1930s it was recognized that a highly generalizable disinfectant spectrum of biological effects exists for most and probably all disinfectants. Marshall and Hrenoff described the action of disinfectants as a flexible blending of differential modes or degrees of activity.[70] The initial response band describes a range of dilutions of an agent between 0% and the greatest dilution which can be shown to exert an action on the bacteria. This was called the ineffective band (i.e., no effect zone). The second band reflects a range of relatively high dilutions in which there is some slight stimulation of bacterial multiplication which they called the stimulative band. This is the stimulatory zone of Hotchkiss, Hueppe and others. It is interesting to note that Marshall and Hrenoff, who emphasized the concept of bacteriostasis, stated that this stimulatory range is 'ordinarily narrow and is frequently of no practical significance, although there is no material objection to the application of the spectrum to substances not ordinarily considered as disinfectants, whereupon this band might become broad'.[70] It is not clear why Marshall and Hrenoff believed that the stimulatory range would be broader with non-disinfectants. In fact, this does not appear to be the case.[24] The third or inhibiting band and the fourth or germicidal band merge indistinguishably. It is these third and fourth bands where the interest of the disinfection biologist/chemist has centered. The fifth band is a range of highly toxic concentrations which are referred to as too toxic for reasonable application.

Less than a decade later Miller *et al.*[78] commented on the bacteriological implications of the low dose stimulatory zone as discussed by Marshall and

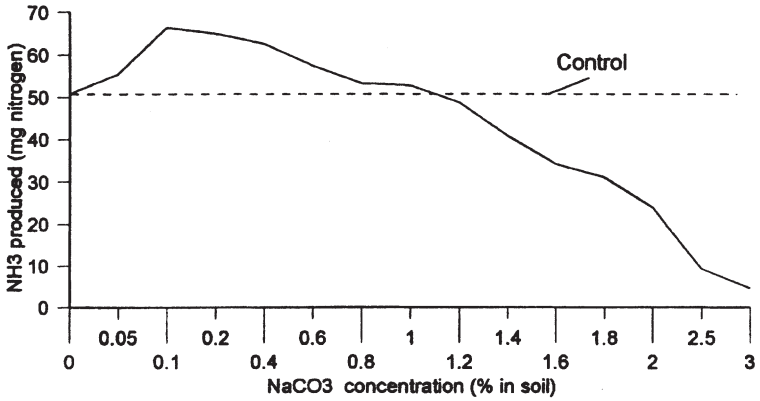
Table 3 Summary of stimulatory response data on the effect of various chloride salts on the growth of bacterium coli (data from Hotchkiss[64])

Chemical	*Stimulatory*	*No. dose*	*No. doses in stim. zone*	*Max. stim. (%)*	*Molar concen. for max. stim.*	*Range of stim. zone*
NaCl	Y	12	4	88	0.25	10
KCl	Y	11	6	50	0.25	50
NH_4Cl	Y	9	6	540	0.25	40
LiCl	Y	8	4	20	0.125–0.025	10
SrCl	Y	9	5	56	0.025	10
$CeCl_2$	Y	8	6	104	0.05	50
$MgCl_2$	Y	6	4	90	0.05	20
$BaCl_2$	Y	8	1	45	0.05	no range
$MnCl_2$	N	2	N/A	N/A	N/A	N/A
$AlCl_3$	N	4	N/A	N/A	N/A	N/A
$CdCl_2$	N	4	N/A	N/A	N/A	N/A
$CoCl_2$	N	3	N/A	N/A	N/A	N/A
$FeCl_2$	N	3	N/A	N/A	N/A	N/A
$FeCl_3$	N	4	N/A	N/A	N/A	N/A
$CeCl_2$	Y	8	2	29	0.00001	5
$HgCl_2$	Y	6	1	17	0.000001	no range
$NiCl_2$	Y	7	5	75	0.0001–0.00005	100
$PbCl_2$	Y	5	1	33	0.00005	no range
$SnCl_4$	Y	8	3	44	0.00005–0.000005	20
$TiCl_3$	Y	5	3	118	0.0005	10
TiCl	N	6	N/A	N/A	N/A	N/A
$ZnCl_2$	Y	6	2	13	0.00005–0.00001	5
		6.5 median	4 median	50 median		35 median

Hrenoff[70] with respect to their practical applications. Miller *et al.* noted that during World War II penicillin was in short supply in England.[78] Consequently doses were reduced in order for the limited supplies to reach all the needy patients. This diminishing of the dose was believed to affect the stimulation rather than the killing of the highly toxic gram-positive staphylococcus bacteria in many individuals (see[79] for a discussion).

Physiological processes: ammonification, nitrification and nitrogen-fixation

Inorganic salts. While the work described above focused principally on the responses of inorganic salts and disinfectants on bacterial colony growth, a parallel line of investigations on the effects of low levels of toxic substances on the capacity of bacteria to undertake ammonification, nitrification, and nitrogen fixation also proceeded. Initial investigations were undertaken by Lipman[80] who was interested in assessing the impact of the vast amount of waste alkali land in states such as California on the capacity of soil bacteria to perform ammonification and nitrification. Since ammonification represents the initial stage in the transformation and simplification of organic soil nitrogen, Lipman decided to assess the effect of various solutions on pure cultures of ammonifiers.[80] He selected the bacterium, *Bacillus subtilis* (obtained from soil in Auburn, California) since it was one of the more efficient ammonifiers as well as being easily isolated and cultured. Four salts (NaCl, KCl, $CaCl_2$, and $MgCl_2$) were tested in duplicate fashion over either six or 16 doses depending on the agent. The KCl, $CaCl_2$, and $MgCl_2$ treatments displayed marked dose-dependent inhibitory responses. However, the NaCl treatment displayed a 21% increase at the lowest (and only stimulatory) dose which was then followed by a dose-dependent decrease. A similar stimulatory ammonification response was reported in a later study by Lipman in the soil for Na_2CO_3.[81] However, in this 1912 experiment Lipman did not utilize pure cultures of bacteria but rather a light sandy soil free from alkali from a walnut grove



$NaCO_3$ Conc. (percent in soil)	NH_3 Produced (mg nitrogen)	$NaCO_3$ Conc. (percent in soil)	NH_3 Produced (mg nitrogen)
0.00	50.77	1.20	48.71
0.05	55.42	1.40	40.95
0.10	66.34	1.60	34.21
0.20	64.87	1.80	31.19
0.40	62.56	2.00	23.99
0.60	57.44	2.50	9.40
0.80	53.24	3.00	4.67
1.00	52.73	-----	------

Figure 2 Ammonification response of bacteria in soil exposed to various concentrations of Na_2CO_3 (data from Lipman[81])

in southern California. The soil had been supplied with humus as a result of a careful system of green manuring and developed a vigorous flora of ammonifying bacteria. This experiment offered a more detailed accounting for the low dose response with seven doses in the stimulatory zone with the response displaying the typical β-curve (Figure 2). The maximum stimulatory response was 30.6%; the range of stimulation was 20-fold while the distance from the maximum stimulation to the NOAEL was approximately 10–12-fold.

Low doses of NaCl, Na_2SO_4 and $MgSO_4$ enhanced nitrification in normal soil.[82] In the case of NaCl, addition of 0.005% caused a 44.5% increase in soil nitrification over controls 4 weeks after initiation of the experimental conditions. As in the case with *B. subtilis* for ammonification in the Lipman study[80] the stimulation was reported only for the lowest doses with the NOAEL about twofold higher than the maximum stimulatory dose. In the case of Na_2SO_4 Brown and Hitchcock reported a low dose stimulatory response of 33.7% at the maximum stimulation.[82] The most notable difference with the NaCl was the broader stimulatory range of nearly fourfold. In the case of $MgSO_4$ there was again a low dose stimulation (67%) (maximum) with a tenfold stimulatory zone. Calcium carbonate, well known as being generally beneficial to plants, was also found to stimulate nitrification over a nearly fourfold range with higher doses becoming markedly inhibitory. These collective findings in normal soil indicate that small quantities of NaCl, Na_2SO_4, and $MgSO_4$ and large amounts of $CaCO_3$ stimulate nitrification. The toxic concentrations were 0.02% for NaCl, 2.0% for Na_2SO_4, and between 1.5–6.0% for $CaCO_3$. No toxic effect was noted for $MgSO_4$. In a series of similar experiments using alkali, soil nitrification was enhanced by small quantities of $NaHCO_3$, Na_2CO_3 and $CaCO_3$ with the salts becoming toxic at 0.03% for sodium carbonate and bicarbonate and at 6% for $CaCO_3$.

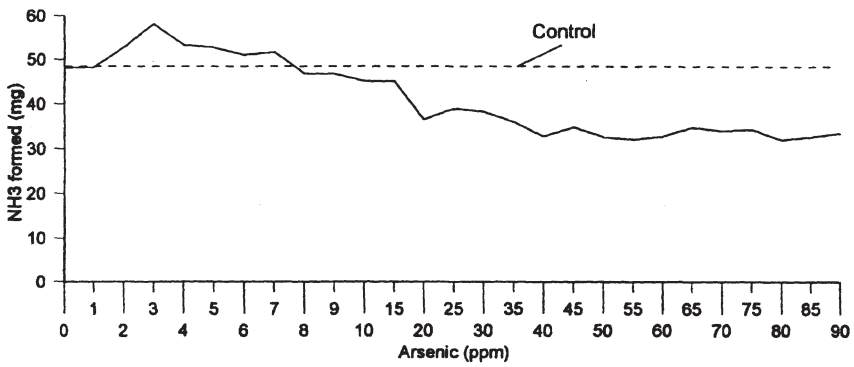
Arsenic compounds Since arsenic was known to occur naturally in many soils at appreciable levels, was used as an insecticide and in some commercial fertilizers and was present in flue gas from various smelters, concern was raised that various arsenic compounds may accumulate in soil and become toxic to vegetation and possibly to grazing animals. In 1913 Greaves assessed the capacity of various arsenic compounds to affect ammonification (Figure 3) and nitrification (Figure 4) in soils.[83] In this study Greaves assessed the capacity of soluble arsenic (sodium arsenate) and various insoluble arsenic compounds such as lead arsenate, Paris green, zinc arsenite and arsenic trisulfide which were used as insecticidal agents on ammonification and nitrification of a sandy loam high in calcium and iron but low in nitrogen.

In the initial experiment with sodium arsenate on ammonification a control plus 26 doses over a 90-fold range (1–90 p.p.m.) was employed (Figure 3).[83] Stimulatory responses were observed at low doses (2–7 p.p.m.) with the highest ammonification stimulatory dose (i.e., 3 p.p.m.) displaying a 21% increase over controls. Based on the work of Lipman,[81] Greaves suggested that the enhancement effect may be due to both the sodium and arsenic ions separately.[83]

In the case of the so-called insoluble forms, arsenic trisulfide had a marked stimulatory effect on ammonification at the lowest five doses of the 29 treatments as compared to the controls (Figure 4).[83] The dose response displayed a typical β-curve with the stimulatory range over eightfold, but with the maximum stimulatory response of 18.9% approximately 20-fold from the NOAEL. Both zinc arsenite and lead arsenite yielded an apparent minor stimulatory response (2–4%) only at the lowest dose for ammonification. It was unfortunate that the very broad dosage range tested using 29 doses did not extend below the 20 p.p.m. level in order to assess whether stimulatory responses would occur. All Paris green treatments were inhibitory for the ammonification.

With respect to nitrification, sodium arsenate was assessed over a 1–100 p.p.m. dose range using 36 doses (Figure 5).[83] Stimulatory doses were encountered at three successive doses (75, 80 and 85 p.p.m.). The maximum stimulatory response occurred at 85 p.p.m. (60.8% over controls) with inhibition evident at 100 p.p.m. Each of the four insoluble arsenic compounds which were assessed over 29 doses (20–1120 p.p.m.) displayed a marked stimulatory effect (Table 4). The stimulatory zone varied according to the compound; over 20–240 p.p.m. for lead arsenate, 280–720 p.p.m. for Paris green, 480–680 p.p.m. for zinc arsenate and 20–280 p.p.m. for arsenite trisulfide. The nature of the stimulatory response in each case followed the β-curve with maximum stimulatory responses ranging from 39–78% depending on the agent.

The findings of Greaves[83] are remarkable for the consistency of the observed responses between replicate samples, and the nature of the dose range employed. Nonetheless, it should be emphasized that the nature of the stimulation and inhibition observed were not defined with respect to the specific organisms affected (i.e., the soil microorganisms were not isolated or identified). Such a study complements the highly specific cultured bacterial assessment of Lipman with *B. subtillis*



Arsenic (ppm)	mg NH_3 Formed	Arsenic (ppm)	mg NH_3 Formed	Arsenic (ppm)	mg NH_3 Formed
0	48.02	9	46.76	50	32.60
1	48.24	10	45.06	55	32.12
2	52.72	15	45.06	60	32.78
3	58.12	20	36.54	65	34.82
4	53.36	25	38.90	70	34.00
5	52.68	30	38.22	75	34.40
6	51.02	35	36.00	80	32.12
7	51.66	40	32.78	85	32.68
8	46.76	45	34.80	90	33.60

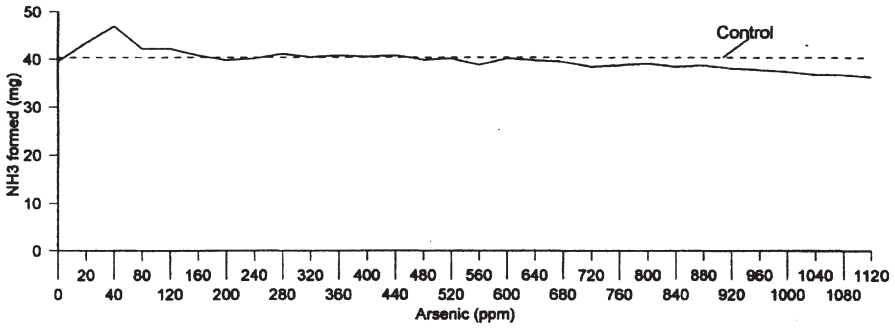
Figure 3 Data from Greaves[83] pertaining to the ammonia produced in 100 grams of soil containing different amounts of arsenic in the form of sodium arsenate

with respect to ammonification.[80] Later experiments by Greaves and Carter with the more toxic sodium arsenite effected a progressive increase in the number of bacteria in soil over a rather broad range of arsenic levels in soil (i.e., 1–200 p.p.m. arsenic) by 7–8-fold in sandy, unmanured and manured soil.[84] Despite the striking increase in bacterial numbers in all three soil types, the arsenic treatment enhanced ammonification only in the manured soil at low doses while affecting inhibitory responses for ammonification with the other two soils and nitrification with all three soil types.

When administered to soil in the form of sodium arsenate, arsenic trisulfide, or zinc arsenite, arsenic stimulates the nitrogen-fixation rate of the soil. This stimulating response was most apparent with lead arsenate, least with zinc arsenite, while Paris green did not display stimulation at any concentration (Table 5).[85] The range of the stimulatory responses was highly variable depending on the agent, with only one concentration stimulatory for sodium arsenate and zinc arsenite, while lead arsenate and arsenic trisulfide were stimulatory in a range of 2.5 and 16.0-fold, respectively.

These data were obtained in experiments using the same type of soil Greaves had employed in previous studies (i.e., sandy loam high with calcium iron and other minerals but low in nitrogen with a moisture content of 18%). The experimental samples were incubated for 18 days at 28–30°C with total nitrogen then determined.

This type of investigation is of considerable value since it permits a direct comparison of the series of arsenic compounds over a broad dose range for three different functions, (i.e., ammonification, nitrification and nitrogen fixation). For example, the nitrogen fixing organisms displayed optimal stimulation at 200 p.p.m. for lead arsenate while the nitrifiers and ammonifiers required much smaller quantities (e.g., 20–40 p.p.m.). With respect to



Arsenic (ppm)	mg NH_3 Formed	Arsenic (ppm)	mg NH_3 Formed	Arsenic (ppm)	mg NH_3 Formed
0	39.44	360	40.80	760	38.75
20	43.40	400	40.46	800	39.10
40	46.92	440	40.79	840	38.41
80	42.16	480	39.79	880	38.72
120	42.16	520	40.11	920	38.08
160	40.80	560	38.77	960	37.74
200	39.78	600	40.12	1000	37.40
240	40.13	640	39.77	1040	36.71
280	41.08	680	39.44	1080	36.70
320	40.44	720	38.41	1120	36.38

Figure 4 Data from Greaves[83] pertaining to ammonia produced in 100 grams of soil containing different amounts of arsenic in the form of arsenic trisulfide

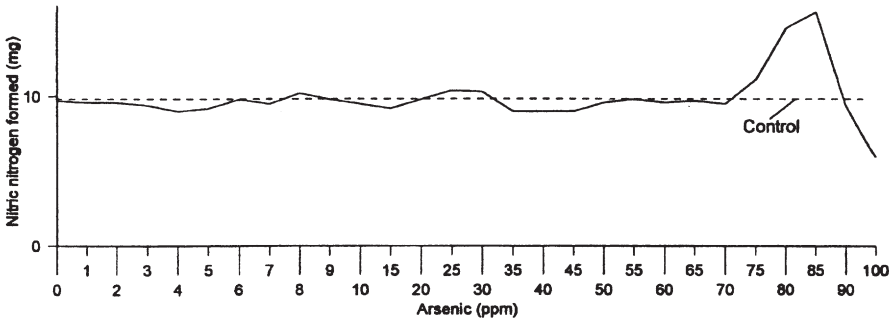
arsenic trisulfide the nitrogen-fixing organisms required the highest concentration for optimal stimulation followed by the nitrifying and ammonifying organisms, respectively. In contrast, the nitrifying organisms required the highest concentration of zinc arsenate for optimal stimulation while much smaller quantities were required for maximum stimulation of nitrogen fixation and ammonification.

While these initial investigations compared the response in the soil, follow-up experiments isolated several bacterial strains for direct assessment on their nitrogen-fixing capacity. One type of *Azotobacter* was isolated which was stimulated in studies with sodium arsenate. The general conclusion was that the stimulation was most likely the result of the organism more efficiently utilizing its source of carbon in the presence of the arsenic perhaps via the liberation of phosphous from its insoluble compounds.

Uranium In a similar fashion to arsenic investigations a number of investigators have reported a variety of stimulatory actions at low doses by uranium compounds (see[86] for a review). Of particular interest is that work reported by Stoklasa and Penkava[87] which indicates that low doses of uranyl nitrate ($UO_2(NO_3)_2$) enhance nitrogen-fixation by bacteria. In their experiment nine concentrations were evaluated along with a concurrent control. The dose-response revealed a β-curve-like relationship, with the maximum stimulation of 36.3%, the range of stimulation at least fourfold and the distance from the maximum stimulatory response to the NOAEL approximately twofold (Table 6).

Academic leadership in low dose bacterial research

The understanding of how dose affects the responses of bacterial physiology and reproduction



Arsenic (ppm)	mg Nitric Nitrogen Formed	Arsenic (ppm)	mg Nitric Nitrogen Formed	Arsenic (ppm)	mg Nitric Nitrogen Formed
0	9.7	10	9.5	60	9.6
1	9.6	15	9.2	65	9.7
2	9.6	20	9.8	70	9.5
3	9.4	25	10.4	75	11.1
4	9.0	30	10.3	80	14.5
5	9.2	35	9.0	85	15.6
6	9.8	40	9.0	90	9.4
7	9.5	45	9.0	100	6.0
8	10.2	50	9.6	----	-----
9	9.8	55	9.8	----	-----

Figure 5 Data from Greaves[83] pertaining to the nitric nitrogen (nitrification) produced in 100 grams of soil containing different amounts of arsenic in the form of sodium arsenate

has a long history and one that is relatively unambiguous. From the time of the initial claims of Hueppe there has been the repeated documentation that low concentrations of toxic substances stimulate colony growth in bacteria. While a variety of investigators participated in the unfolding of this research perhaps the most significant contributions came from the laboratory of Professor Winslow who was initially at MIT and then later at Yale University where the vast majority of his work in this area was performed. It is also important to recognize that Professor Winslow helped shape and develop a number of PhD students through whom his research ideas were developed. In fact, the next generation of researchers (i.e., his students) also contributed to the knowledge of how low dose exposures affect stimulatory responses in low doses areas. For example, Frederick W Fabian who completed his PhD under the direction of Professor Winslow in 1929 went on to become a professor of bacteriology at Michigan State University and directed his students in the low dose area as well.[69]

Summary

Of significance is that the general conclusion that low doses of toxic agents are stimulatory to bacterial systems was uncontroversially incorporated into the leading bacteriological texts over four decades. Even the field of disinfectant biology and infectious disease recognized that chemical disinfectants were stimulatory to bacterial agents in low doses.[70,71] However, given the emphasis on the killing of bacteria and on maintaining disinfectant residuals the appreciation of the biological significance of low dose stimulatory responses was overshadowed by the immediate and continuing public health significance of assuring safe drinking water and other consumer products and/or activities. Nonethe-

Table 4 Data from Greaves[83] pertaining to nitric nitrogen produced in 100 grams of soil containing different ammounts, and different forms, of arsenic

	Nitric nitrogen produced (mg) by different forms of arsenic			
Arsenic added (p.p.m.)	*Lead arsenate*	*Paris green*	*Zinc arsenite*	*Arsenic trisulfide*
0	10.5	10.4	10.2	9.5
20	17.9	10.8	10.5	10.2
40	18.7	9.4	10.8	11.3
80	14.5	9.6	10.0	13.2
120	13.5	10.0	9.9	11.3
160	13.5	9.7	10.3	11.9
200	12.5	9.9	9.3	10.2
240	11.5	10.4	10.3	10.2
280	9.5	13.4	9.7	11.6
320	10.2	15.4	9.4	9.7
360	10.3	14.4	9.9	9.2
400	9.8	13.4	10.8	9.6
440	9.7	14.0	9.0	10.0
480	9.6	14.1	14.0	8.8
520	10.5	13.4	14.4	8.3
600	10.3	13.4	15.0	6.3
640	10.0	12.4	16.8	5.4
680	9.8	10.4	12.1	6.5
720	9.5	11.0	10.0	5.0
760	9.4	9.9	9.0	4.0
800	9.3	9.7	8.6	4.0
840	8.5	9.4	7.4	3.9
880	8.4	9.4	6.8	2.2
920	8.2	9.4	6.0	2.2
960	8.0	8.4	6.8	2.0
1000	7.9	7.4	7.0	1.0
1040	7.8	7.4	6.0	1.0
1080	7.5	5.4	5.2	1.0
1120	7.1	3.4	5.3	0.8

Table 5 Quantity of nitrogen (milligrams) fixed in 100 gm of soil during 18 days with varying amounts and different forms of arsenic (data from Greaves[85])

	Quantity of nitrogen fixed (mg) by different forms of arsenic				
Arsenic (p.p.m.)	*Sodium arsenate*	*Lead arsenate*	*Paris green*	*Arsenic trisulfide*	*Zinc arsenite*
0	18.2	16.1	15.2	9.8	9.1
20	22.4	16.0	13.7	11.2	11.9
40	14.0	16.4	13.0	14.0	9.7
80	14.0	18.9	14.0	15.4	9.6
120	15.0	21.0	8.8	16.2	10.5
160	14.4	21.0	8.3	16.4	9.7
200	14.0	21.7	7.4	14.0	8.4
240	12.6	16.8	6.7	12.8	8.4
280	0	16.1	6.0	11.2	8.4
320	0	16.0	6.0	11.2	9.0
360	0	16.6	6.0	9.8	9.1
400	0	16.8	5.2	9.8	9.1

less, the broad spectrum of dose-response relationships for chemical toxicants for bacteria unmistakably recognized and explicitly incorporated the concept of low dose stimulation and its broad generalizability across the chemical classes regardless of whether the agent was an essential element, an inherent nonessential toxicant or a disinfectant agent.

Fungi

Introduction

The toxicological assessment of the role of dose on the physiology, growth, reproduction and toxic responses of fungi has a long history starting with the publication in 1869 by Raulin[4] who observed that zinc and iron sulfates enhanced the growth of

Table 6 Fixation of nitrogen by bacteria as influenced by uranyl nitrate (data from Stoklasa and Penkava[87])

Uranium content (gram atoms/liter nutrient solution)	*Nitrogen increases (per liter of solution)*
0	183.0
0.000,003,125	200.1
0.000,006,25	223.2
0.000,012,5	256.3
0.000,025	172.4
0.000,05	132.4
0.000,1	81.0
0.000,2	52.2
0.000,5	2.5
0.001	1.3

Aspergillus niger. He concluded that these stimulatory effects were manifestations of essentiality functions of these salts rather than non-nutritive stimulatory responses. While differentiating the underlying cause of low dose stimulation and the interpretation of Raulin would come to strongly influence the area of low dose stimulatory experiments of fungi, other low dose research bearing directly upon the hormetic hypothesis from independent settings was concurrently conducted. These independent contributions involved research principally to define quantitative relationships of disinfection agents on fungi (see below), bacteria,[77] and plants[15] using similar experimental methodologies and study designs. The other principal independent research direction for low dose fungal experimentation assessed the generalizability of the hypothesis of Schulz that low doses of poisons are stimulatory while high doses are inhibitory. Finally, even though yeast are classified as fungi, they will be addressed in a separate section due to their historical significance to the area of hormetic effects.

Fungal toxicity studies

The establishment of fungal dose-response data near the turn of the century was addressed principally by three investigators, including Stevens[32] at the University of Chicago, and Clark[88–90] and Duggar[91] both from Cornell University. Clark's research was completed in partial fulfilment of his MS degree. Duggar was identified by Clark as an instructor in plant physiology at Cornell to whom he acknowledged assistance in his work. Duggar in turn acknowledged the assistance of Professor Pfeffer at the University of Leipzig where his work was undertaken.[91] In later years Duggar became a professor at the University of Wisconsin and chaired a National Research Council assessment on the biological effects of radiation. He also became a faculty advisor to the well-known geneticist Alexander Hollander, who later helped start the Environmental Mutagen Society and became its first elected President.[92]

The above three researchers offered complementary insights into the effects of numerous agents over wide dosage ranges on fungal germination and/or mycelium growth and fruiting. In the case of Stevens, five fungal strains were employed to assess the effects of 24 agents on spore germination in experiments using up to eight doses. The germinating medium was distilled water for all molds except penicillin. Of the 24 agents only alcohol (i.e., with three fungal strains) and NaCl (i.e., with two fungal strains) displayed what the author considered to be convincing evidence of stimulation. Precise quantitative data were not presented which precludes the readers from being able to render their own judgment.

In the case of Clark, five fungal species were assessed with 37 compounds over 23 doses/experiment using three endpoints (i.e., spore germination, mycelium growth and the fruiting of spores).[88,89] The author utilized a beet extract for the growth medium thereby precluding any direct comparison with the results of Stevens[32] and Duggar[91] who appear to have used tap water. Of the 37 compounds tested, seven (KOH, NH_4OH, HCl, HNO_3, ethanol, H_2O_2 and HCHO) were reported to have enhanced mycelium growth. No convincing stimulatory evidence on germination or fruiting was reported. While some consideration was given to defining the dosage range (e.g., HCl) and the magnitude of stimulation (e.g., H_2O_2) and fungal stressor agents affected (e.g., ethanol, H_2O_2, HCl), the reporting of results was often incomplete.

The third paper, that of Duggar,[91] had a different goal than that of Stevens[32] and Clark[88–90] which, in part, was to define the toxicity threshold of harmful substances in molds using the research methodology of Kahlenberg and True.[15] Duggar was also interested in a broad range of other questions including one directly relevant to the present assessment. More specifically, Duggar asked: 'May a chemical irritant or poison, which is not primarily a food substance, thus function as a stimulus?[91] Using two fungal species, 18 agents were tested over 4–5 doses with usually a tenfold difference between doses. In the case of *Aspergillus flavus* all 18 compounds were found to be stimulatory while ten of the 18 agents were stimulatory in the *Sterigmatocystis* (Table 7). The *Aspergillus* model was generally much more responsive in cases where both species displayed stimulation of germination (e.g., ethanol, methanol, phenol, HCl, HNO_3, tartaric acid, $MnCl+FeSO_4$). The only case where *Sterigmatocystis* displayed noticeably more germination

Table 7 Comparison of low dose stimulatory response for 18 agents with *Aspergillus* (A) and *Sterigmatocystis* (S) by Duggar[91]

	Comparisum between A and S[a]	*Approximate range of stimulation (-fold range)*		*Maximum stimulation dose to toxic threshold (-fold range)*	
		A	S	A	S
Ethanol	75 *vs* 5–20	1000	10	>100	10
Methanol	10–75 *vs* 1	1000	<10	>1000	<10
Phenol	80 *vs* 1–5	100	~10	<10	10
Strychain	3–5 *vs* –	1000	–	CD	–
Ether	2–5 *vs* –	1000	–	>10 000	–
Camphor	50 *vs* 30–70	1000	CD	CD	CD
Petroleum	1–3 *vs* –	10	–	CD	–
Vaseline	50–75 *vs* 25–50	1000	⩾10	CD	⩾10
Clove oil	1 *vs* –	10	0	<10	0
HCl	20–30 *vs* 5–10	10	100	⩾10	~100
HNO_3	5–25 *vs* 1–3	10	<10	<10	<10
Acetic acid	20 *vs* 20	1000	⩾1000	⩾10	<100
Tauteric acid	25–50 *vs* 1–5	100	⩾1000	>100	CD
Oxalin acid	1–5 *vs* 25	10	⩾100	<10	10
$Cu(NO_2)_2$	5–10 *vs* –	100	~10	0	0
$CuSO_4$	25 *vs* 15–30	1000	~10	?	<10
$MnCl+FeSO_4$	30 *vs* ~1	1000	~10	⩾100	<10
$ZnSO_4$	1–5 *vs* –	10	0	⩾10	*0*

[a]Maximum stimulatory response (%) over controls, CD = Cannot Determine

than *Aspergillus* was with oxalic acid. In addition, the range of stimulatory concentrations were generally wider in *Aspergillus* than that of *Sterigmatocystis*. Even though a number of agents were tested in common between the Stevens and Duggar studies, their use of different fungal strains precluded direct comparison.

The collective findings of these three papers should have provided a foundation to assess how dose may affect the response of fungi. Nonetheless, it would appear that these papers, while extraordinarily useful to the broader debate on dose-*versus* stimulation and/or toxicity, had little influence on the issue of whether a stimulatory response is the result of fulfiling an essentiality function or is rather acting in a manner more consistent with the hormetic hypothesis articulated in the Arndt-Schulz Law. Note that even later reports such as that of McCallan and Wilcoxon[93] which explicitly report low dose stimulation of either germination and/or growth by hydrogen sulfide in four species of fungi provided no reference citation to these earlier stimulation responses in fungi by low doses.

Essentiality versus *non-nutritive stimulation*

One of the first reports concerning chemically induced stimulatory responses was that of Raulin who noted that certain metallic salts especially those of zinc induced a considerably enhanced fungal growth rate.[4] Based on his findings, Raulin concluded that the stimulatory response reflected the essentiality of the metallic salts for the growth of the fungi. This became a contentious point over the subsequent years as seen in the influential research of Richards[22,50] who claimed that Raulin misinterpreted his findings. Richards[50] argued that it was not necessary to invoke the claim for essentiality since later work revealed that a simpler nutrient solution than was thought necessary at the time of the Raulin's experiment is sufficient for completely normal fungal development. Since these stimulatory salts were not then believed to be acting as essential nutrients, Richards[22,50] hypothesized that the action of metallic salts noted by Raulin as well as others should be considered as a response to a chemical irritation which stimulates or accelerates the metabolic activity of the fungus. He claimed that the result of this enhanced metabolism over time will be a greater amount of metabolic product (i.e., dry substance) as compared with the controls. Richards drew the above conclusion based on experiments with three fungal strains (i.e., *Aspergillus niger, Penicillium glaucuum* and *Botrytis*) and the effect of ten agents including zinc, iron, cobalt, and nickel sulfates, NaF, LiCl, $NaSiO_3$, cocaine, morphine, and amygdalin. The estimates of the amount of stimulus induced was based on the increase in dry weight of the mycelium mass in the culture as compared with controls not receiving the stimulating agent. Despite the acceleration of the mycelium development, the treatment decreased the production of conidia, thereby leading to the tentative conclusion that a stimulation of one phase of development does not imply a stimulatory response for all functions.

The inverse relationship between growth stimulation reported by Richards[50] is consistent with the discussion by Copeland.[11] See the section on plant hormesis for a broader discussion of this issue.

The above findings set the stage for an important conceptual advance with respect to the biological significance of low dose stimulation on the physiological economy of the affected organism. This increased physiological response (i.e., the enhancement of dry weight) represented, according to Richards, an increase in the 'economic coefficient' of the organism.[50] That is, the stimulatory response permitted the organism to make use of the enhanced sugar production to increase the mycelium weight (i.e., dry substance). This conclusion was based on a series of experiments (see[94]) assessing the effects of $ZnSO_4$, $FeSO_4$ and LiCl on the growth rate of *Sterigmatocystis* and *Penicillium*. In follow-up research under the direction of Richards at Barnard College, Watterson concluded that the stimulus caused the fungus to convert more of the food into body weight and less into waste products such as oxalic acid while the respiration rate remained similar to controls after adjusting for body weight.[94]

The basic thrust of this work was carried forth by Latham[12,95] under the continued direction of Professor Richards, this time using chloroform as the chemical stimulant. As in the earlier work, small quantities of the stimulatory agent, in this case the chloroform, stimulated the growth of *Sterigmatocystis nigra* and *Penicllium*. Likewise, the enhanced growth was associated with less sugar consumption, thereby indicating an enhanced economic coefficient. The authors concluded that the chloroform was acting as a stimulant since it could not serve as a carbon source for the fungal growth.

The effects of the stimulatory response on nitrogen-fixation were also assessed by Latham[94] using zinc sulfate as the stimulant (see[50]) rather than chloroform due to greater methodological difficulties working with gaseous agents. Based on a series of experiments it was shown that *Sterigmatocystis* does fix free nitrogen and that the amount of nitrogen fixed decreases if the culture is subjected to a growth stimulatory dose.

The research of Richards[22,50] and the activities of his students (Watterson,[94] Latham[12,95]) played a significant role in establishing the legitimacy of the hypothesis that low doses of a wide range of agents enhanced not only fungal growth but metabolic efficiency. This interpretation was strengthened by the independent findings of Ono in experiments with sulfates of zinc, nickel, iron, and cobalt as well as NaF, lithium nitrate, potassium arsenate, copper sulfate and mercuric chloride.[96] Consistent with the findings of Richards' laboratory, Ono also noted that the enhanced vegetative growth reduced the development of fungal spores.[96] Likewise, Richter reported that zinc sulfate in low amounts enhanced the dry weight while at higher doses such stimulation was prevented.[97] Javillier,[98] retesting the original work of Raulin[4] and a report of Coupin,[99] under extreme care showed that dilute amounts of zinc sulfate were indeed highly stimulatory to the growth of *Sterigmatocystis*. As with Ono,[96] Javiller[98] also noted similar low dose enhancement of growth in yeasts, although the optimal doses in yeast were notably lower than observed in fungi.

Testing the Arndt-Schulz hypothesis

The Arndt-Schulz Law seems to have played no role in the initial two major research themes involving fungal responses (i.e., Stevens, Duggar, and Clark with respect to defining dose response relationships and the area of essentiality *versus* stimulation lead by Richards). However, in 1927 Niethammer, from the Plant Physiology Institute of the German University in Prag, set forth to assess the ability to generalize the Arndt-Schulz hypothesis that 'all' poisons are stimulatory in low doses using the fungal model, *Aspergillus niger*.[7] While she clearly asserted that this was the case for numerous agents, the broad generalizability of the concept was challenged. The approach of Niethammer was more critical than past attempts highlighted by the need to assess a broad range of doses able to induce both stimulatory and inhibitory responses on the dose-response continuum.[7] This lack of adequate consideration for toxic zone responses in the overall study design had characterized the earlier work of Richards and his students and therefore limited its application to directly assessing the hypothesis of Arndt-Schulz. Niethammer then highlighted the research of Bertrand and Javillier,[100] where the broad dosage range was employed and was effective in inducing both stimulatory and inhibitor zone responses thereby supporting the Arndt-Schulz maxim.[7] In addition, she noted the important observation of Richter[97] in which the stimulatory response to $ZnSO_4$ occurred after several days and then disappeared by the 6th day after treatment, thereby articulating a role for temporal factors in the dose-response relationship. Recognition of the significance of temporal factors is in accord with the more recent findings of Stebbing,[79] Townsend with plants,[40] Branham with yeast,[101] and Watkins and Winslow in bacteria.[102]

As a result of these considerations, Niethammer[7] incorporated a broad dosage range (i.e., often up to seven doses along with a concurrent control) for each inorganic and organic agent tested. She also assessed the responses over varying time intervals such as 5 or 7 days and using growth media with

varying amounts of sucrose (i.e., 2, 5, 10, 20% sucrose). By using such a wide range of chemical agents, and dosages along with varying the lengths of observations and varying growth conditions, she was able to better address the issue of generalizability of the Arndt-Schulz Law than previous attempts. Based on this more comprehensive array of experiments a low dose stimulatory and high dose inhibitory response occurred with respect to fungal culture weight for numerous inorganic agents including zinc sulfate, nickel sulfate, silver nitrate, lead nitrate, chromium sulfate and potassium and calcium chloride. This judgment was based on the mean values of three to four singular observations per treatment level. Similarly, stimulatory responses were seen for an array of organic agents including uspulum (i.e., an organic mercury compound), formaldehyde, salicylic acid and thymol.

No convincing stimulatory response was reported for $CuSO_4$ when the fungal culture was grown in 20% sucrose solution for 8 days. However, at the second lowest dose (0.00002%) in the seven dose experiment a 6.2% increase in culture weight was reported. Despite the lack of convincing stimulatory response (i.e., also negative in a 5% sucrose solution), Niethammer noted that copper sulfate had been shown to act in a stimulatory fashion under other experimental conditions such as in the stimulation of certain fruits and grains.[7] For example, Bokorny reported that the germination of barley can be stimulated via the addition of small amounts of copper sulfate (0.0025%) into the germinating bed.[103] A similar stimulatory effect was noted by Becker after soaking wheat in 0.05% copper sulfate for 1 h.[104] In addition, Noeldecher demonstrated a stimulating effect of copper on barley which was toxic at higher doses.[105]

According to Niethammer, most of the toxic substances evaluated induced 'typical, recognizable stimulatory effects' and she was successful in defining the dose-response continuum including both stimulatory and inhibitory response zones.[7] However, despite the general consistency with the Arndt-Schulz Law she indicated two principal types of exceptions to this maxim. First, she claims that stimulation only occurred in the presence of 'good' (i.e., ample) nutrition in the cases of uspulun, zinc sulfate and manganese sulfate). In these cases when the sugar solution is equal to or less than 5% only inhibitory effects were seen while at 10% the stimulatory-inhibitory dose continuum is observed. A variation of this response was seen for calcium chloride and potassium chloride when the stimulatory response without toxic effects was seen in 10% sucrose medium.

Based on these supportive data and the above apparent exceptions to full consistency with the Arndt-Schulz Law, Neithammer concluded that there is inadequate basis to claim the universal validity of the Arndt-Schulz Law.[7] More emphatically, she asserted that the belief that every toxin in an appropriately diluted solution can act as a stimulant must be emphatically rejected. Thus, the Arndt-Schulz Law was believed to have been formulated in much too general a fashion since the capacity to enhance stimulation at low doses by toxic substances appears to be affected by various physiological processes and experimental conditions. These observations are strikingly similar to the study published more than 50 years later by Vichi and Tritton[106] concerning how the chemotherapeutic agent adriamycin induced low dose stimulatory responses and how this response can be affected by culture conditions, thereby leading to the concept of optimization of conditions with respect to hormetic responses.

Summary

Taken collectively the three lines of research concerning fungal responses to low doses of toxic substances display remarkable agreement. These findings provided a broad scientific foundation to assess fungal responses within the context of the hormetic hypothesis. For example, data are available on numerous fungal species, dozens of chemical agents, various experimental conditions and over broad dosage ranges. In addition, the concept of optimization of stimulatory conditions as developed by Neithammer was so far ahead of her time that it is not even adequately appreciated today. Furthermore, the attempts to resolve the issue of essentiality *versus* non-nutritive stimulation by Richards foreshadowed this same debate in the area of plant research during the decades of the 1920s–1930s.[49,60] Likewise, the notion of an economic coefficient as developed by Richards and the enhancement of metabolic efficiency resulting from the non-nutritive stimulation was again very insightful and not apparently appreciated and adopted by low dose researchers in other areas. In fact, this concept, as with the perspective of experimental optimization by Niethammer, seems to have been missed by subsequent researchers either with respect to adopting its premise, building upon it or re-evaluating its premise in light of more sophisticated experimental techniques. Nonetheless, the general conclusion emerged that low doses of toxic agents were capable of inducing stimulatory

responses especially of fungal vegetative growth, and spore germination, even if the evolving dose-response model required the refinement of the original Arndt-Schulz maxim.

Yeast

Introduction

The principal impetus for the assessment of the modern concept of hormesis originated over a century ago by a group of microbiologists interested in determining the effects of various chemical agents on the process of fermentation following the early work of Pasteur (see[107] for a review). These investigations, which were initially undertaken by PhD students (Hoffmann,[108] Thol,[109] Gottbrecht[110]) at the university in Greifswald, determined the rate and amount of CO_2 production in mixtures of yeast cells and sugar solutions to which these substances were added. Based on research presented in the above cited dissertations and the subsequent reports of Hugo Schulz,[1,2] it was shown that high dilutions of numerous agents (e.g., thallium tartrate, mercuric chloride, iodine, bromine, arsenic acid, chromic acid, formic acid, and salicylic acid) caused a marked increase in CO_2 production while being inhibitory at higher concentrations.

Arsenic on fermentation

In the decade following the initial observations of Schulz and his associates, Buchner and Rapp assessed the effects of potassium arsenite on yeast fermentation based on its antiseptic nature and its potential utility for testing the protoplasmic nature of the agents present in yeast juice.[111–115] They found that its effect on fermentation was frequently irregular while also not being an efficient antiseptic in the concentrations which could effectively be employed. For example, even 2% of arsenious oxide, added as a potassium salt, had in many instances a marked inhibitory effect on the total fermentation.

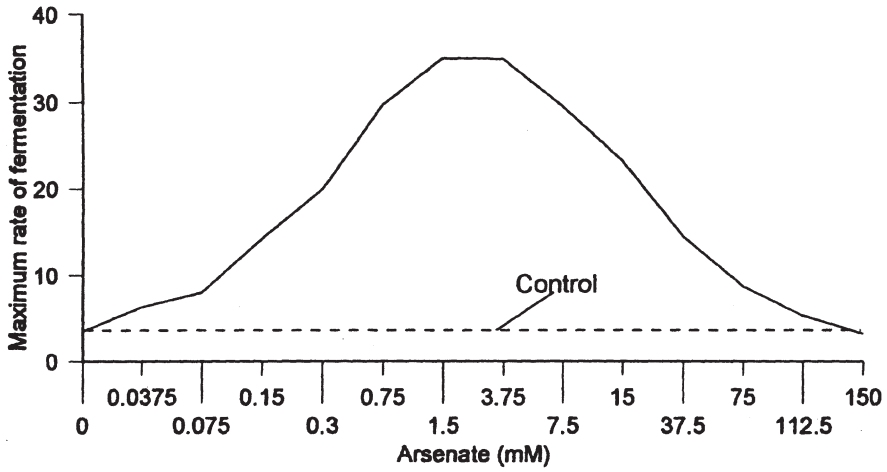
Despite these problematic aspects of the use of arsenite with respect to their original research goals, Buchner and Rapp made the critical observation that the addition of a suitable amount of arsenite caused a substantial increase in fermentation rate during the initial 16 h after treatment. While an optimum stimulatory concentration of arsenite was found, further increases in concentration lead to inhibition. These earlier studies revealed that at optimum concentrations the arsenite could enhance the yeast fermentation rate by threefold.

While research concerning the effects of arsenite on fermentation by yeast was initiated in the 1890s, it was not until 1906 when Harden and Young assessed the effect of arsenate on yeast fermentation based on its similarity with the chemical properties of phosphate.[116] Previous research with inhibitory and accelerating agents of yeast-juice fermentation revealed that phosphate induces substantial stimulation (5–10-fold) at low doses but as the concentration is progressively increased inhibitory responses become evident (see[117] for a review).

In the initial assesment of the role of arsenate on the fermentation of yeast-juice, it was shown that as an equivalent quantity of arsenate is substituted for the phosphate a comparable acceleration is produced. However, the precise response was at least quantitatively different (i.e., somewhat greater) than that of phosphate with the rate of increase 8–12 times over controls. In addition, the high fermentation rate continued for a considerable time without change; this prolonged stimulatory response was inconsistent with the earlier diminishing stimulatory response with the phosphorus treatment. Increasing the concentration beyond the optimum produces a rapid inhibition of fermentation, a phenomenon believed to result from some secondary effect on the fermenting complex (e.g., possible formation of compounds unable to combine with sugar and thereby incapable of affecting the fermentation process). Figure 6 provides the results of a typical example of the dose-response relationship of arsenate to the fermentation rate.

Having observed both arsenite and arsenate stimulation of yeast fermentation, Harden speculated that arsenic salts may affect the stimulation by either replacing the phosphate in the reaction or the arsenic salt may enhance the action of the hexosephosphate by increasing the rate of the phosphate produced thereby increasing the fermentation rate.[117] Subsequent experiments supported the hypothesis that the arsenic enhanced the rate at which the phosphate is produced from the hydrolytic action on the hexosephosphate by the action of the enzyme hexosephosphatase.

While Branham[101] concluded that the action of arsenates and arsenites by Harden and Young[116] represented the same low dose stimulatory phenomenon of the earlier work of Schulz[2] and colleagues, this is difficult to determine without the incorporation of a temporal comparison. In the Harden and Young research it was noted that the stimulatory response was evident as early as 15 min after the treatment commenced.[116] These findings suggest that the arsenic produces a direct stimulatory response without causing initial toxicity. Thus, the initial increased activity does not appear to be related to an over compensation following a disruption in homeostasis as defined by Stebbing.[79]



Arsenate (mM)	Maximum Rate of Fermentation	Arsenate (mM)	Maximum Rate of Fermentation
0	3.5	3.75	34.9
0.0375	6.3	7.5	29.5
0.075	8	15	23.2
0.15	14.2	37.5	14.5
0.3	19.9	75	8.7
0.75	29.7	112.5	5.3
1.5	35	150	3.2

Figure 6 Dose-response relationship of arsenate to the rate of fermentation (data from Harden[117])

In similar fashion, the tenfold increase in fermentation rate (by arsenate) far exceeds the up to twofold increased response typically observed in the hormetic dose-response framework.

More consistent with the typical hormetic response have been the data reported by Branham.[101,107] In her study, chemicals were assessed for their capacity to affect CO_2 production in Brewer's yeast (*Saccharomyces cerevisiae*) over a 5 h period. The study involved the testing of 16 agents, over variable numbers of dilutions (i.e., concentrations) ranging from a low of two (i.e., ethyl alcohol) to a high of eight (i.e., copper sulfate, metaphen and mercurochrome). Only ethyl alcohol had less than four doses. For each dilution there were six tubes along with concurrent controls for each chemical. In addition, measurements were taken over 15, 30 min plus 1, 2, 3, 4, and 5 h. The data, however, were not provided for the initial two time (i.e., 15 and 30 min) measurements. The dosage range over which effects were evaluated varied according to the chemical tested (i.e., from 2.0–250-fold). Of the 16 agents evaluated 14 displayed a notable stimulatory response. The magnitude of the maximum stimulatory response of the 14 agents was generally below approximately twofold with the notable exception of NaCl where the increase was 3–3.5-fold depending on the timing of the measurement.

Of particular interest was the general temporal trend in which many of the agents (e.g., $HgCl_2$, phenol, lysol, formaldehyde, iodine, hexylresorsanol) were clearly inhibited during the initial period of observation after which they responded in a strong stimulatory fashion. It was not uncommon to observe inhibitory responses of up to 50% by the

first hour after treatment with this followed by such a powerful subsequent stimulatory response that by 5 h the treatment response exceeded (i.e., was stimulated) the control by approximately 50%. Figure 7 illustrates this process for phenol.

Colchicine

From the late 1920s Professor Oscar Richards at Yale University undertook a series of investigations concerning the effects of various substances on the growth of yeasts. Of particular interest were reports on the effects of low levels of deuterium[118] and colchicine[119] since both exhibit a stimulatory response. With respect to the colchicine, Richards assessed the influence over a broad dose range (e.g., 1 p.p.m. to 4.5%) on the growth of the yeast *Saccharomyces cerevisiae.* The growth of the yeast was determined by counting the number of yeast cells by several methods (e.g., hemocytometer, photoelectric nephelometer). The low levels of colchicine induced no significant changes from the controls. However, as the concentration increased there was a notable increase in the growth response maximizing at 1% after which there was an abrupt decrease in the stimulatory responses with 4.5% approaching the control value. The dose-response followed closely to that of the β-curve. The maximum stimulating response was approximately 50% over the controls.

The stimulation of yeast colony growth was the result of the colchicine group growing directly to a maximum crop in a single cycle of growth while the control group required two growth cycles to achieve a similar growth yield.[119] At the end of the first growth cycle, there is a selective killing of the larger buds by toxic waste products in the medium and a decrease in the available food supply. If additional food is provided just prior to the end of the first cycle the retarded growth is prevented and the population will advance to a maximum crop, that is, it will only have one cycle. Based in part on this information, Richards hypothesized that the colchicine is likely acting either as a food source or by making the medium less unfavorable via a buffering type of mechanism.

Deuterium

Experiments assessing the effects of heavy (i.e., deuterium) water on the growth of yeasts revealed that both concentrations of the heavy water (i.e., 1:1, 1:3 ratios of heavy water to distilled water) enhanced the growth of the yeast as measured by dry weight. The findings indicate that the optimum dose occurred at about 1/4,000 p.p.m. while concentrations above and below were less stimulatory.[118] Other experiments by Barnes provided evidence that low doses of deuterium enhanced longevity in the yeast.[120,121]

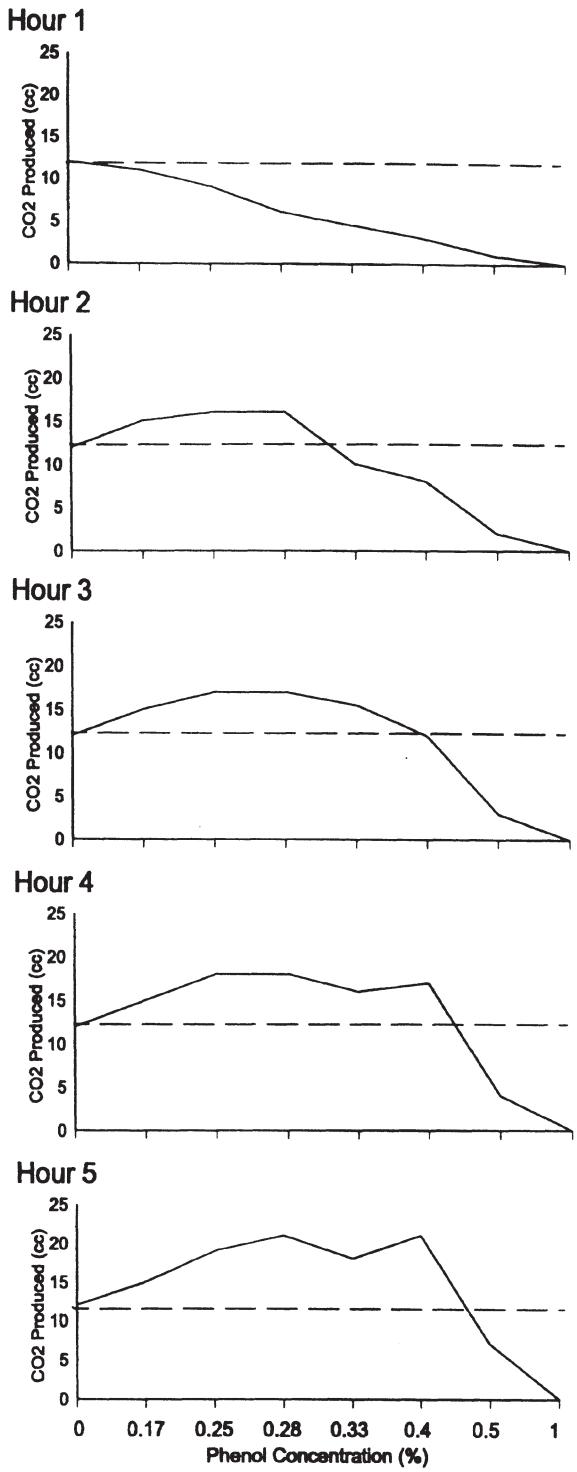


Figure 7 Effect of varying dilutions of phenol upon carbon dioxide production over time by 10% yeast and 1% sucrose (data from Branham[101])

Thallium

Other experiments by Professor Richards[122] confirmed the earlier research of Gottbrecht[110] and Schulz[123] that thallium at low doses was stimulatory to yeast. He became interested in this question principally because different brands of the nutrient media component, asparagine, provided differential yields of yeasts when all other conditions were similar. As suggested above, this difference was shown by Richards to be due to a thallium impurity which either enhanced or inhibited yeast growth depending on the concentration.

Summary

The research on yeast has played a highly significant role in the history of hormesis since the experiments performed by Schulz[1,2] utilized it as the experimental model. Despite the initial findings of Schulz and the controversial broad generalization of his theory of low dose stimulation by toxic substances, yeast did not play a major role in subsequent development and refinement of the Arndt-Schulz maxim. In fact, considerably more research on the occurrence of low dose induced stimulatory responses was observed in plants, bacteria and other fungi. This appears to be based on the nature of the investigators and their selection of experimental model rather than any known differential capacity to display hormetic effects across species. Despite this lower number of published hormetic findings in the early literature with yeasts, there are nonetheless several important areas of research beyond the initial discoveries of Schulz that warrant attention. Most notably are the observations that arsenate and arsenite are markedly stimulatory of yeast fermentation and the effect of disinfectants on yeast. Both areas are of particular interest since the arsenic induces the stimulatory response in the absence of an apparent initial inhibitory response while the disinfectants commonly induce an inhibition of growth followed by a substantial overcompensation as time progresses. Such findings clearly suggest that the hormetic dose-response curve is likely to be consistent with a variety of biochemical mechanisms.

Discussion

This paper assessed the early history of hormetic research from its origin in the late 1880s until the 1930's. This historical assessment is likewise interwoven with the evolution of experimental methods, improvements in study design and statistical analysis. While there is little doubt that the rigor of experimental biology has advanced far beyond the early discourses of the likes of Pasteur and Koch, it is important not to forget that each generation of scientists represents not a break from the past but an extension of past achievements. Thus, the subsequent observations of other scientists offered fundamental progress for improved understandings however imperfect and incomplete their findings were. In general this conceptual perspective is equally relevant to the field of hormesis.

What has emerged in the present analysis is the copious documentation of the concept that low doses of numerous toxic substances are often stimulatory to biological models. These observations have at times been reported in studies that have been poor by today's standards and yet at other times the data are based on experiments that are quite substantial even today. The remarkable feature in this process has been the independent and usually unexpected nature of the observations of stimulation. Quite importantly, these initial observations of low dose stimulatory effects were not dismissed as experimental methods improved over the decades. The fundamental reproducibility of the hormetic phenomenon was consistently confirmed and was never a stumbling block to its acceptance.

Hormesis has a long and very rich history which has been written in the pages of the most prestigious journals of the late 1880s to 1930s. It has been the object of research at the most influential universities in Europe and the US, the subject of numerous dissertations, and the principal focus of a European scientific journal, *Zell-Stimulations-Forschungen* (i.e., Cell Stimulation Research) published from 1924–1930.[124] The concept of hormesis has also had an impressive scientific lineage being traced back to highly productive students of Nobel Prize winners in biology (e.g.. Hueppe was the student of Koch) and chemistry (e.g. Kahlenberg was the student of Ostwald). Such research was not reserved to just the students of Nobel Laureates. Charles Richet, who was awarded the Nobel prize in 1913 for his clarification of the immunological concept of anaphylaxis,[125] published several highly significant papers concerning the low dose stimulatory effects of numerous toxic substances.[126,127]

Yet despite the copious examples presented in this review obtained from the most mainstream of scientific journals and texts and leading scientists, the belief that hormetic responses were real and reproducible became essentially lost to the very community of scientists who had discovered this principle and lead to its progressive refinement. Given its long and substantial history how did hormesis come to be the forgotten concept, ignored by leading retrieval systems such as Index Medicus, Current Contents, Excerpta Medica which have not

included either the term hormesis nor synonyms (e.g. Arndt-Schulz law, Hueppe's Rule, hormoligosis),[128] and omitted from leading toxicological texts of the latter decades of the 20th century? While it is true that hormetic effects continued to be reported in the peer-reviewed literature hormesis did become significantly marginalized within the scientific community to the point where it became at best a footnote and at worse an object of ridicule and disbelief. An accompanying article explores the history of the development of the dose-response relationship, the issues that framed its development and how the concept of hormesis became excluded from its formulation and application in the field of toxicology and risk assessment.

Acknowledgements

This article was previously published in *Toxicologic Pathology* 1999; **27**:195–216 and has been reproduced in *Human and Experimental Toxicology* with the approval of *Toxicologic Pathology* and the Society of Toxicologic Pathologists. Please note that Figures 2–6 represent modifications of Tables 4–7 and 12 in the original publication. This work has been sponsored in part by an award to the University of Massachusetts (Edward J Calabrese, PhD, Principal Investigator) by the Texas Institute for Advancement of Chemical Technology, Inc. (TIACT). The text is based upon Appendix 7 of *Hormesis: the Stimulatory Action of Sub-Inhibitory Amounts of Toxins*, a report prepared for TIACT.

References

1 Schulz H. Zur Lehre von der Arzneiwirdung. *Virchows Archiv fur Pathologische Anatomie und Physiologie fur Klinische Medizin* 1887; **108:** 423–445.
2 Schulz H. Uber Hefegifte. *Pflugers Archiv fur die gesamte Physiologie des Menschen und der Tiere* 1888; **42:** 517–541.
3 Reveil (no first initial). Recherches de physiologie vegetale. In: *De l'action des poisons sur les plantes.* Paris, 1865.
4 Raulin J. Etudies chimiques sur la vegetation. *Annales des Sciences Naturelles, Botanique et Biologie Vegetale* 1869; **11:** 93–299.
5 Storp F. Ueber den Einfluss von Kochsalz- und Zincsulfathaltigen Wassen auf Boden und Pflanzen. *Landwirtschaftliche Jahrbucher* 1883; **12:** 795–844.
6 Hofmann P. *Ueber die Gultigkeit des Arndt-Schulzschen biologischen Grundgesetzes bei der Wirkung von Bakteriengiften.* Inaugural Dissertation, tierarztl. Fakultat. Munchen, 1922 (Translated from the German by University of Massachusetts translation services).
7 Niethammer A. The stimulation effect of toxins on fungus and the Arndt-Schulz axion. *Biochemische Zeitschrift* 1927; **184:** 370–382. (Translated from the German by University of Massachuestts translation service).
8 Hueppe F. Principles of Bacteriology. Translated from the German by E.O. Jordan. The Open Court Publishing Company: Chicago, IL, 1896.
9 Suzuki S. On the action of vanaiun compounds on plants. *Bulletin of the College of Agriculture of Japan* 1902–1903a; **5:** 513–515.
10 Suzuki S. On the action of highly diluted potassium iodide on agricultural plants. *Bulletin of the College of Agriculture of Japan* 1902–1903b; **5:** 199–201.
11 Copeland EB. Chemical stimulation and the evolution of carbon dioxide. *Botanical Gazette* 1903; **35:** 81–183.
12 Latham ME. Stimulation of Sterigmatocystis by chloroform. *Bulletin of the Torrey Botanical Club* 1905; **32:** 337–351.
13 Hayes WJ. Factors influencing toxicity. In: *Toxicology of Pesticides.* The Williams & Wilkins Co.: Baltimore, 1975.
14 Servos JW. Physical Chemistry from Ostwald to Pauling. Princeton University Press: Princeton, NY, 1990.
15 Kahlenberg L, True RH. On the toxic action of dissolved salts and their electrolytic dissociation. Preliminary paper. *Journal of the American Medical Association* 1896; July 18.
16 Kahlenberg L, True RH. On the toxic action of dissolved salts and their electrolytic dissociation. *Botanical Gazette* 1896; **22:** 81–124.
17 Porter TM. The Rise of Statistical Thinking, 1820–1900. Princeton University Press: Princeton, NY, 1986.
18 Heald FD. On the toxic effect of dilute solutions of acids and salts upon plants. *Botanical Gazette* 1896; **22:** 125–153.
19 True RH . The toxic action of a series of acids and of their sodium salts on *Lupinus albus. American Journal of Science IV,* 1900; **9:** 183.
20 Copeland EB, Kahlenberg L. The influence of the presence of pure metals upon plants. *Wisconsin Academy of Sciences, Arts and Letters* 1899; **12:** 454–474.
21 Nageli C. Ueber oligodynamische Erscheinungen in labenden Zellen. *Denkschr. Schweiz Naturgesell.* 1893; **33:** 1–43.
22 Richards HM. Die Beeinflussung des Wachsthums einiger Pilze durch chemische Reinze. *Jahrbucher fur wissenschaftliche Botanik* 1897; **30:** 665–688.
23 Guenther E. Beitrag zur mineralischen Nahrung der Pilze. Inaugural Dissertation, Erlangen. *Ref. Botanisches Centralblatt* 1897; **74:** 194.
24 Calabrese EJ, Baldwin LA. A quantitatively-based methodology for the evaluation of chemical hormesis. *Human and Ecological Risk Assessment* 1997; **3:** 545–554.
25 Jacobi B. Ueber den Einfluss verschiedener Substanzen auf die Athmung und Assimilation suberser Pflanzen. *Flora* 1899; **86:** 289–327.
26 Klebs G. Einige Ergebnisse der Fortpflanzungs-Physiologie. *Berichte der Deutschen Botanishen Gesellschaft* 1900; **18:** 201–215.

27 True RH, Gies WJ. On the physiological action of some of the heavy metals in mixed solutions. *Bulletin of the Torrey Botanical Club* 1903; **30:** 390–402.
28 True RH, Oglevee CS. The effect of the presence of insoluble substances on the toxic action of poisons. *Botanical Gazette* 1905; **39:** 1–21.
29 Fitch R. The action of insoluble substances in modifying the effect of deleterious agents upon the fungi. *Annales de Chimie et de Physique* 1906; **4:** 313–322.
30 Jensen GH. Toxic limits and stimulation effects of some salts and poisons on wheat. *Botanical Gazette* 1907; **43:** 11–44.
31 Schreiner O, Reed HS. The toxic action of certain organic plant constituents. *Botanical Gazette* 1908; **45:** 73–102; 271.
32 Stevens FL . The effect of aqueous solutions upon the germination of fungus spores. *Botanical Gazette* 1898; **26:** 377–406.
33 Schreiner O, Reed HS. Some factors influencing soil fertility. *US Department of Agriculture Bureau of Soil,* Bulletin No. 40, 1907.
34 Suzuki S, Aso K. On the physiological action of iodine and fluorine compounds on agricultural plants. Part. A. On the influence of iodide on oats. Part. B. On the influence of potassium iodide on radish. Part C. On the influence of sodium fluoride on oats. *Bulletin of the College of Agriculture of Japan* 1902–1903; **5:** 473–479.
35 Aso K. On a stimulating action of calcium fluoride on phaenograms. *Bulletin of the College of Agriculture, Tokyo.* 1906/1908; **7:** 90.
36 Suzuki S. On the poisonous action of potassium ferrocyanid on plants. *Bulletin of the College of Agriculture of Japan* 1902–1903; **5:** 203–205.
37 Townsend CO. The correlation of growth under the influence of injuries. *Annals of Botany* 1897; **11:** 509–532.
38 Stebbing ARD. Hormesis–the stimulation of growth by low levels of inhibitors. *Science of the Total Environment* 1982; **22:** 213–234.
39 Calabrese EJ, Baldwin LA. The dose determines the stimulation (and poison): Development of a chemical hormesis database. *International Journal of Toxicology* 1997; **16:** 545–559.
40 Townsend CO. The effect of ether upon the germination of seeds and spores. *Botanical Gazette* 1899; **27:** 458–466.
41 Townsend CO. The effect of hydrocyanic acid gas upon grains and other seeds. *Maryland Agricultural Experiment Station Bulletin* 1901; **75:** 183–198.
42 Townsend CO. The effect of hydrocyanic acid gas upon the germination of seeds. *Scientific American Supplement* 1899; **48:** 20010–20011.
43 Townsend CO. The effect of hydrocyanic acid gas upon grains and other seeds. *Botanical Gazette* 1901; **31:** 241–264.
44 Moore W, Willaman JJ. Studies in greenhouse fumigation with hydrocyanic acid: physiological effects on the plant. *Journal of Agricultural Research* 1917; **11:** 319–339.
45 Clayton EE. Hydrogen cyanide fumigation. *Botanical Gazette* 1919; **67:** 483–500.
46 Evans WH. Copper sulphate and germination. *US Department of Agriculture Division of Vegetable Physiology and Pathology,* Bulletin. No. 10, 1896.
47 Rumm C. Ueber die Wirkung der Kapferpraparate bei Bekampfung der sogenannten Blattfallkrenkheit der Weinrebe. *Botanische Zeitung* 1893; **51:** 163–164.
48 Fairchild DG. Bordeaux mixture as a fungicide. *US Department of Agriculture Division of Vegetable Pathology,* Bulletin No. 6, 1894.
49 Brenchley WE. Inorganic Plant Poisons and Stimulants. Cambridge University Press: London, 1914.
50 Richards HM. The effect of chemical irritation on the economic coefficient of sugar. *Bulletin of the Torrey Botanical Club* 1899; **26:** 463–479.
51 Lipman CB, Wilson FH. Toxic inorganic salts and acids as affecting plant growth. *Botanical Gazette* 1913; **55:** 409–420.
52 Haywood JK. Injury to vegetation by smelter fumes. *United States Department of Agriculture Bureau of Chemistry,* Bulletin No. 89, 1905.
53 Haywood JK. Injury to vegetation and animal life by smelter fumes. *United States Department of Agriculture Bureau of Chemistry,* Bulletin No. 113 1908.
54 Haywood JK. Injury to vegetation and animal life by smelter fumes. *United States Department of Agriculture Bureau of Chemistry,* Bulletin No. 113 revised, 1910.
55 Johnson B. On the influence of arsenic on the germination of seeds. *Experimental Station Record* 1896–1897; **8:** 232.
56 Harter LL. The variability of wheat varieties in resistance to toxic salts. *United States Department of Agriculture, Bureau of Plant Industry* Bulletin No. 79, 1905.
57 Brenchley WE. Inorganic Plant Poisons and Stimulants, Second Edition. Cambridge University Press: London, 1927.
58 Lindsay J. The elimination of algae in lochs and ponds. *Transactions of the Edinburgh Field Nat. Micros. Society* 1913; **6:** 422–431.
59 Allison RV, Bryan OC, Hunter JH. The stimulation of plant response on the raw peat soils of the Florida Everglades through the use of copper sulphate and other chemicals. *Florida Agricultural Experimental Station Bulletin* 1927; **190:** 35–80.
60 Brenchley WE. The essential nature of certain minor elements for plant nutrition. *The Botanical Review* 1936; **2:** 173–196.
61 Kearney TH, Cameron FK. Some mutual relations between alkali soils and vegetation. *United States Department of Agriculture,* Report No. 71. US Gov. Printing Office: Washington, DC, 1902.
62 Hune, (no first initial) (1909). Die begunstigende Reizwirkung kleinster Mengen von Bakteriengiften auf die Bakterienvermehrung. *Centralblatt fur Bakteriologie I. Abt. Orig.* 1909; **48:** 135.
63 Delepine AS, Greenwood A. The action of some metals upon certain water and other bacteria. *Royal Sanitation Institute Journal* 1914; **35:** 317–327.
64 Hotchkiss M. The Influence of Various Salts Upon the Growth of Bacterium communis. Doctoral Dissertation, Yale University, 1923.
65 Hotchkiss M. Studies on salt action. VI. The stimulating and inhibitive effect of certain cations upon bacteria growth. *Journal of Bacteriology* 1923; **8:** 141–162.

66 Winslow C-EA, Dolloff AF. Relative importance of additive and antagonistic effects of cations upon bacterial viability. *Journal of Bacteriology* 1928; **15:** 67–92.
67 Winslow C-EA, Haywood ET. The specific potency of certain cations with reference to their effect on bacterial viability. *Journal of Bacteriology* 1931; **22:** 49–69.
68 Fabian FW, Winslow C-EA. The influence upon bacterial viability of various anions in combination with sodium. *Journal of Bacteriology* 1929; **18:** 265–291.
69 Fabian FW, Bryan CS. The influence of cations on aerobic sporogenesis in a liquid medium. *Journal of Bacteriology* 1933; **26:** 543–558.
70 Marshall MS, Hrenoff AK. Bacteriostasis. *Journal of Infectious Diseases* 1937; **61:** 42–54.
71 Salle AJ. Fundamental Principles of Bacteriology. McGraw-Hill Book Co., Inc.: New York, 1939, pp. 166–197.
72 Clifton CE. Introduction to Bacterial Physiology. McGraw-Hill Book Company, Inc.: New York, 1957, pp. 317–338.
73 Thimann K. The Life of Bacteria, Second Edition. The Macmillan Co.: New York, 1963, pp. 154–192.
74 Lamanna C, Mallette MF. Basic Bacteriology Its Biological and Chemical Background, Third Edition. The Williams & Wilkins Company: Baltimore, MD, 1965.
75 Falk IS. The role of certain ions in bacterial physiology. A review (Studies of Salt Action. VII). *Abstracts of Bacteriology* 1923; **7:** 33–50; 87–105; 133–147.
76 Winslow C-EA, Hotchkiss M. Studies on salt action. V. The influence of various salts upon bacteria growth. *Proceedings of the Society for Experimental Biology and Medicine* 1922; **19:** 314–315.
77 Eisenberg P. Untersuchgen uber spezigische desingektionsvorgange. II Mitt. Uber die wirkung von salzer und ionen auf bakterien. *Centralblatt fur Bakteriologie I. Abt. Orig.* 1918; **82:** 69–208.
78 Miller WS, Green CA, Kitchen H. Biphasic action of penicillin and other similar sulphonamides. *Nature* 1945; **155:** 210–211.
79 Stebbing ARD. A theory for growth hormesis. *BELLE Newsletter* 1997; **6:** 1–11.
80 Lipman CB. Toxic and antagonistic effects of salts as related to ammonification by Bacillus Subtilis. *Botanical Gazette* 1909; **48:** 105–125.
81 Lipman CB. Toxic effects of 'Alkali Salts' in soils on soil bacteria. I. Ammonification. *Centrlblatt fur Bakteriologie* 1912; **32:** 5864.
82 Brown PE, Hitchcock EB. The effects of alkali salts on nitrification. *Soil Science* 1917; **4:** 207–229.
83 Greaves JE. The influence of arsenic upon the biological transformation of nitrogen in soils. *Biochemical Bulletin* 1913; **3:** 2–16.
84 Greaves JE, Carter EG. Influence of sodium arsenite on microflora of soil. *Botanical Gazette* 1924; **77:** 63–72.
85 Greaves JE. Stimulating influence of arsenic upon the nitrogen-fixing organisms of the soil. *Journal of Agricicultural Research* 1913; **6:** 389–416.
86 Buchanan RE, Fulmer EI. Physiology and Biochemistry of Bacteria. Vol. II. Effects of Environment Upon Microorganisms. The Williams and Wilkins Company: Baltimore, MD, 1930.
87 Stoklasa J, Penkava J. Biologie des Uraniums. *Biochemische Zeitschrift* 1928; **194:** 15–77.
88 Clark JF. Electrolytic dissociation and toxic effect. *Journal of Physical Chemistry* 1899; **3:** 263–316.
89 Clark JF. On the toxic effect of deleterious agents on the germination and development of certain filamentous fungi. *Botanical Gazette* 1899; **28:** 289–327, 378–404.
90 Clark JF. On the toxic properties of some copper compounds with special reference to bordeaux mixture. *Botanical Gazette* 1902; **33:** 26–48.
91 Duggar BM. Physiological studies with reference to the germination of certain fungous spores. *Botanical Gazette* 1901; **31:** 38–66.
92 Wassom JS. Origins of genetic toxicology and the environmental mutagen society. *Environmental and Molecular Mutagenesis* 1989; **16**(Supplement): 1–6.
93 McCallan SEA, Wilcoxon F. The fungicidal action of sulphur. II. The production of hydrogen sulphide by sulphured leaves and spores and its toxicity to spores. *Contributions to the Boyce Thompson Institute* 1931; **3:** 13–38.
94 Watterson A. The effect of chemical irritation on the respiration of fungi. *Bulletin of the Torrey Botanical Club* 1904; **31:** 291–303.
95 Latham ME. Nitrogen assimilation of *Sterigmatocystis nigra* and the effect of chemical stimulation. *Bulletin of the Torrey Botanical Club* 1909; **36:** 285.
96 Ono N. Notes on the stimulating effect of certain substances upon the growth of algae and fungi. *Botanical Magazine of Tokyo* 1900; **14:** 75–78.
97 Richter A. Zur Frage der chemischen Reizmittel. *Centralblatt fur Bakteriologie* 1901; **7:** 417–429.
98 Javillier M. (1907). Sur l'influence favorable del petites doses de zinc sur la vegetation de *Sterigmatocystis nigra* V. Tgh. *Comptes Rendus* 1907; **145:** 1212–1215.
99 Coupin H. Sur la nutrition du sterigmatocystis nigra. *Comptes Rendus* 1903; **136:** 392–394.
100 Bertrand G, Javillier M. Influence du manganese sure le developpement de l'Aspergillus niger. *Comptes Rendus Academie des Sciences* 1911; **152:** 225–228.
101 Branham SE. The effects of certain chemical compounds upon the course of gas production by baker's yeast. *Journal of Bacteriology* 1929; **18:** 247.
102 Watkins JH, Winslow C-EA. Factors determining the rate of mortality of bacteria exposed to alkalinity and heat. *Journal of Bacteriology* 1932; **24:** 243–265.
103 Bokorny T. Uber den Einfluss verschiedener Substanzen auf die Keimung der Pflanzensamen. Wachstums Forderung durch einige. *Biochemische Zeitschrift* 1913; **51:** 1–48.
104 Becker A. Ueber den Einfluss der Samenbehandlung mit Reizchemikalien auf die Keimung und das Wachstum. *Landwirtschafliches Jahrbuch* 1926; **63:** 501–561.

105 Noeldechen J. Uber die Reizwirkung von Metallsalzen auf die Keimung von Gerste. *Kuehn-Archive* 1925; **9:** 264–309.
106 Vichi P, Tritton TR. Stimulation of growth in human and murine cells by adriamycin. *Cancer Research* 1989; **49:** 2679–2682.
107 Branham SE. An improved technique for the comparison of antiseptics by yeast fermentation. *Journal of Infectious Diseases* 1929; **44:** 142–149.
108 Hoffman G. Experimentalle Untersuchungen uber die Wirkung der Aneisensaure. Inaugural Dissertation, Greifswald, 1884.
109 Thol W. Ueber den Einfluss organischer, nichte aromatischer Sauren auf Gahrung und Faulnis. Inaugural Dissertation, Greifswald, 1885.
110 Gottbrecht C. Experimentalle Untersucheunge uber die Wirkung Thalliums. Inaugural Dissertation, Greifswald, 1886.
111 Buchner E, Rapp R. Alkoholische Gahrung ohne Hefezellen III. *Berichte der Deutschen Chemischen Gesellschaft* 1897; **30:** 2668–2678.
112 Buchner E, Rapp R. Alkoholische Gahrung ohne Hefezellen IV. *Berichte der Deutschen Chemischen Gesellschaf* 1898; **31:** 209–217.
113 Buchner E, Rapp R. Alkoholische Gahrung ohne Hefezellen V. *Berichte der Deutschen Chemischen Gesellschaf* 1898; **31:** 1084–1090.
114 Buchner E, Rapp R. Alkoholische Gahrung ohne Hefezellen VI. *Berichte der Deutschen Chemischen Gesellschaf* 1898; **31:** 1090–1094.
115 Buchner E, Rapp R. Alkoholische Gahrung ohne Hefezellen IX. *Berichte der Deutschen Chemischen Gesellschaf* 1899; **32:** 2086–2094.
116 Harden A, Young WJ. The influence of sodium arsenate on the fermentation of glucose by yeast-juice. (preliminary notice). *Proceedings of the Chemical Society* 1906; **22:** 283–284.
117 Harden A. Alcoholic Fermentation. Longmans, Green and Co.: New York, 1911.
118 Richards OW. The effects of deuterium on the growth of yeast. *Journal of Bacteriology* 1934; **28:** 289–294.
119 Richards OW. Colchicine stimulation of yeast growth fails to reveal mitosis. *Journal of Bacteriology* 1938; **36:** 187–195.
120 Barnes TC. Further observations on the physiological effect of the heavy hydrogen isotope on spirogyra. *American Journal of Botany* 1933; **20:** 681–682.
121 Barnes TC. A possible physiological effect of the heavy isotope of H in water. *Journal of the American Chemical Society* 1933; **55:** 4432–4433.
122 Richards OW. The stimulation of yeast growth by thallium, a 'bios' impurity of asparagine. *Journal of Biological Chemistry* 1932; **96:** 405–418.
123 Schulz H. Die Wirkung der Thallinsalze auf Faulnis und Garung. *Centr Med Wissensch* 1886; **24:** 113–115.
124 *Zell-Stimulations-Forschungen.* **Vol. 1–3.** M Popoff, M. and Gleisberg W (eds.). Paul Parey: Berlin, 1924–1930.
125 Sourkes T. Nobel Prize Winners in Medicine and Physiology: 1901–1905. Abeland-Schuman: New York, 1966.
126 Richet C. De l'action de doses minuscules de substances sur la fermentation lactique. *Archives Internationales de Physiologie* 1905; **3:** 203–217.
127 Richet C. De l'action de doses minuscules de substance sure le fermentation lacique-troisieme memoire–Periodes d'acceleration et ralentissement. *Archives Internationales de Physiologie* 1906–1907; **4:** 18–50.
128 Oberbaum M, Cambar J. Hormesis: Dose-dependent reverse effects of low and very low doses. In: Ultra High Dilutions. PC Endler and J Schultze (eds.) Kluver Academic Publishers: Dordrecht, The Netherlands, 1994.

This article has been reproduced in *Human & Experimental Toxicology* with the approval of *Toxicologic Pathology* and the Society of Toxicologic Pathologiests

Regulatory Toxicology and Pharmacology 61 (2011) 73–81

Contents lists available at ScienceDirect

Regulatory Toxicology and Pharmacology

journal homepage: www.elsevier.com/locate/yrtph

The hormesis database: The occurrence of hormetic dose responses in the toxicological literature

Edward J. Calabrese *, Robyn B. Blain

Department of Public Health, Environmental Health Sciences Division, School of Public Health an Health Sciences, Morrill I, N344, University of Massachusetts, Amherst, MA 01003, USA

ARTICLE INFO

Article history:
Received 4 March 2011
Available online 15 June 2011

Keywords:
Hormesis
Biphasic
Inverted U-shaped
J-shaped
Hormetic
Plasticity
Dose–response
Adaptive response

ABSTRACT

In 2005 we published an assessment of dose responses that satisfied *a priori* evaluative criteria for inclusion within the relational retrieval hormesis database (Calabrese and Blain, 2005). The database included information on study characteristics (e.g., biological model, gender, age and other relevant aspects, number of doses, dose distribution/range, quantitative features of the dose response, temporal features/repeat measures, and physical/chemical properties of the agents). The 2005 article covered information for about 5000 dose responses; the present article has been expanded to cover approximately 9000 dose responses. This assessment extends and strengthens the conclusion of the 2005 paper that the hormesis concept is broadly generalizable, being independent of biological model, endpoint measured and chemical class/physical agent. It also confirmed the definable quantitative features of hormetic dose responses in which the strong majority of dose responses display maximum stimulation less than twice that of the control group and a stimulatory width that is within approximately 10–20-fold of the estimated toxicological or pharmacological threshold. The remarkable consistency of the quantitative features of the hormetic dose response suggests that hormesis may provide an estimate of biological plasticity that is broadly generalized across plant, microbial and animal (invertebrate and vertebrate) models.

1. Introduction

Hormesis is a dose–response phenomenon in which opposite effects are observed at low, compared to high, doses for the same measured parameter. This will result in either an inverted U-shaped or a J-shaped dose–response curve. The concept of hormesis has received considerable interest in the toxicological, pharmacological and general biomedical areas over the past 10–15 years (Calabrese, 2008; Calabrese and Baldwin, 2001a, 2003b). For example, in the entire decade of the 1980s the Web of Science database reported about 10–15 citations per year for the terms hormesis or hormetic. However, in 2010 alone the number of citations was 3269 with publications in over 100 journals covering a broad range of biomedical disciplines. In 1997 Calabrese and Baldwin (1997a) reported on the creation of a relational retrieval hormesis database along with *a priori* entry criteria and numerous study parameters on which data would be entered. In 2005, Calabrese and Blain published the results of a detailed assessment of the database which contained nearly 5000 dose responses (Calabrese and Blain, 2005). The findings indicated that hormetic dose responses were observed in a broad range of biological models (i.e., from plant to human), occurring over a diverse set of biological endpoints and across a wide range of chemical classes. The analysis also revealed that the stimulation in the low-dose zone was typically modest with the maximum stimulation being generally 30–60%. The overall findings were important since they demonstrated that hormetic dose responses were reproducible and broadly generalized. Several related publications extended these findings, providing a frequency estimate of hormesis within the toxicological and pharmacological literature (Calabrese and Baldwin, 2001b, 2003a; Calabrese et al., 2006, 2008). These studies also revealed that the hormetic dose response was far more common than the threshold and linear dose response models in direct comparisons using the same *a priori* entry and evaluative criteria. The present paper extends the 2005 study of Calabrese and Blain by presenting an updated analysis of the hormesis database that has approximately 9000 dose responses. The updated analysis strengthens the basic findings of the original paper (Calabrese and Blain, 2005) with respect to the conclusion that hormesis is highly generalized with no apparent restriction to biological model, endpoint, or chemical classes.

1.1. Database entry criteria

The hormesis database was created using a specific set of *a priori* criteria designed to:

* Corresponding author. Fax: +1 413 545 4692.
E-mail address: edwardc@schoolph.umass.edu (E.J. Calabrese).

0273-2300/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.yrtph.2011.06.003

(1) identify probable cases of hormesis,
(2) assess the quantitative features of the hormetic dose response, and
(3) assess the generalizability of the hormetic phenomenon according to biological model, endpoint, and chemical class and physical stressor agents.

The hormesis database inclusion criteria have been previously published (Calabrese and Blain, 2005) and will only be discussed briefly. They include: (1) a minimum of 10% stimulation (i.e., inverted U-shaped dose response) or a 3% depression of response (i.e., J-shaped dose response). (2) If the 10% or 3% selection criteria were not satisfied, a dose response could have been entered into the database if the response achieved statistical significance in hypothesis testing. (3) The study had to employ an adequate concurrent control for comparison. As a general rule, dose responses that satisfied the above three criteria, satisfied hypothesis testing requirement, and provided the most detailed assessment of response both above (toxic zone) and below (hormetic stimulatory zone) the threshold were prioritized for selection into the database. Such dose responses offered the best opportunity to provide more detailed quantitative assessment of the broader dose response continuum, as well as a more robust assessment of overall dose responses.

1.2. Database scoring methodology

Each dose response entered into the hormesis database was scored according to Calabrese and Baldwin (1997a,b) and Calabrese and Blain (2005) in which numerical values were awarded based on the quality of the study design, response magnitude, statistical analysis, and reproducibility of the response in order to evaluate the capacity of an experiment for demonstrating hormesis (i.e., strength of evidence). Points were awarded for (1) the number of doses below the zero equivalent point (ZEP; location where response crosses the control value), (2) estimation of a ZEP (i.e., curve crosses the control value or was such that the curve would estimate where it would cross the threshold), (3) the number of statistically significant responses below the ZEP, (4) the magnitude (percent of control value) of the stimulatory response, and (5) reproducibility of the data by other studies with data provided (in other publications or within the same publication). The points are summed and the evidence of hormesis is awarded as indicated in Table 1.

This scoring methodology rewarded studies that explicitly considered below zero equivalent point (ZEP) or NO(A)EL doses in their study designs with the balance of the points being skewed in favor of response over design. While only a single point is awarded for dose responses that identify or could estimate a ZEP, the points for response progressively increase from one for a minimal entry starting at 10% (or 3%) up to four points for 400% (or 0%) of the control. Because increases of 400% or greater than the control may represent a different phenomenon than hormesis, the points were arbitrarily capped at four for responses 400% of the control or greater. Statistical significance was also emphasized in the scoring since hypothesis testing considers the sample size, variability, and magnitude of response in a reliable and nonbiased manner.

1.3. Description of the database

The hormesis database contains 2527 citations with 8962 dose responses (i.e., endpoints). The articles have been obtained mainly through extensive searches through numerous journals,

Table 1
Summary of criteria with assigned point values used in the quantitative evaluation of hormesis.

Doses below ZEP (n)	Point value	ZEP determined/estimated	Point value
Study design criteria			
1	1	Yes	1
2	2	No	0
3	3	–	–
4	4	–	–
≥5	5	–	–
Doses statistically significant (n)	Point value	Reproducibility	Point value
Response criteria			
1	2	Yes	3
2	4	No	0
3	8	–	–
≥4	16	–	–
Inverted U-shaped curve	J-shaped curve		Point value[a]
Magnitude of response (percentage control value)			
≥110% ⩽ 125%	⩽97% ⩾ 92%		0.5
>125% ⩽ 150%	<92% ⩾ 84%		1
>150% ⩽ 200%	<84% ⩾ 68%		2
>200% ⩽ 400%	<68% ⩾ 5%		3
>400%	<5%		4
Total point range		Hormesis evidence category	
Summary of total point ranges			
1–2		No–low	
>2–8		Low	
>8–12		Low–moderate	
>12–16		Moderate	
>16–20		Moderate–high	
>20		High	

[a] The point value is multiplied by the number of experimental doses falling within the corresponding percentage range. For example, if an experiment has three doses exhibiting stimulatory responses within the 125–150% range with the curve approaching the ZEP and two of the responses achieve statistical significance and one does not, the total number of points would be: 3 × 1 = 3 (Doses below the ZEP; study design criteria); 1 (estimated ZEP; study design criteria); 4 (response criteria, statistically significant); 3 × 1 = 3 (magnitude of response), for a total of 11 points. The 11 points would achieve the categorization of hormesis evidence of "low–moderate".

Table 2
Articles in the hormesis database by year of publication.

Publication year	Number of articles	Percent of total (2527)
<1930	39	1.5
1930–1939	30	1
1940–1949	45	2
1950–1959	95	4
1960–1969	158	6
1970–1979	429	17
1980–1989	505	20
1990–1999	706	28
2000–2010	520	20.5

cross-referencing journal citations, MEDLINE, Web of Science and other database searches using multiple words such as hormesis, hormetic stimulation, inverted U-shaped, J-shaped, biphasic, bell-shaped, and others. The articles have occurred in a diverse array of journal publications (nearly 500 different journals), although several books (51 citations), theses and dissertations are included. Information on the citation, chemical, biological model, study duration, treatment, endpoint, and dose–response curve are entered into the database (see Calabrese and Blain, 2005 for specific entry fields). Queries can be conducted using any of the fields. The query system was employed to yield the descriptive assessments offered in this article.

1.4. General

The database is arranged so that each citation is associated with studies (i.e., experiments). Data on gender and different species are considered separate studies as are different variations in study design. Each study may have examined several endpoints (e.g., body weight, survival). Each endpoint has a dose response associated with it. Any search performed with the database provides the number and percent of total for citations, studies, and endpoints. The database includes experimental findings from 1899 to the present (Table 2). However, the majority (85%) of the articles in the database were published after 1970.

Reproducibility was difficult to implement because of the uncertainty over what constitutes a bona fide case. It was decided that reproducibility should only be claimed in cases where the follow-up study was essentially identical to the original study. This would explain why only nine citations (with a total of 15 responses) were determined to be reproducible between publications. Reproducibility reported within the same article occurred with 258 dose responses in 133 different articles. However, a dose response is only considered reproducible when the article provided the results of the separate experiments. Cases where data were combined and reported as averages or where the study author claimed that the results were reproducible and only the results from a representative study were provided were not considered reproducible in the database due to lack of data for confirmation.

2. Study design considerations

2.1. Agent

Nearly 2000 different agents from approximately 245 different chemical classes have been entered into the database, about twice as many as reported in 2005. While 81% (i.e., 7216) of the 8962 dose responses entered into the database used a chemical agent, 19% (i.e., 1746) employed radiation or radioactive material. Table 3 provides the chemical classes and physical agents with the greatest frequency in the database. Calabrese and Blain (2004) have examined the hormetic response of metals in greater detail elsewhere.

2.2. Model

In the 2005 publication, plants and animals were equally represented in the database. However, in the last 5 years more animal studies have been entered with animal models predominating

Table 3
Most prevalent chemical classes and physical agents in the hormesis database.

Chemical class/physical agents	Number of dose responses	% of total (8962)
Inorganics (including elements and metals)	1717	19
Radiation/radionuclides	1746	19
Organics	5499	62
Organophosphate/pesticides/herbicides/fungicides	573	6
Hormones/amino acids/fatty acids/enzymes/protein/neurotransmitters/neuropeptides/peptides/steroids	425	5
Alcohols/phenols	410	5
Carboxylic acids	349	4
Heterocyclic compounds	233	3
Chlorinated hydrocarbons/halogenated hydrocarbons/chlorinated furanone	202	2
Organometals	171	2
Nitrate/nitrile/nitro compounds/nitrofurans/nitrosamines/nitroso compounds/nitrosoureas	163	2
Hydrocarbons/PAHs/aromatic hydrocarbons	152	2
Amides/amines/imides/imines	143	2
Antibiotics/antifungals/antivirals/antiprotozoals/antiseptics	132	1.5
PCB/PBB	102	1
Antineoplasics	89	1
Azo compounds/azo dyes/ azoic dye fragment/dye intermediates	89	1
Plant extracts/alkaloids/alkaloid derivatives/pyrrolizines alkaloids	79	1
Dioxins	73	1
Flavanols/Flavones/flavanones/flavonoids	67	1
Carbamates	59	1
Polymer/polyamine/polynucleotide/polysaccaride	55	1
Aldehydes	47	0.5
Ester/ethers	42	0.5
Mycotoxins	38	0.4
Sulfonic acids	35	0.4
Miscellaneous[a]	631	7

[a] Miscellaneous chemical class refers to complex chemicals or chemicals that could not be placed in a specific chemical class.

Table 4
Experimental models by year of publication and test system.

Year of publication	Experimental model								
	Animal			Plant			Bacteria		
	Total	In vitro	In vivo	Total	In vitro	In vivo	Total	In vitro	In vivo
Before 1970	323 (7%)[a]	104 (32%)[b]	219 (68%)	981 (31%)	350 (36%)	631 (64%)	151 (19%)	142 (94%)	9 (6%)[c]
1970–1979	604 (12%)	252 (42%)	352 (58%)	912 (29%)	150 (16%)	762 (84%)	78 (10%)	78 (100%)	–
1980–1989	1119 (23%)	568 (51%)	551 (49%)	451 (14%)	131 (29%)	320 (71%)	352 (44%)	352 (100%)	–
1990–1999	1639 (34%)	763 (47%)	876 (53%)	452 (46%)	139 (31%)	313 (69%)	91 (11%)	88 (97%)	3 (3%)
2000–2010	1183 (24%)	823 (70%)	360 (30%)	370 (12%)	103 (28%)	267 (72%)	134 (17%)	132 (99%)	2 (1%)
Total	4868 (10%)	2510 (52%)	2358 (48%)	3166 (100%)	873 (28%)	2293 (72%)	806 (100%)	792 (98%)	14 (2%)

[a] Number in parentheses is the percent of the total row (e.g., 323/4868 = 7%).
[b] Number in parentheses is the percent for that year group (e.g., 104/323 = 32%).
[c] Bacteria experiments were considered in vivo when the bacteria was injected into a host (e.g., rat).

the database (54% animals as compared to 35% plants). Rats (1085 dose responses) and mice (1218 dose responses) followed by humans (794 dose responses) are the most commonly used animal models, which is similar to that observed in 2005. Although no particular plant species predominates wheat was the most common (188 dose responses). In addition, fungi (167 dose responses), algae (219 dose responses), and yeast (110 dose responses) were also broadly represented. Ninety-four (4%) of the citations compared the effects of an agent on a certain endpoint across different ages. The conclusion remains that hormesis occurs in all different developmental and age related stages.

Most of the dose responses in the database (i.e., 5889) either did not specify the gender studied or gender does not apply to the model (e.g., plants or cell lines). Males (2196 dose responses) and females (1804 dose responses) were used in approximately the same number of studies. Because males and females were used together in some in vivo studies but the results were not separated by sex, the total number of dose responses by sex will not equal the total number of dose responses in the database because the study will be counted under both male and female results. If results were separated by sex, they are reported as separate studies. The sexes were compared in 128 (5%) of the citations in the database.

2.3. Test system

In contrast to the 2005 publication, where in vivo experiments predominated, there are similar amounts of dose responses in the database conducted in vivo (4698; 52%) and in vitro (4284; 48%) in the present analysis. This may be due to the increase in animal studies conducted in vitro (Table 4). Animal models were conducted at a similar rate in vivo and in vitro, while plants were conducted more often in vivo (Table 4). Study durations varied greatly (i.e., from a few minutes to a few years) because of differences between in vivo and in vitro studies. Even in vivo studies can vary greatly from a single injection or a few seconds of radiation to a lifetime of repeated exposures. Because of the variety of methods employed, the type of control used also varied greatly; however, the control used was an adequate control to study the effect.

2.4. Time course

Because hormesis may be related to an adaptive response, it may only be observed at certain times after or during exposure. Studies that examine an effect over different time periods demonstrate that in many cases, the hormetic effect is observed only at certain time points, while in other cases the hormetic effect is consistent over the time points measured. There are 1324 dose responses in the database that examined an endpoint at more than a single time point. Only one of the repeated measurements is entered into the database. While 36% of the 1324 dose responses only had two measurements, 64% of the 1324 endpoints had three or more measurements obtained.

2.5. Hypothesis testing

A dose response was considered to have hypothesis testing if the study authors provided statistical results comparing the treatment group to the control. While only 45% of the dose responses had hypothesis testing that fit the criteria of the database, there were instances where statistical analysis was performed but the study authors did not provide comparisons between treatment and control. Mutagenicity studies have often used a twice above control value as an indication of a positive effect instead of hypothesis testing. However, mutagenicity studies had a similar use of hypothesis (47%; i.e., 422 of 892 dose responses) testing as the general database (45%). Table 5 demonstrates an increase in hypothesis testing after 1970 compared to before 1970 with the greatest use of hypothesis testing occurring in more recent years.

2.6. Transgenerational

The database includes 168 dose responses from studies that examined transgenerational effects. Each generation fitting the criteria is entered into the database as a separate study. There were 45 citations that examined generations indicating that more than one generation from a citation had a hormetic effect. Sometimes studies examined the transgenerational effects, but only reported results for the second generation and not the generation exposed. Therefore, the number of generations is considered only one in the database. This occurred in 5% of the dose responses, however, the majority (83%) of the studies examined two generations.

2.7. Subjects

The number of dose responses in which subjects were presented for inclusion in the database is 5576 (62% of the dose responses) (Table 6). Because some of the dose responses include a different number of subjects for different doses (e.g., controls had twice as many subjects or the endpoint was only measured in survivors and each group had a different number), the total (i.e., 5915) will not be equal to the number of dose responses (i.e., 5576) that included the number of subjects. The number of subjects was considered the number that the study authors provided in the tables or figures and could include the number of experiments, the number of cultures, or the number of individuals. Table 6 demonstrates that there are only slightly more in vivo studies than in vitro studies, but that in vivo studies were more likely to have >10 subjects per group.

Table 5
The number of dose responses with hypothesis testing by publication year.

Publication year	Number of dose responses	Number with hypothesis testing	Percent of yearly total
Before 1970	1470	212	14
1970–1979	1625	690	43
1980–1989	1950	729	37
1990–1999	2213	1308	59
2000–2010	1704	1136	67
Total	8962	4075	45

Table 6
Number of subjects per treatment group.

Number of subjects per treatment group	Number of dose responses	In vitro	In vivo
⩽10	3632 (61%)[a]	2165 (60%)[b]	1467 (40%)
11–50	1483 (25%)	297 (20%)	1186 (80%)
51–100	355 (6%)	74 (21%)	281 (79%)
101–1000	371 (10%)	68 (18%)	303 (82%)
>1000	74 (1%)	22 (30%)	52 (70%)
Total	5915 (100%)	2626 (44%)	3289 (56%)

[a] Number in parentheses is the percent of the total row (e.g., 3632/5915 = 61%).
[b]Number in parentheses is the percent for that number of subjects grouping (e.g., 2165/3632 = 60%).

Table 7
The number of dose responses by endpoint.

Endpoint type	Number of dose responses	Percent of total dose–response relationships (8962)	Plants (3166)	Animals (4868)	Bacteria (806)
Growth	3353	37	2197 (69%)[a]	922 (19%)	186 (23%)
Metabolic[b]	1996	22	598 (19%)	1157 (24%)	199 (25%)
Mutagenic[c]	892	10	51 (2%)	498 (10%)	342 (42%)
Immune response	581	6.5	–	581 (12%)	–
Survival	568	6	60 (2%)	415 (8.5%)	74 (9%)
Reproduction[d]	534	6	179 (6%)	342 (7%)	4 (0.5%)
Neurological	285	3	0 (0%)	285 (6%)	0 (0%)
Behavioral[e]	266	3	–	265 (5%)	–
Cancer	161	2	–	161 (3%)	–
Longevity	152	2	1 (0.03%)	148 (3%)	1 (0.1%)
Disease[f]	68	0.8	60 (2%)	8 (0.2%)	–
Damage[g]	61	0.7	19 (0.6%)	42 (0.9%)	–
Developmental[h]	45	0.5	1 (0.03%)	44 (0.9%)	–

[a] Number in parentheses is the percent of the total for that specific model (e.g., 2197/3166 = 69%).
[b] Examples: DNA repair, enzyme activity, hormone levels, ROS production, ATP response, oxygen uptake, or urine volume.
[c] Examples: number of revertants, micronucleus frequency, incidence of bent humeral bristles, chromosome aberrations, drug resistance, or DNA integrity.
[d] Examples: fecundity, hatching rate, eggs/female, number of young, number of resorptions, seed germination, or number of flowers.
[e] Examples: distance travelled, flinches/min, number of bites, rearings, or head dips, or number of correct choices.
[f] Disease in plants refers to rot or spoilage of fruit, diseased plants, or number of weeds; disease in animals refers to infection (e.g., kidney infection) or parasites.
[g] Damage in plants refers to disintegrating roots, decay, number of holes caused by insects or oxidation; damage in animals refers to cell rounding, ALT or LDH release, lesions, lipid peroxidation, or hyperplasia.
[h] Examples: adult eclosing rate or malformations.

2.8. Endpoints

The dose–response relationships are divided into two different sections for endpoint (endpoint type – e.g., growth; and endpoint parameter – e.g., body weight). Thirteen endpoint types were selected for generalized search capacity (Table 7). Since the 2005 publication, a neuroscience endpoint was added and the database was re-evaluated by endpoint parameter to add the neurological endpoint type. Although growth was the endpoint type associated with the most dose responses, there were numerous dose responses in each of the endpoint types. In addition, studies may examine multiple endpoint parameters that predict the same or closely related process (e.g., cell proliferation), but it is estimated via a different endpoint parameter (e.g., DNA synthesis, tritiated thymidine uptake, cell numbers, etc.) and each that meets the criteria is listed as a separate endpoint within the study. As is expected, the endpoint type measured the most frequently in plants was growth and mutagenic was the most frequently studied endpoint in bacteria. Animals had a more uniform distribution by endpoint type.

2.9. Hormetic curve

The majority of the dose–response relationships were inverted U curves (82%). Table 8 describes the width of the 5668 dose responses in which a range could be determined. The majority of the dose responses displayed a stimulatory response range less than 10-fold wide. However, the response range could be highly variable with a low percentage (7%) displaying a stimulatory range that exceeded 1000-fold. There was only a slight variation in the stimulatory range between the models with bacteria having a greater proportion of the dose responses having a stimulatory range between 10- and 100-fold (Table 8). Although stimulatory ranges of ⩾1000-fold occurred at the lowest frequency for all endpoints, this range occurred at a greater frequency in neurological and immune responses than the other endpoints (Table 9). The

Table 8
Width of stimulation range by model.

Width (-fold)	Number of dose–response relationships	Number in plants	Number in animals	Number in bacteria
⩾1 < 10	2450 (43)[a]	922 (54%)[b]	1290 (48%)	197 (35%)
⩾10 < 100	2054 (36)	572 (34%)	572 (21%)	233 (42%)
⩾100 < 1000	760 (13)	149 (9%)	484 (18%)	113 (20%)
⩾1000	404 (7)	58 (3%)	326 (12%)	18 (3%)
Total	5668[c]	1701 (100%)	2672 (100%)	561 (100%)

[a] The number in parentheses is the % of total dose–response relationships.
[b] The number in parentheses is the % of total for that model.
[c] The value of 5668 differs from the total number dose–response relationships (i.e., 8962) because calculation of the range was not possible for all dose responses for various reasons (e.g., data presentation precludes exact determination of range).

Table 9
Width of stimulatory range by endpoint.

Endpoint type	Number of dose responses with ranges available	Width of the stimulatory range			
		⩾1 < 10	⩾10 < 100	⩾100 < 1000	⩾1000
Growth	1971	1074 (54%)[a]	587 (30%)	201 (10%)	109 (6%)
Metabolic	1213	455 (38%)	435 (36%)	208 (17%)	115 9%)
Mutagenic	755	173 (23%)	402 (53%)	145 (19%)	35 (5%)
Immune response	304	100 (33%)	118 (39%)	51 (17%)	35 (11%)
Survival	465	277 (60%)	149 (32%)	32 (7%)	7 (2%)
Reproduction	313	163 (52%)	116 (37%)	24 (8%)	10 (3%)
Neurological	202	68 (34%)	58 (29%)	42 (21%)	34 (17%)
Behavioral	166	52 (31%)	82 (49%)	24 (14%)	8 (5%)
Cancer	75	18 (24%)	43 (57%)	9 (12%)	5 (7%)
Longevity	84	43 (51%)	36 (43%)	5 (6%)	0 (0%)
Disease	9	1 (11%)	5 (56%)	2 (22%)	1 (11%)
Damage	35	11 (31%)	16 (46%)	4 (11%)	4 (11%)
Developmental	37	15 (41%)	7 (19%)	13 (35%)	2 (5%)

[a] Number in parentheses is the percent of the total for that specific endpoint (e.g., 1074/1971 = 54%).

higher frequency of mutagenic endpoints with a stimulatory response between 10- and 100-fold supports the finding that bacteria were more likely to have a stimulatory range between 10- and 100-fold. Although the majority of dose responses had a stimulatory range less than 10-fold, there were some endpoint types (i.e., mutagenic, immune response, behavioral, cancer, disease, and damage) that had a greater frequency of dose responses with stimulatory ranges between 10- and 100-fold.

The relative proportions of maximum stimulatory responses for J-shaped curves were more evenly distributed than in inverted U-shaped curves (Table 10). The maximum stimulatory response range in inverted U-shaped curves was between 110% and 150% of the control; 79% of the dose responses had a maximum stimulatory response less than 200% of the control regardless of model or endpoint (Tables 10–12). The majority of J-shaped curves had a maximum stimulatory response between 50% and 100% of the control (Table 10). While this range is maintained regardless of the model employed (Table 10), certain endpoints had a greater tendency to have % of control values ⩽50% of the control (i.e., immune response, disease, and developmental) while other endpoints (i.e., survival and cancer) were more evenly distributed across the maximum stimulatory response ranges (Table 11).

Information on the distance from the maximum stimulatory response and the dose where the curve would cross the zero equivalence point (i.e., response equal to control; ZEP) is provided in Table 13. To calculate the point, the curve must have peaked and crossed the control value again. This occurred in 5331 (59%) of the dose responses. In the majority (62%) of cases the maximum stimulatory response is within a factor of five from the ZEP, which occurs regardless of biological model used (Table 13) or endpoint examined (Table 14). In plants, the majority of the dose responses are in vivo regardless of the distance from the maximum stimulatory response and the ZEP. In animals, however, as the distance between the maximum stimulatory response and the ZEP increased, the more likely the study was to be in vitro (Table 13).

2.10. Strength of evidence

Calabrese and Baldwin (1997a,b) provided a numerical scoring system to determine the strength of evidence for assessing to what extent the dose response was consistent with the hormetic dose–response model, which is described above in the scoring methodology section. Table 15 indicates that the majority of the responses (57%) had low evidence of hormesis. After entering nearly 9000 dose responses, this is more likely a limitation of the scoring system than in the quality of the dose responses. Although there are some dose responses with only one or two doses in the hormetic region, Table 16 indicates that there are as many curves with three, four, or five doses in the hormetic range. The low evidence results are due to a limited maximum stimulatory response in the predominantly inverted U-shaped curves and the lack of hypothesis testing in approximately 50% of the dose responses. In plants, the studies were generally conducted in vitro and this did not change with the number of doses below the ZEP. In animals, however, the in vivo studies were more likely to have fewer doses below the ZEP. As the number of doses below the ZEP increased, more of the animal studies were found to be in vitro (Table 16).

3. Discussion

Since the publication in 2005 3400 new dose responses have been added to the database, approximately 40% of the database (Calabrese and Blain, 2005). Despite the substantial entry enlargement of the database the quantitative aspects of the hormetic curve (i.e., maximum stimulatory response, stimulatory range,

Table 10
Maximum stimulatory response by model.

Maximum stimulatory response (% control)	Number of dose–response relationships	Number in plants	Number in animals	Number in bacteria
J-shaped curve				
⩽100 > 75	583 (37%)[a]	47 (26%)[b]	314 (30%)	219 (64%)
⩽75 > 50	526 (33%)	63 (35%)	372 (35%)	91 (27%)
⩽50 > 25	277 (18%)	32 (18%)	224 (21%)	21 (6%)
⩽25	193 (12%)	40 (22%)	144 (14%)	9 (3%)
Total	1579 (100%)	182 (100%)	1054 (100%)	340 (100%)
Inverted U-shaped curve				
⩾100 < 110	42 (0.6%)	11 (0.4%)	30 (0.8%)	1 (0.2%)
⩾110 < 150	4379 (59%)	1913 (64%)	2147 (56%)	253 (54%)
⩾150 < 200	1443 (20%)	572 (19%)	754 (20%)	98 (21%)
⩾200 < 500	1191 (16%)	397 (13%)	694 (18%)	73 (16%)
⩾500 < 1000	190 (3%)	62 (2%)	102 (3%)	22 (5%)
⩾1000	138 (2%)	29 (1%)	87 (2%)	19 (4%)
Total	7383 (100%)	2984 (100%)	3814 (100%)	466 (100%)

[a] The number in parentheses is the % of total dose–response relationships.
[b] The number in parentheses is the % of total for that model.

Table 11
Maximum stimulatory response in a J-shaped curve by endpoint.

Endpoint type	Number of dose responses with ranges available	Maximum stimulatory response			
		⩽100 > 75% of control	⩽75 > 50% of control	⩽50 > 25% of control	⩽25% of control
Growth	42	13 (31%)[a]	13 (31%)	6 (14%)	10 (24%)
Metabolic	157	66 (42%)	60 (38%)	25 (16%)	6 (1%)
Mutagenic	830	406 (49%)	284 (34%)	93 (19%)	47 (6%)
Immune response	44	8 (18%)	6 (14%)	14 (32%)	16 (36%)
Survival	43	7 (16%)	9 (21%)	16 (37%)	11 (26%)
Reproduction	55	17 (31%)	19 (35%)	13 (24%)	6 (11%)
Behavioral	81	10 (12%)	44 (54%)	18 (22%)	9 (11%)
Cancer	153	33 (22%)	45 (29%)	42 (27%)	33 (22%)
Disease	67	2 (3%)	12 (18%)	17 (25%)	36 (54%)
Damage	48	6 (12.5%)	21 (44%)	15 (31%)	6 (12.5%)
Developmental	37	8 (22%)	5 (13.5%)	11 (30%)	13 (35%)

[a] Number in parentheses is the percent of the total for that specific endpoint.

Table 12
Maximum stimulatory response in an inverted U-shaped curve by endpoint.

Endpoint type	Number of dose responses with ranges available	Maximum stimulatory response					
		⩾100 < 110% of control	⩾110 < 150% of control	⩾150 < 200% of control	⩾200 < 500% of control	⩾500 < 1000% of control	⩾1000% of control
Growth	3311	25 (0.8%)[a]	2175 (66%)	588 (18%)	423 (13%)	68 (2%)	32 (1%)
Metabolic	1839	5 (0.3%)	895 (49%)	395 (21%)	399 (22%)	79 (4%)	66 (4%)
Immune response	537	1 (0.2%)	272 (51%)	145 (27%)	100 (19%)	11 (2%)	8 (1%)
Survival	525	2 (0.4%)	397 (76%)	72 (14%)	45 (9%)	6 (1%)	3 (0.6%)
Reproduction	479	0 (0%)	292 (61%)	81 (17%)	76 (16%)	11 (2%)	19 (4%)
Neurological	265	0 (0%)	143 (54%)	67 (25%)	41 (15%)	9 (3%)	5 (2%)
Behavioral	185	0 (0%)	69 (37%)	41 (22%)	68 (37%)	5 (3%)	2 (1%)
Longevity	150	7 (5%)	82 (55%)	39 (26%)	22 (15%)	0 (0%)	0 (0%)
Mutagenic	62	1 (2%)	36 (58%)	10 (16%)	12 (19%)	1 (2%)	2 (4%)

[a] Number in parentheses is the percent of the total for that specific endpoint.

number of doses below the ZEP, etc.) have remained consistent. The data, in general, reveal that the literature basis is both quite extensive and broadly distributed across the range of biologically based sub-disciplines that utilize dose–response relationships. Previously it had been concluded that studies with more doses/concentrations resulted from the use of less expensive biological models such as plants. While this may have been true in earlier years when in vitro cell lines were hard to maintain and obtain, the shorter duration and ease of study designs have lead to more use of in vitro models, which in the database have more often been animal models. Although it is not clear whether the increase in in vitro animal models entered in the database is an artifact of subject interest, it is noted that there were more endpoints using in vitro animal models that demonstrate hormesis published in the last 10 years than any other time point.

The point system was developed along with the database in order to provide a framework to assess whether hormesis was a viable toxicological hypothesis. Although the point system has limitations, the database has provided the opportunity to assess the quantitative features of the dose–response curve. As was noted in the 2005 publication, the majority of the inverted U-shaped curves are 110–150% of the control with nearly 80% of the maximum responses less than 2-fold greater than the control. Therefore, the stimulatory responses, mainly in inverted U-shaped

Table 13
Distance from the maximum stimulatory response to the ZEP[a] by model.

Distance (-fold)	Number of dose–response relationships	Number in plants	In vitro studies in plants	In vivo studies in plants	Number in animals	In vitro studies in animals	In vivo studies in animals	Number in bacteria
⩾1 < 5	3331 (62%)[b]	1065 (69%)[c]	757	308	1905 (60%)	923	982	307 (58%)
⩾5 < 10	862 (16%)	257 (17%)	166	91	492 (16%)	191	301	99 (19%)
⩾10 < 100	837 (16%)	175 (11%)	98	77	543 (17%)	213	330	103 (19%)
⩾100	304 (6%)	57 (4%)	37	20	222 (12%)	65	157	21 (4%)
Total	5331 (100%)[d]	1554 (100%)	1058	496	3162 (100%)	1392	1770	530 (100%)

[a] ZEP = zero equivalent point (i.e., the highest dose showing a response equal to the control response).
[b] The number in parentheses is the % of total dose–response relationships.
[c] The number in parentheses is the % of total for that model.
[d] The value of 5331 differs from the total number of dose–response relationships (i.e., 8962) because calculation of the distance was not possible for all dose responses for various reasons (e.g., data presentation precludes exact determination of ZEP).

Table 14
Distance to the maximum stimulatory response to the ZEP by endpoint.

Endpoint type	Number of dose responses with distance available	Distance from maximum stimulatory response to the ZEP			
		⩾1 < 5	⩾5 < 10	⩾10 < 100	⩾100
Growth	1832	1167 (64%)[a]	308 (17%)	252 (14%)	105 (6%)
Metabolic	1128	621 (55%)	205 (18%)	232 (21%)	70 (6%)
Mutagenic	731	475 (65%)	107 (15%)	122 (17%)	27 (4%)
Immune response	323	157 (49%)	49 (15%)	64 (20%)	53 (16%)
Neurological	193	91 (47%)	37 (19%)	48 (25%)	17 (9%)
Survival	439	323 (74%)	61 (14%)	46 (10%)	9 (2%)
Reproduction	296	208 (70%)	43 (15%)	0 (0%)	13 (3%)
Behavioral	151	105 (70%)	25 (17%)	17 (11%)	4 (3%)
Cancer	78	60 (77%)	7 (9%)	9 (12%)	2 (3%)
Damage	36	28 (78%)	4 (11%)	3 (8%)	1 (3%)
Developmental	36	27 (75%)	5 (14%)	2 (6%)	2 (6%)
Longevity	82	64 (78%)	9 (11%)	8 (10%)	1 (1%)

[a] Number in parentheses is the percent of the total for that specific endpoint.

Table 15
Evidence of hormesis as used in the hormesis database[a].

Evidence of hormesis	Number of dose–response relationships	Percent of total dose–response relationships (8962)
High	859	10
Moderate–high	403	4
Moderate	907	10
Low–moderate	1616	18
Low	5137	57
No–low	40	0.4

[a] For a description of the quantitative methodology used to derive evidence of hormesis see Calabrese and Baldwin (1997a) and/or Table 1.

Table 16
The number of doses below the ZEP.

Number of doses	Number of dose responses (% of the 8962 dose responses)	Number of dose responses in plants	In vitro plant studies	In vivo plant studies	Number of dose responses in animals	In vitro animals studies	In vivo animal studies
1	1417 (16%)	472	142	330	804	362	442
2	1946 (22%)	697	198	499	1045	483	562
3	1975 (22%)	722	181	541	1056	513	543
4	1439 (16%)	491	132	359	805	433	372
5	979 (11%)	349	90	259	520	267	253
6+	1206 (13%)	435	130	305	638	452	186

curves, observed in the hormesis database still exhibit a modest magnitude and width in the majority of the cases. These features are biologically significant since they occur across biological models, endpoints, and chemical class/physical agent.

Although the majority of the dose responses in the database displayed low evidence of hormesis, the findings reflect both the strength of the data supporting hormesis as well as the evaluative criteria applied to the study design and response data and their evaluation (e.g., statistical analysis). The hormesis–frequency database (Calabrese and Baldwin, 2001b) provides an absolute judgment on whether the evaluative criteria of hormesis were satisfied or not. If such criteria were satisfied, then hormesis was judged as present. In contrast, the hormesis database, which is the subject of the present paper, makes no absolute judgment on the existence of

hormesis; instead, it applied different evaluative criteria, which result in dose responses being characterized according to the degree to which they are consistent with the hormetic–biphasic dose response. Given that these databases were constructed for different purposes and used different evaluative criteria, it is important to note that when all 245 dose–responses that satisfied the evaluative criteria (i.e., hormesis) in the hormesis–frequency database were assessed using the scoring system employed on the dose responses in the hormesis database, the distribution of the ranked scores were similar (Calabrese and Blain, 2005). This is significant since it was our strong general impression before conducting this inter database comparison that the entry criteria of the hormesis–frequency database (Calabrese and Baldwin, 2001b) was considerably more stringent than the hormesis database, but this is not the case.

The overall findings indicate that the quantitative features of the hormetic dose response are consistent across biological model and endpoint. These observations suggest that this feature has been highly conserved within an evolutionary context. It also suggests that the hormetic dose response may provide a quantitative estimate of biological plasticity that is broadly generalizable (Calabrese and Mattson, 2011; Calabrese, 2010).

Conflict of interest statement

The authors declare that there are no conflicts of interest. This manuscript has not been published previously and is not under consideration for publication elsewhere.

Acknowledgments

Effort sponsored by the Air Force Office of Scientific Research, Air Force Material Command, USAF, under Grant number FA9550-07-1-0248. The US Government is authorized to reproduce and distribute for governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsement, either expressed or implied, of the Air Force Office of Scientific Research or the US Government. The funding source had no involvement in study design, collection, analysis and interpretation of data; in writing the report and in the decision to submit the manuscript for publication.

References

Calabrese, E.J., Mattson, M.P., 2011. Hormesis provides a generalized quantitative estimate of biological plasticity. J. Cell Commun. Signal 5, 25–38.
Calabrese, E.J., 2010. Hormesis is central to toxicology, pharmacology and risk assessment. Hum. Exper. Toxicol. 29, 249–261.
Calabrese, E.J., 2008. Hormesis: why it is important to toxicology and toxicologists. Environ. Toxicol. Chem. 27, 1451–1474.
Calabrese, E.J., Baldwin, L.A., 2003a. The hormetic dose response model is more common than the threshold model in toxicology. Toxicol. Sci. 71, 246–250.
Calabrese, E.J., Baldwin, L.A., 2003b. Toxicology rethinks its central belief – hormesis demands a reappraisal of the way risks are assessed. Nature 421, 691–692.
Calabrese, E.J., Baldwin, L.A., 2001a. Hormesis: U-shaped dose responses and their centrality in toxicology. Trends Pharmacol. Sci. 22, 285–291.
Calabrese, E.J., Baldwin, L.A., 2001b. The frequency of U-shaped dose responses in the toxicological literature. Toxicol. Sci. 62, 330–338.
Calabrese, E.J., Baldwin, L., 1997a. A quantitatively-based methodology for the evaluation of chemical hormesis. Hum. Ecol. Risk Assess. 3, 545–554.
Calabrese, E.J., Baldwin, L., 1997b. The dose determines the stimulation (and poison): development of a chemical hormesis database. Inter. J. Toxicol. 16, 545–559.
Calabrese, E.J., Blain, R., 2004. Metals and hormesis. J. Environ. Monit. 6, 14N–19N.
Calabrese, E.J., Blain, R., 2005. The occurrence of hormetic dose responses in the toxicological literature, the hormesis database: an overview. Toxicol. Appl. Pharmacol. 202, 289–301.
Calabrese, E.J. et al., 2008. Hormesis predicts low-dose responses better than threshold models. Int. J. Toxicol. 27, 369–378.
Calabrese, E.J. et al., 2006. Hormesis outperforms threshold model in NCI anti-tumor drug screening data. Toxicol. Sci. 94, 368–378.

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Toxicology and Applied Pharmacology 202 (2005) 289–301

Toxicology and Applied Pharmacology

www.elsevier.com/locate/ytaap

Review

The occurrence of hormetic dose responses in the toxicological literature, the hormesis database: an overview

Edward J. Calabrese*, Robyn Blain

Environmental Health Sciences, University of Massachusetts, Amherst, MA 01003, United States

Received 19 May 2004; accepted 29 June 2004
Available online 27 October 2004

Abstract

A relational retrieval database has been developed compiling toxicological studies assessing the occurrence of hormetic dose responses and their quantitative characteristics. This database permits an evaluation of these studies over numerous parameters, including study design and dose–response features and physical/chemical properties of the agents. The database contains approximately 5600 dose–response relationships satisfying evaluative criteria for hormesis across over approximately 900 agents from a broadly diversified spectrum of chemical classes and physical agents. The assessment reveals that hormetic dose–response relationships occur in males and females of numerous animal models in all principal age groups as well as across species displaying a broad range of differential susceptibilities to toxic agents. The biological models are extensive, including plants, viruses, bacteria, fungi, insects, fish, birds, rodents, and primates, including humans. The spectrum of endpoints displaying hormetic dose responses is also broad being inclusive of growth, longevity, numerous metabolic parameters, disease incidences (including cancer), various performance endpoints such as cognitive functions, immune responses among others. Quantitative features of the hormetic dose response reveal that the vast majority of cases display a maximum stimulatory response less than two-fold greater than the control while the width of the stimulatory response is typically less than 100-fold in dose range immediately contiguous with the toxicological NO(A)EL. The database also contains a quantitative evaluation component that differentiates among the various dose responses concerning the strength of the evidence supporting a hormetic conclusion based on study design features, magnitude of the stimulatory response, statistical significance, and reproducibility of findings.

Keywords: Hormesis; Inverted U-shaped; J-shaped; Dose–response; Risk assessment; Database; Biphasic; Adaptive; Low dose

Contents

* Corresponding author. Environmental Health Sciences, University of Massachusetts, Morrill I, N344, Amherst, MA 01003. Fax: +1 413 545 4692.
E-mail address: edwardc@schoolph.umass.edu (E.J. Calabrese).

0041-008X/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.taap.2004.06.023

Introduction

The concept of hormesis has generated considerable interest in recent years. While this interest is broadly based in the biological and biomedical sciences (Calabrese and Baldwin, 2001, 2003), it is also of considerable importance to the area of environmental risk assessment (Kaiser, 2003). This is principally because the hormetic dose–response model challenges the linear at low-dose model employed by regulatory agencies such as the EPA and FDA in cancer risk assessment activities.

Of particular importance is the need to consider in detail the toxicological evidence that allows the hormetic hypothesis to be assessed. Within this framework, it is necessary to have the capacity to explore the generalizability of the hormesis concept. Two separate databases assessing the concept of hormesis have been developed. One was designed to address the frequency of hormesis (i.e., Hormesis-Frequency Database) within the toxicological literature employing rigorous a priori entry and evaluative criteria (Calabrese and Baldwin, 2003). An assessment of this database, which evaluated over 21,000 toxicological articles from the mid-1960s to the present in three different journals (*Life Sciences*, *Bulletin of Environmental Contamination and Toxicology*, and *Environmental Pollution*), determined that hormetic responses were observed in nearly 40% of dose responses in which the a priori entry criteria were satisfied. The second hormesis database (i.e., Relational Hormesis Database, hereafter called the Hormesis Database), which complements the Hormesis-Frequency Database, is based upon a different set of a priori evaluative criteria and was designed:

(1) to identify likely cases of hormesis with the intention of assessing the quantitative features of the hormetic dose response;
(2) to assess the generalizability of the hormetic hypothesis according to biological model, endpoint, and nature of the chemical and physical stressor agents inducing hormetic responses;
(3) to assess historical aspects of toxicological research designed to assess various features of the hormetic dose response;
(4) and to create a resource of possible hormesis articles in a relational retrieval system to assess a variety of toxicologically based hypotheses and/or questions.

The following paper provides a detailed description of the Hormesis Database and how it permits unique insight into the nature of the dose response in the low-dose zone.

Database entry criteria

The Hormesis Database inclusion criteria were a minimum of 10% stimulation of response (i.e., inverted U) in at least one dose when a depression is expected (e.g., growth) or a 3% (rounded down from 3.3%) depression (i.e., J-shaped) in response when stimulation is expected (e.g., cancer). The J-shaped dose–response criteria were structured to be proportionally similar to that of the inverted U-shaped dose response. The reason for the different values between the inverted U and J-shaped dose response (3 vs.

Table 1
The 20 most cited journals in the hormesis database

Journal title	Number of citations
Bulletin of Environmental Contamination and Toxicology	121
Radiation Research	100
Environmental and Experimental Botany (formerly Radiation Botany)	92
Environmental Pollution (including Series A)	84
Stimulation Newsletter	59
International Journal of Radiation Biology	56
NTP Toxicity Report Series	45
Environmental Toxicology and Chemistry	29
American Journal of Botany	25
Biochemical Pharmacology	23
Life Sciences	19
Nature	18
Ecotoxicology and Environmental Safety	17
Molecular Pharmacology	14
Toxicology and Applied Pharmacology	14
Science	13
Physiologia Plantarium	11
Journal of Biochemistry	11
Carcinogenesis/Journal of Bacteriology/Soil Science/ Journal of Immunology/Toxicological Sciences	Each were cited 10 times

Table 2
Description of the information entered into the hormesis database

Citation	Chemical	Biological model	Study design	Treatment	Endpoint	Dose–response curve	Strength of hormesis
Author, title, and source	substance name and synonyms, CAS number, chemical class, and water solubility	taxonomy, gender, and age where appropriate or known	length of study, length of exposure to test article, controls used, test system, statistics performed and if measurements were taken over time	dose/concentration, number of subjects/ treatment group, results by dose/concentration, results of hypothesis testing of dose/ concentration compared to the control, and percent of control	type, specific parameters, comparisons made between ages, sexes, or species	shape, N/LOAEL, number of doses/concentrations in the stimulatory range, the number significantly different from the control, the amplitude of the maximum response, distance of the stimulatory range, distance between the maximum stimulatory dose and the ZEP, and reproducibility	based on criteria described in Calabrese and Baldwin, 1997, Hum. Ecol. Risk Assess. 3:545–554

10%) was that maximum stimulatory response boundaries for the two different curves differ by 3-fold. That is, the stimulation could increase in absolute terms by 300% from 100% to 400% for the inverted U, while the J-shaped curve could decrease from 100% to 0.0%. If the 10% or 3% selection criteria were not satisfied, a dose response could have been entered into the database if the response achieved statistical significance in hypothesis testing. In addition, the study had to employ an adequate concurrent control for comparison.

Description of the database

The Hormesis Database contains the findings of approximately 5600 dose responses obtained from approximately 1450 articles where a stimulatory response consistent with the above noted criteria has been observed. The articles have been obtained mainly through extensive searches through numerous journals, cross-referencing journal citations, MEDLINE, and other database searches using multiple words such as hormesis, stimulation, inverted U-shaped, J-shaped, biphasic, and others. The studies have been principally found in a diverse array of journal publications (over 300 different journals; the 20 most cited journals are provided in Table 1 below) although several books (42 citations), theses, and dissertations (six citations) are included. Information entered into the database is specified in Table 2 below. Queries can be conducted using any of these fields. The query system was employed to yield the descriptive assessment offered in this article.

Database scoring methodology

Each dose response entered into the database was scored according to the design of Calabrese and Baldwin (1997 provides detailed description) in which numerical values were given to various specific components of the study design, response magnitude, statistical analysis, and reproducibility of the response to evaluate the capacity of an experiment for demonstrating hormesis (i.e., strength of evidence). Points were awarded for (1) the number of doses below the NO(A)EL up to a total of 5 points, (2) experimental determination or estimation of a NO(A)EL, (3) the number of statistically significant responses below the NO(A)EL, (4) the magnitude (percent of control value) of the stimulatory response, and (5) reproducibility of the data by other studies with data provided (in other publications or within the same publication). The scheme by which specific point values are applied is given in Table 3.

This scoring methodology was designed to take into account and to reward studies that explicitly considered below NO(A)EL doses in their study designs, but also to ensure that the balance of the points be strongly skewed in favor of response over design. A point was awarded if a

Table 3
Summary of criteria with assigned point values used in the quantitative evaluation of hormesis

Study design criteria			
Doses below NO(A)EL (*n*)	Point value	NO(A)EL determined	Point value
1	1	Yes	1
2	2	No	0
3	3	–	–
4	4	–	–
≥5	5	–	–

Response criteria			
Doses statistically significant (*n*)	Point value	Reproducibility	Point value
1	2	Yes	3
2	4	No	0
3	8	–	–
≥4	16	–	–

Magnitude of response (percentage control value)		
Inverted U-shaped curve	J-shaped curve	Point value[a]
≥100% ≤ 125%	≤97% ≥ 92%	0.5
> 125% ≤ 150%	<92% ≥ 84%	1
> 150% ≤ 200%	<84% ≥ 68%	2
> 200% ≤ 400%	<68% ≥ 5%	3
> 400%	<5%	4

[a] The point value is multiplied by the number of experimental doses falling within the corresponding percentage range. For example, if an experiment has three doses exhibiting stimulatory responses within the 125% to 150% range and two of the responses achieve statistical significance and one does not, the total number of points would be: 3 × 1 = 3 (study design criteria); 4 (response criteria, statistically significant); 3 × 1 = 3 (magnitude of response), for a total of 10 points. The 10 points would achieve the categorization of hormesis evidence of "low–moderate" (Table 3A).

NO(A)EL or Zero Equivalent Point [ZEP, place where response crosses the control (100%) value] was identified or could be estimated since this was important with respect to identifying below NO(A)EL doses but also in identifying the quasi threshold of the dose response. The points for response were progressively increased from a minimal entry starting at 10% (or 3%) up to 400% (or 0%) greater than controls. Because increases greater than 400% of the control may represent a different phenomenon than hormesis, which tends to display more limited maximum stimulatory responses, the points were capped at 4 for responses 400% of control or greater. Points were still applied to these responses due to the likelihood that a response of this magnitude is real and more likely to be reproducible.

Some dose responses also have hypothesis testing applied to the data. In these cases, additional points were provided for each response displaying statistical significance (up to 4 doses for a total of 16 points). The

Table 3A
Summary of total point ranges for hormesis evidence categories used in the quantitative evaluation of articles for evidence of hormesis

Total point range	Hormesis evidence category
1–2	No–low
>2–8	Low
>8–12	Low–moderate
>12–16	Moderate
>16–20	Moderate–high
>20	High

Table 4
Articles in the hormesis database by year of publication

Publication year	Number of articles	Percent of total articles (1443)
<1930	37	2.6
1930–1939	30	2.1
1940–1949	42	2.9
1950–1959	86	6.0
1960–1969	140	9.7
1970–1974	166	11.5
1975–1979	166	11.5
1980–1984	150	10.4
1985–1989	160	11.1
1990–1994	204	14.1
1995–1999	224	15.5
2000–2003	39	2.7

Table 5
Most prevalent chemical classes and physical agents

Chemical class	Number of responses	% of total (5632)
Inorganics (including elements and metals)	1308	23
<1970	254 (19%)	
1970–1979	240 (18%)	
1980–1989	352 (27%)	
1990–1999	392 (30%)	
Radiation/radionuclides	1654	29
<1970	447 (27%)	
1970–1979	697 (42%)	
1980–1989	254 (15%)	
1990–1999	251 (15%)	
Organics	–	
Organophosphate/pesticides	285	7
PBBs/PCBs	86	1.5
Carboxylic acids	232	4
Chlorinated hydrocarbons/ halogenated hydrocarbons	105	2
Alcohols/phenols	302	5
Heterocyclic compounds	113	2
Antibiotics/antivirals	73	1
Hydrocarbons/PAHs/aromatic hydrocarbons	86	2
Carbamates	35	1
Amides/amines/imides	56	1
Hormones/amino acids/ enzymes/neuropeptides/peptides/steroids	161	3
Nitrosamines/nitrosoureas	28	0.5
Dioxins	42	1
Antineoplastics	33	0.5
Alkaloids	26	0.5
Mycotoxins	22	0.5
Miscellaneous[a]	269	5

[a] Miscellaneous chemical class refers to complex chemicals or chemicals that could not be placed in a specific chemical class.

Table 6
The number of dose responses by experimental models

Experimental model	Number of dose–response relationships	Percent of total dose–response relationships (5632)
Plant	2596	46.0
Animal	2601	46.2
Bacteria	348	6.2
Protozoa	31	0.6
Microcosm	23	0.4
Phytoplankton	17	0.3
Virus	10	0.2
Mesocosm	5	0.1

magnitude of points was also weighted more heavily for statistical significance than magnitude of response. This is because hypothesis testing takes into account sample size, variability, and magnitude of response in a reliable and nonbiased manner. Thus, response criteria were much more heavily weighted than study design criteria with studies displaying statistical significance receiving the greatest consideration. The summation of the points was then categorized from low (>2–8) to high (>20) to the consistency of the dose response with the hormetic hypothesis (Table 3A).

General

The database is such that each citation is associated with studies (i.e., experiments). Each experiment may have examined several endpoints (e.g., body weight, survival, serum chemistries). Each endpoint will have a dose response associated with it. Any search performed with the database provides the number and percent of total for citations, studies, and endpoints.

The database includes experimental findings from 1899 to present (Table 4). Twenty-three percent (i.e., 335) of the articles were published before 1970 with 21 publications before 1920. From 1970, the results are given in 5-year intervals to discern potential time-related trends within the experimental findings. In general, entries are not provided for articles published since 2000 due to the lack of completion of the 5-year interval, although in several instances available data are provided. Since the articles reviewed reflected numerous factors including availability of specific journals at the UMASS/Amherst library, efficiency and availability of interlibrary requests, and specific area of interests (e.g., immunotoxicology), no evaluative judgment as yet should be placed on the relative frequency of entries per time period or journal citation frequency.

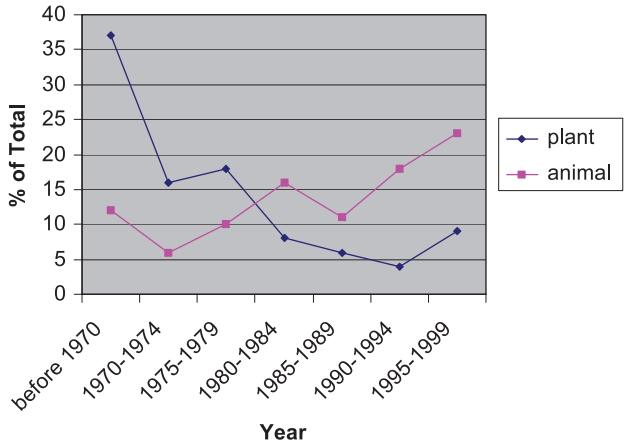


Fig. 1. Percent of total plant or animal hormetic dose responses by time period.

Reproducibility of findings is critical in establishing causality. However, the concept of reproducibility was difficult to implement because of the uncertainty over what constituted a bona fide case. It was decided that reproducibility could only be claimed in cases where the follow-up study was essentially identical as the original. This would explain why only seven citations (with a total of 7 dose responses) were determined to be reproducible between publications. Reproducibility reported within the same article occurred with 197 dose responses in 92 different

Table 7
Experimental models by year of publication and test system

Year of publication	Experimental model								
	Plant			Animal			Bacteria		
	Total	In vitro	In vivo	Total	In vitro	In vivo	Total	In vitro	In vivo
Before 1970	948 (37%)[a]	342 (36%)[b]	606 (64%)	295 (11%)	89 (30%)	206 (70%)	126 (36%)	117 (93%)	9 (7%)[c]
1970–1974	415 (16%)	46 (11%)	369 (89%)	164 (6%)	71 (43%)	93 (57%)	30 (9%)	30 (100%)	–
1975–1979	453 (17%)	87 (19%)	366 (81%)	255 (10%)	100 (39%)	155 (61%)	24 (7%)	24 (100%)	–
1980–1984	208 (8%)	46 (22%)	162 (78%)	434 (17%)	177 (41%)	257 (59%)	34 (10%)	34 (100%)	–
1985–1989	162 (6%)	27 (17%)	135 (83%)	275 (11%)	122 (44%)	153 (56%)	60 (17%)	60 (100%)	–
1990–1994	112 (4%)	32 (29%)	80 (71%)	473 (18%)	152 (32%)	321 (68%)	41 (12%)	41 (100%)	–
1995–1999	230 (9%)	62 (27%)	168 (73%)	584 (22%)	252 (43%)	332 (57%)	31 (9%)	28 (90%)	3 (10%)
2000–2003	66 (3%)	5 (8%)	61 (92%)	121 (5%)	92 (76%)	29 (24%)	2 (0.5%)	2 (100%)	–
Total	2594 (100%)	647 (25%)	1947 (75%)	2601 (100%)	1055 (41%)	1546 (59%)	348 (100%)	336 (97%)	12 (3%)

[a] Number in parentheses is the percent of the Total row (i.e. 948/2594).
[b] Number in parentheses is the percent for that year group (i.e. 342/948).
[c] Bacteria experiments were considered in vivo when the bacteria was injected into a host (e.g. rat).

Table 8
The number of repeat measures obtained during time course studies

Number of repeat measures	Number of dose–responses (% of the 923 dose–responses with time courses)
2	274 (30%)
3–5	448 (49%)
>5	201 (22%)

Table 10
The number of generations observed

Number of generations	Number of dose–responses (% of the 151 dose–responses)
1	9 (6%)
2	124 (82%)
3–5	12 (8%)
>5	6 (4%)

articles. However, it should be noted that a dose response is only considered reproducible when the article provides the results of the separate experiments. Cases where the data are combined and reported as the average of several studies or where the study authors claim that the results were reproducible and only the results from a representative study that were provided were not considered reproducible in the hormesis database due to the lack of data for confirmation.

Study design considerations

Agent

There have been about 900 different agents from approximately 75 different chemical classes that have been entered into the database. While 71% (i.e., 3978) of the 5632 dose responses entered into the database used a chemical agent, 29% (i.e., 1654) employed radiation or radioactive material. Table 5 provides the chemical classes and physical agents with the greatest frequency in the database. Calabrese and Blain (2004) have examined the hormetic response of metals in greater detail elsewhere.

Model

Plant and animal models are similarly represented in the database (approximately 45% for each; Table 6). Rats (562 dose-responses) and mice (642 dose-responses) have been the most commonly used animal models. Human models (generally in vitro) have also been commonly employed (i.e., 340 dose responses). Although no particular plant species predominated, many common species were noted including: wheat (187 dose-responses); fungi (127 dose-responses); and algae (139 dose–responses). Although plant and animal models were equally represented in the database, there has been a shifting over time in the frequency of cited studies from plant to animal models as is indicated in Table 7 and Fig. 1. More than one species were examined and compared in 344 (24%) citations.

The age of the models varied widely depending on the model system. Although there are too many different specific ages in the database that would be biologically meaningful to list in a table, the essential conclusion is that hormetic effects occurred in all different stages of development from in utero to elderly conditions. Seventy-five (5%) of the citations compared the effects of an agent on a certain endpoint across different ages.

Most of the dose responses in the database (i.e., 3919) either did not specify the gender studied or gender did not apply to the model (e.g., plant or cell lines). Males (1155 dose responses) and females (1057 dose responses) were used in approximately the same number of studies. Because males and females were both used together in some in vivo studies, but the results were not separated by sex, the total number of dose responses by sex will not equal the total number of dose responses because the study will list both the males and females together. If males and females are both examined in in vitro studies, the sex is listed as not applicable (N/A) unless the results are reported separately. The sexes were compared in 102 (7%) of the citations in the database.

Test system

The majority of the dose responses were obtained from in vivo (63%) experiments while the remainders (37%) were from in vitro experiments. The % of in vitro studies varied slightly by year and did not vary greatly between animal and

Table 9
The number of dose responses with hypothesis testing by publication year

Publication year	Number of dose–responses	Number with hypothesis testing	Percent of yearly total
Before 1970	1382	202	15
1970–1974	624	252	40
1975–1979	746	313	42
1980–1984	685	357	52
1985–1989	508	186	37
1990–1994	640	319	50
1995–1999	855	547	64
2000–2003	190	150	79

Table 11
Number of subjects per treatment group

Number of subjects per treatment group	Number of dose responses (% of total; 3250)	In vitro	In vivo
≤10	1904 (59%)[a]	907 (48%)[b]	997 (52%)
11–50	1018 (31%)	184 (18%)	834 (82%)
51–100	247 (8%)	28 (11%)	219 (89%)
101–1000	263 (8%)	15 (6%)	248 (94%)
>1000	38 (1%)	4 (11%)	34 (89%)

[a] Number in parentheses is the percent of the total dose–responses.
[b] Number in parentheses is the percent of the dose–responses for that number of subjects grouping.

Table 12
The number of dose responses by endpoint

Endpoint type	Number of dose–responses	Percent of total dose–response relationships (5632)	Plants (2596)	Animals (2601)	Bacteria (348)
Growth	2632	46.7	1859 (73%)[a]	595 (23%)	143 (41%)
Metabolic[b]	1382	24.5	465 (18%)	725 (28%)	168 (48%)
Reproduction[c]	420	7.5	169 (7%)	236 (9%)	6 (2%)
Immune response	431	7.8	–	431 (17%)	–
Survival	270	4.8	30 (1.2%)	194 (7%)	29 (8%)
Longevity	139	2.5	1 (0.04%)	137 (5%)	1 (0.3%)
Cancer	148	2.6	–	148 (6%)	–
Disease[d,e]	57	1.0	49 (2%)[d]	8 (0.3%)[e]	–
Damage[f,g]	36	0.6	18 (0.7%)[f]	18 (0.7%)[g]	–
Behavioral[h]	44	0.8	–	46 (2%)	–
Mutagenic[i]	59	1.0	5 (0.2%)	52 (2%)	1 (0.3%)
Developmental[j]	14	0.2	–	14 (0.5%)	–

[a] Number in parentheses is the percent of total for that specific model.
[b] Examples: DNA repair, enzyme activity, hormone levels, ROS production, ATP response, oxygen uptake, or urine volume.
[c] Examples: fecundity, hatching rate, eggs/female, number of young, number of resorptions, seed germination, fruit weight/plant, number of pods, seeds/plant, or number of flowers.
[d] Disease in plants refers to rot or spoilage of fruit, diseased plants, or the number of weeds.
[e] Disease in animals refers to infection (e.g., kidney infection) or parasites.
[f] Damage in plants refers to disintegrating roots, decay, number of holes caused by insects or oxidation.
[g] Damage in animals refers to cell rounding, ALT or LDH release, lesions, lipid peroxidation, or hyperplasia.
[h] Examples: distance traveled, flinches/min, number of bites, rearings, or head dips, or number of correct choices.
[i] Examples: micronucleus frequency, incidence of bent humeral bristles, chromosome aberrations, drug resistance, or DNA integrity.
[j] Examples: adult eclosing rate, or malformations.

plant models; there was no significant change in the % of in vitro studies until 2000 in animal models (Table 7). Because both in vivo and in vitro systems may be used, the study duration can vary from a few minutes to a few years depending on the system and endpoint. The exposure duration can vary just as extensively with exposures of a single injection or a few seconds of radiation to a lifetime of repeated exposures. Because of the variety of methods employed and both in vivo and in vitro studies included, the number of different controls employed and the exposure routes used are numerous.

Time course

Because hormesis may be related to an adaptive response, it may only be observed at certain times after or during exposure. Studies that examine an effect over different time periods demonstrate that in many cases, the hormetic effect is observed only at certain time points, while in other cases the hormetic effect is consistent over the time points measured. There are 923 dose responses in the database that examined an endpoint at more than a single time point. Table 8 indicates the time-course studies by the number of repeat measurements that were obtained. However, only one dose response was selected of the several comprising the dose–time-response relationship. The selected dose response typically was one that had the greatest conformity to the quantitative features of the hormetic dose response. Generally, this meant the dose response with the highest stimulation in the low dose range along with evidence of a high dose inhibitory response.

Table 13
Dose–response relationships by width of stimulation range

Width (-fold)	Number of dose–response relationships	Percent of total dose–response relationships (2516)[a]
≥1 < 10	1293	51.4
≥10 < 100	877	34.9
≥100 < 1000	232	9.2
≥1000	114	4.3

[a] The value of 2516 differs from the total number of dose–response relationships entered into the database (i.e. 5632) because calculation of the range was not possible for all dose responses for various reasons (e.g., data presentation precludes exact determination of range).

Table 14
Width of the stimulation range by model

Width (-fold)	Number of dose responses	Number in plants	Number in animals	Number in bacteria
≥1 < 10	1293	628 (59%)[a]	554 (45%)	89 (54%)
≥10 < 100	877	333 (31%)	469 (38%)	60 (37%)
≥100 < 1000	232	77 (7%)	138 (11%)	13 (8%)
≥1000	114	26 (2%)	85 (7%)	3 (2%)
Total	2516	1064 (100%)	1222 (100%)	165 (100%)

[a] The number in parentheses is the % of total for that model.

Hypothesis testing

A dose response was considered to have hypothesis testing if the study authors provided statistical results comparing the treatment group to the control. While only 41% of the dose responses had hypothesis testing that fit the criteria of the database, there were instances where statistical analysis was performed but the study authors did not provide comparisons between treatment and control. Table 9 demonstrates a fairly consistent frequency of statistical analyses from 1970 to 1994; there were fewer dose responses with hypothesis testing before 1970 and a greater percent after 1995.

Transgenerational

The database includes 151 dose responses from studies that examined transgenerational effects (Table 10). Each generation fitting the criteria is entered as a separate study in the database. Sometimes, studies examined the transgenerational effect, but only reported the second generation and not the generation exposed. Therefore, the number of generations is considered only one in the database.

Subjects

The number of subjects was present for inclusion in the database in 3250 (58%) of the dose responses (Table 11). Because some dose responses included a different number of subjects for each dose (e.g., controls had twice as many subjects or the endpoint was measured only in survivors and each group had a different number), the total will not be equal to the number of dose responses (i.e., 3250) that included the number of subjects. The number of subjects was considered the number that the study authors provided in the tables or figures and could include the number of experiments, the number of cultures, or the number of individuals. Table 11 demonstrates that the strong majority of the studies were performed in vivo, when the number of subjects was >10 subjects per treatment group.

Table 16
Dose–response relationships by maximum stimulatory response

Maximum stimulatory response (% control)	Number of dose–response relationships	Percent of total dose–response relationships[a]
J-shaped curve		
≤100 > 75	104	23.8
≤75 > 50	121	27.7
≤50 > 25	118	27.1
≤25	93	21.3
Inverted U-shaped curve		
≥100 < 110	25	0.5
≥110 < 150	3028	58.3
≥150 < 200	1032	19.8
≥200 < 500	840	16.2
≥500 < 1000	147	2.8
≥1000	119	2.3

[a] Total number of dose–response relationships for J-shaped curves = 436; total number of dose–response relationships for inverted U-shaped curves = 5196.

Endpoints

The dose–response relationships are divided into two different sections for endpoint (endpoint type, e.g., growth and endpoint parameter, e.g., body weight). Twelve endpoint types were selected for generalized search capacity. The 12 types are listed in Table 12 with the number and percent of dose–response relationships. Although growth (e.g., increase in cell number, body weights, or organ weight) was the type of endpoint measured most frequently, there were experiments that

Table 15
Width of stimulatory range by endpoint

Endpoint type	Number of dose–responses with ranges available	Width of the stimulatory range			
		≥1 < 10	≥10 < 100	≥100 < 1000	≥1000
Growth	1155	681 (59%)[a]	355 (31%)	82 (7%)	37 (3%)
Metabolic	618	283 (46%)	231 (37%)	79 (12%)	25 (4%)
Reproduction	183	107 (57%)	63 (35%)	11 (6%)	2 (2%)
Immune response	212	61 (29%)	70 (33%)	38 (18%)	43 (20%)
Survival	158	89 (56%)	62 (39%)	5 (3%)	2 (1%)
Longevity	60	28 (47%)	27 (45%)	5 (8%)	0
Cancer	66	18 (27%)	36 (55%)	9 (14%)	3 (5%)
Disease	3	0	2 (75%)	1 (25%)	0
Damage	13	3 (23%)	7 (54%)	1 (8%)	2 (16%)
Behavioral	22	12 (55%)	8 (36%)	2 (9%)	0
Mutagenic	24	9 (38%)	13 (54%)	1 (4%)	1 (4%)
Developmental	4	2 (50%)	2 (50%)	0	0
Total	2516	1293	877	232	114

[a] The number in parentheses is the % of total for that endpoint.

Table 17
Maximum stimulatory response by model

Maximum stimulatory response (% control)	Number of dose–response relationships	Plants	Animals	Bacteria
J-shaped curve				
≤100 > 75	104	13 (14.4%)[a]	90 (26.2%)	1 (50%)
≤75 > 50	121	25 (27.8%)	96 (27.9%)	0
≤50 > 25	118	20 (22.2%)	97 (28.2%)	1 (50%)
≤25	93	32 (35.6%)	61 (17.7%)	0
Total	436	90 (100%)	344 (100%)	2 (100%)
Inverted U-shaped curve				
≥100 < 110	25	9 (0.4%)	16 (0.7%)	0
≥110 < 150	3028	1594 (63.6%)	1225 (54.2%)	162 (46.4%)
≥150 < 200	1032	486 (19.4%)	454 (20.1%)	81 (23.4%)
≥200 < 500	840	334 (13.3%)	422 (18.7%)	66 (19.1%)
≥500 < 1000	147	55 (2.2%)	70 (3.1%)	19 (5.5%)
≥1000	119	28 (1.1%)	70 (3.1%)	18 (5.2%)
Total	5196	2506 (100%)	2257 (100%)	346 (100%)

[a] Number in parentheses is the percent of total for that model.

Table 18
Maximum stimulatory response in a J-shaped curve by endpoint

Endpoint type	Number of dose–responses with ranges available	Maximum stimulatory response			
		≤100 > 75 % of control	≤75 > 50 % of control	≤50 > 25 % of control	≤25 % of control
Metabolic	37	17 (47%)[a]	14 (39%)	6 (14%)	0
Reproduction	29	12 (41%)	8 (28%)	8 (28%)	1 (3%)
Immune response	28	4 (14%)	5 (18%)	5 (18%)	14 (50%)
Survival	31	7 (23%)	5 (16%)	10 (32%)	9 (29%)
Cancer	146	33 (23%)	44 (30%)	41 (28%)	28 (19%)
Disease	57	2 (3.5%)	9 (16%)	14 (25%)	32 (56%)
Damage	36	3 (8%)	13 (36%)	14 (39%)	6 (17%)
Mutagenic	52	21 (40%)	20 (38.5%)	10 (19%)	1 (2%)
Developmental	10	2 (20%)	1 (10%)	5 (50%)	2 (20%)
Total	436	104 (24%)	121 (28%)	118 (27%)	93 (21%)

[a] The number in parentheses is the % of total for that endpoint.

Table 19
Maximum stimulatory response in an inverted U-shaped curve by endpoint

Endpoint type	Number of dose–responses	Maximum stimulatory response					
		≥100 <110 % of control	≥110 < 150 % of control	≥150 < 200 % of control	≥200 < 500 % of control	≥500 < 1000 % of control	≥1000 % of control
Growth	2630	16 (1%)[a]	1709 (65%)	481 (18%)	343 (13%)	53 (2%)	28 (1%)
Metabolic	1345	0	630 (47%)	286 (21%)	290 (22%)	72 (5%)	67 (5%)
Reproduction	391	0	232 (59%)	61 (16%)	70 (18%)	12 (3%)	16 (4%)
Immune response	403	1 (0.3%)	201 (50%)	115 (29%)	76 (19%)	5 (1%)	3 (1%)
Survival	239	0	155 (65%)	46 (19%)	31 (13%)	5 (2%)	2 (1%)
Longevity	139	7 (5%)	74 (53%)	36 (26%)	22 (16%)	0	0
Behavioral	36	0	21 (56%)	8 (20.5%)	7 (20.5%)	0	0
Total	5196	25 (0.5%)	2963 (58%)	1006 (20%)	830 (16%)	146 (3%)	116 (2.5%)

[a] The number in parentheses is the % of total for that endpoint.

Table 20
Dose–response relationships by distance from the maximum stimulatory response to the ZEP[a]

Distance (-fold)	Number of dose–response relationships	Percent of total dose–response relationships (2345)[b]
≥1 < 5	1625	69.3
≥5 < 10	350	14.9
≥10 < 100	278	11.9
≥100	92	3.9

[a] ZEP = zero equivalent point (i.e. the highest dose showing a response equal to the control response).
[b] The value of 2345 differs from the total number of dose–response relationships entered into the database (i.e., 5358) because calculation of the distance was not possible for all dose responses for various reasons (e.g., data presentation precludes exact determination of ZEP).

examined a parameter within each of the 12 endpoint types. In many cases, the endpoint parameters may predict the same or closely related process (e.g., cell proliferation), but it is estimated via a different predictive means (e.g., DNA synthesis, tritiated thymidine uptake, cell numbers, etc.) and each is listed separately as an endpoint parameter.

Hormetic curve

The majority of the dose–response curves were inverted U curves (92%). Table 13 describes the width of the 2516 dose responses in which a range could be determined. Because any stimulation of adequate magnitude was considered for entry into the database and the curve did not have to provide a return to control values, the curve did not pass the 110% (or 97%) point twice in many cases for a range to be determined. In general, the majority of the dose responses displayed a stimulatory response range less than 10-fold wide. However, the response range could be highly variable with a low percentage displaying a stimulatory range that exceeded 1000-fold. While the width of the stimulatory range did not noticeably vary by biological model (Table 14), it did vary by endpoint (Table 15). Nearly one fifth of immune responses displayed a stimulatory range ≥1000-fold, a value far exceeding all other endpoints with a robust comparison number.

The relative proportions of maximum stimulatory responses for J-shaped curves were fairly evenly distributed while the maximum stimulatory response for inverted U-shaped curves were generally within 150% of the controls with nearly 80% of the dose responses being less than 200% of the control (Table 16). There were slight differences noted in the maximum stimulatory response for J-shaped curves between plants and animals. With respect to J-shape dose responses plants were more likely than animal models to have a maximum stimulatory response less than 25% of control while animals were more likely

Table 21
Distance from the maximum stimulatory response to the ZEP by model

Distance (-fold)	Number of dose–responses	Number in plants	Number in animals	Number in bacteria
≥1 < 5	1625	681 (71%)[a]	805 (68%)	114 (72%)
≥5 < 10	350	163 (17%)	154 (13%)	26 (16%)
≥10 < 100	278	85 (9%)	168 (14%)	15 (9%)
≥100	92	29 (3%)	60 (5%)	4 (3%)
Total	2345	958 (100%)	1187 (100%)	159 (100%)

[a] The number in parentheses is the % of total for that model.

Table 22
Distance from the maximum stimulatory response to the ZEP by endpoint

Endpoint type	Number of dose–responses with ranges available	Distance from the maximum stimulatory response to the ZEP			
		≥1 < 5	≥5 < 10	≥10 < 100	≥100
Growth	1063	733 (69%)[a]	185 (17%)	107 (10%)	38 (4%)
Metabolic	578	372 (64%)	95 (16%)	97 (17%)	14 (2%)
Reproduction	171	133 (78%)	21 (12%)	14 (8%)	3 (2%)
Immune response	194	103 (53%)	23 (12%)	35 (18%)	33 (17%)
Survival	154	131 (85%)	13 (8%)	8 (5%)	2 (1%)
Longevity	54	43 (80%)	4 (7%)	6 (11%)	1 (2%)
Cancer	69	55 (80%)	6 (9%)	7 (10%)	1 (1%)
Damage	10	7 (70%)	2 (20%)	1 (10%)	0
Behavioral	20	17 (85%)	0	3 (15%)	0
Mutagenic	30	26 (87%)	2 (7%)	1 (3%)	1 (1%)
Total	2345	1625	350	278	92

[a] The number in parentheses is the % of total for that endpoint.

Table 23
Evidence of hormesis[a] as used in the hormesis database

Evidence of hormesis	Number of dose–response relationships	Percent of total dose–response relationships (5632)
High	551	9.8
Moderate–high	250	4.4
Moderate	566	10.0
Low–moderate	1040	18.5
Low	3185	56.6
No–low	39	0.7

[a] For a description of the quantitative methodology used to derive evidence of hormesis see Calabrese and Baldwin (1997), *Hum. Ecol. Risk Assess*. 3:545–554.

than plants to have a maximum stimulatory response between 75% and 100% of the controls (Table 17). The maximum stimulatory response distribution in inverted U-shaped curves was fairly consistent across biological models (Table 17). There were also some variations across endpoint types in J-shaped curves (Table 18). Although there were also some variations across endpoint types for inverted U-shaped curves, the majority of the maximum stimulatory responses for all endpoint types were less than 200% of controls (Table 19).

Information on the distance from the maximum stimulatory response and the dose where the curve would cross the zero equivalence point (i.e., response equal to control; ZEP) is provided in Table 20. To calculate this point, the curve must have peaked and crossed the control value again. Generally, more than 70% of the cases in which this occurred were within a factor of five from the ZEP, which occurred regardless of biological model used (Table 21) or endpoint examined with the notable exception of immune responses which were 53% (Table 22).

Strength of evidence

Calabrese and Baldwin (1997) provided a numerical scoring system to determine the strength of evidence for assessing to what extent the dose response was consistent with the hormetic dose–response model, which is described above in the scoring methodology section. Table 23 separates the dose–responses by the evidence of hormesis using the Calabrese and Baldwin system. Table 24 indicates that the points are evenly distributed throughout the years, but between 1995 and 1999 when there was a notable increase in statistical analysis, there was also the greatest number of high scores.

Table 25
Comparison of the scores for the dose–responses in the hormesis frequency database with the general hormesis database

	Frequency database (Calabrese and Baldwin, 2001)		Hormesis database (present paper)	
	Number	%	Number	%
Total	245	100	5632	100
Performed hypothesis testing	87	36	2309	41
Low	130	53	3185	57
Low–moderate	65	27	1040	19
Moderate	28	11	566	10
Moderate–high	12	5	250	4
High	10	4	551	10

Comparison of hormesis-frequency database with hormesis database (present paper)

Although the majority of the dose responses entered into the database had low evidence of hormesis, the results reflect both the strength of the data supporting hormesis as well as the evaluative criteria applied to the study design and response data and their evaluation (e.g., statistical analysis). The Hormesis-Frequency Database (Calabrese and Baldwin, 2001) makes an absolute judgment on whether the evaluative criteria of hormesis were satisfied or not. If such criteria were satisfied, then hormesis was judged as present. In contrast, the Hormesis Database, which is the subject of the present paper, makes no absolute judgment on the existence of hormesis; rather it applies different evaluative criteria, which result in dose responses being characterized according to the degree to which they are consistent with the hormetic–biphasic dose response. Despite the fact that these databases were constructed for different purposes and used different evaluative criteria, it is important to note that when all 245 dose–responses that satisfied the evaluative criteria (i.e., hormesis) in the

Table 24
Evidence of hormesis by publication year

Publication year	High	Moderate–high	Moderate	Low–moderate	Low
Before 1970	127 (23%)[a]	51 (20%)	156 (28%)	262 (25%)	765 (24%)
1970–1974	61 (11%)	32 (13%)	64 (11%)	130 (13%)	337 (11%)
1975–1979	59 (11%)	48 (19%)	56 (10%)	158 (15%)	423 (13%)
1980–1984	41 (7%)	28 (11%)	75 (13%)	126 (12%)	408 (13%)
1985–1989	53 (10%)	20 (8%)	45 (8%)	81 (8%)	308 (10%)
1990–1994	53 (9%)	24 (10%)	70 (12%)	104 (10%)	383 (12%)
1995–1999	129 (23%)	39 (16%)	84 (15%)	135 (13%)	465 (15%)
Total	551 (100%)	250 (100%)	566 (100%)	1040 (100%)	3185 (100%)

[a] The number in parentheses is the % of total for that evidence of hormesis.

Hormesis-Frequency Database were assessed using the scoring system employed on the dose responses in the Hormesis Database, the distribution of the ranked scores were similar (Table 25). This indicates a high degree of evaluative concordance between the two different but complementary evaluative methodologies. This is important since it was our strong general impression before conducting this inter database comparison that the entry criteria of the Hormesis-Frequency Database (Calabrese and Baldwin, 2001) was considerably more stringent than the Hormesis Database, but this is not the case.

Discussion

The present assessment provides insight on the general robustness of published studies used to establish support for the hormetic dose–response hypothesis. The data reveal that the underlying literature basis is both quite extensive and broadly distributed across the range of biologically based sub-disciplines that utilize dose–response relationships. Dose responses receiving the highest numerical score in the evaluation scheme will possess the strongest features of study design most likely accompanied with statistically significant findings. Consequently, those studies providing the most support for the hormetic dose–response hypothesis are usually those with the strongest study designs and statistical power.

The heightened requirement for study designs with many doses has resulted in the consideration of biological models which are generally less expensive. This requirement has yielded a database consisting of nearly 50% of its dose–response relationships based on plants. More expensive investigations that involve lifetime studies with rodents using two or three dosages are generally inadequate to evaluate hormesis. Likewise, it is not uncommon for investigators to omit hypothesis-testing procedures of the hormetic response or fail to provide results in their analysis of data. Such studies would also have a markedly reduced likelihood of receiving a high score in the current system. Despite these factors, the range and frequency of the biological models and endpoints studied is very substantial and permits biological insights into factors affecting the nature of the hormetic dose–response relationship that appear to be well founded based on the consistency of their occurrence in the database.

When the present Hormesis Database was initiated, it was created and designed with the intent of providing a type of evaluative framework to assess whether hormesis was a viable toxicological hypothesis. However, the information comprising the database is useful in a number of additional ways. Most notably, it has permitted an opportunity to assess some of the quantitative features of the hormetic dose response. These findings provide insight on the magnitude of the stimulatory response, the width of the stimulatory response, and the relationship between the maximum stimulatory response and the NO(A)EL. These findings indicate that the maximum stimulatory response is generally approximately 30–60% greater than the concurrent control with nearly 80% of the maximum responses being less than two-fold greater than the control. Nearly 90% of the dose–response relationships display a stimulatory width that is less than 100-fold of the dose range immediately below the NO(A)EL. Thus, the stimulatory responses seen within the vast majority of cases within the hormetic database are typically of both modest magnitude and width. These dose response characteristics are biologically significant since they occur independent of biological model, endpoint, and chemical class and physical agent.

The toxicological mechanisms to account for these quantitative features of the hormetic dose response are not well understood. Calabrese and Baldwin (2001) hypothesized that the stimulatory response resulted from a modest overcompensation response to an initial disruption of homeostasis. While they provided substantial evidence to support this conclusion, only 17% of dose–response studies within the Hormesis Database include a temporal component that could be used to examine the overcompensation hypothesis.

The width of the stimulation range, which is typically quite narrow (i.e. < a factor of 10 from the NO(A)EL), may also extend well beyond a 1000-fold range depending on the experimental conditions. One possible contributory explanation relates to the heterogeneity of the population studied. The more homogeneous the population is, the narrower the range of the stimulatory response. Simulations of dose–response relationships using a wide range of assumed sensitivities revealed that one could readily contract or expand the width of the stimulatory zone (Calabrese, unpublished data). These activities suggest that the principal reason for the relatively narrow range of stimulation is due to the highly homogeneous nature of the sample populations under study. Since many experimental models are highly homogenous and experimental conditions are uniform except for the variable being tested, it creates the framework for more uniform (i.e., narrow stimulatory dose range) responses.

A major potential use of the hormetic database is its capacity to identify examples of hormetic dose–response relationships that can be selected for further evaluation. That is, the database permits one to efficiently explore a wide range of topic areas very efficiently. For example, the database could be useful in facilitating an assessment of age and/or gender on the occurrence of hormetic responses. Likewise, the database could enhance an evaluation of hormetic response by chemical class and endpoints of interest. This is particularly noteworthy since it is not easy to identify examples of hormesis in the biological and biomedical literature with routine searches. While the hormesis database should not be considered an exhaustive accumulation of data-based articles on hormesis, it is comprehensive however, thereby providing a very useful

starting point in project development. Finally, the Hormesis Database is a dynamic entity that is continuously being expanded. Thus, it is likely that such new entries will alter to some extent the summary percentages for different categories within the database. Nonetheless, the size of the current database is substantial and should provide reasonable stability for many of the categories presented.

Acknowledgments

Effort sponsored by the Air Force Office of Scientific Research, Air Force Material Command, USAF, under grant number F49620-01-1-0164. The U.S. Government is authorized to reproduce and distribute for governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsement, either expressed or implied, of the Air Force Office of Scientific Research or the U.S. Government.

References

Calabrese, E.J., Baldwin, L., 1997. A quantitatively-based methodology for evaluation of chemical hormesis. Hum. Ecol. Risk Assess. 3 (4), 545–554.

Calabrese, E.J., Baldwin, L.A., 2001. The frequency of U-shaped dose–responses in the toxicological literature. Toxicol. Sci. 62, 330–338.

Calabrese, E.J., Baldwin, L.A., 2003. The hormetic dose response model is more common than the threshold model in toxicology. Toxicol. Sci. 71 (2), 246–250.

Calabrese, E.J., Blain, R., 2004. Metals and hormesis. J. Environ. Monit. 6, 14–19.

Kaiser, J., 2003. Sipping from a poisoned chalice. Science 302, 376–379.

International Journal of Toxicology, 27:369–378, 2008

ISSN: 1091-5818 print / 1092-874X online
DOI: 10.1080/10915810802503735

Hormesis Predicts Low-Dose Responses Better Than Threshold Models

Edward J. Calabrese
Department of Public Health, Environmental Health Sciences Division, University of Massachusetts, Amherst, Massachusetts, USA

Edward J. Stanek, III
Department of Public Health, Biostatistics and Epidemiology Division, University of Massachusetts, Amherst, Massachusetts, USA

Marc A. Nascarella
Department of Public Health, Environmental Health Sciences Division, University of Massachusetts, Amherst, Massachusetts, USA

George R. Hoffmann
Department of Biology, College of the Holy Cross, Worcester, Massachusetts, USA

This study evaluated characteristics of the concentration-response relationships of chemicals from the U.S. National Cancer Institute (NCI) Yeast Anticancer Drug Screen database with respect to the threshold and the hormetic dose-response models. The database reported concentration-response studies of 2189 chemicals from a broad range of chemical classes. The biological end point was growth in 13 strains of yeast (*Saccharomyces cerevisiae*), most of which contain genetic alterations affecting DNA repair or cell cycle control. The analysis was limited to studies that satisfied a priori entry criteria for evaluation, including having two or more concentrations in the nontoxic zone (below a Benchmark Dose). The mean growth response compared to untreated controls of these doses was significantly greater than 100% in all 13 yeast strains, ranging from ∼105% to ∼111%. Under a threshold model, one would expect values more closely approximating 100%. Moreover, the distribution of responses below the BMD_5 for chemicals was shifted upwardly from the expectations of a threshold model for all strains. These results indicate that for the chemicals and yeast strains studied, the responses are more consistent with a hormetic model than a threshold model, and they strengthen previous results presented by Calabrese et al. (2006, *Toxicol. Sci.* 94:368–378). Taken together, the analyses provide strong evidence for hormesis, a phenomenon with a broad range of biomedical and toxicological implications.

Keywords Antitumor, Biphasic, Hormesis, Low Dose, Threshold, Yeast

Critical analysis and validation of low-dose-zone predictions by dose-response models, especially the threshold model, are important because this model has long been used as the default model for toxicological risk assessment in the United States for all noncarcinogenic chemicals entering commerce and food products (Beck et al. 2001). In some countries the threshold model may also be considered in assessing risks for carcinogenic agents (Seeley et al. 2001). The use of different default models has important implications in many areas, including the establishment of limits for chemical exposures. A better understanding of hormesis can also be important for the prevention of degenerative diseases and the development of safe, effective regimens for the treatment of cancer and other disorders. The implications of hormesis for a broad range of biomedical sciences and for risk assessment in toxicology and radiation biology have recently been reviewed (Calabrese 2008; Arumugam et al. 2006).

In a recent study Calabrese et al. (2006) used the U.S. National Cancer Institute (NCI) Yeast Anticancer Drug Screen database to compare the ability of a hormetic model and a

Received 25 April, 2008; accepted 7 August, 2008.
This effort was sponsored by the Air Force Office of Scientific Research, Air Force Material Command, USAF, under grant number FA9550-07-1-0248. The U.S. Government is authorized to reproduce and distribute for governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsement, either expressed or implied, of the Air Force Office of Scientific Research or the U.S. Government.
Address correspondence to Edward J. Calabrese, Department of Public Health, Environmental Health Sciences, Morrill I, N344, University of Massachusetts, Amherst, MA 01003, USA. E-mail: edwardc@schoolph.umass.edu

threshold model to predict responses to prospective antitumor drugs at concentrations below estimated toxic thresholds. A Benchmark Dose (BMD) methodology was used to estimate toxic thresholds. The NCI database contains 2189 chemicals, with each agent studied over five concentrations with approximately half-log concentration spacing (1.2 to 100 μM). Thirteen yeast strains, including strains with unique genetic alterations that affect cell cycle regulation or DNA repair functions and their wild-type counterparts, were used to test the 2189 chemicals under similar conditions. Analysis of patterns of response below a broad range of BMDs (2.5% to 12.5%) showed that the patterns were more consistent with a hormetic model than with a threshold model. These findings were consistent with earlier analyses of below-threshold responses in published literature, showing the hormetic dose-response model to be more broadly applicable than the threshold model in assessments made across a range of environmental and pharmaceutical agents, biological models, and end points (Calabrese and Baldwin 2001, 2003). A limitation of the pattern analysis used in Calabrese et al. (2006) is that it did not evaluate the magnitude of the hormetic effect below the toxic threshold or the consistency of individual chemical responses with hormesis. The current investigation extends the analyses applied to the U.S. NCI Yeast Anticancer Screen database in Calabrese et al. (2006) by estimating the average magnitude of the hormetic effect for each strain and the distribution of the hormetic effect over chemicals.

METHODS

A detailed description of the NCI database, experimental design, and evaluation methods is given in Calabrese et al. (2006). Briefly, data from stage 2 of the NCI testing procedure were evaluated on 2189 compounds considered prospective antitumor agents based on preliminary testing. Each agent was tested at 1.2, 3.7, 11, 33, and 100 μM.

The chemicals were tested in 13 strains, 11 of which contain mutations in genes that can affect susceptibility to toxicants and radiation by altering the capacity for DNA repair or cell cycle controls (Simon 2001; Holbeck and Simon 2007). The genetic alterations of interest are *bub3*, *mec2*, *mgt1*, *mlh1*, *rad14*, *rad18*, *rad50*, *rad52*, *sgs1*, and overexpression of *CLN2*. None of the strains are wild type in a strict genetic sense, but the strains designated "wild type," also called SPY50644 (*MATa erg6Δ::LEU2 pdr1Δ::LEU2 pdr3Δ::hisG::URA3::hisG ade2 ade3 leu2 ura3*), and SPY50780 (*MAT α erg6Δ::TRP1 pdr1Δ::LEU2 pdr3Δ::hisG ade2 ade3 leu2 trp1 ura3*) carry the wild-type alleles of the genes of interest in the other strains (Holbeck and Simon 2007; http://dtp.nci.nih.gov/yacds/index.html; http://dtp.nci.nih.gov/yacds/spy50644.html; http://dtp.nci.nih.gov/yacds/spy50780.html). With the exception of strain *rad50*EPP$^+$, the strains have enhanced susceptibility to chemicals owing to mutations in *erg6*, *pdr1,* and *pdr3* (Dunstan et al. 2002). The *erg6* mutation confers enhanced permeability by depleting membrane ergosterol, whereas the *pdr1* and *pdr3* mutations eliminate transcription factors required for the expression of chemical-efflux transporters and thereby confer pleiotropic drug resistance.

The responses in the NCI database were obtained from the growth of the yeast strain exposed to the compound relative to the growth of the same yeast strain in a solvent (i.e., DMSO) control. Yeast cells in the exponential phase of growth were inoculated into synthetic complete medium containing 2% glucose and the test chemical. The initial cell density was 10^4 cells per well containing 200 μl of medium. Each agent was assessed four times at the same five concentrations in each yeast strain. Chemicals were tested in 96-well plates, with 80 chemicals at the same concentration on one plate. The remaining 16 peripheral wells were used as controls, of which 4 were unexposed controls, 8 solvent controls, and 4 controls containing cycloheximide. The assay was deemed invalid if growth occurred in the presence of cycloheximide. All concentrations of a drug were incubated over the same 12-h period on different plates such that there were five plates run on the same chemical at the same time. The chemical location in the 96-well plate was systematic rather than randomly allocated. Employing a different source of chemical on each day and different daily yeast cultures maximized variability in response.

The response data consisted of a ratio of the optical density (OD) of the response well for the treatment divided by the mean of the OD readings of eight solvent-control wells for each concentration. OD readings were at 600 nm. This process was repeated on a second day, and the ratios from the 2 days were averaged. We refer to the average response as the replication response. Two replication responses were produced for each concentration in each strain. Although the data on the NCI website give the average of the two response values and the difference between the two values, the original OD values are not available due to a computer malfunction at the Fred Hutchinson Cancer Research Center (FHCRC) facility. In light of the loss of the original data, special efforts were made to ensure that a detailed understanding of the entire research methodology and all quantitative methods was obtained prior to undertaking our analyses and that our written understandings were confirmed by the Principal Investigator (Julian Simon) of the original yeast study (Calabrese et al. 2007).

Identifying Chemicals with High Concentration Inhibition

The concentration-response patterns for chemicals with concentrations below an estimated BMD were evaluated in order to estimate the average magnitude of response below the toxic threshold. To achieve this goal within a five-concentration study design, a priori entry criteria were created. Evidence of toxicity (i.e., a response of $\leqslant$80% of the control, called the inhibition response) at the highest concentration (100 μM) was required, in addition to having either two or three concentrations

(1.2 and 3.7 μM or 1.2, 3.7, and 11.0 μM) below the BMD, excluding concentrations used to determine the BMD. The BMD_5 was estimated by a linear interpolation on the log scale between the concentration immediately above and below the 95% response, similar to Calabrese et al. (2006). Concentrations used to derive the BMD were not used in the assessment of the predictions of the threshold and hormetic dose-response models.

Setting the Benchmark Dose (BMD)

The BMD_5 is the concentration at which the response is estimated to have decreased 5% below the control value (Crump 1984). The BMD_5 was chosen for the primary analysis because 5% is near the midpoint for the standard deviation of the yeast strains (which ranged from 3.0% to 7.5% for the 13 strains).

A BMD_5 was calculated for each of the 28,457 concentration-response experiments (2189 chemicals and 13 strains) using the average of two replications as the response. The BMD_5 was estimated through the following procedure:

1. The largest concentration with an average response below 95% was identified. Let this concentration be C_{below}, and let the associated response be R_{below}.
2. If the average response at the next smallest concentration was at least 95%, then let this concentration be C_{above}, and let the associated response be R_{above}. The BMD_5 is estimated by linear interpolation on the log concentration scale: $BMD_5 = \exp[\log(C_{\text{above}}) + (0.95 - R_{\text{above}})(\log(C_{\text{below}}) - \log(C_{\text{above}}))/(R_{\text{below}} - R_{\text{above}})]$.
3. If the average response at the next lowest concentration below C_{below} was less than 95%, then let this concentration be C_{below} with response R_{below}, and return to step 2.

Selection of Studies and Concentrations below the BMD_5

We identified all studies where there was one or more concentration below the concentration corresponding to C_{above}, which we refer to as the low-concentration range. We then constructed a set of data that included only the studies with one or more concentration in the low concentration range, and we analyzed responses from these studies for different chemicals for each strain of yeast. We report analyses for sets of chemicals with two concentrations below C_{above} separately from analyses for sets of chemicals with three concentrations below C_{above}. Our analysis does not include an assessment of chemicals with only one concentration below the estimated BMD_5 because such data are less robust than those for agents with two or three concentrations below the BMD_5 and because the only available concentration (i.e., 1.2 μM) is so close to the estimated toxic threshold that responses for some of these agents are apt to reflect residual (i.e., carryover) toxicity (Calabrese and Baldwin 2003). Data supporting the basis of this exclusion were published in Calabrese et al. (2006).

Evaluating Responses for Concentrations in the Low-Concentration Range

The response at each concentration was reported as the mean response constructed from two replications. The response for individual replications was not recorded, but the absolute value of the difference in response between the replications was reported. We added and subtracted one half of this difference to the mean response to form the basic replicate responses at each concentration.

For each yeast strain we use a linear mixed model to predict the average response in the low concentrations range for chemicals. In our models, response is reported as percent of control, as indicated in the methods. The model is specified in the context of a large population of possible chemicals, indexed by $s = 1, \ldots, N$, where for each chemical, measures of response indexed by $k = 1, \ldots, r$ are made at each low concentration indexed by $t = 1, \ldots, d$. In the yeast study, we fit separate models for chemicals with $d = 2$ low concentrations and for chemicals with $d = 3$ low concentrations. At each concentration, there are two replicates, such that $r = 2$. We represent a model for the kth measure of response at concentration t for chemical s as

$$Y_{stk} = \mu_{st} + E_{stk}$$

where $E(Y_{stk}) = \mu_{st}$ corresponds to the expected response at concentration t for chemical s. We represent $var(E_{stk}) = \sigma^2_{st}$. We allow the response variance to depend on the chemical and the concentration.

Our primary interest is in the average response for low concentrations, which we define for chemical s as $\mu_s = \frac{1}{M}\sum_{t=1}^{M} \mu_{st}$. Using this definition, we define a concentration effect for chemical s as $\delta_{st} = \mu_{st} - \mu_s$. We use the mean response over all chemicals, which is given by $\mu_s = \frac{1}{M}\sum_{t=1}^{M} \mu_{st}$ to define the chemical effect as $\alpha_s = \mu_s - \mu$. With these definitions, we can represent response for the kth replication of chemical s at concentration t by

$$Y_{stk} = \mu + \alpha_s + \delta_{st} + E_{stk} \quad [1]$$

The only random term in the model is the replication error term, E_{stk}.

The mixed model that we fit is parallel to model (1). The difference is that we make the additional assumption that the chemicals included in the analysis correspond to a simple sample from the larger set of chemicals, and the concentrations below the BMD_5 are a simple random sample of concentrations from that range. We distinguish the chemicals in the population from a selected chemical by the subscript i, and index the selected concentration by the subscript j. Representing the chemical effect and the corresponding concentration effects as random

TABLE 1
Number of chemicals satisfying the BMD_5 entry criteria according to strain

Strain	No. chemicals with 3 concentrations below BMD_5	No. chemicals with 2 concentration below BMD_5	No. chemicals with 1 concentration below BMD_5	Total no. of chemicals
CLN2oe	263	389	332	984
SPY50780	258	489	387	1134
bub3	215	433	408	1056
mec2	228	426	393	1047
mgt1	226	461	369	1056
mlh1	189	482	412	1083
rad14	178	479	411	1068
rad18	142	387	369	898
rad50	157	360	383	900
*rad50*EPP$^+$	74	215	321	610
rad52	157	338	379	874
sgs1	111	382	441	934
Wild type	253	394	311	958

effects, the resulting model is the mixed model given by

$$Y_{ijk} = \mu + C_i + D_{ij} + E_{ijk} \quad [2]$$

TABLE 2
Frequency description of chemicals by the number of strains where the response at 100 μM was less than or equal to 80% of the control and there is at least one concentration below the BMD concentration interval

No. strains with concentrations below the BMD_5	Frequency	Percent	Cumulative frequency	Cumulative percent
0	253	12.49	253	12.49
1	225	11.11	478	23.60
2	117	5.78	595	29.38
3	122	6.02	717	35.41
4	113	5.58	830	40.99
5	88	4.35	918	45.33
6	112	5.53	1030	50.86
7	114	5.63	1144	56.49
8	106	5.23	1250	61.73
9	125	6.17	1375	67.90
10	154	7.60	1529	75.51
11	168	8.30	1697	83.80
12	210	10.37	1907	94.17
13	118	5.83	2025	100.00

Note. The 164 chemicals with no inhibitory response in any strain are not included in this table.

where $E(C_i) = 0$, $E(D_{ij}) = 0$, $\text{var}(C_i) = \sigma_C^2$ and $\text{var}(D_{ij}) = \sigma_D^2$, with $\text{var}(E_{ijk}) = \sigma_e^2$, corresponding to the average replication variance over concentrations and chemicals (as indicated in Stanek and Singer [2004]). This model is a nested mixed model and is fit using SAS Proc Mixed (Littell et al. 2006).

There are three variance components in model (2), and each has meaning. The first variance component σ_C^2 is the variance of the distribution of mean response for chemicals. The size of this component provides evidence of the variability of response across chemicals. Relative to the average response for a chemical, σ_D^2 provides a measure of how variable concentration-specific responses are from the chemical mean. This variance is formed by averaging the chemical-specific concentration variances over chemicals in the population. The third variance component, σ_e^2, is the replication variance, averaged over chemicals and concentrations.

We fit this model to the yeast data for chemicals that have concentrations below the BMD_5 and estimated the mean response in addition to variance components for each strain of yeast. In addition, we used the results to predict the average response for each of the sample chemicals using the best linear unbiased predictor (BLUP), and we generated confidence intervals for these predictors.

RESULTS

Using the criterion of $r_{inhib} = 80$, concentration-response relationships were identified that met the criterion of a high-concentration inhibition ($\leqslant$80% of control value at the highest concentration tested). The results indicate that 77% (21,977 of 28,457 studies) satisfied the criterion for inhibition at the highest concentration.

TABLE 3
Summary of average response by strain for BMD_5 with three low concentrations based on a mixed model treating chemicals as random effects

Strain	*N*	Mean	Standard serror	Chemical *SD*	Concentration *SD*	Replication *SD*
CLN2oe	263	106.52	0.65	9.80	0.20	8.76
SPY50780	258	111.47	0.97	15.13	0.00	9.81
bub3	215	108.99	0.86	11.87	0.36	9.09
mec2	228	109.67	1.08	15.74	0.00	9.87
mgt1	226	108.80	0.99	14.29	0.36	9.50
mlh1	189	108.04	1.12	14.65	0.43	10.87
rad14	178	110.32	1.11	14.29	0.00	9.65
rad18	142	111.11	1.33	15.28	0.00	10.59
rad50	157	106.87	1.14	12.99	0.00	14.46
*rad50*EPP$^+$	74	111.62	1.69	12.31	1.09	15.08
rad52	157	109.47	1.24	15.12	0.00	8.27
sgs1	111	111.08	1.50	14.86	0.42	11.97
Wild type	253	109.00	0.86	13.04	0.00	10.51

The concentration-response studies that were eligible showed some differences among strains. The strain designated "wild type" had the lowest proportion of chemicals satisfying the inhibition entry criterion at 67%, and the *rad50* strain had the highest value (83%). The result for the wild-type strain is consistent with the fact that the genetic defects in most of the other strains would be expected to confer increased sensitivity to the toxic effects of some chemicals. In the case of *rad50*, which is deficient for recombinational repair by homologous recombination and nonhomologous end joining (Holbeck and Simon 2007), the enhanced sensitivity would encompass many agents that cause DNA damage.

Before exploring the low-concentration zone, we tabulated numbers of chemicals that gave a toxic response at the highest concentration. Nearly half the chemicals (1093 of 2189) had an inhibitory response equal or below $r_{inhib} = 80$ at their highest concentration (100 μM) for all the strains. In contrast, 164 chemicals had no concentration-response studies with an inhibitory response equal or below $r_{inhib} = 80$ at their highest concentration in any strain. There were 338 chemicals for which an inhibitory response was not apparent at the highest concentration for only one strain. Of these 338 chemicals, 229 (67.8%) occurred with the *rad50*EPP$^+$ strain, which has wild-type membrane permeability and an efficient efflux trans-

TABLE 4
Summary of average response by strain for BMD_5 with two low concentrations based on a mixed model treating chemicals as random effects

Strain	*N*	Mean	Standard error	Chemical *SD*	Concentration *SD*	Replication *SD*
CLN2oe	389	104.96	0.52	9.22	0.00	9.34
SPY50780	489	109.79	0.66	13.93	0.00	8.87
bub3	433	107.64	0.59	11.23	0.00	10.03
mec2	426	108.60	0.77	14.94	0.00	11.20
mgt1	461	108.41	0.68	13.76	0.00	9.75
mlh1	482	109.16	0.74	15.42	0.00	9.60
rad14	479	109.16	0.63	12.89	0.00	9.22
rad18	387	108.11	0.72	13.13	0.00	10.93
rad50	360	107.01	0.72	12.39	0.00	11.65
*rad50*EPP$^+$	215	109.30	1.15	15.33	0.00	14.12
rad52	338	109.32	0.86	14.93	0.00	10.68
sgs1	382	108.85	0.74	13.29	0.00	11.88
Wild type	394	109.61	0.74	13.79	0.00	10.16

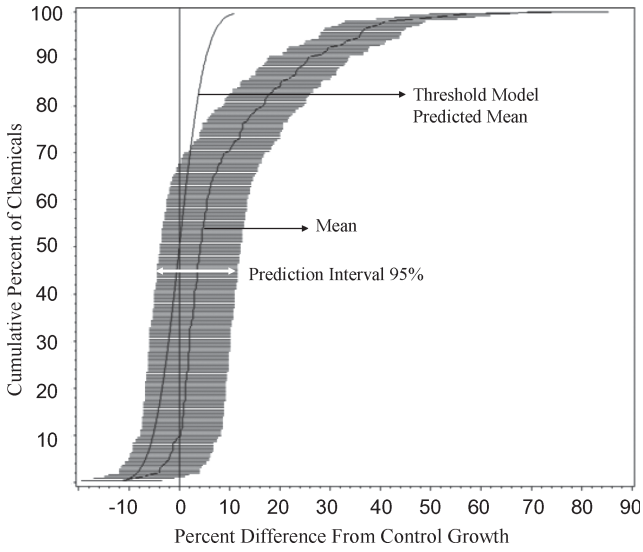


FIGURE 1
Distribution of predicted mean response and 95% prediction interval values of the 253 chemicals satisfying the a priori entry criteria for the wild-type strain with three responses below the BMD_5. These findings are compared to expectations for a threshold model.

porter system, thereby reducing susceptability relative to the other strains tested.

Of the 21,977 studies that satisfied the high concentration toxicity requirement, 12,602 (57.3%) had at least one concentration below the BMD_5. There were 2451 studies that met the inhibition criterion at the highest concentration and had three concentrations below the BMD_5, whereas 5235 studies had two concentrations below the BMD_5, and 4916 studies had one concentration below the BMD_5.

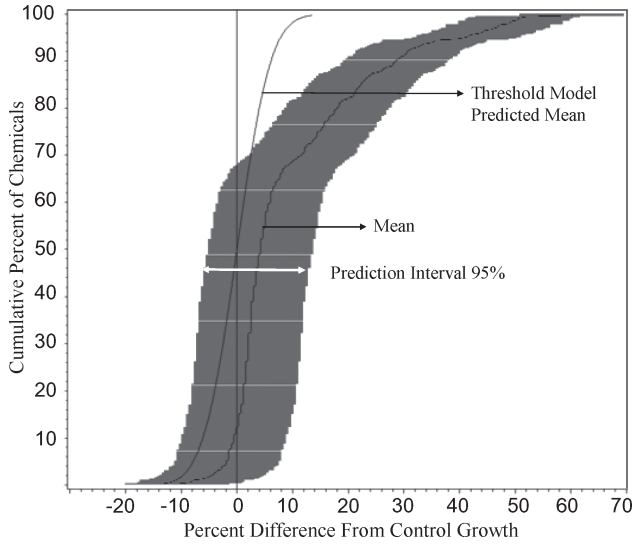


FIGURE 2
Distribution of predicted mean response and 95% prediction interval values of the 394 chemicals satisfying the a priori entry criteria for the wild-type strain with two responses below the BMD_5. These findings are compared to expectations for a threshold model.

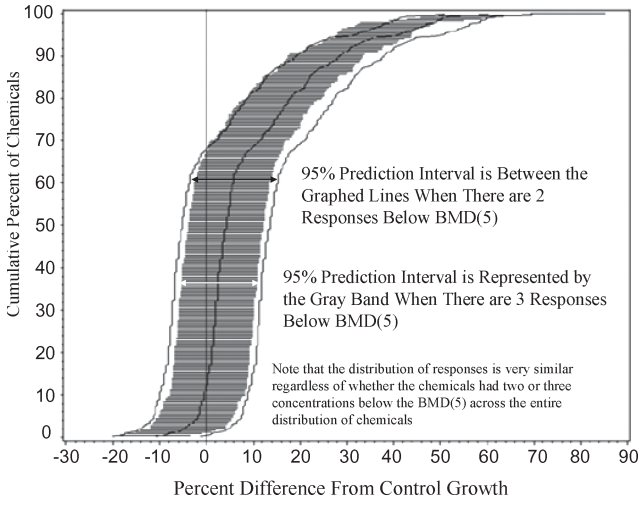


FIGURE 3
Comparison of the response distributions when 2 ($N = 394$ chemicals) or 3 ($N = 253$ chemicals) chemical concentrations are below the BMD_5.

Among the studies that showed evidence of high-concentration inhibition, the strain with the largest number of studies with concentrations below the BMD_5 is SPY50780, whereas *rad50*EPP$^+$ had the lowest number (Table 1). The low number for *rad50*EPP$^+$ is consistent with this strain's chemical efflux pump and diminished permeability reducing susceptibility to toxicants. The high number for SPY50780 is consistent with its wild-type repair and cell cycle control having led to fewer concentration responses being eliminated for reasons of excessive toxicity. The strain-specific responses reflect the entry criteria whereby responses are eliminated for nontoxicity at high concentration or excessive toxicity at low concentration.

We were also interested in differences among chemicals for numbers of studies where there was one (or more) concentration

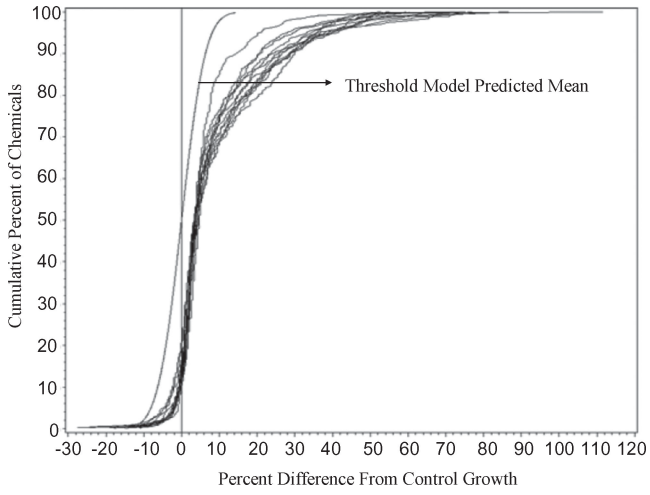


FIGURE 4
Summary of the predicted means for each yeast strain in the below-BMD_5 zone when there are three responses below the BMD_5 compared to the hypothetical threshold-model prediction. The data are in Table 5.

below the BMD_5 (Table 2). We only used concentration-response curves that met the high-concentration inhibition criterion. Of the 2189 chemicals in the database, 2025 chemicals had one (or more) concentration-response with an inhibition at the highest concentration. Of these, 118 chemicals had an inhibition for all strains, along with concentrations below the BMD_5 range.

Table 3 summarizes the mean responses of each yeast strain for all chemicals with a BMD_5 value between 33 and 100 μM. For these chemicals, there are three concentrations below the BMD. The estimate, standard error, and variance components were evaluated by fitting model (2) to each strain. There was considerable variation among strains for numbers of chemicals with three concentrations below the BMD_5, with *CLN2oe*, SPY50780, and wild type having the most and *rad50*EPP$^+$ having the fewest. The mean values for the 13 yeast strains range from 106.5% to 111.6%, compared to 100% for the control group's cell proliferation. Table 3 provides information on variation for each yeast strain indicated by the *SD* for chemical, concentration, and replication separately and their integration in the standard error value. Notice that the estimated variance of response for concentrations about the mean response, $\hat{\sigma}_D$, is zero for most strains. The estimate of zero is most likely due to the small number of observations per chemical and the similar response for concentrations relative to the variability in replicated response at a concentration. For all strains except *rad50* and *rad50*EPP$^+$, the standard deviation between chemicals is larger than the response variance. A similar assessment was made for chemicals with BMD_5 values within the concentration range beginning at 11 μM but less than 33 μM. These chemicals have two concentrations (i.e., 1.2 and 3.7 μM) below the BMD and gave results consistent with those in Table 3 with respect to the mean increase across strains (i.e., ~5% to 9.5%) and the standard error (Table 4).

Figure 1 shows a graph of the predicted responses of 253 chemicals with three concentrations below the BMD for the wild-type strain. The figure was constructed by first ordering from smallest to largest the predictor of each chemical (considered to be a realized random effect in model (2)), and then plotting the predictor along with the 95% prediction interval (PI) in increasing order. The abscissa represents the response (in percent) under or over 100. The ordinate has been scaled to 100%, representing 100% of the chemicals in the analysis. Each of the 253 horizontal lines in the plot represents a 95% prediction interval for a chemical. The dark line at the center of the prediction intervals connects the predictors for adjacent chemicals.

The continuous line on the left (black line with a value of 0 at 50%) is a plot of the mean response for the chemicals under the assumption that a threshold model holds for each chemical in the analysis. The departure of that line from the vertical line at 0 (on the abscissa) is due to replication and concentration variability. It is calculated by multiplying the percentile from a standard normal distribution by $\sqrt{\frac{1}{d}(\hat{\sigma}_D^2 + \frac{\hat{\sigma}_e^2}{r})}$, the estimated standard deviation of the chemical mean. Note that the threshold model predicts that some chemicals will display responses greater than 100% (upper end of the distribution) and that others will display responses less than 100% (lower end of the distribution). Such deviations from 100% are ascribable to chance.

A comparison of the actual data for the wild-type strain with the predicted responses of the threshold model indicates a consistent shift to the right across the entire distribution of 253 chemicals with three concentrations below the BMD_5 (Figure 1). These findings indicate that the threshold model does not match

TABLE 5

Percentiles of the responses by yeast strain in the below-BMD_5 zone when there are three responses below the BMD_5

Strain	P1	P10	P20	P30	P40	P50	P60	P70	P80	P90	P99
bub3	−6.8	−0.1	2.3	3.2	4.5	5.4	6.9	8.4	13.2	23.6	49.6
CLN2oe	−6.9	0.3	1.2	2.1	3	4.2	5.4	9.9	16.8	25.5	53.8
mec2	−6	−0.7	0.9	1.8	2.8	3.7	4.7	6.9	16.5	32.5	65.4
mgt1	−6.8	−0.3	1.2	2.2	3.1	4	4.7	6.5	12.7	29.1	57.9
mlh1	−14	−2	0.4	1.3	2.2	3.1	4	8.3	15.3	25.7	58.8
rad14	−11	−1.1	1	2.3	3.5	4.9	7.5	14.1	20.9	30.5	54.9
rad18	−9.1	−0.4	1.4	2.7	3.9	5.8	8.2	11.9	19.3	34.2	59.5
rad50	−14	−2.7	−0.5	0.9	2	3.4	5	8.6	13.1	20	57
*rad50*EPP$^+$	−18	−0.5	2.8	4.7	5.3	9.7	13.6	15.9	18.9	28.6	56.2
rad52	−15	−1.1	0.8	1.7	2.4	3.6	4.6	6.8	19.8	33.2	55.7
sgs1	−9.1	0.8	1.7	2	3.2	4.4	7.4	14.9	20	31.7	50.7
SPY50780	−6.9	0.3	1.2	2.1	3	4.2	5.4	9.9	16.8	25.5	53.8
Wild type	−6.9	0.3	1.2	2.1	3	4.2	5.4	9.9	16.8	25.5	53.8
Threshold	−9.4	−5	−3.1	−1.6	−0.3	0	2.3	3.9	4.9	6.2	9.4

Note. Responses are percent difference from control growth. The hypothetical distribution based on a threshold model is shown in the last row of the table. These data are shown graphically in Figure 4.

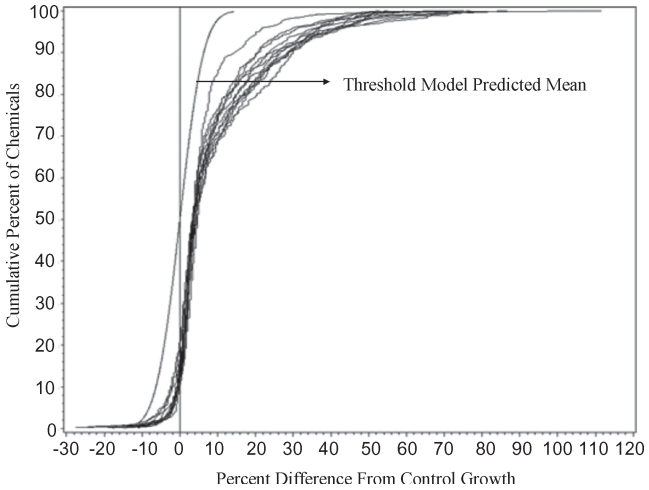


FIGURE 5
Summary of the predicted means for each yeast strain in the below-BMD_5 zone when there are two responses below the BMD_5 compared to the hypothetical threshold-model prediction. The data are in Table 6.

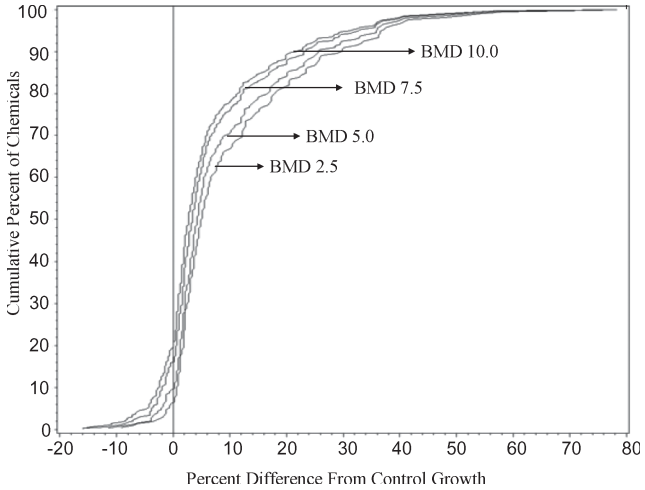


FIGURE 6
Assessment of the impact of the BMD criterion (2.5, 5.0, 7.5, or 10.0) on the cumulative distribution of mean responses for 253 chemicals with three responses below the BMD that satisfied a priori entry criteria for the wild-type strain.

the responses in the low-concentration (i.e., subtoxic) zone for this set of data. The upshift in response occurs across the entire distribution of chemicals. All chemicals gave responses compatible with a hormetic dose-response relationship. Below the 65th percentile, the overall upshift in response is about 4% to 5%. However, there is a marked and progressive increase in the upshift starting around the 65th to 74th percentile of the chemicals.

A similar comparison is presented in Figure 2 for the wild-type strain for 394 chemicals with two concentrations below the BMD_5. The overall shape and quantitative features of the plotted data are very similar to Figure 1. Figure 3, which superimposes Figures 1 and 2 on each other, demonstrates that the findings are quantitatively very similar. Thus, the hormetic dose response predominated regardless of whether there were two or three concentrations below the BMD_5, situations that reflect the grouping of agents that differ in toxic potency by about threefold. As in the previous case, a response compatible with hormesis was observed for all chemicals assessed. Similar results for each of the remaining 12 yeast strains are given in Figure 4/Table 5 and Figure 5/Table 6. All responses are upshifted from the estimates based on a threshold dose-response model regardless of whether there were two or three concentrations below the BMD_5.

TABLE 6
Percentiles of the responses by yeast strain in the below-BMD_5 zone when there are two responses below the BMD_5

Strain	P1	P10	P20	P30	P40	P50	P60	P70	P80	P90	P99
bub3	−12	0	1.7	2.5	3.8	4.6	6.7	8.4	12.9	20.4	44.6
CLN2oe	−10	−2.2	−0.2	1	2.6	3.4	4.6	5.8	7.8	14.9	36
mec2	−6	−0.7	0.6	1.5	2.4	3.7	5	7.2	15.1	24.7	66.8
mgt1	−7.5	0	1.4	2.3	2.7	3.6	4.9	6.7	12.9	25.4	60
mlh1	−8.3	−0.6	0.8	1.7	2.2	3.5	4.5	8.1	17.2	29.1	71
rad14	−6.1	−0.3	1	1.9	2.8	4.1	5.9	10.8	17	29	46.3
rad18	−7.7	−0.5	1.2	2.1	2.9	3.8	5.5	8.4	13.6	23.8	50.6
rad50	−9	−2.4	0	1.3	2.1	3.3	5.4	7.8	12.3	21.1	52
*rad50*EPP$^+$	−7.9	−1.3	1	2	3.3	4.5	6.6	10.3	16.9	26	60.2
rad52	−7.8	−0.7	1.1	1.5	2.4	3.7	5	9.5	17.9	29	68
sgs1	−11	−0.2	1.5	1.9	2.7	4	5.2	9.8	15.2	26.5	56.1
SPY50780	−6.4	−0.5	0.9	1.8	2.7	4.1	5.9	10.9	20.4	30	48.6
Wild type	−8.1	−0.6	1.1	2	2.9	4.2	6	11.7	18.3	28.4	50.9
Threshold	−11	−6.3	−4.1	−2.6	−1.3	0	1.3	2.6	4.1	6.3	11.3

Note. Responses are percent difference from control growth. The hypothetical distribution based on a threshold model is shown in the last row of the table. These data are shown graphically in Figure 5.

A similar response trend is seen for each of the 13 strains of yeast regardless of the BMD criterion (i.e., 2.5, 5.0, 7.5, or 10.0) and number of concentrations examined. Although the findings using different BMD criteria all support the occurrence of hormesis, there is an increase in the strength of the evidence for hormesis as the BMD criterion decreases from 10 to 2.5 (Figure 6). The most likely explanation is that the BMD_{10} reflects a low level of toxicity. This residual toxicity diminishes as the BMD criterion decreases toward 2.5. All 13 yeast strains gave responses supportive of the hormesis model, regardless of their diverse genetic alterations.

DISCUSSION

The present analysis indicates that the distribution of predicted responses in the low-concentration range is upshifted in a manner that is inconsistent with the threshold dose-response model. Although this was the case across the entire distribution of chemicals, it was most striking for the upper 20% to 40% of chemicals. The upshift was consistent across the 13 strains and did not appear to depend on the inherent toxicity of the chemicals, as indicated by comparisons of chemicals with two or three concentrations below the BMD_5. Figure 7 demonstrates the consistency in the wild-type strain of this upshifted distribution of responses relative to the threshold model predictions for chemicals with two or three concentrations below the BMD_5. These findings are consistent across all strains. The general upshift in the predicted responses is consistent with a hormetic dose-response model. The findings confirm and extend an earlier report of Calabrese et al. (2006) in which a different mode of analysis of the same database showed that the hormetic dose-response model better predicted below-threshold responses than did the threshold dose-response model. Taken together, the two analyses provide a more substantial and integrative perspective of hormesis in this large yeast database.

The data reveal that about 20% to 40% of the chemicals, depending on the yeast strain, show strong evidence of hormesis on an individual basis. The remaining 60% to 80% give responses that also support the hormetic model, in that they consistently differ from the prediction of the threshold model in the direction expected for hormesis. The strong evidence of hormesis for some chemicals suggests that beyond a general stress response there may be effects dependent on chemical structure. However, the NCI study protocol used only one time point for the measurement of cell proliferation; therefore one cannot discern whether differences in response magnitude are more likely to be related to chemical structure or to temporal factors that can affect adaptive responses.

The methodology that we used adjusts for the likelihood that variability apparent with small sample sizes will regress toward the average variation with repeat testing/sampling. This approach, commonly referred to as best linear unbiased prediction (BLUP) or empirical Bayes, provides more accurate predictors of the true chemical mean response than the simple mean. The regression towards the mean affects only chemicals whose predictor differs from the mean, not the mean itself.

The present methodology compared the distribution of the predicted response in the low-concentration region to results anticipated on the basis of the threshold dose-response model with the same concentration and response variances. Although the actual data were consistently upshifted compared to expectations of the threshold model, some treatments yielded results not only <110% of control responses but also less than 100%. Although such low values would not normally be considered to display a hormetic effect, we regard chemicals with such responses as consistent with hormesis because they displayed a clear upshift in response compared to the threshold-model distribution. Thus, the below-BMD treatments were shown to be upshifted across the entire population of chemical agents.

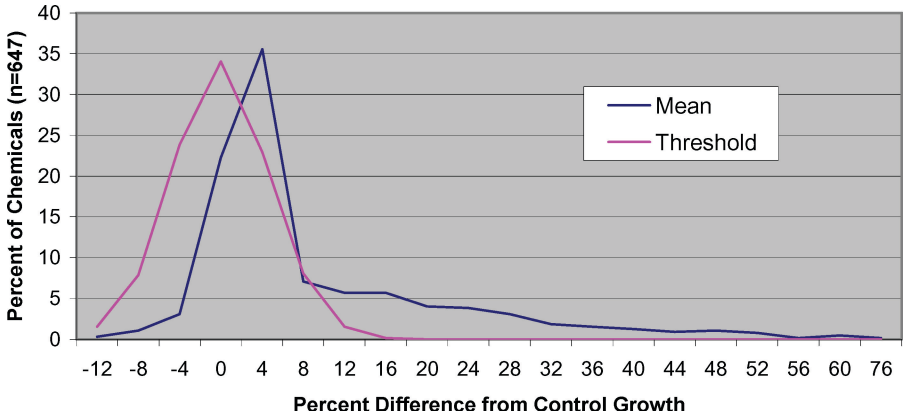


FIGURE 7

Histogram of mean wild-type (WT) yeast strain responses and expected threshold responses for 647 chemicals with two or three concentrations below the BMD_5.

We had hoped that the use of genetically altered yeast strains with well-characterized genetic alterations affecting DNA repair processes and cell cycle control might provide mechanistic insight into the hormetic stimulatory process, but the 13 strains were similar to one another in showing responses consistent with hormesis. However, further analysis of specific chemicals and yeast strains and their interactions may provide an opportunity to better define hormetic response pathways. The fact that the distribution of predicted responses consistently supported a hormetic model suggests that the hormetic response is general and not strongly affected by the diverse genetic changes built into the yeast strains of the NCI screening program.

These findings clearly indicate that elevated responses (over 100%) are to be expected for chemicals at concentrations in the low-concentration range in yeast, a conclusion consistent with the hormetic dose-response model. The threshold dose-response model was adopted in the 1930s and became accepted by national regulatory agencies for safety assessments without thorough critical evaluation with respect to the low-dose zone (Calabrese 2005a). In addition to the present findings, a large volume of peer-reviewed literature and experimental data are more consistent with a hormetic model than the threshold model for predicting responses to low-dose exposures (Calabrese and Blain 2005; Calabrese 2005b). These results further support calls for a reexamination of the threshold dose-response model and its use by regulatory agencies and suggests that the hormetic dose-response model may be useful for interpreting responses in the low-dose zone. Finally, these data suggest that antitumor drugs have the potential to enhance tumor cell proliferation at low concentrations. A recent comprehensive review of the published literature concerning human tumor cell lines is consistent with this perspective (Calabrese 2005c).

REFERENCES

Arumugam, T. V., M. Gleichmann, S.C. Tang, and M.P. Mattson. 2006. Hormesis/preconditioning mechanisms, the nervous system and aging. *Ageing Res. Rev.* 5:165–178.

Beck, B. D., T. M. Slayton, E. J. Calabrese, L. Baldwin, and R. Rudel. 2001. The use of toxicology in the regulatory process. In: *Principles and methods of toxicology*, ed. A. W. Hayes, 23–76. Philadelphia: Taylor & Francis.

Calabrese, E. J. 2005a. Historical blunders: How toxicology got the dose-response relationship half right. *Cell. Mol. Biol.* 51:643–654.

Calabrese, E. J. 2005b. Paradigm lost, paradigm found: The re-emergence of hormesis as a fundamental dose response model in the toxicological sciences. *Environ. Pollut.* 138:378–411.

Calabrese, E. J. 2005c. Cancer biology and hormesis: Human tumor cell lines commonly display hormetic (biphasic) dose responses. *Crit. Rev. Toxicol.* 35:463–582.

Calabrese, E. J. 2008. Hormesis: Why it is important to toxicology and toxicologists. *Environ. Toxicol. Chem.* 27:1451–1474.

Calabrese, E. J., and L. A. Baldwin. 2001. The frequency of U-shaped dose-responses in the toxicological literature. *Toxicol. Sci.* 62:330–338.

Calabrese, E. J., and L. A. Baldwin. 2003. The hormetic dose response model is more common than the threshold model in toxicology. *Toxicol. Sci*. 71:246–250.

Calabrese, E. J., and R. Blain. 2005. The occurrence of hormetic dose responses in the toxicological literature, the hormesis database: an overview. *Toxicol. Appl. Pharmacol.* 202:289–301.

Calabrese, E. J., J. W. Staudenmayer, E. J. Stanek III, and G.R. Hoffmann. 2006. Hormesis outperforms threshold model in NCI anti-tumor drug screening data. *Toxicol. Sci.* 94:368–378.

Calabrese, E. J., J. W. Staudenmayer, E. J. Stanek III, and G. R. Hoffmann. 2007. Hormesis and high throughput studies: Crump's analysis lacks credibility. *Toxicol. Sci.* 98:602–603.

Crump, K. S. 1984. A new method for determining allowable daily intakes. *Fundam. Appl. Toxicol.* 4:854-871.

Dunstan, H. M., C. Ludlow, S. Goehle, M. Cronk, P. Szankasi, D. R. H. Evans, J. A. Simon, and J. R. Lamb. 2002. Cell-based assays for identification of novel double-strand break-inducing agents. *J. Natl. Cancer Inst.* 94:88–94.

Holbeck, S. L., and J. Simon. 2007. The FHCRC/NCI yeast anticancer drug screen. In: *Yeast as a tool in cancer research,* ed. J. L. Nitiss and J. Heitman, 347–374. New York: Springer.

Littell, R. C., G. A. Milliken, W. W. Stroup, R. D. Wolfinger, and O. Schabenberger. 2006. SAS for Mixed Models, second edition. Cary, NC: SAS Institute.

Seeley, M. R., L. E. Tonner-Navarro, B. D. Beck, R. Deskin, V. J. Feron, G. Johanson, and H.M. Bolt. 2001. Procedures for health risk assessment in Europe. *Reg. Toxicol. Pharmacol.*, 34:153–169.

Simon, J. A. 2001. Yeast as a model system for anticancer drug discovery. *Emerg. Therap. Targets* 5:177–195.

Stanek, E. J., and J. M. Singer. 2004. Predicting random effects from finite population clusted samples with response error. *J. Am. Stat. Assoc.* 99:1119–1130.

TOXICOLOGICAL SCIENCES **62,** 330–338 (2001)

The Frequency of U-Shaped Dose Responses in the Toxicological Literature

Edward J. Calabrese[1] and Linda A. Baldwin

Department of Environmental Health Sciences, School of Public Health and Health Sciences, University of Massachusetts, Amherst, Massachusetts 01003

Received November 17, 2000; accepted April 17, 2001

Hormesis has been defined as a dose-response relationship in which there is a stimulatory response at low doses, but an inhibitory response at high doses, resulting in a U- or inverted U-shaped dose response. To assess the proportion of studies satisfying criteria for evidence of hormesis, a database was created from published toxicological literature using rigorous *a priori* entry and evaluative criteria. One percent (195 out of 20,285) of the published articles contained 668 dose-response relationships that met the entry criteria. Subsequent application of evaluative criteria revealed that 245 (37% of 668) dose-response relationships from 86 articles (0.4% of 20,285) satisfied requirements for evidence of hormesis. Quantitative evaluation of false-positive and false-negative responses indicated that the data were not very susceptible to such influences. A complementary analysis of all dose responses assessed by hypothesis testing or distributional analyses, where the units of comparison were treatment doses below the NOAEL, revealed that of 1089 doses below the NOAEL, 213 (19.5%) satisfied statistical significance or distributional data evaluative criteria for hormesis, 869 (80%) did not differ from the control, and 7 (0.6%) displayed evidence of false-positive values. The 32.5-fold (19.5% vs 0.6%) greater occurrence of hormetic responses than a response of similar magnitude in the opposite (negative) direction strongly supports the nonrandom nature of hormetic responses. This study, which provides the first documentation of a data-derived frequency of hormetic responses in the toxicologically oriented literature, indicates that when the study design satisfies *a priori* criteria (i.e., a well-defined NOAEL, ≥ 2 doses below the NOAEL, and the end point measured has the capacity to display either stimulatory or inhibitory responses), hormesis is frequently encountered and is broadly represented according to agent, model, and end point. These findings have broad-based implications for study design, risk assessment methods, and the establishment of optimal drug doses and suggest important evolutionarily adaptive strategies for dose-response relationships.

Key Words: **hormesis; compensatory responses; overcompensation; U-shaped; J-shaped; dose response; low doses; risk assessment; extrapolation.**

The occurrence of hormesis in the toxicological sciences has a long and controversial history (Calabrese and Baldwin, 2000a,b,c,d,e). Evidence supporting the existence of hormesis is substantial, with numerous reproducible examples suggesting potential broad generalizability (Calabrese *et al*., 1999). However, little information exists concerning the frequency of hormesis within the toxicological literature; that is, how often one would expect to observe hormesis given appropriate study design parameters. Two databases were previously created from the published literature to quantify aspects of hormetic responses in toxicological studies. In the case of Davis and Svendsgaard (1994), an attempt was made to estimate the incidence of hormetic responses based on the frequency of deviation from control responses independent of study design, NOAEL (no observed adverse effect level), and statistical significance. The second database (Calabrese and Baldwin, 1997a,b) focused on describing the quantitative features of the hormetic dose response and issues relating to generalizability rather than frequency in the toxicological literature.

Taking into consideration the limitations of the previous databases and incorporating suggestions by Crump (2001), a new database was created to assess the proportion of studies in the toxicological literature satisfying criteria for evidence of hormesis consistent with the definition of Stebbing (1998). Rigorous *a priori* entry criteria were established based on study design characteristics to identify data sets with the potential to detect a hormetic effect. Data sets meeting these criteria, independent of outcome, were entered into the database. Subsequent application of *a priori* evaluative criteria identified those dose-response relationships satisfying requirements for evidence of hormesis.

METHODS

Journal selection. Because a broad range of experimental models, end points, and agents, including mixtures, was desired, two environmentally-oriented toxicological journals (*Environmental Pollution,* 1970–1998; *The Bulletin of Environmental Contamination and Toxicology,* 1966–1998) and one pharmacologically oriented toxicological journal (*Life Sciences*, 1962–1998) were selected. Use of these journals ensured broad coverage of the toxicological literature without truncated end-point selection associated with more specialized journals. This was viewed as a desirable and necessary journal selection strategy at this stage of project development, as it would offer

The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Office of Scientific Research or the U.S. Government.

[1] To whom correspondence should be addressed. Fax: (413) 545-4692. E-mail: edwardc@schoolph.umass.edu.

greater opportunity to address issues of generalizability. Furthermore, inclusion of approximately 30 years of articles from each journal ensured the opportunity to incorporate independent peer review over prolonged periods, studies reflecting changes in toxicological funding priorities (thereby enhancing the range of chemicals, end points, and hypotheses assessed), improvements in study design, analysis, and technical developments as the field evolved, and assessment of historical trends if needed.

Screening protocol. All articles were initially screened in ascending chronological order beginning with volume 1, number 1 of each journal through 1998, with the exception of *Life Sciences*. Due to the increasingly large number of articles published per year in this journal (by the end of 1979 approximately 6000 articles had been screened with an annual publication rate increasing to over 600 articles), a decision was made to limit additional screening to 6 years, approximately equally spaced over the remaining 19 years of publication (1982, 1985, 1988, 1992, 1995, 1998). During the initial screening, exclusion and entry criteria described below were applied to all dose-response relationships reported in tabular or graphical form in each article. Dose-response relationships meeting the entry criteria were later examined with evaluative criteria described below for satisfying or not satisfying evidence of hormesis. The initial screening and the subsequent application of evaluative criteria were performed by the two authors; the results of the application of evaluative criteria were examined a second time by one of the authors.

Exclusion criteria. Only studies with experimental data were considered. Review articles, abstracts, non-English language articles, epidemiologic studies, and field studies were excluded. Studies lacking any of the following conditions were excluded: (1) a concurrent control; (2) the capacity to achieve responses greater than (or less than, depending on end point) the control response (e.g., studies where the end point was survival and the control response was 100% or where the end point was tumor incidence and the control response was zero); (3) at least two doses below the NOAEL; and (4) at least one dose showing *a priori* criteria-based inhibition.[2]

NOAEL designation. The NOAEL designation represents a unique dose that can be satisfied by only one dose. In the hormesis database this dose is satisfied by definitional determinants such that this dose represents the highest dose not differing from the control and having defined decrements at immediately higher doses. Any dose lower than this designated NOAEL that displays a response below that of the control would be interpreted as displaying either variability or error. As a result of this definition of NOAEL and applying it consistently throughout the database, possible subjective reinterpretation and designation of the NOAEL dose was prevented. The implications of this scheme were to allow for the inclusion of negative variability/error in the dose-response relationship below a designated NOAEL to permit false-positive estimation. If this approach had not been followed, some dose-response relationships could have been eliminated from satisfying entry criteria, ultimately resulting in a higher proportion of studies satisfying the evaluative criteria.

Residual bias may occur as a result of the NOAEL designation used in this assessment. Some doses that are characterized as NOAELs may in fact display evidence of low/modest toxic responses. However, if the decrement does not achieve a certain designated level (e.g., statistical significance, percent decrement), a determination could be made for that dose being the NOAEL. Thus, it is possible to inappropriately designate a bona fide LOAEL (lowest observed adverse effect level) as a NOAEL. This concern is widely recognized in regulatory toxicology and is one of the reasons why the NOAEL has been broadly criticized with respect to its no adverse effect designation. This possible limitation has led to proposals for application of statistical procedures, such as the benchmark dose (BMD), to estimate the NOAEL. If a NOAEL is actually a LOAEL in the current hormesis database, this would have implications for detection of hormesis at lower doses in the dose response spectrum. In fact, it could limit the potential detection to possibly one dose under certain study design scenarios. Again, even this one dose may still actually represent a type of LOAEL, if in fact it too had low residual deficits. This suggests that for dose responses in the present hormesis database where the NOAEL reflects a dose with a slight/modest toxic response, a false-negative potential for hormesis estimation may exist.

A decision was made in the development of the criteria to include as NOAELs for evaluative purposes doses that could satisfy evaluative criteria for evidence of hormesis. Although it is possible that one could have eliminated NOAELs within an evaluative designation, this approach was rejected, since the NOAEL, when it exceeds the control value, could be considered as being in the hormetic zone. This is because the designation of the NOAEL is not a perfect representation of the zero equivalent point (i.e., the highest dose with a response equal to the control response), but could err on either side of the control for real biological effect purposes. For this reason, it was decided that it would be unfair to bias a determination against a hormetic perspective. It should be noted that it was argued above that mischaracterization of a LOAEL with a NOAEL could lead to false-negative representation. However, allowing a NOAEL to be positively identified as a hormetic response is not a misrepresentation.

Entry criteria. The entry criteria were designed to ensure consistency with the U (or inverted U) shape of the hormetic dose-response relationship. That is, all studies needed to have sufficient evidence to demonstrate the occurrence of high-dose inhibition based on statistical and/or quantitative criteria, a NOAEL, and doses below the NOAEL that were to be evaluated for the potential of a low-dose stimulatory response based on statistical and/or quantitative criteria. Studies satisfying these general criteria were placed into one of three entry criteria tiers (T1, T2, T3) presented in Table 1: T1 includes dose-response relationships subjected to hypothesis testing; T2 was designed to identify dose-response relationships lacking hypothesis testing but reporting standard deviation (SD) or standard error of the mean (SEM) information, thereby providing information on the distribution of the data. T3 was designed to identify dose-response relationships defined only by data points reflecting mean/median values with no reference to variation.

Evaluative criteria. All dose-response relationships meeting the entry criteria were then subjected to evaluative criteria for evidence of hormesis (Table 1). An outcome satisfying criteria for evidence of hormesis is considered indicative of a dose-response relationship demonstrating stimulation at low doses and inhibition at higher doses. (See Fig. 1 for examples of data sets satisfying evaluative criteria.)

Where no hypothesis testing was performed, a difference of at least two SD or two SEM between the control and treatment group was considered indicative of potential statistical significance. Although our intent was to standardize all such data to conform to a similar distribution (i.e., SEM), this was not possible because of considerable variability in the nature and specificity of the information provided (e.g., out of a total of 196 dose-response relationships in category T2, 66 distributions were reported as SEM, 60 were reported as SD, and 70 were not identified). Consequently, we used the distribution provided in

[2] For purposes of this study, the NOAEL was defined as the highest dose with a response not statistically significantly different with respect to adverse responses from the control in studies where hypothesis testing was performed; in studies lacking hypothesis testing and in studies where hypothesis testing was performed but statistical significance was not observed with respect to adverse effects, the NOAEL was defined as the highest dose with a response $\geq$ 90% of the control for inverted U-shaped dose-response relationships or as the highest dose with a response $\leq$ 110% of the control for U- or J-shaped dose-response relationships. Inhibition was defined as occurring when: (1) the response for at least one dose higher than the NOAEL was statistically significantly different from the control in studies where hypothesis testing was performed; (2) the response for at least one dose higher than the NOAEL showed no 2$\times$ SD/SEM overlap with the control response in studies where only data distribution was reported; or (3) in the absence of statistical significance or nonoverlapping distributions, the response for at least two doses higher than the NOAEL was $<$ 90% of the control for inverted-U shaped dose-response relationships or $>$ 110% of the control for U- or J-shaped dose-response relationships.

TABLE 1
Summary of *a Priori* Entry and Evaluative Criteria with Descriptions of Outcomes Satisfying Evidence of Hormesis

Category	T1	T2	T3
Entry criteria			
Hypothesis testing	Yes	No	No
Data distribution reported	Not relevant	Yes	No
Minimum no. doses below NOAEL[a]	2	2	2
Minimum no. doses above NOAEL[b]	1 dose with a statistically significant response or 2 doses with responses < 90% of control	1 dose with a response showing no 2 × SD/SEM overlap with control or 2 doses with responses < 90% of control	2 doses with responses < 90% of control
Evaluative criteria[c]			
Outcome satisfying evidence of hormesis[d]			
Including and/or below the NOAEL responses	At least 1 stimulatory dose with a statistically significant response or at least 3 doses with responses ≥ 110% of control	At least 1 stimulatory dose with a response showing no 2 × SD/SEM overlap with control or at least 3 doses with responses ≥ 110% of control	At least 3 doses with responses ≥ 110% of control

Note. T1, dose-response relationships subjected to hypothesis testing; T2, designed to identify dose-response relationships lacking hypothesis testing but reporting SD or SEM, thus providing data distribution information; T3, designed to identify dose-response relationships defined only by data points reflecting mean/median values with no reference to variation.

[a]NOAEL, no observed adverse effect level. For the purposes of this study the NOAEL was defined as the highest dose with a response not statistically significantly different from the control with respect to adverse effects in studies where hypothesis testing was performed; in studies lacking hypothesis testing and in studies where hypothesis testing was performed but statistical significance was not observed with respect to adverse effects, the NOAEL was defined as the highest dose with a response ≥ 90% of the control for inverted U-shaped dose-response relationships or as the highest dose with a response ≤ 110% of the control for U- or J-shaped dose-response relationships.

[b]For the purposes of this study inhibition must be demonstrated as follows: at least one dose higher than the NOAEL with a response statistically significantly different from the control in studies where hypothesis testing was performed; at least one dose higher than the NOAEL with a response showing no 2 times SD/SEM overlap with the control in studies where data distribution was reported; or at least 2 doses higher than the NOAEL with responses < 90% of the control.

[c]Please note that these descriptions apply to inverted U-shaped dose-response relationships; in the case of J- (or U-) shaped dose-response relationships the evaluative response criterion value including and/or below the NOAEL is ≤ 90% of control and the evaluative response criterion above the NOAEL is > 110% of control.

[d]Consistent with the U- (or inverted U-) shape of the hormetic dose-response relationship.

the paper, recognizing that the comparison between studies would lack the intended uniform comparability.

In cases where data were graphically represented, on some occasions error bars were depicted for treatment data points, but not for the control. In those cases the dose responses were considered indicative of potential statistical significance if the error bars (SD/SEM × 2) of the treatment did not cross the control value.

In order to avoid exclusion of potentially relevant data below the NOAEL and to enhance the rigor of evaluative criteria, dose-response relationships with at least three doses with responses ≥ 110% of control (or with responses ≤ 90% for J- or U-shaped curves), i.e., alternative quantitative criteria were considered satisfying evidence of low-dose stimulation in the absence of statistical significance or potential statistical significance as determined by data distribution.

In order to avoid exclusion of potentially relevant data due to absence of a statistically significant or potentially statistically significant inhibitory response at high doses, dose-response relationships with at least two doses with responses < 90% of control (or > 110% for J- or U-shaped curves) were considered satisfying evidence of inhibition in the absence of statistical significance or potential statistical significance as determined by data distribution.

Assessment of false-positive responses. An indication of the frequency of false-positive responses (i.e., to what extent the positive findings could be accounted for by chance or random variation) was obtained by assessing the responses of treatment doses below the NOAEL and comparing the proportion of negative findings to positive findings. This is based on the assumption that if chance or random variation was responsible for the positive findings (i.e., a hormesis designation) then the number of negative responses should approximate the number of positive responses. It should be noted that although the NOAEL dose was included when assessing dose-response relationships with the evaluative criteria (Table 1), only treatment doses below the NOAEL were evaluated for false-positive responses. The NOAEL by definition cannot display an adverse (or negative) response, and its inclusion in the assessment of false-positives would therefore bias the outcome. By excluding the NOAEL values in this assessment, bias favoring false-positive estimation was minimized. This approach therefore provides a rate of false-positive/negative estimates that could be applied to the total rather than deriving the absolute number by direct estimation.

When the evaluative criteria were based on the response of single doses, the proportion of false-positive findings was derived by dividing the total number of doses below the NOAEL showing significant or potentially significant negative responses by the total number of significant or potentially significant responses of a positive and negative nature for both the hypothesis testing (T1) and distributional data (T2) categories. A similar procedure was employed to estimate false-positive findings when alternative quantitative criteria were used.

Assessment of false-negative responses. An indication of the frequency of false-negative responses was obtained by assessing the proportion of dose-response relationships satisfying the alternative quantitative evaluative criteria to the total number of dose-response relationships not satisfying evaluative criteria in the hypothesis testing category T1. This procedure was also applied

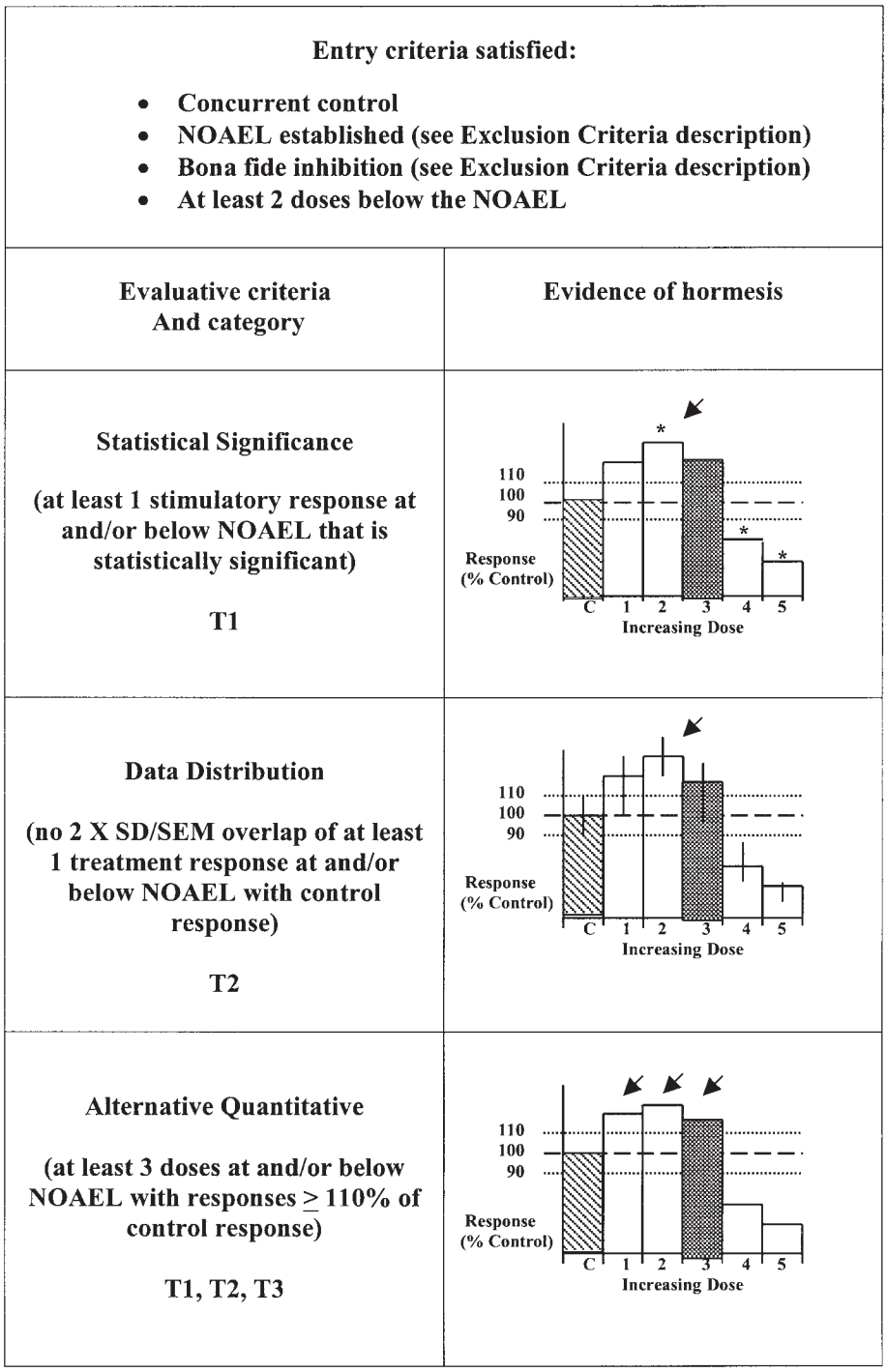


FIG. 1. Hypothetical data sets comprising a control (C, hatched lines) and five treatment groups (doses) satisfying entry criteria; data sets showing evidence of hormesis for each evaluative approach (statistical significance, data distribution, and alternative quantitative) and category. See Table 1 for category descriptions. NOAEL dose is indicated by cross-hatching; arrows indicate response(s) satisfying evaluative criteria for hormesis; *statistical significance at $p \leq 0.05$. Error bars represent the mean ± 2 SD. Please note that these descriptions apply to inverted U-shaped dose-response relationships; in the case of J- (or U-) shaped dose-response relationships, the evaluative response criterion value including and/or below the NOAEL is ≤ 90% of control.

to dose-response relationships that failed to satisfy evaluative criteria for the distributional data category T2.

RESULTS

Frequency of Hormetic Effects

Table 2 presents the results of application of the entry criteria organized by journal and year of publication. Of the 20,285 articles screened, 195 articles (1%) contained 668 dose-response relationships meeting the entry criteria. The number of articles screened was equally divided between the environmentally oriented journals (51.5%; 10,462 articles published in *Environmental Pollution* and *The Bulletin of Environmental Contamination and Toxicology*) and the more pharmacologically oriented journal (48.4%; 9823 articles published in *Life Sciences*). Approximately 1% of the articles in each journal

TABLE 2
Summary of Results of Application of *a Priori* Entry Criteria for Articles, Organized by Journal and Publication Year

Environmental Pollution				*Bulletin of Environmental Contamination and Toxicology*				*Life Sciences*			
Year	Published	Entered	No. d-r	Year	Published	Entered	No. d-r	Year	Published	Entered	No. d-r
								1962	127	0	0
								1963	162	0	0
								1964	217	0	0
								1965	333	1	1
				1966	38	1	4	1966	299	1	1
				1967	46	0	0	1967	350	0	0
				1968	45	0	0	1968	349	0	0
				1969	45	0	0	1969	344	1	1
1970	13	0	0	1970	103	0	0	1970	170	2	3
1971	23	0	0	1971	100	1	4	1971	168	1	2
1972	28	0	0	1972	130	0	0	1972	131	2	4
1973	52	2	17	1973	136	1	2	1973	247	1	1
1974	54	2	6	1974	234	1	1	1974	444	6	7
1975	58	0	0	1975	259	4	9	1975	459	0	0
1976	53	2	6	1976	246	3	11	1976	438	2	4
1977	88	0	0	1977	235	0	0	1977	494	5	8
1978	80	1	1	1978	247	2	10	1978	643	9	26
1979	83	0	0	1979	428	3	5	1979	593	4	9
1980	83	0	0	1980	312	1	3				
1981	86	0	0	1981	270	3	11				
1982	83	2	9	1982	248	2	6	1982	741	5	9
1983	65	2	3	1983	213	0	0				
1984	109	0	0	1984	210	1	2				
1985	107	3	50	1985	254	2	11	1985	623	11	27
1986	51	3	14	1986	283	1	2				
1987	138	2	66	1987	322	0	0				
1988	180	4	11	1988	260	3	5	1988	610	8	13
1989	159	2	3	1989	271	6	12				
1990	161	4	25	1990	276	3	17				
1991	115	2	8	1991	283	1	1				
1992	133	0	0	1992	277	5	27	1992	596	13	33
1993	150	1	2	1993	271	0	0				
1994	152	1	11	1994	277	3	7				
1995	168	0	0	1995	267	6	26	1995	658	12	20
1996	162	1	4	1996	289	4	14				
1997	127	1	1	1997	279	4	14				
1998	297	3	14	1998	250	1	2	1998	627	11	42
Total	3058	38	251	Total	7404	62	206	Total	9823	95	211

Note. Environmental Pollution was divided into two series during the years 1980–1986 (Series A, Ecological and Biological; Series B, Chemical and Physical). Series A was selected for screening. *Life Sciences* was divided into two parts during the years 1970–1973 (Part 1, Physiology and Pharmacology; Part 2, Biochemistry, General, and Molecular Biology). Part 1 was selected for screening. Due to the increasingly large number of articles published per year in this journal additional screening was limited to 6 years (1982, 1985, 1988, 1992, 1995, 1998); d-r, dose-response relationships meeting *a priori* entry criteria.

contained dose-response relationships meeting the entry criteria (*Environmental Pollution*, 1.2% $^{38}/_{3058}$; *The Bulletin of Environmental Contamination and Toxicology*, 0.8% $^{62}/_{7404}$; *Life Sciences*, 1.0% $^{95}/_{9823}$). The number of dose-response relationships meeting the entry criteria was approximately equally divided among the three journals (*Environmental Pollution*, 37.5% $^{251}/_{668}$; *The Bulletin of Environmental Contamination and Toxicology*, 30.8% $^{206}/_{668}$; *Life Sciences*, 31.5% $^{211}/_{668}$).

Figure 2 presents the results of application of the evaluative criteria to the 668 dose-response relationships satisfying the entry criteria organized by category. Two hundred forty-five (245) dose-response relationships (36.7% of 668) from 86 articles (0.4% of 20,285) satisfied the requirements for evidence of hormesis. Eighteen articles containing 118 dose-response relationships were in *Environmental Pollution*; 28 articles containing 68 dose-response relationships were in *The Bulletin of Environmental Contamination and Toxicology*; and 40 articles containing 59 dose-response relationships were in *Life Sciences*.

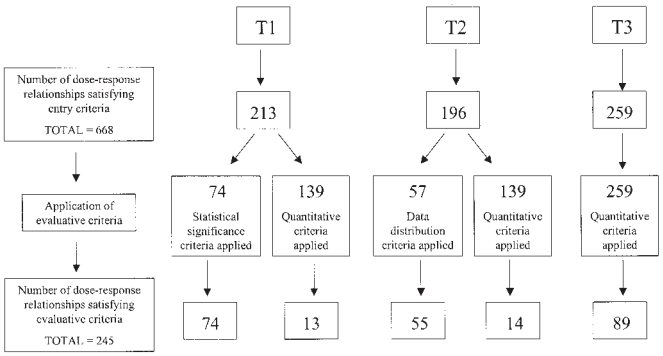


FIG. 2. Summary of application of the evaluative criteria to the 668 dose-response relationships satisfying the entry criteria, organized by category. See Table 1 for category descriptions.

Table 3 presents the results of the 245 dose-response relationships satisfying evaluative criteria for evidence of hormesis organized by experimental model, end point, and agent, including mixtures. A total of 73 different agents and mixtures from a broad range of chemical classes is represented.

Assessment of False-Positive Responses

Tables 4 and 5 present the results of the assessment of false-positive responses for each entry/evaluative criteria category. The collective findings indicate that the false-positive rate from the various categories was approximately 4%. When the false-positive/negative values were totaled, it yielded a net 1.4% false-positive estimate. This would reduce the 36.7% hormesis frequency to 35.3% (Table 6). These findings indicate that the methodology was not very susceptible to false-positive/negative error.

Assessment of False-Negative Responses

There were 139 dose-response relationships in T1 that satisfied entry but failed to satisfy evaluative criteria for hormesis. Thirteen of these 139 dose-response relationships satisfied the alternative quantitative evaluative criteria. This value provides an estimate of the false-negative rate for hormesis of 9.3% within the hypothesis testing criteria (category T1, Fig. 2). A similar procedure applied to the distributional data revealed a false-negative rate of 10.1% (category T2, Fig. 2). A direct comparison of dose-response relationships satisfying evaluative criteria for both hypothesis testing and alternative quantitative criteria revealed that such dose responses were approximately twice as likely to satisfy the evaluative criteria for hypothesis testing than for alternative criteria (i.e., of the 75 dose-response relationships satisfying hypothesis testing criteria, 38 also satisfied alternative criteria). That is, it is twice as difficult to have three doses at and/or below the NOAEL with responses $\geq$ 110% of the control response as to have one of these doses with a responses statistically significantly greater than the control. These findings not only strongly support the use of the methodology to estimate false-negative rates, but also indicate that the actual false-negative rates are likely to be higher than estimated. These results also suggest that the findings provided in the alternative criteria for hormetic estimates are considerably more rigorous than the hypothesis testing and distributional methods.

DISCUSSION

The findings indicate that in studies satisfying entry criteria, 36.7% satisfied the evaluative criteria for a hormetic response. Although the above assessment indicates that the study findings cannot be accounted for by false-positive responses or by

TABLE 3

Summary of the Dose-Response Relationships Satisfying *a Priori* Entry and Evaluative Criteria for Evidence of Hormesis Organized by Experimental Model, Endpoint, and Agent, Including Mixtures

	No. of dose-response relationships		
	Satisfying entry criteria	Satisfying evaluative criteria	%
Experimental model			
Plant	309	138	45 (138/309)
Vertebrate	266	83	31 (83/266)
Invertebrate	48	9	19 (9/48)
Microbe	43	15	35 (15/43)
Protozoan	2	0	0 (0/2)
Total	668	245	
Endpoint analyzed			
Metabolic	231	97	42 (97/231)
Growth	183	72	39 (72/183)
Reproductive	124	43	35 (43/124)
Molecular	64	11	17 (11/64)
Behavioral	36	17	47 (17/36)
Physiologic	17	3	18 (3/17)
Survival	13	2	15 (2/13)
Total	668	245	
Agent and Mixtures			
Effluents	116	67	58 (67/116)
Pesticides	114	30	26 (30/114)
Metals	85	30	35 (30/85)
Petroleum products/ constituents	24	15	62 (15/24)
Alcohol production wastes	19	16	84 (16/19)
Polycholorinated biphenyls	17	11	65 (11/17)
Solvents	10	6	60 (6/10)
Miscellaneous	283	70	25 (70/283)
Total	668	245	

Note. % = no. satisfying evaluative criteria/ no. satisfying entry criteria. Metabolic endpoints include enzyme activities, photosynthesis rate, respiration rate, protein synthesis, etc. Physiologic endpoints include muscle contraction/ relaxation, blood pressure, heart rate, etc. Pesticides include insecticides, herbicides, fungicides and ectoparasiticides. Miscellaneous includes a variety of agents and mixtures (e.g., pharmaceutical products, receptor agonists and antagonists, detergents, etc.).

TABLE 4
Assessment of False Positive Responses in Cases Where *a Priori* Evaluative Criteria Were Based on the Response of Single Doses below the NOAEL

Criteria			Number treatment doses below NOAEL			
Entry	Evaluative	Category	Total	Positive hormesis evidence	Chance positive evidence	False positive rate
Hypothesis testing	Statistical significance	T1	551	129	1	1/130 = 0.008 (0.8%)
Data distribution	Data distribution	T2	538	84	6	6/90 = 0.067 (6.7%)
			1089	213	7	7/220 = 0.032 (3.2%)

Note. Positive hormesis evidence, treatment doses below NOAEL with statistically significant or potentially significant stimulatory responses. Chance positive evidence, total number of treatment doses below NOAEL with statistically significant or potentially significant inhibitory responses. False positive rate, false positive/positive + false positive.

random variation, there are fundamental limitations in the current study methodology that are likely to yield a tendency for false-negative conclusions (values lower than actual hormesis estimates). The false-negative rate was nearly three times greater than the false-positive rate (i.e., 9.7% vs 3.5%). The false-negative criteria were established as being twice as rigorous as the false-positive estimation procedure. Finally, while false-positive evaluation was able to be applied to all possible instances for positive responses, this was not the case for the 170 negative dose responses in the alternative quantitative criteria for which no validation procedure is available. Given these three factors, it is likely that the 36.7% estimate of hormetic dose-response frequency is conservative and is likely somewhat higher.

In addition, the study did not take temporal factors into consideration. Numerous investigations exist that demonstrate stimulatory responses occur only following a disruption in homeostasis, that is, after an initial decrement in response (Stebbing, 1998; Calabrese, 2001). If responses were not taken at multiple times during the experiment, possible stimulatory responses could be missed, leading to false-negative conclusions.

It is interesting to note that of 1089 treatment doses below the NOAEL using hypothesis testing and distributional data entry criteria (Table 4), 213 (19.5%) of the treatment doses were determined to satisfy hormesis evaluative criteria. Only seven treatment doses (7/1089 = 0.6%) were significantly below the control. This suggests that hormetic responses in these categories occurred approximately 30-fold (19.5%/0.6% = 32.5) more frequently than a response of similar magnitude in the opposite (negative) direction. This finding, which employs the treatment doses below the NOAEL as the unit of comparison, provides striking support for the position that hormetic effects cannot be attributed to chance.

The data further revealed that the general occurrence for hormetic dose responses was widely incorporating of biological model, end points, and chemical classes. These findings represent the first attempt to assess the frequency of hormetic responses within the context of a biological/toxicological model based on study design, dose response, and statistical features. The results are particularly noteworthy, as they directly challenge the long-held view that hormetic responses should be seen as statistical exceptions, paradoxical findings, or otherwise unexpected events.

TABLE 5
Assessment of False Positive Responses in Cases Where *a Priori* Evaluative Criteria Were Based on the Response of Multiple Doses below the NOAEL

Criteria			Number dose responses with ≥ 3 doses below NOAEL			
Entry	Evaluative	Category	Total	Positive hormesis evidence	Chance positive evidence	False positive rate
Hypothesis testing	Alternative quantitative	T1	49	7	2	2/9 = 0.222 (22.2%)
Data distribution	Alternative quantitative	T2	68	10	1	1/11 = 0.09 (9%)
Alternative quantitative	Alternative quantitative	T3	104	40	0	0 (0%)
			221	75	3	3/78 = 0.038 (3.8%)

Note. The values of 110% and 90% refer to inverted U-shaped dose-response relationships; the values would be reversed in the case of U- (or J-) shaped dose-response relationships. Positive hormesis evidence, total number of dose-responses with at least 3 doses below the NOAEL with a response ≥ 110%. Chance positive evidence, total number of dose-responses with at least 3 doses below the NOAEL with a response ≤ 90%. False positive rate, false positive/positive + false positive.

TABLE 6
Adjustment for Potential False Positive Responses of the Number of Dose-Response Relationships Satisfying Evidence of Hormesis

Criteria						
Entry	Evaluative	Categories	Unadjusted positives/total	Est. false positive potential	No. false positive	Adjusted positives/total
Hypothesis testing	Statistical significance	T1	74/213	0.8%	1 (0.8%×74)	73 (74–1)
Hypothesis testing	Alternative quantitative	T1	13/139[a]	22.2%	3 (22.2%×13)	10 (13–3)
Data distribution	Data distribution	T2	55/196	6.7%	4 (6.7%×55)	51 (55–4)
Data distribution	Alternative quantitative	T2	14/139[a]	9.0%	1 (9.0%×14)	13 (14–1)
Alternative quantitative	Alternative quantitative	T3	89/259	0%	0	89 (89–0)
			245/668 =36.7%			236/668 = 35.3%

Note. The unadjusted positives/total value (i.e., 245/668 = 36.7%) includes correction for potential false negatives (i.e., positive responses with alternative quantitative criteria in categories T1 and T2).

[a]This is based on the assumption that these categories adjust for false negative responses by employing alternative quantitative evaluative criteria in cases where the data do not satisfy statistical significance (T1) or potential statistical significance as indicated by data distribution (T2). The estimated potential false negative responses are 9.3% (13/139) for dose-response relationships satisfying hypothesis testing entry criteria and 10.1% (14/139) for dose-response relationships satisfying distributional data entry criteria.

Although the above findings suggest that hormetic responses are quite common if assessed with the appropriate study design criteria, only 1% of the more than 20,000 published articles contained data meeting the study design criteria for entry into the database. This emphasizes the fact that very few published studies have the potential for detecting hormetic responses in the low-dose region of dose-response relationships. In fact, the criteria used in the present study ignored temporal features. If adequate temporal features were required, the proportion of studies satisfying entry criteria would have been far less than the 1%. Yet, if hormetic effects are to be adequately characterized, multiple appropriately spaced doses need to be assessed over multiple periods. The dual combination of multiple doses and periods places extraordinary demands on the investigation and are generally ignored, at least in part, thereby affecting the opportunity to assess hormetic effects. Thus, it is not surprising that hormetic effects have been considered exceptions or paradoxical responses, as our findings indicate that only 1 of 100 studies has the appropriate dosage design needed to assess this hypothesis.

Although there are multiple reasons why entry criteria were not satisfied, the most likely reason is that the hormetic evaluation has high study design criteria requirements, especially with respect to the number of doses and doses below the NOAEL. At a minimum, four doses plus a concurrent control are required, with two of the four doses being below the NOAEL. Historically, there has been a strong emphasis on high-dose evaluation, as these responses are often more definitive and publishable for defining the NOAEL. These factors minimize the proportion of experiments that emphasize below-NOAEL responses. Likewise, there has been the long-standing belief that responses below the NOAEL are most likely due to normal variation and not reproducible treatment effects. It is this very central assumption of modern experimental and regulatory toxicology that the present findings challenge. Yet, it is this historically controlling assumption that has strongly influenced past toxicological study designs and contributes to the observation that 99% of studies do not satisfy the entry criteria for hormetic responses.

The selection of the three journals noted in the Methods section was designed to achieve a broad representation of biological models, agents tested, and end points assessed. Although this approach was generally successful in achieving these goals, there were several important omissions or underrepresentations in certain categories. For example, the number of microbiological models was minimally represented; likewise, studies involving various types of radiation were also minimal. Nonetheless, areas such as microbiological responses and effects of radioactivity have been extensively documented and are represented in the earlier and separate database developed by Calabrese and Baldwin (1997a,b). Such underrepresentation in the present study is believed to be a result of the journal selection rather than a biological restriction of the hormetic response.

The findings presented here add to and strengthen the earlier reports on the potential widespread generalizability of hormesis. They provide a useful complement to the Calabrese and Baldwin (1997a,b) hormesis database, which includes several thousand examples of dose-response relationships satisfying quantitative criteria for assessing hormesis, as well as study replication and mechanistic findings that account for the biphasic nature of the dose and temporal responses of the hormetic phenomenon. Although numerous examples of apparent hormetic responses exist independent of chemical, biological model, and end point, the previous database cannot address the issue of frequency of occurrence of hormetic responses. The current study addresses this limitation and suggests that hormetic responses are commonly encountered if the study design is appropriate.

The present findings have important implications for the

design, conduct and interpretation of toxicological investigations as well as the potential to alter current concepts of NOAEL and challenge findings of risk assessment modeling activities commonly used for regulatory practices that assume linearity in low-dose areas. More specifically, for hormetic effects to be properly assessed, it is important that consideration be given to animal model and end-point selection. For example, an assessment of end points such as mutagenicity, carcinogenicity, and teratogenicity within a hormetic framework cannot be made using models with zero or negligible background/control incidence. It is also important to establish in a reliable manner the NOAEL for end points of interest and to include multiple and carefully spaced doses below the NOAEL. Furthermore, it may be necessary to include a temporal component within the study design if the hormetic mechanism represents an overcompensation response (Hart and Frame, 1996; Morré, 2000; Stebbing, 1998). The above suggestions are not trivial recommendations, as they require the commitment of substantial additional resources. Nonetheless, these features are necessary to more properly determine the nature of the dose response in the low-dose zone.

These findings address fundamental aspects of the nature of the dose response in the low-dose zone and suggest the need to incorporate U-shaped features in future modeling aspects of biological responses. Although the current investigation has focused on toxicologically derived data, sufficient data exist within the original hormesis database to indicate that this phenomenon is operational and similarly significant across the broad spectrum of biological, pharmacological, and other biomedical disciplines.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Kenny Crump (ICF Clements) for his guidance and technical assistance. This effort was sponsored by the Air Force Office of Scientific Research, Air Force Material Command, USAF, under grant F49620-98-1-0091. The U.S. Government is authorized to reproduce and distribute for Governmental purposes notwithstanding any copyright notation thereon.

REFERENCES

Calabrese, E. J. (in press). Overcompensation stimulation: A mechanism for hormetic effects. *Crit. Rev. Toxicol*.

Calabrese, E. J., and Baldwin, L. A. (1997a). A quantitatively-based methodology for the evaluation of chemical hormesis. *Hum. Ecol. Risk Assess*. **3,** 545–554.

Calabrese, E. J., and Baldwin, L. A. (1997b). The dose determines the stimulation (and poison): Development of a chemical hormesis database. *Int. J. Toxicol*. **16,** 545–559.

Calabrese, E. J., and Baldwin, L. A. (2000a). Chemical hormesis: Its historical foundations as a biological hypothesis. *Hum. Exp. Toxicol*. **19,** 2–31.

Calabrese, E. J., and Baldwin, L. A. (2000b). Radiation hormesis: Its historical foundations as a biological hypothesis. *Hum. Exp. Toxicol*. **19,** 41–75.

Calabrese, E. J., and Baldwin, L. A. (2000c). Radiation hormesis: The demise of a legitimate hypothesis. *Hum. Exp. Toxicol*. **19,** 76–84.

Calabrese, E. J., and Baldwin, L. A. (2000d). Tales of two similar hypotheses: The rise and fall of chemical and radiation hormesis. *Hum. Exp. Toxicol*. **19,** 85–97.

Calabrese, E. J., and Baldwin, L. A. (2000e). The marginalization of hormesis. *Hum. Exp. Toxicol*. **19,** 32–40.

Calabrese, E. J., Baldwin, L. A., and Holland, C. D. (1999). Hormesis: A highly generalizable and reproducible phenomenon with important implications for risk assessment. *Risk Anal.* **19,** 261–281.

Crump, K. S. (2001). The regulatory implications of hormesis: Is hormesis a universal phenomenon? *Crit. Rev. Toxicol.*, in press.

Davis, J. M., and Svendsgaard, D. J. (1994). Nonmonotonic dose-response relationships in toxicological studies. In *Biological Effects of Low Level Exposures: Dose-Response Relationships* (E. J. Calabrese, Ed.), pp. 67–85. CRC Press, Boca Raton.

Hart, R. W., and Frame, L. T. (1996). Toxicological defense mechanisms and how they may affect the nature of dose-response relationships. *BELLE Newsl.* **5,** 1–16.

Morré, D. J. (2000). Chemical hormesis in cell growth: a molecular target at the cell surface. *J. Appl. Toxicol.* **20,** 157–163.

Stebbing, A. R. D. (1998). A theory for growth hormesis. *Mutat. Res.* **403,** 249–258.

TOXICOLOGICAL SCIENCES **71,** 246–250 (2003)

The Hormetic Dose-Response Model Is More Common than the Threshold Model in Toxicology

Edward J. Calabrese[1] and Linda A. Baldwin

Department of Environmental Health Sciences, University of Massachusetts, Amherst, Massachusetts 01003

Received September 12, 2002; accepted October 15, 2002

The threshold dose-response model is widely viewed as the most dominant model in toxicology. The present study was designed to test the validity of the threshold model by assessing the responses of doses below the toxicological NOAEL (no observed adverse effect level) in relationship to the control response (i.e., unexposed group). Nearly 1800 doses below the NOAEL, from 664 dose-response relationships derived from a previously published database that satisfied *a priori* entry criteria, were evaluated. While the threshold model predicts a 1:1 ratio of responses "greater than" to "less than" the control response (i.e., a random distribution), a 2.5:1 ratio (i.e., 1171:464) was observed, reflecting 31% more responses above the control value than expected ($p < 0.0001$). The mean response (calculated as % control response) of doses below the NOAEL was 115.0% ± 1.5 standard error of the mean (SEM). These findings challenge the long-standing belief in the primacy of the threshold model in toxicology (and other areas of biology involving dose-response relationships) and provide strong support for the hormetic-like biphasic dose-response model characterized by a low-dose stimulation and a high-dose inhibition. These findings may affect numerous aspects of toxicological and biological/biomedical research related to dose-response relationships, including study design, risk assessment, as well as chemotherapeutic strategies.

Key Words: **hormesis; biphasic; risk assessment; dose response; linear; threshold.**

It is widely accepted in essentially all disciplines dealing with dose-response relationships that the threshold model is the overwhelmingly dominant paradigm (Hayes, 2001; Klaassen, 2001). This model can affect numerous aspects of research activities including biological model selection, endpoint measured, and study design. It may also affect the interpretation and modeling of dose-response relationships. The threshold model has long been used by regulatory agencies such as the FDA and EPA in establishing acceptable exposures to non-carcinogens.

Despite its clear dominance, the threshold model has been receiving strong challenges over the past decade. Perhaps the most notable challenge has been from the hormetic dose-response model (i.e., the biphasic model characterized by a low-dose stimulation and a high-dose inhibition). This model appears to be quite common in the biomedical and toxicological literature, with responses highly generalizable according to biological model, endpoint measured, and chemical and physical stressor agents tested (Calabrese and Baldwin, 2001a,b,c,d, 2002a,b; Calabrese and Baldwin, in press; Calabrese *et al.*, 1999).

The present study, which was designed to assess the capacity of the threshold model to predict responses of doses below apparent toxicological thresholds, demonstrated that not only was the threshold model unable to adequately account for the data, but also that the responses were consistent with the hormetic model.

METHODS

A previously described database created from the published toxicological literature, using rigorous *a priori* entry criteria (Calabrese and Baldwin, 2001a), was employed for the evaluation. The *a priori* entry criteria required the dose-response relationships to have a lowest observed adverse effect level (LOAEL), a NOAEL, at least two doses below the NOAEL, and a concurrent control. The database is comprised of 664 dose-response relationships from 195 articles and contains 1791 doses below the NOAEL. All responses were converted to values representing a percentage of their respective control response. In the calculation of responses as percentage of control, fractions were rounded to whole numbers (e.g., the category 100% contains values ranging from 99.6% to 100.4%.

RESULTS

For simplification purposes, the ratio of responses "greater than" (or above) the control response to responses "less than" (or below) the control response are referred to as an A/B ratio (above/below).

This study was sponsored by the Air Force Office of Scientific Research, Air Force Material Command, USAF, under grant number F49620-98-1-0091. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Office of Scientific Research or the U.S. Government. The U.S. Government is authorized to reproduce and distribute for Governmental purposes, notwithstanding any copyright notation thereon.

[1] To whom correspondence should be addressed as above at the Morrill I Science Center, N344. Fax: (413) 545-4692. E-mail: edwardc@schoolph.umass.edu.

TABLE 1

Responses (Represented as % Control Response) of 1791 Doses below the NOAEL from 664 Dose-Response Relationships Evaluated by Model, Endpoint, and Agent

	Number of responses				
	Total	<100%	100%	>100%	A/B ratio
Experimental model					
Plant	773	162	54	557	3.4/1.0
Vertebrate	795	229	79	487	2.1/1.0
Invertebrate	111	50	9	52	1.0/1.0
Microbe	108	22	14	72	3.3/1.0
Protozoan	4	1	0	3	3.0/1.0
Total	1791	464	156	1171	2.5/1.0
Endpoint analyzed					
Metabolic	637	168	49	420	2.5/1.0
Growth	425	87	26	312	3.6/1.0
Reproductive	329	99	33	197	2.0/1.0
Molecular	231	64	31	136	2.1/1.0
Behavioral	91	24	10	57	2.4/1,0
Physiologic	51	14	5	32	2.3/1.0
Survival	27	8	2	17	2.1/1.0
Total	1791	464	156	1171	2.5/1.0
Agent and mixtures					
Effluents	275	15	11	249	16.6/1.0
Pesticides	263	92	22	149	1.6/1.0
Metals	208	74	26	108	1.5/1.0
Petroleum products/constituents	58	10	1	47	4.7/1.0
Alcohol production wastes	47	0	0	47	—
Polychlorinated biphenyls	51	10	0	41	4.1/1.0
Solvents	23	3	0	20	6.7/1.0
Miscellaneous	866	260	96	510	2.0/1.0
Total	1791	464	156	1171	2.5/1.0

Note. In the calculation of responses as % control, fractions were rounded to whole numbers (e.g., the category 100% contains values ranging from 99.6% to 100.4%).

Doses below NOAEL. The mean response (calculated as the percentage of control) of the 1791 doses below the NOAEL was 115.0% ± 1.5 (SEM) and the median response was 105.0. Eleven hundred seventy-one responses were greater than the control response, 156 were equal to (i.e., rounded to) the control response, and 464 responses were less than the control response. The A/B ratios were 2.5:1 (1171/464). Table 1 summarizes the response distribution by experimental model, endpoint, and agent. The A/B ratios were 3.4:1 (557/162) for plant and 2.1:1 (487/229) for vertebrate models. The mean (± SEM) and median responses for plant models were 118.4% ± 1.2 and 110.0%, respectively. The mean (± SEM) and median responses for vertebrate models were 109.0% ± 0.8 and 103.0%, respectively.

NOAEL doses. The mean response (calculated as % control) of the 664 NOAEL doses was 107.1% ± 2.1 (SEM) and the median response was 98.0. Two hundred seventy-three NOAEL responses were greater than the control response, 37 NOAEL responses were equal to the control response, and 354 NOAEL responses were less than the control response. The A/B ratio was 0.8/1 (273/354). Table 2 summarizes the NOAEL response distribution by experimental model, endpoint, and agent. The A/B ratios were 1.1:1 (157/140) for plant and 0.5:1 (84/163) for vertebrate models. The mean (± SEM) and median NOAEL responses for plant models were 111.5% ± 1.7 and 101.0%, respectively. The mean (± SEM) and median NOAEL responses for vertebrate models were 98.0% ± 1.3 and 95.0%, respectively.

Effect of NOAEL response on the A/B ratio of responses of doses below the NOAEL. The A/B response ratio for doses below the NOAEL was markedly affected by the value of the NOAEL (Table 3). When the NOAEL response was greater than the control response the A/B response ratio of doses below the NOAEL was 6.7:1 (572/85); when the NOAEL response was equal to the control response the A/B response ratio was 2.5:1 (56/22), and when the NOAEL response was less than the control response the A/B response ratio was 1.5:1 (549/356).

TABLE 2

Responses (Calculated as % Control) of 664 NOAEL Doses from 664 Dose-Response Relationships Evaluated by Model, Endpoint, and Agent

	Number of responses				
	Total	<100%	100%	>100%	A/B ratio
Experimental model					
Plant	309	140	12	157	1.1/1.0
Vertebrate	262	163	15	84	0.5/1.0
Invertebrate	49	36	3	10	0.3/1.0
Microbe	42	14	7	21	1.5/1.0
Protozoan	2	1	0	1	1.0/1.0
Total	664	354	37	273	0.8/1.0
Endpoint analyzed					
Metabolic	232	118	13	101	0.9/1.0
Growth	177	80	6	91	1.1/1.0
Reproductive	128	70	12	46	0.6/1.0
Molecular	64	45	3	16	0.3/1.0
Behavioral	33	20	1	12	0.6/1.0
Physiologic	17	13	1	3	0.2/1.0
Survival	13	8	1	4	0.5/1.0
Total	664	354	37	273	0.8/1.0
Agent and mixtures					
Effluents	115	27	2	86	3.2/1.0
Pesticides	111	59	9	43	0.7/1.0
Metals	84	50	7	27	0.5/1.0
Petroleum products/ constituents	24	12	0	12	1.0/1.0
Alcohol production wastes	16	6	1	9	1.5/1.0
Polychlorinated biphenyls	17	8	1	8	1.0/1.0
Solvents	9	4	1	4	1.0/1.0
Miscellaneous	288	188	16	84	0.4/1.0
Total	664	354	37	273	0.8/1.0

Note. In the calculation of responses as % control, fractions were rounded to whole numbers (e.g., the category 100% contains values ranging from 99.6% to 100.4%).

TABLE 3
Relationship of the NOAEL Response (Calculated as % Control) to the A/B Ratio of Responses below the NOAEL

	Number of responses				
Response	Total	<100%	100%	>100%	A/B ratio
>110	421	36	13	372	
110	16	5	2	9	
109	20	1	2	17	
108	22	4	0	18	
107	19	6	0	13	
106	21	1	2	18	6.7 (572/85)
105	20	4	1	15	
104	26	2	3	21	
103	46	6	2	38	
102	43	14	7	22	
101	37	6	2	29	
100	100	22	22	56	2.5 (56/22)
99	44	16	4	24	
98	58	16	6	36	
97	57	15	6	36	
96	58	16	10	32	
95	104	37	11	56	
94	70	27	6	37	1.5 (549/356)
93	99	37	6	56	
92	100	29	9	62	
91	30	9	2	19	
90	86	34	14	38	
<90	294	120	21	153	

Note. In the calculation of responses as % control, fractions were rounded to whole numbers (e.g., the category 100% contains values ranging from 99.6% to 100.4%).

Effect of position of dose less than the NOAEL on response and A/B ratio. Table 4 summarizes the responses and A/B ratios of doses below the NOAEL, based on the position of the dose from the NOAEL (i.e., first dose, second dose, or third and greater doses less than the NOAEL). Although the mean and median responses are similar as the position of the dose relative to the NOAEL decreases, the response decreases (i.e., the highest response was observed in the dose closest to the NOAEL, the next highest in the second dose, and the lowest in the third and greater doses from the NOAEL). In contrast, the A/B ratio was observed to increase as the position of the dose from the NOAEL increased.

DISCUSSION

The findings indicate that the responses of doses below the NOAEL, where there is a transition between adaptation and toxicity, are nonrandomly distributed, with the strong majority having values greater than the control response. This observation is further supported by the mean response of 115.0% ± 1.5 (SEM) of below NOAEL responses. These findings challenge the long-held belief that treatment-related effects below the NOAEL are unexpected and, if observed, are simply normal variation.

Despite the fact that the data are at variance with predictions of the threshold model, it is important to emphasize that there is considerable overlap in the low-dose stimulatory zone between the threshold and hormetic model predictions. That is, for those doses with responses greater than the control value, it is not possible to distinguish between the two models. This is because the magnitude of the low-dose stimulation of the hormetic model is quite modest (i.e., 30–60% greater than controls at maximum) and consistent with a response often regarded as normal variability depending on the model employed and endpoint measured. Thus, the principal manner by which the two models can be differentiated is the nonrandom distribution of responses greater than the control response as seen in this study. In fact, it is the high degree of overlap between the two models that makes it very difficult to differentiate the threshold and hormetic models when only data from a single dose response are considered.

An assessment of dose-response relationships without evidence of hormesis suggests that this may be due, at least in part, to responses occurring at doses less than the identified NOAEL. Even though the NOAEL, by definition, does not differ in a statistically significant manner from the control, it is still quite possible that doses less than the NOAEL, especially that dose closest to the NOAEL, may display a distinct but lesser degree of toxicity than the NOAEL. If this were true, it would affect the capacity to detect hormetic responses in such dose-response relationships. This concept was evaluated by

TABLE 4
Effect of Position of Dose below the NOAEL Dose on Response (Calculated as % Control) and A/B Ratio

	Number of responses						
Position below NOAEL	Total	<100%	100%	>100%	Mean ± SEM (%)	Median (%)	A/B ratio
1st dose	664	191	41	432	118.7 ± 3.7	107	2.26
2nd dose	664	173	58	433	113.5 ± 1.3	106	2.50
≥3rd doses	463	100	57	306	111.8 ± 1.1	105	3.06

Note. Below the NOAEL dose, less than the NOAEL dose. In the calculation of responses as % control, fractions were rounded to whole numbers (e.g., the category 100% contains values ranging from 99.6% to 100.4%).

TABLE 5

Evidence Supporting Residual Toxicity in Dose-Response Relationships Not Showing Evidence of Hormesis When the NOAEL Response Is Equal To or Less Than 95% of the Control Response

		No. doses with responses <100% control response	
		Position of dose from NOAEL	
NOAEL value	No. of NOAEL doses	1st dose	2nd dose
95	32	17	14
94	18	11	7
93	28	16	13
92	24	13	9
91	5	4	3
90	26	14	13
<90	75	46	35
	208	121 (121/208=58%)	94 (94/208=45%)

assessing the responses of doses from dose responses not satisfying our functional definition of hormesis, in which the NOAEL was equal to or less than 95% of the control. In such circumstances it was hypothesized that the "residual (i.e., carryover) toxicity" response would progressively diminish with the lower doses. This suggests that the likelihood of observing a response greater than the control would increase as the dose decreased (at least within the optimal hormetic response zone). As seen in Table 5, this is the trend that is observed. There was a 25–30% increase (i.e., 121 vs. 94) in the number of responses equal to or greater than the control response when comparing the responses of the first and second doses less than the NOAEL. This observation suggests the occurrence of residual toxicity in such dose-response relationships. Since nearly 70% of vertebrate toxicology studies assessed here had NOAELs less than the control, it suggests the possibility of residual toxicity in a certain percentage of such dose-response relationships, a factor that could significantly underestimate the frequency of hormesis in the vertebrate toxicological literature. This observation also suggests why toxicologists may have mistakenly dismissed a hormetic hypothesis in favor of a threshold hypothesis.

An alternative interpretation to the residual toxicity hypothesis could involve the assumption that the population studied is highly heterogeneous, being comprised of a variety of subgroups with differential susceptibility. Under such a scenario, it is possible that one could account for such responses of doses less than the NOAEL based on subgroup-specific responses rather than the residual toxicity hypothesis. Numerous examples exist within the hormesis database in which hormetic responses occurred when the NOAEL was equal to or less than 95% of the control response. This suggests the possibility of population heterogeneity as a factor explaining the dose-response relationship (Calabrese and Baldwin, 2002a). In addition, it is possible that differences in study design, especially those relating to dose spacing below the NOAEL, could also be a critical determinant affecting the interpretation of the dose response. Further assessment in this area is necessary to clarify the factors affecting the nature of the dose response in the sub-NOAEL zone.

The threshold model is not only challenged by the nonrandom distribution of responses of doses below the NOAEL, but is further weakened by the data supporting the residual toxicity hypothesis in circumstances in which hormesis was not even observed. These findings raise the critical question of how the field of toxicology could have accepted the threshold model over the past century (Klaassen, 2001). This is particularly important since the concept of the dose response is the most central feature in toxicology.

The use of the NOAEL to provide a quasi-estimate of the threshold has both strengths and limitations. Most notably, it provides a statistically based framework by which a consistent comparison across the large number of dose responses may be evaluated. However, the precision by which the NOAEL provides a close approximation of the actual threshold is affected by the quality of the study design, especially with respect to the number of doses used and the nature of the dose spacing. Since the *a priori* entry criteria for the present study required a LOAEL, NOAEL, at least two doses below the NOAEL, and a concurrent control, the study designs were generally quite robust with respect to the number of doses, such that concerns dealing with adequacy of the NOAEL to provide a reasonable estimate of the threshold were minimized.

The findings are broadly generalizable according to endpoint measured, since the below NOAEL stimulatory response was the most dominant response for all general endpoint response categories and chemical classes (Table 1). These findings, along with the average magnitude of stimulation, are consistent with the published literature dealing with hormesis (Calabrese and Baldwin, 2001a,b,c,d, 2002b; Calabrese and Baldwin, in press; Calabrese *et al.*, 1999).

The implications of the findings are striking and challenge the fundamental teachings of the dose response and textbook treatment of this concept. They may affect study designs assessing dose-response relationships, risk assessment procedures for carcinogens and noncarcinogens, strategies for chemotherapeutic applications, and the selection of biological model and endpoints measured.

ACKNOWLEDGMENTS

We thank Ralph Kodell for his thoughtful and constructive review of the manuscript.

REFERENCES

Calabrese, E. J., and Baldwin, L. A. (2002a). Hormesis and high-risk groups. *Reg. Toxicol. Pharmacol.* **35,** 414–428.

Calabrese, E. J., and Baldwin, L. A. (in press). Hormesis: The dose-response revolution. *Annu. Rev. Pharmacol. Toxicol.*

Calabrese, E. J., and Baldwin, L. A. (2002b). Radiation hormesis and cancer. *Hum. Ecol. Risk Assess.* **8,** 327–353.

Calabrese, E. J., and Baldwin, L. A. (2001a). The frequency of U-shaped dose responses in the toxicological literature. *Toxicol. Sci.* **62**, 330–338.

Calabrese, E. J., and Baldwin, L. A. (2001b). Hormesis: U-shaped dose responses and their centrality in toxicology. *Trends Pharmacol. Sci.* **22,** 285–291.

Calabrese, E. J., and Baldwin, L. A. (Eds.) (2001c). Scientific foundations of hormesis. *Crit. Rev. Toxicol.* **31,** 351–695

Calabrese, E. J., and Baldwin, L. A. (2001d). U-shaped dose responses in biology, toxicology, and public health. *Annu. Rev. Public Health* **22,** 15–33.

Calabrese, E. J., Baldwin, L. A., and Holland, C. D. (1999). Hormesis: A highly generalizable and reproducible phenomenon with important implications for risk assessment. *Risk Anal.* **19,** 261–281.

Hayes, A. W. (Ed.) (2001). *Principles and Methods of Toxicology*, 4th ed. Taylor and Francis, Philadelphia.

Klaassen, C. D. (Ed.) (2001). *Casarett and Doull's Toxicology*, 6th ed. McGraw-Hill, New York.

Human & Experimental Toxicology (2000) 19, 41–75

www.nature.com/het

Radiation hormesis: its historical foundations as a biological hypothesis

EJ Calabrese*,[1] and LA Baldwin[1]

[1]*Department of Environmental Health Sciences, School of Public Health, N 344 Morrill Science Center, University of Massachusetts, Amherst, MA 01003, USA*

This paper represents the first systematic effort to describe the historical foundations of radiation hormesis. Spanning the years from 1898 to the early 1940's the paper constructs and assesses the early history of such research and evaluates how advances in related scientific fields affected the course of hormetic related research. The present effort was designed to not only address this gap in current knowledge, but to offer a toxicological basis for how the concept of hormetic dose-response relationships may affect the nature of the bioassay and its role in the risk assessment process.
Human & Experimental Toxicology (2000) **19,** 41–75

Keywords: hormesis; low dose; stimulation; β-curve; radiation

Introduction

Since 1980 there have been two books concerning radiation hormesis,[1,2] various international symposia directly related to this topic,[3–7] and a substantial number of articles. However, none of these attempts to describe and assess the concept of radiation hormesis has addressed, except in a very limited fashion, the historical foundations of this concept. In fact, we have been unable to uncover any attempt to assess this topic, even in the earlier decades of the 20th century, despite a substantial effort to uncover such possible efforts. This paper therefore is designed to provide a comprehensive and critical review of the historical foundations of radiation hormesis, with particular emphasis on ionizing radiation. The timeframe of the paper encompasses the late 1890's to approximately 1940. A parallel type of evaluation was recently published concerning the historical foundations of chemical hormesis[8] and how it became marginalized within the toxicological community.[9]

At the onset of this paper it is important to define the term hormesis. Hormesis is a concept that describes the nature of dose-response relationships in biological systems as displaying a stimulatory response at low doses and an inhibitory response at higher doses. Recently Calabrese and Baldwin[10,11] have attempted to quantitatively define this relationship with respect to the dose range of the stimulatory response, the maximum stimulatory response and the relationship of the maximum stimulatory response to the traditional toxicological No Observed Adverse Effect Level (NOAEL). Although this proposed scheme is consistent with the vast majority of data currently assessed on this topic, notable and reliable exceptions do exist to this framework and have recently required a broader delineation of the above defined nature of the hormetic dose-response relationship and its mechanistic underpinnings.[12] The present paper has been guided in this evaluation of hormesis by the above quantitative criteria without regard for whether the low dose stimulatory response is deemed beneficial, harmful or of unknown biological significance.

This paper has opted for a broad search of the biological/radiobiological/toxicological literature including responses to plants, bacteria, fungi, other micro-organisms, invertebrates and vertebrates including human epidemiological/clinical data. This broadly based biologically oriented approach was principally designed to assess to what extent the concept of hormesis may be generalizable. This approach also sought to provide an evaluation of radiation hormesis as a biological hypothesis rather than as an explanatory feature of selected medical practices, such as in low dose radiological practices in traditional medicine or as a possible theoretical framework of the practice of homeopathy. It should also be noted that the term hormesis was not coined until 1943 by Southam and Erhlich[13] who were assessing chemical extracts from cedar wood on fungi. However, the concept of hormesis was

*Correspondence: EJ Calabrese
Received 15 October 1999; accepted 15 October 1999

embodied in terms such as the Arndt-Schulz Law and Hueppe's Rule, which came into widespread, but not universal, use in the early 1900's based initially on the independent work of Schulz[14,15] with yeast and Hueppe[16] with bacteria.

This review of the historical foundations of radiation hormesis will ironically conclude at about the same time the term hormesis was coined. Thus, the concept of low dose stimulation, high dose inhibition has had three specific designations over the century (i.e., Arndt-Schulz Law, Hueppe's Rule, hormesis), yet one underlying concept and these terms have typically been used interchangeably.

The information contained here will provide an assessment of the status of the hormesis hypothesis in the radiation and toxicological communities up to the 1940's. This paper will then serve as a basis (see companion paper[8]) to evaluate how this concept became abandoned by the mainstream leaders of both radiation and toxicology during the middle and later decades of the 20th century. Finally, a third paper[9] will offer a comparative assessment of both chemical and radiation hormesis with respect to differential development of an hormetic hypothesis, the relative strengths and weaknesses of their underlying data, and the differential factors affecting the acceptance of both hypotheses.

Plants

Introduction

The evaluation of the potential for radiation to stimulate plant growth has a long and complex history. Such an evaluation of plant responses to radiation is seen within the context of the type/source of irradiation including X-rays and naturally occurring sources such as radium, cobalt and other elements that emit various types of radiation including gamma, beta and alpha rays. Each type of radiation has a unique history and will be assessed separately.

The present review is designed to assess the historical foundations of the response of plants to radiation especially as it pertains to the nature of low dose responses. In the case of X-rays this historical review encompasses nearly 40 years, spanning the years from 1898 when the first claims of a low dose stimulatory response were reported to the 1940's when the plant research of the former eastern-block countries and Soviet Union became more readily available to western scholarly analysis and evaluation.

The first part of this review evaluates the effects of X-rays on plant growth and in certain instances on seed germination. While 70 different species of plants were evaluated in over 60 published papers for the effects of X-rays during these early decades of the 20th century, several species (i.e., wheat, sunflower, broad bean and rice) have been the object of more intense investigation. Consequently, the following section on X-rays will provide a more detailed evaluation of the response of these four species, since they provide the most comprehensive information on the nature of the dose-response, especially in the low dose range, as well as to the critical issue of reproducibility of findings. The findings of all 64 separate publications reviewed (Table 1) often included multiple experiments with multiple endpoints measured. Consequently, there is substantial information available to provide a general assessment of the effects of X-ray treatments on plant growth. The summarized data provide information on a number of relevant parameters, especially with respect to study design features (e.g., number of doses, dose range, and spacing of doses). For example, of the 63 publications, 18 papers reported experiments with greater than or equal to six doses (i.e., X-ray treatments). Experiments with such a large number of treatment groups offer an excellent opportunity to assess the hormetic hypothesis, especially if optimal dose selection was employed. The table also reveals that the investigators generally used seeds, as the principal object of exposure (i.e., more than two-thirds of the studies), followed by the use of sprouts. Common experimental considerations involved the use of either dry or soaked seeds, with the length of time that the seeds were soaked in water prior to irradiation differing according to the specific experiment. In general, the findings revealed that approximately two-thirds of the publications reported X-ray induced stimulation of plant growth, seed germination or other parameters. As expected, those studies using large numbers of doses, especially in the low dose range, provided the most useful information to assess the hormetic hypothesis and in general were supportive of this hypothesis.

The time span over which the evaluation of X-rays on plant growth is conducted is the period from the late 1890's to the early 1940's. As will be seen, during this period research methods underwent rapid developmental refinement not only with respect to X-ray technology and dosimetry, but also with complementary aspects relating to study design, statistical analysis procedures, and reporting of data. For example, statistical methods such as the chi-square test, the *t*-test of Student and analysis of variance were not developed until 1900, 1908, and 1918, respectively.[76] It was during this period that considerable data emerged to affect judgments on

Table 1 A history and summarization of the effects of X-rays on plant growth from 1896–1941.

Author	*Year*	*Type of radiation*	*No. of doses*	*Dose range*	*Specific doses*	*Plant name*	*Plant part exposed*	*If seed, D S G UK*	*Endpoint*	*Results*	*Conclusion*	*Comments*
Schober[17]	1896	X-ray	NA	NA	NA	Oat	Seedling	NA	Growth	Stimulation	Stimulatory	
Maldiney and Thouvenin[18]	1898	X-ray	NA	NA	NA	*C. alvensis* *L. sativum* *P. millaceum*	Seeds	UK	Germination	Acceleration of seed germination	Stimulatory	Low no. of seeds; low confidence
Perthes[19]	1904	X-rays	NA	NA	NA	*V. faba*	Seeds	UK	Root growth	Retardation	Inhibitory	
Euler[20]	1906	X-rays	NA	NA	NA	Beans, radishes	Seeds	UK	Sprouting, growth, time to blooming	Acceleration of blooming	Stimulatory	
Koernicke[21,22]	1904, 1905	X-rays	3	16–26 H	11, 20, 36 H	*V. faba* *V. sativa* *B. napus*	Seeds	D, S, G	Growth, germination	20 H small acceleration for *V. faba* and large acceleration for *B. napus;* germination stimulated	Stimulatory	Study included methodology advances
Guilleminot[23]	1907	X-rays radium	16	10–20,000 R	NA	Gilly flower (methiola)	Seeds	D	Growth	Accelerated growth at 5000 & 7500 R; higher doses inhibitory	Stimulatory	Stimulatory responses thought to be too small to be significant
Schmidt[24]	1910	X-rays	5	0.05–1 HED	1/20, 1/10,¼, ½, 1 HED	Peas	Seeds	S	Height, leaf size, pod size	Considerable stimulation at low doses; bottom four doses stimulated	Stimulatory	Stimulation so great as to have commercial importance
Wetterer[25]	1912, 1913	X-rays	4	5–40 H	5, 10, 20, 40 H	Sunflower	Seeds	S	Growth	Retardation proportional to dose	Inhibitory	
Promsy and Drevon[26]	1912	X-rays	NA	NA	NA	Lentils, rye, beans, white lupine, kidney beans	Seeds	G	Growth	At 15°C strong retardation; at 30-40 °C mostly stimulation	Stimulatory	Temperature dependent response
Schwarz[27]	1913	X-rays	4	30–120 s	30, 60, 90 and 120 sec; 5 min	*V. faba*	Seeds Seedlings	D	Growth–height	120 s produced most favorable results; 5 min retardation	Stimulatory	Small no. of experiments; lack of controlled conditions
Miege and Coupe[28]	1914	X-rays	NA	NA	NA	*R. lepidium*	Seeds	Growth–weight & height	45% increase leaf weight; 59% increase		Stimulatory	

(Continued)

Table 1 *(Continued)*

Author	*Year*	*Type of radiation*	*No. of doses*	*Dose range*	*Specific doses*	*Plant name*	*Plant part exposed*	*If seed, D S G UK*	*Endpoint*	*Results*	*Conclusion*	*Comments*
Koernicke[29,30]	1915-1920	X-rays	10	1/100–5 H	1/100, 1/60, 1/40, 1/20, 1/10, ½,1.5,2.5, 3.5,5 H	*V. faba* *P. multiforma* *P. vulgaris* *L. albus* *S. arrensis* *P.somniferum* *T. vulgare* *B. napus* *Z. mays* *A. sativa*	Seeds	D, S	Growth	Growth at 1/60–1/30 H was accelerated for air dried seeds; seeds soaked 1-2 days were stimulated; seedlings stimulated	Stimulatory	220-3000 seeds per experiment; acceleration only when air dried and germinated seeds irradiated
Yamada[31]	1917	X-rays	4	3–11 H	3,5,7,11 H	*O. sativa*	Seeds	S (168 hr)	Germination and growth	Growth stimulated (8.3%) at 3H	Stimulatory at lowest doses	Higher two doses showed a modest decrease
Nakamura[32]	1919	X-rays	3	5–15 min	5, 10, 15 min	*O. sativa*	Seeds	D	Crop yield	Plants in 5 min group were stimulated	Stimulatory at the lowest dose	
Sierp and Robbers[33]	1923	X-rays	NA	NA	NA	*A. sativa*	Sprouts		Separate plant organs	Early stimulation, later retardation	Stimulatory/ inhibitory	
Lallemand[34]	1922	X-rays	7	1/12–20 H	1/12, ¼,1/3, 1/2,5,10 20 H	*P. miliaceum* *L. sativum* *B. napus* *T. sativum* *L. asulenta* *P. velgavis* Onion bulbs	Seeds	D, S	Growth after 14th day	Weakest doses did not sitmulate lentil, wheat and kidney beans; modest, high doses did not stimulate lentil & kidney bean	Inhibitory	
Weber[35]	1922	X-rays	NA	NA	NA	Lilac	Buds		Sprouting	Stimulation	Stimulatory	
Altmann *et al.*[36]	1923	X-rays	> 4	1–12 H	1,3,6,12 H	*P. vulgaris*	Seeds	D, S, G	Multiple growth endpoints	Transient stimulation; stimulation depends on stage of development	Stimulatory	Triplicated experiment
Komuro[37]	1923	X-rays	7	40–150 H	40, 50, 60, 80, 100, 120,150 H	*V. faba*	Seeds	S	Growth	Germination reduced	Inhibitory	Exp. 1 and 2 replicates
Komuro[37]	1923	X-rays	8	20–155 H	20, 30, 50, 60, 80, 100, 120, 155 H	*V. faba*	Seeds	S	Growth	Growth accelerated	Stimulatory at 20 H	Exp. 3
Komuro[38]	1924	X-rays	3	5–15 H	5,10,15 H	*O. sativa*	Seeds	S	Germination and growth	Germination and growth accelerated	Stimulatory at 10 and 15 H	Exp. 1

(Continued)

Table 1 *(Continued)*

Author	*Year*	*Type of radiation*	*No. of doses*	*Dose range*	*Specific doses*	*Plant name*	*Plant part exposed*	*If seed, D S G UK*	*Endpoint*	*Results*	*Conclusion*	*Comments*
Komuro[38]	1924	X-rays	3	5–15 H	5,10,15 H	*O. sativa*	Seeds	D	Germination and growth	Germination stimulated early	Stimulatory at both doses	Exp. 2
Komuro[38]	1924	X-rays	3	7–15 H	7,10,15 H	*O. sativa*	Seeds	D	Germination and growth	Germination and growth accelerated	Stimulatory at 7 and 10 H	Exp. 4
Komuro[38]	1924	X-rays	3	5–15 H	5,10,15 H	*O. sativa*	Seeds	D	Germination	Germination stimulated	Stimulatory at all doses	Exp. 6
Czepa[39]	1924	X-rays	NA	0.5–25 and 150 H, 300 H	H NA	*V. faba* *V. sativa* *P. vulgaris* Lettuce	Seeds		Rate of germination; growth	One experiment showed stimulation at 25-50 H	Stimulatory /inhibitory	Authors concluded no stimulation occurred
Martius[40]	1924	X-rays	NA	NA	NA	NA	NA		NA	Failed to show stimulation	No stimulation	
Geller[41]	1924	X-rays	NA	NA	NA	NA	NA		NA	Stimulation at some doses	Stimulatory	Experiments covered 1920-1923
Gambarov[42]	1924	X-rays	6	1–10 HED	1,2,3,4,5, 10 HED	*V. faba*	Seeds	S	Root length; time of lateral root development	No stimulation	Inhibitory	Measured daily for 12 days
Ancel[43]	1924	X-rays	NA	NA	NA		Seeds		Germination rate	Large differences across treatments	No stimulation	Author did not relate to treatment
Tushanakova (as reported in Breslavets[44])	1924	X-rays	4	5–20 min	5,10,15,20 min	Sedge, soya, blue lupine, tomato, melon			Fruit maturation	Soya & lupine: matured earlier; melons: fruited earlier; Tomato: increased no. of fruit	Stimulatory	Showed compensation stimulation; sedge stimulation too small to be useful

(Continued)

Table 1 *(Continued)*

Author	*Year*	*Type of radiation*	*No. of doses*	*Dose range*	*Specific doses*	*Plant name*	*Plant part exposed*	*If seed, D S G UK*	*Endpoint*	*Results*	*Conclusion*	*Comments*
Iven[45]	1925	X-rays	9	1/250–22 HED	1/250, 1/100, 1/10,1/2, 1,5,10,18, 22 HED	*V. faba*	Seeds	D	Germination and leaf development	Stimulation noted	Stimulatory	Stimulation followed Arndt-Schulz Law
Ancel[46,47]	1925 1925	X-rays	NA				Seeds			Negative	No stimulation	
Kol'tsov and Kol'tsov[48]	1925	X-rays	NA	NA	NA	Peas Wheat	Seeds	D, G	Height Flowering	Experiment with peas give more definite evidence of stimulation	Stimulatory	
Ancel[49]	1926	X-rays	4	40-150 H	40,70,100, 150 H	Lentils Beans	Seeds	D	Individual plant parts	No stimulation	Inhibitory	
Ancel[50]	1926a	X-rays	1		8 H	Lentils	Buds		Bud sprouts	Enhanced growth	Stimulatory	Seen as compensation to injury
Bersa[51]	1926	X-rays	NA		0.5 H	*V. faba* S. alba			Stem and root lengths	*V. faba:* 26% increase in root length; *S. alba:* rootlets & hypocotyls showed stimulation	Stimulatory	Limited power due to small n (n = 10)
Johnson[52,53]	1926 1928	X-rays	2	5–10 H	5, 10 H	Sunflower	Seeds	D, S	Growth, germination, sprouting	No stimulation; plants from soaked seeds bloomed earlier	Inhibitory	
Doroshenko[54]	1929	X-rays	3	5–20 min 50–80 min	5,10,20 min; 20, 30, 40, 60, 75, 80 min	*A. bizantine* Millet Winter rye			Growth yield	*Avena bizantine*: stimulated at low doses; winter rye: stimulated; millet: inhibited	Stimulatory /inhibitory depending on species and dose	Higher doses cause inhibition; low doses cause stimulation
Sprague and Lenz[55]	1929	X-rays	NA		2H (exp1) 1H (exp2)	Irish cobbler Green mt.	Tubers		Yield	Treatment effect; increased	Stimulatory	
Patten and Wigoder[56]	1929	X-rays	NA	1/20 HED 3 HED		Beans Mustard Barley	Seeds		Growth	Mustard most rapid growth at 1/20 HED; barley inhibited; beans not clear	Stimulatory	Data not presented

(Continued)

Table 1 *(Continued)*

Author	*Year*	*Type of radiation*	*No. of doses*	*Dose range*	*Specific doses*	*Plant name*	*Plant part exposed*	*If seed, D S G UK*	*Endpoint*	*Results*	*Conclusion*	*Comments*
Cattell[57]	1931	X-rays	NA	150–1100 r	100, 250, 400, 700, 1100 r	Wheat	Sprouts		Growth of roots, leaves, coleoptiles	Coleoptiles and leaves slightly enhanced at low doses	Stimulatory /inhibitory outcomes	200,000 measurements
Johnson[58-60]	1931, 1933	X-rays	NA	NA	NA	Two thistles, several *Solenaceae*, *V. faba*	Seeds	S (40 hr)	Fresh and dry weights	Only sunberry plant showed stimulation with weak rays	Stimulatory /inhibitory	Grew plants for 25 days
Chekhov[61]	1932	X-rays	NA	NA	NA	Barley, rye, lentils, oats	Seeds	D, G	Germination and growth	Weak doses had stimulatory effect on development	Stimulatory	Like Arndt-Schulz Law
Shull and Mitchell[62]	1933	X-rays	5	38–190 r	38,76,118, 152,190 r	Corn, oats, sunflower, wheat (3 varieties)	Seeds	D, S	Growth	Coleoptiles enhanced 5–26%	Stimulatory	Optimum response determined
Benedict and Kersten[63]	1934	X-rays	NA	NA	NA	Wheat	Seeds		Enzyme activity	Increased enzyme activities	Stimulatory	
Francis[64]	1934	X-rays	6	565–13,560 r	565,1130, 1695, 3390, 6780, 13,560 r	Wheat	Seedlings		Growth	No stimulation	Inhibitory	
Breslavets and Afanas'eva[65]	1935	X-rays	7	250–8000 r	250, 500, 750, 1000, 2000, 4000, 8000 r	Rye	Sprouts	Sprouted seeds S	Height, number of stems/plant, number of ears	Height retarded; increased no. of stems/plant at low doses; ears stimulated at low doses	Stimulatory	Used hard and soft X-rays; Arndt-Schulz Law
Breslavets and Afanas'eva[66]	1935	X-rays	7	250–8000 r	250, 500, 750, 1000, 2000, 4000, 8000 r	Rye	Seeds	S	Height, number of stems/plant, ear size	Stimulated at low doses	Stimulatory	Growth at day 27

(Continued)

Table 1 *(Continued)*

Author	*Year*	*Type of radiation*	*No. of doses*	*Dose range*	*Specific doses*	*Plant name*	*Plant part exposed*	*If seed, D S G UK*	*Endpoint*	*Results*	*Conclusion*	*Comments*
Tsuryipa (as reported in Breslavets[44])	1935	X-rays	NA	NA	NA	Wheat, oats, cotton	Seeds		Vegetative and reproductive endpoints	Wheat; greater impact on reproductive than vegetative; oats: increased sprouting, 60% increased yield, ripening shortened by 10 days	Stimulatory	
Long and Kersten[67]	1936	X-rays	5	1–5 s	1,2,3,4,5 s	Soya	Seeds	UK	Weight	Marked increase	Stimulatory	Field conditions 12,751 plants
Frolov[68]	1936	X-rays	NA	NA	NA	Soya, wheat, flax			Yield	Wheat: accelerated at 60% in yield; soya: not stimulated; flax: stimulated only early growth	Stimulatory	Results with soya contrast with Long & Kesten
Johnson[69]	1936, 1936a	X-rays	NA	NA	NA	Tulip	Bulbs		Leaves Blooms	Leaves lengthened flowers not affected	Stimulatory/ inhibitory	
Johnson[69]	1936	X-rays	1	NA	1500 r	Wild potato	Tubers		Tuberization of potato	No stimulation	Inhibitory	Favourable in pilot study but not in follow-up; experiments over five seasons with 17,000 tubers
Saeki[70]	1936	X-rays	6	50–1200 MAM/21^2 at 30 KV	50, 100, 200, 400, 800, 1200	*O. sativa*	Seeds Seedlings	G	Growth Germination Yield	$\leqslant$200 MAM/21^2 stimulatory depending on endpoint	Stimulatory	
Zankevich and Brunst[71]	1937	X-rays	7	250–10 000 r	250, 500, 750, 1500, 3000, 6000, 10 000 r	Tobaco Poppies Flax Rhubarb			Growth	500 and 750 r stimulatory to tobacco and flax	Stimulatory	

(Continued)

Table 1 *(Continued)*

Author	*Year*	*Type of radiation*	*No. of doses*	*Dose range*	*Specific doses*	*Plant name*	*Plant part exposed*	*If seed, D S G UK*	*Endpoint*	*Results*	*Conclusion*	*Comments*
Bless[72]	1938	X-rays	8 (radish) 6 (beans) 10 (lettuce)	0.25–7 sec (radish & lettuce) 0.25–6 sec (beans)	0.25, 0.50, 0.75, 1, 2, 3, 4, 5, 6, 7 s	Radish, beans, lettuce	Seeds	D, S	Growth	Radish: 10% increase in pod wt– not thought to be significant. Bean: no significant effect with dry seeds; 25% increase with spouted seeds. Lettuce: 60% increase in 5 min dry seed group; later repeat showed 2—30% increase	Stimulatory	Radish: 20 plants per treatment; Bean: 25 plants per treatment; Lettuce: 10 plants per treatment
Zaurov[73]	1937	X-rays	5	125–4000 r	125, 500, 1000, 2000, 4000 r	Indian hemp	Seeds		Stem length Growth	Insignificant increase in stem length at 500 r; insignificant increase in growth at 2000 r	Stimulatory	
Breslavets[74]	1937	X-rays	10	50–1000 r	50, 100 200, 250, 350, 450, 550, 650, 750, 1000 r	Peas	Seeds	D	Growth Pods	Excellent stimulation of growth at 350 and 1000 r; increase in pods	Stimulatory	
Breslavets[74]	1937	X-rays	6	250–16,000 r	125, 250, 500, 750, 1000, 2000, 4000, 8000, 16,000 r	Spring wheat		D, S sprouts	Growth	Dry and soaked seeds did not respond; sprouts not stimulated (50–8000 r)	Inhibitory	
Wort[75]	1941	X-rays	7	19–22 r	19, 38, 57, 76, 114, 152, 228 r	Spring wheat Winter wheat	Seeds	D	Growth	Enhanced growth	Stimulatory	Attempted to closely follow Shull & Mithell 1933
Breslavets (as reported in Breslavets[44])	1942	X-rays	9	200–8000 r	200, 300, 400, 600, 750, 1000, 2000, 4000, 8000 r	Meadow grasses	Seeds		Germination No. of roots Root length Stem length	Stimulation of all parameters	Stimulatory	Typical of β-curve

Abbreviations: H = Holzknecht's Unit; HED = Erythema dose; designated by Seitz and Wintz as the Hauteinheitsdosis; S = soaked; D = dried; G = germinating; UK = unknown

Table 1 *(Continued)*

Summary of data
70 species of plants were tested in a total of 64 publications
43 publications were conducted with seeds, 7 with sprouts, and 1 with bulbs
Of the 43 publications conducted with seeds, 15 involved air-dried seeds, 15 water-soaked seeds, and 5 germinating seeds
Of the total 64 publications, 44 showed stimulation, 17 inhibition, and 3 both stimulation and inhibition

how X-rays affected plant growth culminating in interim conclusions provided in U.S. National Academy of Science publications on this topic.

X-rays

Early studies on rice During the early decades of the 20th century several authors investigated the capacity of X-rays to stimulate the germination of rice seeds and the growth of rice seedlings. Six such studies have been typically cited in review papers as providing support to the radiation hormesis hypothesis.[31,32,37,38,70] Four of the six papers which utilized *Oryza sativa* as the plant species, included three doses and a control; the study of Yamada[31] employed four doses and a control, while Saeki[70] used six doses plus a control. Four of the six studies defined the X-ray dose in H units and they were quite similar in dose range (i.e., 3, 5, 7 and 11 H; 5 to 15 H; and 3, 5, and 7 H). The latter study by Saeki[70] defined dose as MAM/21[2] at 30 KV (i.e., 50–1200 MAM/21[2] at 30 KV). It should be noted that the international roentgen (r) as a radiation unit was in general use since 1928. Its equivalence to other units previously used is as follows: (1) skin erythema dose (SED) is considered to be equivalent to about 600 r and equivalent to 1 S.-N unit as introduced by Sabournud and Noire. The Holzknecht (H) unit has two values. As initially given by Holzknecht it equalled 1/3 SED (200 r), but later is equalled to 0.25 N or 125 r. Kienbock divided his scale into unites of X (i.e., Kienbock units) and considered 10 × =1 S.-N=5 H. Thus, X is about 60 r (see Hudson;[77] Taliaferro and Taliaferro[78]).

The principal difference in earlier studies involved how the seeds were handled prior to and after ovulation. In general, the seeds were either air-dried or steeped (soaked) in water for variable time periods prior to irradiation [e.g., Yamada[31] for 168 h and Komuro[38] for 12 h]. In some studies germination was considered or growth or both parameters. In general, the data indicate that air dried seeds were stimulated by the X-ray treatments.[37,38] In the 1924 study of Komuro several experiments indicated a consistent acceleration of germination, especially at 5 and 10 H.[38] The number of seeds in each of these experiments was modest, ranging from 10–25 per treatment group. Nonetheless, the integration of the three experiments indicates that the acceleration was considerable and approached twofold at the 10 H dose. Statistical analyses were not conducted on the data by the authors. Komuro also claimed that soaked rice seeds were also stimulated by low doses of X-rays. These findings were, however, generally marginal increases and are not as reasonably established as with dry seeds.[38]

With respect to growth, the findings of Yamada[31] and Nakamura[32] provide support for the hypothesis that crop yield could be enhanced by X-ray treatment. However, their conclusions were directly challenged by Komuro[37] based on the inadequacy of the control group of these two investigations and especially in light of his generally better study design and non-stimulatory response as far as yield was concerned. However, Komuro reported that the plants displayed 'precocious' growth, meaning that they developed more quickly and were able to be transplanted earlier.[37] Later investigations by Komuro[38] and Saeki[70] supported the hypothesis of X-ray treatment enhanced crop yield in studies with more powerful designs (e.g., six doses in Saeki[70]) and with the magnitude of enhancement being generally in the 10–30% range depending on the endpoint measured.

Other early research on non-rice species In research that both preceded and was contemporaneous with the Japanese work on rice, Koernicke[21,30] assessed the effects of X-rays on the germination and growth of multiple species of plants. Early findings[21] gave some hint that X-rays may enhance germination during certain experimental conditions. More specifically, Koernicke[21] reported a reproducible acceleration of germination in air dried seeds of *Vicia faba*, a phenomenon that was not observed in soaked seeds (see next section).

This initial research of Koernicke[21] was noteworthy not only for the stimulatory response, but also because it involved methodological advances, including the use of multiple species as well as larger numbers of seeds in the investigations. Nonetheless, the study was still limited to only three doses (16, 20 and 24 H) with only the 20 H dose providing evidence of a stimulatory response.

In follow-up experimentation published over a decade later, Koernicke[29] extended his research to

include ten species, employing air dried seeds, water soaked seeds (1, 2 or several days), germinated seeds into radicals, and potted seedlings. The range of doses was markedly increased, with now ten doses (5 H to 1/100 H) in contrast to the 16–24 H dose ranging study. The sample size was also increased to include from 200–3000 seeds per experiment. In general, air dried seeds and those soaked for 1 or 2 days that were more strongly irradiated germinated sooner than weakly or non-irradiated seeds. Other stimulatory growth was reported for seedling responses to low doses of X-rays (1/60 to 1/20 H). These findings of Koernicke[29] were generally consistent with the more limited study of Schwarz,[27] who observed that irradiated air dried *V. faba* seeds resulted in enhanced growth (5 doses) by 3 weeks. While the magnitude of the enhancement was approximately twofold, the sample size was only three plants per treatment.

Other findings published in the early years of the 20th century provided support for the premise that X-rays could stimulate either germination and/or growth. Most notable were those of Euler,[20] Guilleminot,[23] Schmidt,[24] Promsy and Drevon,[26] Miege and Coupe,[28] and Pfeiffer and Simmermacher.[79] These early studies were distinguished by the wide range of species tested, the use of up to 16 doses by Guilleminot[23] and five doses by Schmidt.[24] Nonetheless, most of these investigations had important limitations, including small sample size, such as only ten seeds/treatment,[28] lack of statistical analysis and often inadequately controlled environmental conditions.

Nonetheless, the first two decades of the 20th century witnessed the recognition that low doses of X-rays, especially to seeds in an air dried but also water soaked state, had the potential to have their germination accelerated. Growth was also stimulated depending on the study as measured by enhanced early development, shorter time to blooming[20] and increase in height and weight.[24,28] There was also the progressive improvement in the standardization and reporting of X-ray exposures, and in the quality of the study design. While these studies lacked the capacity to derive definite conclusions about the capacity of X-rays to stimulate germination and/or growth, the data clearly support the hypothesis that stimulation could occur and that follow-up research was necessary to resolve the question. Such findings ushered in an expanded level of research on this topic that would continue over the next several decades.

Vicia faba Perhaps the most tested plant in X-ray stimulation studies is *Vicia faba*, the broad bean. By 1936 fifteen studies were found in the open literature concerning whether X-rays could stimulate seed germination or growth of this plant. It was believed that a more careful consideration of the responses of *V. faba* were warranted since this would more effectively speak to the issues of robustness of the database, endpoint variation, reproducibility and dose range studied.

Of the fifteen studies, six were reported as clearly providing no support to a stimulation hypothesis. Of the remaining nine studies which produced some evidence of an X-ray-induced stimulatory response, one was criticized by other investigators for either lack of controlled conditions (see Komuro's criticism[37] of Schwarz[27]) while another study suggesting stimulation could not be replicated (see Ancel's criticism[43] of Altmann *et al.*[36]). The stimulatory study of Bersa[51] was also criticized as having too small a sample size (n=10) to draw firm conclusions, while Patten and Wigoder[56] presented evidence of a stimulatory response in an abstract-like note without research methods. Of the five remaining articles providing evidence of stimulatory responses, Koernicke[21,22] and Jungling[80] report only one dose in the stimulatory zone, thereby not providing an adequate characterization of the possible stimulatory zone. Of the remaining two studies, Koernicke[30] and Iven[45] ultilized ten (1/20–25 HED) and nine (1/250–5½ HED) X-ray doses, respectively, plus controls. In the case of Koernicke,[30] stimulation was reported at 1/12, 1/8, and 1/5 HED, while in the Iven[45] report the stimulatory range was from 1/250–½ HED. In her major review of the effects of X-rays on plants Breslavets[44] indicated that both of these studies provide support for the Arndt-Schulz Law.

Of particular interest was the fact that Iven[45] provided repeat measures data that revealed that the growth stimulation which appeared within 10–20 days following treatment and then regressed to become equal with the control values. Thus, as Johnson[58] noted, the stimulatory effect with the low dose X-ray treatment was a transitional one. Johnson[81] concluded that Iven[45] was reporting an acceleration of growth following retardation, 'a phenomenon commonly reported after radiation'. Such an interpretation was consistent with the views of Stebbing,[82] that hormesis represents an overcompensation to a disruption in homeostasis. This overcompensation phenomenon was carefully documented for u.v. radiation on fungal growth by Smith.[83,84]

Of further note is that several of the more strongly designed studies which display no evidence of stimulation and/or clear inhibition utilized doses in the inhibitory area of the dose-response of Koernicke[30] and Iven[45] or even apparently higher. For

example, Gambarov[42] employed doses of 1–10 HED, Czepa[39] used doses at 2.5–125 HED. Consequently, the fact that they were negative does not conflict with the observations of Koernicke[30] and Iven[45] who reported stimulatory responses at lower levels. In her review of the *V. faba* data, Breslavets[44] offers four explanations of why the array of papers presented a confusing picture of stimulatory and inhibitory responses: (1) *V. faba* was viewed as an inappropriate biological model because its threshold for stimulation was too low. It was believed to be so radiosensitive that even with normally weak doses a retardation response would ensue; (2) Insufficiently accurate measurement of dose, especially those early studies in which dosage was measured in skin erythemas; (3) The *V. faba* experiments also employed inadequate numbers of seeds. This was principally due to the large size of the seeds coupled with the use of the limited field in the Coolidge tube thereby providing an important barrier for conducting such experiments; and (4) These studies were also criticized for their use of a generally small range of doses. According to Breslavets,[44] the most significant flaw in many of the experiments may have been the *a priori* bias of the investigator. Much was made of the remarks of Seide[85] and Johnson[58] who displayed obvious bias against the theory of X-ray-induced stimulation by ignoring or discounting data inconsistent with their views. On the other hand, Breslavets[44] noted (without being specific) that investigators supportive of the theory may have at times designed experiments that could lead to this favorable (i.e., stimulatory) response.

While many of the conclusions of Breslavets[44] such as low dose sensitivity, poor sample size and limited dose range are valid in their criticisms of the early studies on X-ray-induced changes in *V. faba*, the present analysis indicates that the general pattern of response is consistent with the Arndt-Schulz Law. However, at the time the research was conducted there appears to be considerable confusion over the nature of the low dose exposure dose-response relationship. This is reflected in the major review by Johnson[81] who was accused of bias against the theory of radio-stimulation and in the writings of Breslavets[44] who was a supporter of the low dose stimulatory theory. However, *in toto*, an analysis of the body of data on *V. faba* up through the 1930's is remarkably in agreement with those seen for the sunflower and wheat responses in which analysis of reported studies was consistent with the hormetic perspective.

Wheat Another plant species commonly used to evaluate the effects of X-rays on plant growth has been wheat. However, in contrast to other plants evaluated such as rice which assessed as early as the first decade of the 20th century, research with wheat did not occur until the 1930's. In the assessment of the X-ray plant research with wheat ten studies were identified. Of these ten, four involved exposure to seeds while six involved exposure to seedlings. Attention will be directed here to the responses of seedlings due to the more substantial nature of their research protocols. Research concerning X-rays on seeds will not be followed due to the fact that one of the four papers did not address growth endpoints and two foreign articles require translations.

Of the six studies assessing the effects of X-rays on wheat seedling growth, three studies utilize high doses (i.e., $\geqslant$550 R) and reported dose dependent growth inhibition.[58,64,86] In contrast, two studies providing low doses displayed low dose stimulatory responses.[62,75] Figure 1 indicates the dose-response relationship of the X-ray treatments for multiple endpoints including wet and dry weights.[75] In each case a marked stimulatory response was observed consistent with the hormetic dose-response curve. Similar findings using low dose X-ray exposures were noted for other species tested for corn, wheat, oats, and sunflower.[62] The final article, which covered 150–1100 R, bridged the gap of the higher end of the low dose area and high dose exposure zone.[57] The findings of Cattell[57] displayed suggestive evidence of a weak stimulatory response at the lower doses for coleoptiles, and strong inhibition at the higher end of the doses administered consistent with the hormetic dose-response relationship.

The quality of these post 1930 studies represents substantial progress over those of the early decades

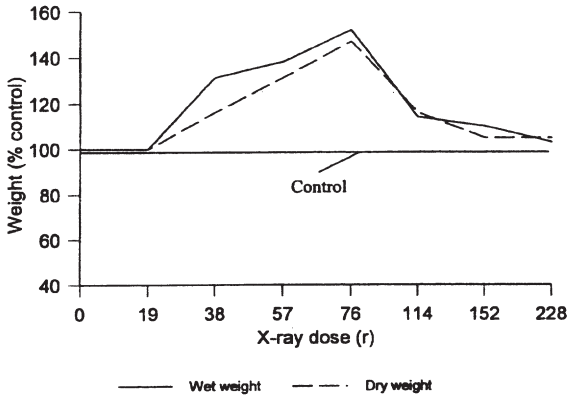


Figure 1 Wet and dry weights (% control) of Marquis spring wheat 56 days after exposure to various doses of X-rays. Exposure was conducted on 24-h seedlings (data from Wort[75])

of the 20th century in terms of study design and adequacy of sample size. For example, the report of Wort[75] involved seven doses plus an unexposed control with 35 plants per group. This experiment, which was replicated, also included a repeated measures component over three consecutive weeks. Wort[75] also provided data from two identical studies using 57 and 9 month old seeds in order to assess the effect of seed age.

Despite their generally strong study designs, the reports of Wort[75] and others during this time period lacked important and more recently emphasized features such as random allocation of subjects (e.g., seedlings) to group and formal statistically-based hypothesis testing techniques. Despite these limitations, the findings of X-rays on wheat seedlings were remarkably consistent with the Arndt-Schultz Law, a phenomenon also clearly mirrored in studies with other plants such as rice, sunflower and broad bean, which were assessed over a wide dose range.

Sunflower One of the most influential figures in the US affecting the acceptance of the Arndt-Schultz Law (i.e., hormesis) was Edna Johnson at the University of Colorado, Boulder. She was perhaps the first American scientist to publish research findings on the topic of X-ray stimulation of plant growth and did so over a span of several decades (mid 1920's to late 1940's). She published a series of original research papers that displayed better design features and attention to detail than most of the previous efforts. In these more credible articles up through the 1930's she consistently found no convincing evidence to support the hypothesis of a direct stimulation of plant growth by X-rays. So substantial was her research in this area that she was invited to author a major review of the topic under the auspices of the NRC and the oversight of such prestigious individuals as Gino Failla, Charles Packard and Benjamin Duggar.

Despite the influence of Johnson on the topic of radiation hormesis on plant growth, a paper published by Shull and Mitchell[62] had the potential to challenge the basis of her denial of evidence of radiation hormesis. In this paper Shull and Mitchell[62] hypothesized that past studies used doses that were far in excess of a potentially stimulating dose range. Consequently, they undertook a series of investigations with corn, wheat (three varieties), oats and sunflower to assess whether X-ray exposures over a broad but lower dose range could be stimulating to recently germinated seeds. While Shull and Mitchell[62] reported stimulatory responses for all species of plants tested, the most significant feature of their work was their inclusion of sunflower since Johnson had studied the response of this same species in three different published papers. Despite the fact that both research groups used sunflower, there were some differences in the research methodologies employed. In two of Johnson's papers[52,53] she irradiated seeds soaked in distilled water, while the third paper[58] utilized 7 day old seedlings. In the Shull and Mitchell[62] paper the X-rays were applied to very recently germinated seeds that had been soaked in distilled water. Thus, the first two reports of Johnson[52,53] were most directly relevant (although not a perfect match) for the Shull and Mitchell study.[62] The doses of radiation used by Johnson in studies one and two ranged from 100–1000 R,[52,53] while the dose range used by Shull and Mitchell ranged from 38–380 R.[62] In the Shull and Mitchell report stimulation was observed over 38–190 R; inhibition was reported at the highest dose (i.e., 380 R).[62]

The follow-up study of Shull and Mitchell[62] should have been used to clarify the alleged discrepancy with the earlier work of Johnson.[52,53,58] However, Shull and Mitchell never attempted to do so.[62] Only limited reference was made to Johnson's work, and even in such instances the discussion was not directed towards the principal issue of low dose stimulation. Why they did not seek to clarify an obvious and important issue is unknown. However, it should be emphasized that Johnson knew Shull, and specifically states in her acknowledgment that she expressed appreciation to Professor CA Shull for assistance during the progress of her studies as at doctoral student at the University of Chicago and as a new faculty member at the University of Colorado. It is possible that Shull did not want to challenge the position of a former student. Similarly, in her influential review for the NRC, Johnson[69] summarizes the paper of Shull and Mitchell[62] but never links it to her work, nor attempts to clarify the obvious discrepancy between her high dose inhibition and the low dose stimulation of Shull and Mitchell.[62]

Despite the central role that Johnson had in affecting the direction of scientific attitudes to radiation hormesis in the US and the potential significance of the Shull and Mitchell paper,[62] no other reviewer has brought forth the proposition offered here as to the scientific reason why Johnson[53,58] did not observe stimulation and why it may not have been resolved.

The work of Johnson continued to be cited in the most prestigious reviews on the topic of radiation stimulation of plant growth. For example, Sax's reviews in 1955 and 1963 cited the work of Johnson[81] and Shull and Mitchell[62] favorably without resolving their apparent conflicting conclusions.[87,88] The book entitled 'Plants and X-rays' by

LB Breslovets[44] directed considerable space to both Johnson[81] and Shull and Mitchell,[62] yet again without an attempt to resolve their apparent conflicting conclusions. Furthermore, Packard, co-editor of the 1936 NRC report in which Johnson strongly emphasized the lacking support for the Arndt-Schulz Law for X-rays on plant growth, reported her incorrect conclusions that X-ray treatment does not stimulate plant growth, citing her 'extensive summary of this topic.'[89]

Summary This section of the historical development of the radiation hormesis hypothesis has considered the effects of X-rays on plant material [i.e., seeds (dry, soaked, germinating) or seedlings (sprouts)]. Due to the substantial diversity of articles, plant species tested, exposure techniques and experimental protocols employed, it was decided that the most effective way to provide clarity to this array of information was to be guided by the premise that the review would focus greatest attention on those plant species which were tested most substantially. This would permit the greatest likelihood of having the broad array of doses applied as well as the most substantial sample sizes and capacity to review independent replication of earlier findings. To that end, reseach on rice, sunflower, broad beans and wheat were selected. Despite the wide range of experimental protocols and perspectives from different investigative teams, the most striking observation is that at low doses of X-rays (as defined for each plant species), a stimulatory growth response was observed, while at high doses inhibitory responses occurred. The dose-response range was similar to the β-curve of the hormesis phenomenon and of course, therefore, consistent with the Arndt-Schulz Law. The present analysis is also important because the reviews of the literature that address these early findings never resolved the obvious challenge of how to properly integrate stimulatory and inhibitory responses within a dose-response continuum. Even the reviews of Sax,[87,88] who helped usher in the modern age of plant cytogenetics, were more descriptive than explanatory. As noted earlier, the review of Breslavets[44] which was quite analytical for the time, ultimately blamed investigator bias as the most important factor affecting proper interpretation of the low dose effects area. Despite potential investigator bias, there is little doubt that the clear weight of evidence should have supported the conclusion that the dose-response relationship supports the theory of hormesis. Nonetheless, it seems clear that the scientific community of the 1930's and 1940's had not resolved the issue of low dose X-ray effects on plant growth. The Arndt-Schulz hypothesis was earlier criticized by fair-minded scientists because of studies using inadequate sample sizes along with poor replication of findings. Such criticisms of weak studies were then contrasted with more convincing high dose studies which unequivocally noted dose dependent inhibition. The combination of the legitimate criticism of weak studies, suggesting stimulatory responses and clear findings indicating inhibitory responses at high doses, lead investigators such as Johnson[69] to relegate the Arndt-Schulz Law to a scientific irrelevancy. The substantial criticism of Johnson had its impact on American leaders in the field of radiation (Failla, Hollender, etc.) even though such criticism lacked a proper perspective. Nonetheless, such a flawed perspective (see Packard[89] for his continued reaffirmation of the flawed conclusions of Johnson[69]) delayed the acceptance of hormesis as a legitimate biological hypothesis. Such criticism as reported in a NAS document is comparable to the harsh attack on the Arndt-Schulz Law by AJ Clark in his 1937 publication, 'Handbook of Pharmacology', in which 15% of this book is explicitly devoted to challenging the Arndt-Schulz hypothesis.[90] Thus, the theory of hormesis has had strong opponents who occupied influencial positions in the scientific community at precisely the same time.

Radium

Studies by Gager Perhaps the first claim that radium exposure could stimulate plant processes such as seed germination and seedling growth was reported in 1908 by Gager in a nearly 300 page report documenting some 93 experiments.[91] As a result of the magnitude of this study, its claims of radium-induced stimulation and the long-term advocacy of Gager[92,93] of the low dose stimulatory hypothesis, this paper will receive a detailed assessment. These experiments addressed a wide range of questions including the effects of radium on seeds (either dry or soaked) (i.e., 31 experiments), plants grown in soil (eight experiments), plants grown in water treated with radium (nine experiments), carbohydrate synthesis in plants (ten experiments), respiration (i.e., aerobic and anaerobic) (six experiments), 12 miscellaneous areas and experiments on yeast fermentation. In general, negative findings were typically noted for soaked seeds, plant growth in treated water and experiments on anaerobic respiration. Limited suggestive evidence of stimulation was reported in some experiments using dry seeds, plants grown in soil, and studies of aerobic respiration. The most consistently reported stimulatory responses occurred with yeast fermentation.

Effects of radium on seeds (dry or soaked)
Studies by Gager with seeds involved 14 experiments with *Lipinus alba*, four with Timothy, three with *Phaseolus*, two with oats, and one each with wheat, alfalfa, buckwheat, *Linum* (flax), *Brassica*, and corn.[91] These experiments were generally characterized by a single treatment and concurrent control with modest numbers of seeds treated. Typically, the sample size was ten or less, but on occasion up to 20 seeds in a treatment were used. The duration of the experiments was typically for one to several weeks. The radium was often in the form of a sealed glass tube of $RaBr_2$ with the radium tube lying against the hilum edges of the seeds. The radiation intensity was variable depending on the experiment, ranging from a low of 7000× to 1.8×10^6. Such values meant that the preparation was that much stronger than an equal weight of uranium. However, at the time of the experiment no universally recognized unit of radioactivity had been formulated. Note that in 1910 the International Congress for Radiological Electricity proposed a unit (i.e., the curie) of radium emanation (i.e., radon gas) as the amount of emanation in an enclosed container which is in equilibrium with one gram of metallic radium.

Of the 31 experiments with seeds, four displayed evidence of stimulation including two with Timothy and two with *L. alba*. Gager summarized his findings with Timothy by stating that 'when Timothy grass seeds were exposed to radium of weak activity (7000×) an initial retardation was followed by apparent recovery after an interval of five days.[91] At the end of this period the exposed seeds averaged even taller than those of the control culture'. Examination of the experimental procedure revealed that Gager did not indicate the number of seeds in either the treatment or controls, nor were individual or group averages presented.[91] Thus, even though Gager stated that the seedlings had a 'decidedly' larger average growth than the controls, no data were available to confirm the author's statement. The second experiment with Timothy involved a comparison of seed germination and seedling growth in relationship to the distance of the plants from the source of radiation, which varied from 5, 10, 25, 20 and 25 mm. The control growth (n=not reported) ranged from 9–14 mm in length over the five locations. However, those exposed to the radium displayed a low dose stimulation and high dose inhibition. While these findings are suggestive of stimulation, the limited and inadequate reporting of experimental details does not permit the drawing of a definitive conclusion except that the results warrant more careful follow-up experimentation.

In another experiment, Gager stated that the 'germination of seeds of *L. albus* and the subsequent growth of the radicle was appreciably accelerated' by exposure to 10 000× for 120 h (5 days).[91] In this case Gager provided the sample size (n=8) and the data for the individual control and treatment plants at day 5, the final day of the study. The difference between the two groups was 62.7 *vs* 53.9 mm (16%). A follow-up experiment with *L. albus* employing eight dry seeds/group exposed seeds for different lengths of time (2, 3, 4, 6, and 14 h) to $RaBr_2$ (1.5×10^6) and later planted the seeds in soil. Measurements at 6 and 9 days after treatment indicated that low exposures were associated with enhanced growth. Although measurements continued, the author did not present further data except to conclude that at the end of 5 weeks there were no appreciable differences related to treatment.

The four experiments were the only ones presenting evidence to support the potential for radium to stimulate the growth of plants. In all cases the seeds exposed were dry. Despite the findings and conclusions of the author, the study designs and reporting, even for 1908, were poor. However, even nearly 30 years later the author concluded that 'this 1908 report provided for the first time experimental evidence that radium rays may, under suitable conditions, accelerate the growth of seedlings'.[93] He stated further that these results lead to the broad generalization that radium rays act as a time stimulus to metabolism.

Effects of radium on the growth of plants in soil
In the next set of experiments, Gager assessed the effects of radium in the soil on the germination and growth of oats, *L. albus*, Brassica, peas, beans, wheat and Timothy.[91] As such, there was one experiment for each species, except for oats for which there were two experiments. In general, the author placed seeds into potted soil. The radium source was inserted into the soil at the center of the pot to a depth of 15 cm. Depending on the experiment, seeds were placed in concentric rows around the radium source. In some experiments there was one source (intensity), while in several experiments multiple (up to 3) levels of radium intensity was employed. Thus, in most of the experiments it was possible to have the potential for a dose-response relationship. Of the eight experiments, four displayed evidence of a stimulatory response. However, two of the four studies in which Gager reported stimulatory responses using *Brassica alba* (white mustard) and peas, no measurements were either taken and/or provided. Of the remaining two stimulatory experiments, the one with oats utilized a single $RaBr_2$ intensity, (1.5×10^6) with seeds

planted at three locations from the source of radium. While Gager reported the height of the three treaments no mention was made of the control height, nor was the number of plants employed in the experimental and control groups stated.[91]

The most important experiment involved wheat at two doses of $RaBr_2$ and one dose of radiotellurium. The radium treatment involved exposure to beta and gamma rays while the radiotellurium involved exposure to alpha rays. In this experiment Gager provided information on sample size (n=12) as well as the values for each individual plant at day 4 of growth.[91] All treatment groups displayed greater growth than the controls by approximately 35–45%. As in the case of his results with seeds, Gager was inconsistent in his description of his methodology and reporting of his data.[91] In this present set of eight experiments, only one of the four experiments that Gager claims is stimulatory have adequate data upon which to make a reasonable preliminary determination.[91]

While attention has been directed towards radium, Gager reported on an experiment concerning the effects of alpha rays from polonium on the germination and growth of wheat (n=16).[91] The results indicated an initial slight growth deficit after 4 days (10%), followed by a more vigorous growth in the treated plants (125.3 *vs* 75.5 mm ave.).

Other investigators Based on the research of Gager,[91] there was great interest in assessing the hypothesis that crop production could be enhanced by adding radioactive substances to the soil with or without ordinary fertilizers. This interest was encouraged further by the research of Stoklasa in 1913 on the response of cultures of nitrifying and denitrifying bacteria to the emanation from pitchblend. It was believed that response to the radioactive substance in soil might increase soil fertility by increasing nitrogen circulation. However, a series of reports by Ewart in Australia,[94] Sutton in England,[95] and Ross,[96] Hopkins and Sachs,[97] and Ramsey[98] in the United States did not support the hypothesis that radium treatment of soil was likely to have any commercial agricultural significance. This lack of enthusiasm for the application of radium and/or perhaps other radioactive preparations needs to be seen within the context of commercial interest rather than scientific inquiry. In fact, Hopkins and Sachs, who were clearly not supportive of the commercial application of radium to agriculture, presented data on 36 experiments (with four doses and a concurrent control), nineteen of which offered evidence of stimulatory responses (Figure 2).[97] Rather than being discouraged, the agricultural research community should have been interested in the interspecies differences in response and the nature of the dose-response relationship. However, the lack of a more universal stimulatory response across all species at the same applied dose, the limited magnitude of stimulation and the difficulty in pinpointing the optimum stimulatory zone discouraged further commercial interests.

While the lack of enthusiasm for the commercial application of radium must have adversely affected research interest in this area, a number of papers continued to be published between 1910 and the early 1930's which were supportive of the premise that low dose exposures to radium may affect plant biological processes. Most notably during this period was the continuing work of Stoklasa who reported that various radioactive sources (e.g., naturally occurring radioactive water, pitchblend and radium enclosed in vessels) enhanced seed germination in multiple species,[99–101] growth of cucumbers, mint and tobacco seedlings, growth as evidenced by increase in photosynthesis, dry weight, earlier flowering, and greater seed production,[102] and bacterial metabolism and yeast fermen-

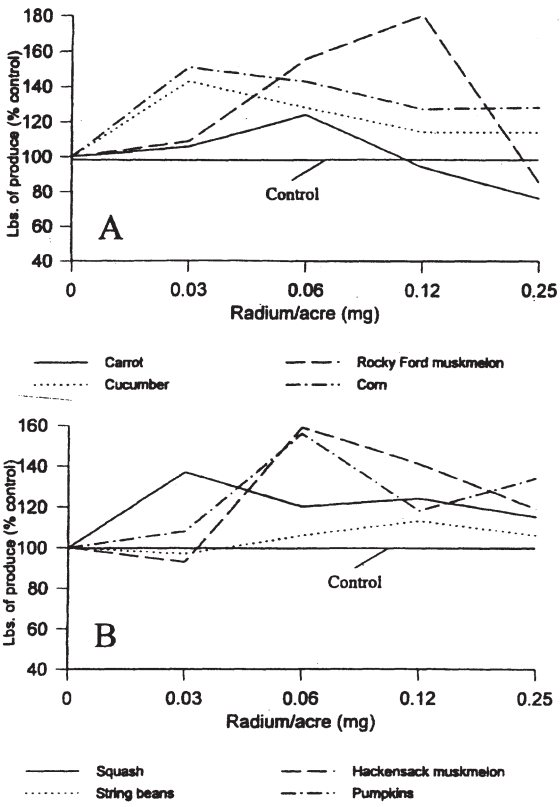


Figure 2 Increase in pounds of produce (% control) of representative crops exposed to various concentrations of radium in the soil (data from Hopkins and Sachs[97])

tation.[102,103] Of relevance to the hormesis hypothesis was that Stoklasa's findings displayed the typical β-curve of a low dose stimulation/high dose inhibition.[100] In further support of the fermentation findings, Kotzareff and Chodat reported a clear β-curve in response to radium exposure.[104] Likewise, the findings of Doumer,[105] Agulhon and Robert,[106] and Montet[107-109] were consistent with the observations of Stoklasa[99,101] that seed germination could be enhanced by exposure to radium sources. It should be noted that an influential paper by Failla and Henshaw reported dose-dependent inhibitory responses in wheat by radium using a very powerful study design.[86] The doses of radium were normalized to that provided by an X-ray exposure. Thus, the lowest dose of radium in X-ray equivalents used in this inhibitory study was 550 R, a dose that is known to be inhibitory in wheat (see wheat section).

The above summary of findings represented the current state of scientific development as of 1936 (see Gager[93]). As could be seen, European researchers continued to study the effects of radium on biological systems especially as related to plant growth and seed germination. As happened in 1915, and again in the late 1930's, the research on radium became a victim of both World Wars I and II with essentially no published findings during these periods.

Major USDA study Two important developments occurred as a result of World War II that were to have a major impact on the assessment of radium on plant growth. The first is that cobalt-60, a gamma source, became readily available as a result of nuclear technology. In fact, the nearly entire focus of gamma rays from radium on plant growth would switch to cobalt-60 from the 1950's onward. Secondly, in 1948 the USDA and a large number (i.e., 13) of state agricultural experiment research stations under a contract with the Atomic Energy Commission (AEC) undertook a large and coordinated study to determine 'whether radioactive material does indeed stimulate plant growth.' This broad goal included the practical aim of whether the farmer would reasonably expect to obtain an increased crop yield by adding one or several naturally radioactive materials to the soil. The impetus for this study was based, at least in part, on reports from Japan of greatly increased crop yields in the vicinity of the bombed areas due to the radioactivity. The study involved three radioactive sources: radium, uranyl nitrate, and alphatron. The alphatron had an alpha ray disintegration rate of 8×10^6/s principally from actinium; the radium source was radium bromide; the radium and uranium sources were used since most of the past studies were with these two agents.[110]

The experimental design involved three doses for the alphatron source and one each for the radium and uranyl nitrate. Each experiment had its own control and each agent was tested on all plant species. In all, the results of 46 experiments on 20 crops (i.e., corn, wheat, barley, oats, clover, soybeans, white beans, red Mexican beans, sugar beets, table beets, carrots, sweet potatoes, spinach, tomatoes, cotton, seed cotton, bright tobacco and peanuts). No information was provided on how the doses were selected. In general, the data did not provide support for the hypothesis that crop yield would be significantly enhanced and provide consistent commercial value. On occasion, there were some instances of 5–10% increases in yield, but it was not possible to determine whether this was a treatment effect or normal variation.[110] European research conducted concurrently or a few years later likewise did not establish any clear effect on yield,[111-115] but according to Kaindl and Linser was not sufficient to deny any stimulatory effect hypothesis.[116] In fact, studies by Linser and Pelikan[117] and Kaindl,[118] using a radium bearing preparation from a French firm and fertilizing at the rate of 10^{-9} grams of radium/kg of soil, reported increases of 16% in yield for buckwheat. However, according to Kaindl and Linser, the generally negative findings of the American research[110] had a dominating influence on the course of both research and further international testing.[116]

Thus, the findings of the USDA had a major impact on the future of US and international research in this area. In retrospect it would appear that the strategy of the USDA was to consider a broad range of plants, but a very limited focus on dose. In fact, such a limited focus on dose and the non-recognition of interspecies differences in response to low doses of radioactive agents was an extremely poor strategy for testing the hormesis hypothesis. Yet, as noted earlier, the essentially negative findings of such an otherwise impressively large study was uncritically accepted as answering the question of low dose stimulation from a US government perspective. While such a conclusion did not end international or US research on the topic of radiation as a plant stimulant, it marked the end of an area for radium with the first stimulatory reports of cobalt-60 on plant growth occurring but a few years later.

Fungi

Introduction

Fungi have long been the object of study concerning the effects of radiation. These studies have en-

compassed the broad spectrum of radiation, including visible and u.v. radiation, X-rays, and radiation from naturally occurring elements such as radium and uranium. In general, such studies have revealed that the typical dose-response relationship was consistent with the monomolecular dose-response (i.e., linear) curve. On occasion, deviations from such a dose-response curve have been reported and usually attributed to factors such as the age of the culture in the study. Despite the broad consistency of the linear and S-shaped dose-response realtionships, low dose stimulation was occasionally reported, although there were disputes about the reproducibility of the findings and/or their interpretation.

U.V. radiation

U.V. radiation induced stimulation of fungal activities has been reported with respect to mycelium growth rate, fruiting structure growth rate and spore production. In the case of mycelium growth rate, Nadson and Philippov reported much greater yeast colony growth around the edges of an irradiated zone, whereas growth in the middle (i.e., higher dose zone) was diminished.[119] The authors believed that the stimulation was due to small amounts of scattered radiation. However, attempts to confirm these observations were unsuccessful as reported by Luyet[120] and Schreiber[121] who obtained no evidence of stimulation with low doses of u.v. irradiation. However, Smith[83] argued that the lack of replication may have been the result of a limitation in study design, since reports by Chavarria and Clark[122] and herself[83] revealed that the key feature in observing the u.v-induced mycelium growth stimulation was the incorporation of an adequate temporal dimension. More specifically, Smith, working with *Fusarium* cultures, reported temporary mycelium growth stimulation which only occurred after a previous toxic or retardation effect.[83] In her nine dose experiment (0.05–15 min exposed) all doses yielded inhibitory effects on mycelium growth at 24 h (Figure 3A). By 48 h, all but the highest dose were displaying compensatory stimulation growth, with two doses greater than the control (Figure 3B). By 72 h, all but the highest dose had exceeded the controls by 15–40% (Figure 3C).

Such observations of Smith,[83] which were later supported by Sperti *et al.*,[123,124] are consistent with the hypothesis that hormesis represents an overcompensation to a disruption in homeostasis. According to these authors, in their experiments, yeast or other cells which become injured can synthesize growth factor agents which stimulate other cells to divide thereby providing a possible mechanistic explanation.[123–133]

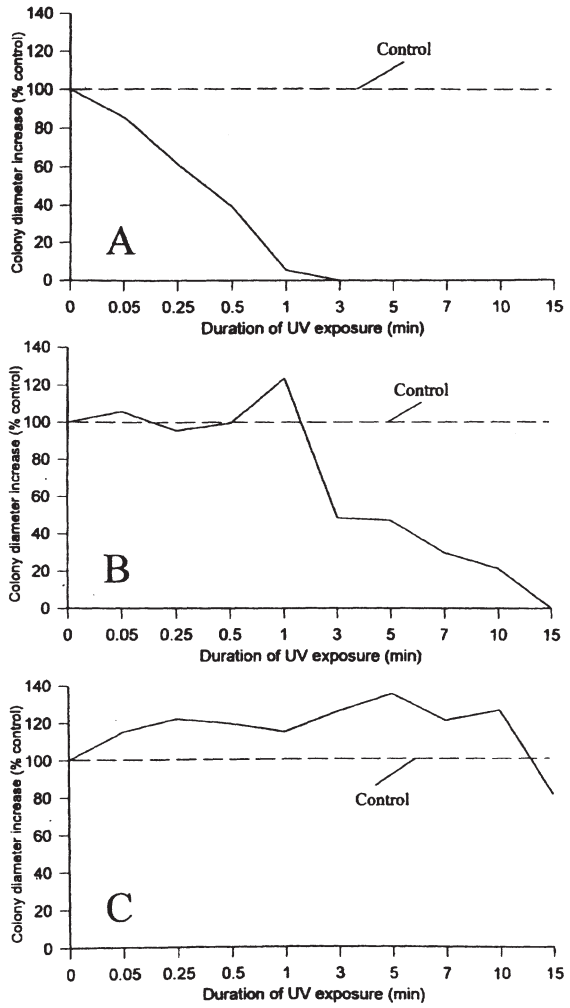


Figure 3 Changes in diameter (% control) of *Fusarium eumartii* Carp. colonies exposed to various durations of ultra-violet (u.v.) irradiation at (**A**) 24 h, (**B**) 48 h, and (**C**) 72 h following irradiation (data from Smith[83])

Stevens, conducting experiments with a large number of fungal species exposed to the full irradiation from a quartz-mercury vapor lamp, observed that the u.v. light may stimulate the formation of reproductive structures.[134–137] Perithecia production was enhanced in cultures of *Glomerella cingulata*,[134,135] *Colletotrichum lagenarium*,[138] and various *Coniothyrium* species.[136] Pycnidia formation was stimulated in *Coniothyrium*.[136] Such stimulatory responses were caused by exposures of less than 1 min at a distance of 20 cm from the lamp.

Consistent with the findings of Stevens is the general observation that long exposures to u.v. irradiation diminish spore production while short exposures stimulate it. Perhaps the earliest report of stimulation of spore production with low doses of radiation was in 1907 by Purvis and Warwick,

working with a *Mucor* culture.[138] They exposed the culture for 10–20 min to direct radiation from a Bach quartz energy vapor lamp located at 30 cm from the culture. The portion of the culture below the center of the opening was killed, but at the edge of the irradiated region spores were stimulated in great numbers. Since that initial discovery, an impressive number of reports were published in which short exposures to u.v. irradiation affected a marked stimulation of spore production in a broad range of fungal species.[83,139–143]

Of particular interest are the findings of Smith[83] since the experiment employed up to ten doses along with a concurrent control. Furthermore, this experiment was conducted at three different temperature settings (21, 25 and 30°C). While the basic trend of an hormetic response was clearly present at each temperature, the temperature had a profound effect on the control number of spores, with the number of control spores increasing as the temperature increased. In addition to the capacity of radiation at low level exposures to increase the number of spores, it may also enhance their formation as seen in the work of Hutchinson and Ashton who reported that sporulation in *Colletotrichum phomoides* was earlier with a brief u.v. exposure duration but delayed with longer duration exposures.[140]

X-rays and naturally occurring radionuclides

Considerably less research on the potential for X-rays and rays emitted from radioactive substances to cause a stimulatory response was conducted in the early decades of the 20th century as compared with u.v. radiation. In the case of X-rays, Lacassagne and Holweck[144] and Wycoff and Luyet[145] reported no evidence of stimulation with low doses of X-rays on yeast. However, Zeller suggested that fermentation may be temporarily increased.[146] In the case of radium, Gager,[91] Kotzareff and Chodat,[104] and Fabre[147] reported that low level exposures were associated with a stimulation of cell division while Ingber[148] reported that small doses of radium may enhance spore production. Likewise, Stoklasa[103] and Kayser and Delaval[149] noted that small doses of radiation enhanced fermentation.

Despite the substantial legitimate criticisms of the presentation by Gager[91] on the stimulatory findings of radium on seeds in culture and for plant growth in soil, the findings provided on alcohol fermentation are his strongest. Of the six experiments, all provide evidence of stimulation and in all cases the data are provided. In general, radium treatments were from 50% to several-fold greater than the control. However, on occassion, the difference between the treated and control was modest (10%), a factor that appeared related to an atypically high value in the control. Nonetheless, of the 14 reported experimental trials involving $RaBr_2$, eight were equal to or greater than twofold that of the control, while five exceeded 40%, and only one was less than 10% (8%). This type of consistency over such a large number of experimental trials provides strong evidence that the radium treatment provided *bona fide* stimulation of alcohol fermentation in yeast. While less extensively evaluated (four experiments) than radium, experiments with radiotellurium which emits alpha rays all displayed stimulatory responses greater than 10% (i.e., 14, 13, 15 and 13).

Summary

Taken collectively, the data as of the mid 1930's supported the conclusion that low doses of u.v. radiation enhanced the growth of the fungal mycelium and spore production. The research with mycelium growth was essentially limited to only three studies.[83,119,122] However, the Smith[83] study was extremely well designed and given heightened credibility as a result of her invitation to singly author a chapter on the effects of radiation on fungi for the National Academy of Sciences in 1936 in which she reaffirmed the hormetic hypothesis.[84] In fact, she linked her observations of the initial reduction in mycelium growth followed by stimulation to several previous reports in different biological models including bacteria,[150] some plant species,[151] and the fungus *Colletotrichum phomoides* using u.v.[140]

The data that u.v. light enhanced spore formation at low doses appears stronger than for mycelium growth since it was more extensively explored by other researchers in addition to Smith. Thus, it appears to be a reproducible and marked response. One major difference with the u.v.-induced stimulation of spores was that the u.v. appeared to act as a direct stimulant, thereby contrasting itself with that observed for the stimulation of growth fungal rate. As in the case of growth stimulation, this stimulatory conclusion was again emphasized by Smith in her 1936 article for the NAS.[84] In contrast to these stimulatory effects induced by u.v. exposure, no general consensus seemed to emerge on the effects of X-rays and naturally radioactive materials on fungal activities. One possible factor that may have affected the broader acceptance of these papers in the US is that each was published in French or German, a factor of uncertain but possible considerable importance in affecting their impact on US scientists.

Algae

The first reports in the literature claiming that u.v.-irradiation accelerated colony development of algae

were given by Meier.[152–156] During the earlier investigations on the lethal effect of 21 wavelengths of the ultraviolet radiation spectrum ranging from 2250–3130 A on a given algal strain, Meier occasionally noted an accelerated increase in cell mass with slightly less exposure than the minimally lethal exposure that destroyed the algal cells.[152–154]

These results lead to follow-up experimentation with the unicellular green algal *Stichococcus bacillaris* to assess whether u.v. radiation could stimulate cell division under various experimental settings. This model offered a variety of attractive experimental features with respect to precise and accurate counting, measurement of the size and method of reproduction. The algae were grown under conditions of regulated temperature and controlled lighting with fluorescent lamps. The algal cultures were grown for 2 weeks following irradiation, at which time a determination of growth rates was made. Separate experiments were conducted at different u.v. wavelengths (i.e., 2352A, 2483A, 2652A, and 2967A) for varying periods of time (i.e., 20 to approximately 300 s exposure depending on the wavelength). The quantity of ergs/s-cm^2 was also specified for each wavelength studied. The growth rate was defined as the final count made 2 weeks after the irradiation divided by the initial count made directly after irradiation. Each growth rate of an irradiated culture was then divided by the growth rate of the control to derive the final growth rates. Duplicate cultures were made of each exposure and control group. The cells of three drops of the culture from each flask were counted and the mean of the three cell counts was used for response determination. Based on the data, there was a strong tendency for a short duration exposure enhancement of growth rates along with a decrease relative to controls at longer durations. While this was the case for each wavelength tested, each wavelength displayed a unique duration of exposure response. Nonetheless, regardless of the unique duration response curve, the maximum stimulatory point for all the tested wavelengths was a duration approximately 65–75% of the toxicity threshold. The magnitude of stimulation varied between approximately 150–225% of the controls with the 2652A wavelength displaying the highest stimulatory response. No statistical analyses of the data were provided. These findings [i.e., magnitude of stimulation (50–125% above controls) and range of stimulation (3–8-fold) depending on the wavelength used] indicate a striking similarity to the recently reported findings with chemical hormesis.[12]

Of significance was that the stimulatory action in the 1939 report of Meier[155] appeared to be sustained with subsequent measurement some 2–3 years later indicating a marked increase in dry weight of the irradiated culture (40 s) for the 2628 A dose, the maximum response group. This and related findings lead Meier-Chase[156] to determine the influence of successive (i.e., repeated) treatments of the algal cells to the original four wavelengths studied (i.e., 2352A, 2483A, 2652A and 2967A). The methodology employed was similar to that used earlier by Meier.[155] However, the time between the successive or repeat exposures varied between the wavelengths used. Likewise, the time or duration of u.v. exposure was different across the wavelengths. However, regardless of the wavelength used, the algal cells were stimulated to approximately 4–5-fold, with the increase appearing as a type of step function with each successive exposure. Follow-up analyses of the algal cells revealed a decrease in length with each stimulatory response along with a general increase in width. Meier-Chase[156] indicated that the decrease in length was predictable because the rate of cell division was so considerably greater in the treated algae that the cells did not have time to achieve the length seen under normal conditions.

The stimulated algal cells were then exposed to lethal doses of u.v. radiation. In all cases the stimulated algal cells were less sensitive to the lethal u.v. doses. In general, the previously stimulated algae required approximately twice as long to display radiotoxic regions as compared to controls.

The findings of Meier[155,156] are striking in their consistency across wavelengths, their repeatability, and their similarity with the copious data available on chemical hormesis. In addition, the follow-up studies display a remarkable similarity to the concept of adaptive response with radiation. However, the long term stimulatory response is more difficult to explain and would require follow-up study.

Protozoans

Experimentation concerning the effects of radiation on protozoans during the early decades of the 20th century was problematic because of their relative insensitivity. Numerous early investigators were unable to induce any notable effects of X-rays on any protozoan species, despite rather prolonged exposures (see Crowther[157] for review). In fact, it was not until the mid 1920's that investigators began to report on the capacity of X-rays to both stimulate[158] and harm protozoan species.[157] Despite the reported apparent stimulation of Markowits[158] with X-rays on paramecia, this section concerns the effects of u.v. radiation on paramecia, since

this received greater attention and is more substantial than other protozoan areas of potential inquiry.

The earliest indication that u.v. radiation may stimulate paramecia was reported by Bovie and Hughes,[159] who noted that the cell division rate of *Paramecium caudatum* could be enhanced or delayed depending on dosage. More specifically, as the duration of exposure increases so does the extent of inhibition. However, and of relevance to the present assessment, the inhibition may be followed by an acceleration of the division rate. These authors hypothesized that acceleration following short periods of inhibition was due to the formation of a 'toxic photoproduct which is gradually removed from the cell' and subsequently 'acts as a stimulant to cell division when the amount becomes very small'. It was not until some 10 years later that the observations of Bovie and Hughes[159] were confirmed and extended by Hinrichs,[160] MacDougall,[161,162] Roskin and Romanowa,[163] and more impressively by Alpatov and Nastiukova.[164] In the case of Hinrichs,[160] cell division rates were assessed over 3 days in paramecia exposed for different durations (1–80 s) and at different distances from the u.v. source (26.5–56.0 cm). In addition, there were differing numbers of paramecia exposed at the same time (i.e., singly, paired, and multiple). Hinrichs summarized her findings by stating that of the 36 experiments conducted, half displayed a u.v.-induced stimulation while the remaining half had a depressive effect.[160] More specifically, in the stimulatory experiments the increase ranged from 7–70.6% over the controls. In these experiments, the exposures were conducted for 5–30 s at 26.5 cm from the u.v. lamp source and for 6–20 s at 37–45 cm from the u.v. source. The number of cases where stimulation occurred was greater in those instances when exposures were conducted at $\geqslant$38.5 cm away from the lamp. In fact, nearly 80% of the exposures conducted $\geqslant$45 cm from the lamp produced a stimulatory division rate, while only 45% of the exposures at closer range showed an increase in the division rate and total offspring produced. Moreover, depression of the rate of division was often observed during the initial 24 h after exposure, with stimulation occurring not until day 2 of observation.

The research of Hinrichs[160] was criticized by Alpatov and Nastiukova[164] because of her use of limited numbers of organisms and the lack of objective means (e.g., hypothesis testing) to guide decisions on stimulation or depression. Despite these legitimate criticisms, it should be emphasized that the value in the work of Hinrichs[160] was that it established an experimentally based framework to test the influence of dose as a function of time of exposure and distance from the u.v. source. Likewise, her findings were consistent with the statements of MacDougall[161] that the cultures of her animal model, *Chilodon uncinatrus*, that were exposed for less than 5 s appeared to be more vigorous and the individuals larger than in the control cultures. It should be noted that in her 1931 paper, MacDougall indicated that her research was being supported by the Committee on Radiation at the NRC.[163]

In their study Alpatov and Nastiukova[164] assessed the effect of u.v. radiation on the division rate of *P. caudatum* with different durations of u.v. exposure while keeping distance from the u.v. source constant. They presented their findings of 14 experiments with typically three doses (i.e., durations of 5, 10 and 20 s) for ten experiments, and longer durations for the remaining four experiments (up to 120 s). The number of organisms involved 20/treatment (i.e., totaling approximately 1000 in the 14 experiments). The findings revealed that at low doses (5–20 s) the division rate of the paramecia was increased, while with the higher durations of exposure (i.e., $\geqslant$40 s) there was a marked decrease. Of significance is that the authors performed statistical testing and claimed that the enhanced responses at the 5–20 s durations were statistically significant.

The collective findings of the stimulatory effects of u.v. radiation on the cell division rate of paramecia up to the mid 1930's were limited to six studies. These studies provide consistent indications that at low doses and/or durations of exposure, the division rate was enhanced, while at high doses (or longer durations) the division rate was diminished. Of these six studies only two provide a quantitative basis for evaluation. In these cases the stronger of the two studies is that of Alpatov and Nastiukova[164] as a result of clearer focus, more powerful study design, enhanced statistical power, and inclusion of hypothesis testing. However, the Alpatov and Nastiukova study[164] was limited to only 24 h of observation after u.v. exposure, whereas the Hinrichs[160] study followed the paramecia for 3 days.

Despite the obvious differences in study design and the various strengths and limitations of the respective studies, it appears that the data clearly suggest that low doses of u.v. radiation can enhance the rate of cell division in Paramecia. The data of Alpatov and Nastiukova[164] were impressive with respect to the dose range employed and statistical power, while those of Hinrichs[160] which were generally consistent with Alpatov and Nastiukova,[164] also offers a dose-temporal relationship. Her

observation of an initial inhibition followed by a stimulatory response are consistent with the over-compensation stimulatory response of Smith,[83] Colley,[150] Townsend,[151] and others. In fact, the low dose stimulatory response reported by Alpatov and Nastiukova[164] was a modest, although statistically significant, response probably because it only included a 24 h period of observation. In the Hinrichs[160] experiment displaying stimulation, the irradiated paramecia had a 5% lower division rate than controls after 24 h. This decrease reversed itself to a stimulatory mode, being some 19 and 38% greater than controls at 2 and 3 days, respectively. Even in the inhibitory response groups the 3rd day displayed a marked acceleration in the division rate over the controls by 40%, although it was insufficient to overcome the earlier inhibitory response.

The findings, while consistent with the concept of low dose stimulation within the context of a compensatory response, would greatly benefit from follow-up experimentation such as a study like that of Alpatov and Nastiukova[164] which included a temporal framework in order to clarify the nature of the dose-response. Nonetheless, these findings, though not conclusive of an hormetic response, were supportive of this relationship as the mid-1930's approached.

It should be noted that the acceleration of division in organisms such as paramecia by small doses of radiation was viewed with skepticism by Kimball nearly two decades later in his generally comprehensive review of the literature of the effects of radiation on protozoa.[165] He cited the well-recognized authority Giese[133] in his review of the effect of radiation on cell division as concluding that 'most of the evidence is of questionable significance' with the effects being small and lacking statistical significance. In some cases Kimball[165] concluded that Giese[133] seemed to accept the validity of several reports of the older literature concerning acceleration by ionizing radiation. Nonetheless, Kimball concluded that 'further investigation seems necessary before accepting stimulation of division by low doses of radiation a real phenomenon.'[165]

The review of Kimball,[165] which was published in a highly authoritative monograph and edited by the renowned Alexander Hollander and therefore given certain enhanced credence, misrepresented the assessment of Giese.[133] Giese's[133] assessment of the acceleration of biological processes by u.v. radiation was presented on pages 263–265 with particular attention directed toward u.v. radiation. In his review, Giese was critical and skeptical of the theory of mitogenic rays (i.e., short u.v. radiation emitted from cells that were hypothesized to stimulate other cells to divide more rapidly).[133] However, he appeared to be supportive of the findings of others when a defined u.v. source induced acceleration of cell division in a variety of models [i.e., Alpatov and Nastiukova;[164] MacDougall;[161,162] Hutchinson and Ashton;[140] Chase;[166,167] Meier;[155] Meier-Chase;[156] Sperti *et al.*;[123,124] Loofbourow *et al.*[128]]. In general, the review of Giese[133] was quite favorable to the stimulatory hypothesis of low doses of radiation with particular focus on u.v. radiation. Consequently, it was unfortunate that the authoritative review of Kimball[165] incorrectly characterized not only the report of Giese[133] but also the broader scientific field and thereby undermined the development of research in this area.

Insects

The evidence associating X-ray exposure with the concept of hormesis in insects during the early decades of the 20th century was extremely limited. In fact, the only research that will be discussed in this context is that of Davey,[168,169] a researcher at General Electric. Despite the limited relevant studies on insects that alleged hormetic effects during this time period, the studies of Davey[168,169] were noted for their unusual quality and remain widely cited references demarcating perhaps one of the first generally convincing earlier experiments presenting evidence consistent with the ionizing radiation hormetic hypothesis.

Perliminary work by Davey[168] explored the effects of a wide range of X-ray doses on the longevity of the confused flour beetle (*Tribolium confusum*). This was initially assessed by comparing the latency period from the time of a single X-ray exposure to death. Davey[169] employed five doses [i.e., 500–8000 milliamperes/min at 25 cm at 50 kilovolts (MAM/25^2 at 50 KV)] and an unexposed control. To the surprise of the author, the lower dose treatments displayed enhanced survival relative to the controls, thereby prompting a follow-up investigation of this stimulatory phenomenon,[169] the frequently cited reference.

The initial report of Davey[168] was unusual in its attention to detail and in its overall intellectual rigor. For example, since the study used mortality as an endpoint, preliminary experiments assessed and eliminated possible confounding factors such as issues of injury due to overcrowding, high temperature due to overcrowding, presence of NO_2 due to high voltage connections of the X-ray tubes, effects of air ionization, humidity and other factors. There was also considerable attention given to the development of an effective, reliable and reproducible quantitative X-ray exposure system. The

author also incorporated the concept of keeping the technicians blind to the study hypothesis as well as attempting to assess uniformity of age distribution of the beetles across exposure and control groups. The sample size was also substantial, making use of several thousand beetles. Furthermore, Davey,[168] while not employing hypothesis testing, did attempt to mathematically model the data using regression techniques. It should be remembered that analysis of variance was not discovered and published until 1918, a year after the report of Davey.[168] Thus, for the numerous reasons cited above, the findings of Davey[168,169] attracted both attention and high regard.

In the initial experiments the dose range studied was 500–8000 MAM/25² at 50 KV, as mentioned above. The findings revealed the typical S-shaped mortality curve with no evidence of a stimulatory response. Subsequent experimentation using 1100 beetles assessed the dose range of 100–500 MAM/25² at 50 KV. This experiment confirmed that the minimum dose needed to kill all the beetles was 500 MAM/25² at 50 KV, but the curves for 100 and 200 MAM/25² at 50 KV displayed a death rate lower than that observed in the controls. It was this finding that was presented in the 1917 paper by Davey.[168]

In the follow-up study of Davey,[169] the effects of X-rays on lifespan were assessed following either a single dose, as in the Davey[168] study, or via low daily X-ray exposures. In the daily exposure experiment, five doses were employed ranging from 6.25–50 MAM/25² at 50 KV daily, with approximately 950 beetles per group. After 5 months nearly all the beetles had died. The mortality rates indicated that the three lowest groups displayed a 25–40% decrease in mortality by 30 days after the start of the study (Figure 4A). The second experiment, using about 850 beetles/group, utilized a single dose involving four doses (100–400 MAM/25² at 50 KV) plus a control (Figure 4B). In contrast to the earlier experiments, the author indicated that these beetles were old, with the controls dying by 40 days. As in the earlier experiments the lowest exposed groups again displayed a reduced mortality rate by 20 days after dosing. According to the author, the 1919 experiments provide a 'direct confirmation' of the previous paper. It is interesting to note that Davey[169] referred to the daily X-ray exposure as 'a series of small 'homeopathic' doses, thereby linking the hormetic findings of his work to the medical practice of homeopathy.

Despite the striking and reproducible findings of Davey,[168,169] it was not until some 40 years later that Cork[170] set forth to reinvestigate the findings of Davey using the same animal model, but using a gamma ray source (Cesium-137) for either single or

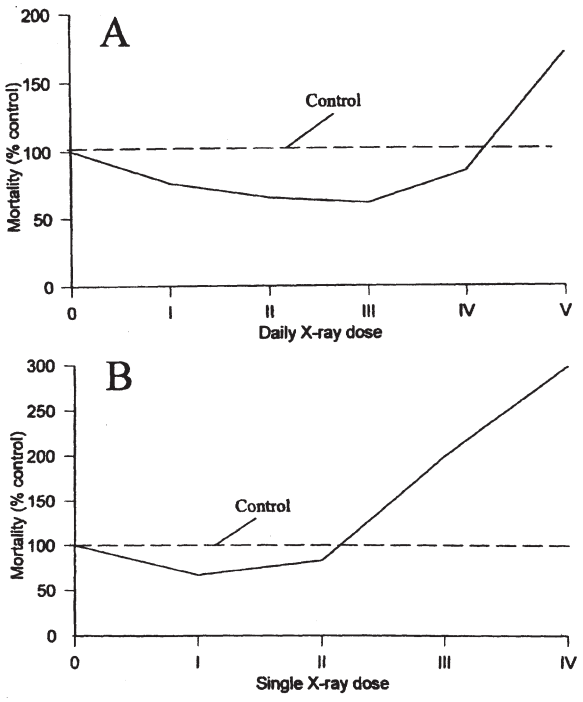


Figure 4 (A) Mortality (% control) of confused flour beetles following 30 days exposure to daily doses of X-rays (dose I=6¼ MAM/25² at 50 KV; dose II=12½ MAM/25² at 50 KV; dose III=25 MAM/25² at 50 KV; dose IV=50 MAM/25² at 50 KV; dose V=100 MAM/25² at 50 KV). (B) Mortality (% control) of confused flour beetles at 20 days after a single exposure to X-rays (dose I=100 MAM/25² at 50 KV; dose II=200 MAM/25² at 50 KV; dose III=300 MAM/25² at 50 KV; dose IV=400 MAM/25² at 50 KV) (data from Davey[169])

chronic daily doses. As in the case of Davey,[168,169] Cork[170] likewise reported a marked extension of the lifespan in a well-designed study with large numbers of beetles.

Avian embryos

Several studies have been published concerning the effects of X-rays on the development of the avian embryo.[171–174] While each of the studies reported a stimulatory response, the paper by Gilman and Baetjer[171] did not present any data but rather descriptive findings and conclusions. The remaining three studies provided markedly more information on research methods and were capable of receiving more detailed attention. In the case of Essenberg[172] the effects of X-rays were assessed for several endpoints: incubation period, time to mating for males and females, and number of eggs produced per month. The author used three treatment groups (30 r, 80 r, 400 r) plus a concurrent control with a total of 600 chicken eggs. It is

assumed (but not stated) that there were 150 eggs/group. No tables or figures were presented, nor were statistical analyses provided. The author claimed that the incubation period varied directly with the X-ray dosage, with the small dosage accelerating development. However, this conclusion appears untenable since the average difference amongst the control and treatment groups is minor (496 h for the controls *vs* 484 h for the 30r group) and no data are presented on variation in response within a group.

The second avian endpoint that the author claimed was accelerated by X-ray treatment of the eggs was 'time to sexual maturity'. In the case of the female, the average control duration was 167 days, while the irradiated eggs required only 134 days (i.e., about 20% accelerated). In the case of the males, the average control was 75 days, while the irradiated males were 69 days (8% acceleration). In both the male and female cases, the author did not provide information on group specific findings, but combined all irradiated groups. Again, no information on variation in response was provided. While it would appear that these findings merit further experimentation, the lack of adequate presentation of the data does not permit a firm conclusion to be drawn. With respect to egg laying, the author reported an acceleration of this process during early weeks followed by a marked reduction, then later accelerations. As in the case of the previous two endpoint assessments, this one also suffers from lack of data presentation thereby precluding any definitive statement.

In contrast to the data presentation limitations of Essenberg,[172] Bless and Romanoff[173,174] offered well-designed and clearly presented studies in which X-rays were administered to 1200 chick eggs across 21 different doses ranging from 1.5–5000r units. For ease of presentation they combined the 21 doses into 7 r-units (8–3000 r). The 24 h blastoderm stage displayed evidence that low doses exerted a stimulatory effect (6–25%) regardless of whether the eggs were exposed in cool beakers, shells, or in preheated shells. Despite the stimulatory response at the blastoderm stage, there was a dose-dependent decrease in the hatchability of eggs.

The studies of Bless and Romanoff[173,174] offer clear evidence that the blastoderm stage is differentially affected by X-rays depending on the dose. However, given the generally negative effect on hatching success, it is uncertain what the biological significance of the stimulation is. Of interest were the poorly reported findings of Essenberg,[172] since it suggested that the developmental processes could be accelerated by low doses of radiation. This finding, while suggestive, represents one area of possible follow-up research some 65 years later.

Salamanders

Morphologenetic stimulation

Stimulation of morphologenetic processes by X-ray treatment has been reported in regions that possess the capacity to form new limbs and when that capacity has not been suppressed by a relatively large dose of radiation. This observation becomes linked to the Arndt-Schulz Law based on reports that stimulation of target tissue is most commonly observed when the target has received less than the intended dose. Under such circumstances the radiation not only does not suppress limb formation, but even stimulates the formation of new limbs. In fact, Brunst[175] reported that animals may grow up to four asymmetrical, but large, hind limbs as well as secondary tails in the salamander. The development of such a radiation-induced secondary tail is what Brunst referred to as the 'zone of stimulation'.[175–177] This zone is characterized by a great mitotic activity in many cells of the narrow boundary zone of the irradiated field. This zone of stimulation represents a very transitory phenomenon and may be easily missed by investigators if they do not adequately sample tissue over time.

In addition to the temporary stimulation there are also cases of late, long continuing stimulation, possibly resulting from stimulatory influences of disintegration products which were referred to as 'necrohormones' orginating from the inhibition zone (see Caspari;[178] Strelin;[179] Zawarzin;[180] Scheremetjewa and Brunst;[181] Brunst and Scheremetjewa[182]). In fact, in the case of irradiation of Triton limbs by Brunst and Scheremetjewa,[182] the beginning of the new regeneration was observed after a period of reduction. Such observations lead to the tentative conclusion that the stimulatory effect can proceed only after a sufficient quantity of disintegration product has accumulated. This interpretation is remarkably similar to the hypothesis of Stebbing[82] that hormesis is an overcompensation to a disruption in homeostasis.

Clinical

Immunological responses and clinical perspectives

There is little question that the concept of 'low dose stimulation, high dose inhibition' as embodied in the Arndt-Schulz Law and subsequently into the concept of hormesis, became the object of clinical verification and application in the early decades of the 20th century in the treatment of human diseases and other conditions by researchers of both traditional and homoepathic perspectives.[183,184] Such attempts of clinical verification and application of

the Arndt-Schulz Law were principally linked to the use of various types of radiation, but especially X-rays. This follows from the timing of the initial reports of Schulz[14,15] in the late 1880's and the discovery less than a decade later of the X-ray by Roentgen. Given the immediate scientific/medical interest in the application of X-rays (i.e., 1000 papers were published on it within 1 year of the discovery!) and the relative ease of creating the condition to produce X-rays, there was little doubt that the testing of the Arndt-Schulz Law in clinical practice would be driven by the X-ray. In fact, by 1897 Leopold Freund became the first person to employ X-rays for therapeutic purposes. He also was the first to report the disappearance of inflammatory symptoms following treatment.[185,186] Such activities of Freund ushered in what was to become the beginning of the medical practice of X-rays for therapeutic application, but also the notion that X-ray treatment can include both beneficial and harmful effects, an hypothesis that was soon to be referred to by the phrase 'depending on the dose.'

As early as 1907, Crane demonstrated that the opsonic index (i.e., a mathematical ratio characterizing the ability of white blood cells to kill specific bacteria[187]) was increased in patients irradiated for infections, an observation that was repeatedly confirmed by well-recognized researchers of that era.[188–193] Such findings lead to the early general conclusion that the bacteriocidal quality of blood was enhanced by small doses of radiation, with the effects peaking some 48–72 h following irradiation. Furthermore, such stimulatory responses on the capacity to opsonize bacteria following low doses of irradiation were consistent with subsequent observations that low doses of X-rays induced reticuloendothelial stimulation likewise at low doses.[194,195] As Pendergrass and Hodes[196] emphasized, these suggestions of beneficial responses applied to small quantities of irradiation, while heavier doses or repeated smaller doses were observed to be harmful, lead to widespread therapeutic applications.

While the effects of low doses of radiation on normal physiological processes such as opsonization and reticuloendothelial stimulation were noted, radiotherapy was also widely employed for the treatment of various inflammatory conditions such as furuncle (boil), carbuncle (suppurating inflammation of the skin and subcutaneous tissues due to Staphylococci), pyrogenic (pus) infections, pneumonia, trachoma, parotitis, nephritis, and numerous other inflammatory conditions (see reviews by Desjardins[197–201]). In the case of pyrogenic infections, the preponderance of the published data indicate that the majority of patients reported rapid and substantial benefit, that is, pain was markedly reduced within a day. Furthermore, the radiotherapy greatly interrupts the predicted progression of the infection, thereby preventing the need for subsequent clinical interventions. The magnitude of the clinical literature, especially in the early decades of the 20th century, was substantial. For example, the 1926 report of Heidenhain[202] reviewed some 855 cases with 76% recovering without surgical intervention. The key factors associated with these initial clinical successes of the therapeutic application of X-rays for inflammatory symptoms was both the striking rapidity of improvement and the low nature of the radiation dose. More specifically, a dose of moderately filtered rays ranging from 50–150 r was demonstrated to be highly effective in a large number of cases.[186]

In the case of pneumonia, the first report of a beneficial response from radiotherapy was given by Musser and Edsall in 1905.[203] This involved the case of a delayed pneumonia resolution in which radiation was followed by immediate resolution and recovery (see Desjardin[197]). Within a year, Edsall, who later became dean of the Harvard Medical School and director of Massachusetts General Hospital, and Pemberton reported beneficial responses from radiotherapy for three additional cases in which moderate irradiation of the lungs was soon followed by recovery.[204] In 1916 the highly regarded Quimbys verified the above mentioned findings with 12 additional cases of delayed resolution.[205] These authors concluded that 'no pathologic process in the body responds quicker to an X-ray exposure that the nonresolution following pneumonia.' Numerous follow-up confirmatory studies over the next several decades were published demonstrating a comparable beneficial effect of radiotherapy on postoperative pneumonia, as well as on pneumonia unrelated to surgical intervention.

The eye disease, trachoma, which involves the sclerotization of eyelids, was first reported to be cured by X-ray treatment by Mayou[206,207] reporting on the findings of 16 patients. These initial striking results were confirmed and extended by numerous investigators (Table 2). Particularly impressive were the findings of Thielemann,[246] Cochard,[250] and Sabbadini.[254] As in the cases of therapeutic application, the beneficial effect is most likely when treatment is administered during the early stages of the disease process.

The issue of what is a low dose has always been problematic. However, in the case of X-ray treatment of inflammatory conditions the guidance offered by Desjardins[197–201] and Borak[186] is informative. They indicate that if the dose needed to cause

Table 2 Studies demonstrating a beneficial effect of low dose X-ray treatment on specific diseases in humans

Furuncle, carbuncle, and other pyogenic infections	*Pneumonia*	*Trachoma*	*Gas-bacillus, peritonitis*
Coyle 1906[208]	Musser and Edsall 1905[203]	Mayou 1902, 1903[206,207]	Kelly 1933, 1936[255,256]
Dunham 1916[209]	Edsall and Pemberton 1907[204]	Stephensen and Walsh 1903[238]	Hubeny and McNattin 1938[257]
Ross 1917[210]	Quimby and Quimby 1916[205]	Bettremieux 1903[239]	Kelly and Dowell
Richards 1922[211]	Krost 1925[226]	Cassidy and Rayne 1903[240]	1936, 1941[258,259]
Lewis 1923[212]	Torrey 1927[227]	Geyser 1903, 1904[241–242]	Altemeier and Jones [260]
Hodges 1924, 1925[213,214]	Heidenhain 1917[228]	Pardo 1904[243]	Bates 1937[261]
Heidenhain and Fried 1924[188]	Heidenhain and Fried 1924[188]	Horniker and Romanin 1905[244]	Faust 1934[262,263]
Pordes 1923, 1923–24,	Kaess 1925[229]	Stargardt 1905[245]	Kelly et al. 1938[264]
1926, 1929[215–218]	Fried 1926[230]	Thielemann 1905[246]	
Holzknecht 1926[219]	Holzknecht 1926[219]	Newcomet 1912[247]	
Gerber 1926[220]	Gadjanski 1927[231]	Jacqueau et al. 1920[248]	
Fraenkel and	Glas 1927[232]	Rollet and Bussy 1927[249]	
Nissjewitsch 1926[221]	Holtz 1929[233]	Cochard 1921[250]	
Solomon and Blondeau 1927[222]	Merritt and McPeak 1930[234]	Meldolesi and Sabbadini 1923[251]	
Carp 1927[223]	McIntire and Smith 1937[235]	Meldolesi 1924[252]	
Light and Sosman 1930[224]	Powell 1938, 1939[236–237]	Lane 1924[253]	
King 1937[225]		Sabbadini, 1926[254]	

erythema of the skin is assumed to be 100%, the dose successful in treating inflammatory conditions has been generally less than 50%, and at times even less than 10%. In fact, they emphasize that the results obtained with doses approaching the SED (skin erythema dose) are less successful than those treatments following the lower dose.

Given the substantial amount of clinical data indicating a beneficial effect of low doses of X-ray treatment on various inflammatory diseases, a number of speculative discussions ensued during the 1930's and 1940's on the possible underlying mechanisms. It has generally been shown that the beneficial X-ray treatment does not have a direct killing effect on the invading bacteria; consequently, the hypothesis that the X-ray treatment was beneficial because it destroyed the known causative agent was discredited. It has also been shown that X-rays act to enhance the bactericidal capacity of the blood as a result of the stimulation of both antibody production and phagocytosis of the reticuloendothelial system. This low dose stimulatory response hypothesis was challenged by Borak[186] who argued that if the stimulatory hypothesis were correct, one would expect that a beneficial effect should be obtained by radiating any region of the body. However, the X-ray treatment works only when the inflamed site is treated. Thus, if a patient has furuncles on both axillae and only one is irradiated, the irradiated region is the only one that will improve. A third hypothetical mechanism involved the enhanced radiosensitivity of leukocytes. This position was challenged by Borak[186] who claimed that the leukocytes do not decrease in cell number unless the blood forming organs are exposed; that if the effect of X-rays were directly related to leukocyte destruction, their effectiveness would be enhanced as the dose increases, yet clinical practice indicates just the opposite. Furthermore, the neutrophils (polymorphonuclear leukocytes) which are major factors in affecting the inflammatory process are relatively insensitive to X-rays. A fourth hypothesis assigns the principal effects caused by X-rays on inflammatory conditions to effects on the blood vessels. This hypothesis argues that the X-rays caused dilation of the capillaries which increase the permeability of the capillary walls, thereby increasing the entrance of antibodies and phagocytes to the inflamed area(s). The enhanced edema results in an increase in tension of the inflamed area. This provides an opening of the lymphatic capillaries. The dilation of the lymph vessel leads to an increase of their resorptive function. In contrast to the X-ray induced effects on blood capillaries, the arteries and veins become narrowed by the same dose due to the swelling of endothelial cells into the lumen. According to Borak,[186] a small dose of X-rays is able to produce dilation of the capillaries and a narrowing of the arteries in the inflammation process since the blood vessels exhibit a greater irritability in an inflammatory condition. Thus, a small dose will produce a further enlargement of the capillaries while reducing the dilated arteries to the normal lumen size.

Marked success was reported by Kelly and Dowell[259,265] in the treating of patients with gas bacillus infections and/or acute peritonitis. Such success had been initially reported by Kelly[255] as early as 1931 based on a presentation at the Radiological Society of North America. These authors used doses of 75 r per day for two days

(150 r total). These findings were substantiated by Dowdy and Sewell,[266] Merritt *et al.*,[267] and Cantril and Buschke.[268] Prior to the 1930's the mortality rate for gas gangrene had been $\geqslant 50\%$ along with substantial amputations. However, with the adoption of X-ray therapy the mortality rate and the need for tissue removal markedly decreased (Figure 5).

This brief review of the clinical literature concerning the beneficial aspects of X-ray therapy is based on numerous studies over the initial four decades of the 20th century. The clinical research was conducted at the most prestigious medical institutions in Europe and the United States and was published in the most mainstream and leading journals in the field. For example, the critical reviews by Desjardins, Chief Radiologist at the Mayo Clinic, were published in the journals *Radiology*, the *Journal of the American Medical Association*, and the *New England Journal of Medicine*[197-201] Likewise, the review by Borak[186] was published in *Radiology*, that of Pendergrass and Hodes[196] in the *American Journal of Roentgenology*, and that of Taliaferro and Taliaferro[78] in the *Journal of Immunology*.

The findings of the clinical researchers, especially in the early years of the 20th century, were often criticized because of the lack of rigorously designed blind clinical trials that are typically conducted today. However, this criticism was often mitigated by the citation of multiple animal model studies that supported the clinical investigations as well as the sheer magnitude of consistent findings from clinical investigations by multiple independent investigators.

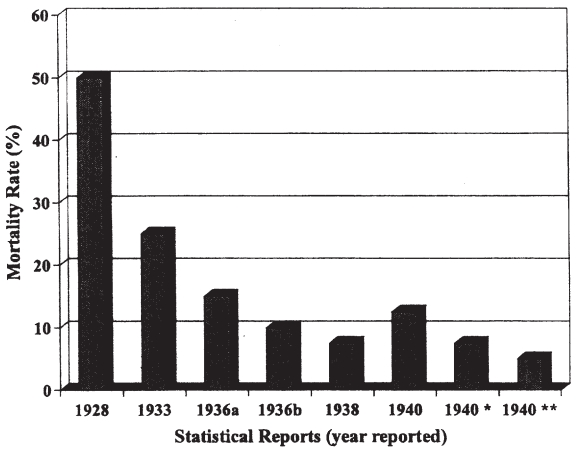


Figure 5 Mortality rate since X-ray therapy was introduced in 1928. Note: mortality associated with patients receiving surgery, serum, and one or more X-ray treatments unless indicated otherwise; (*) indicates mortality associated with patients receiving surgery, serum, and three or more X-ray treatments; and (**) indicates mortality associated with patients receiving three or more X-ray treatments with no surgery or serum treatments. Reports: 1933=Kelly;[255] 1936a=Kelly;[256] 1936b=Kelly and Dowell;[258] and 1938=Kelly *et al.*;[264] 1940=Kelly and Dowell[259]

While the weight of evidence strongly favored a causal relationship of the X-ray treatments and the range of beneficial effects, the issue of whether the response is consistent with the hormetic hypothesis is difficult to resolve within the context of epidemiological studies since often only one dose is evaluated in clinical settings. In the case of the therapeutic use of X-rays to treat a wide range of inflammatory diseases, it appears fairly conclusive that there was a low dose benefit, high dose toxicity, thereby being consistent with the hormetic perspective.

Two papers by Glenn[269,270] provided the capacity to more formally assess the capacity of X-rays to affect immunological parameters with respect to the hormesis evaluation index, and thereby afford the possibility of providing an experimental corroboration of the above cited clinical observations. The initial study by Glenn[269] was of a preliminary nature in assessing the effects of X-rays on the phagocytic capacity of rabbits exposed to hemolytic *Staphyloccus aureus*. Of particular relevance to the hormesis hypothesis was that Glenn used five treatments plus a concurrent control. In this experiment there was a clear low dose stimulation (6.5-fold) followed by a sharp return toward control value as the dose increased. In the follow-up study,[270] nine doses were employed along with the concurrent control. As in the pilot experiment, there was a low dose stimulation of sevenfold followed by a return to control value as the dose increased.

While the collective findings clearly support the perspective that low doses of X-rays have a marked and reproducible therapeutic benefit to patients with various inflammatory diseases, there was still debate even among supportive researchers on how to interpret such findings. More specifically, there were two schools of thought concerning interpretation of the beneficial response. While both agree that functional activity followed low dose X-ray treatment, they markedly differed with respect to the mechanism involved. In the case of Fraenkel and his followers, it was believed that small doses of radiation cause a direct stimulation. In contrast, Holzknecht and Pordes argue that the X-ray treatment causes stimulation via a depressing factor which then releases the cells from a restraining influence.[183,184]

These different perspectives on hormesis have been periodically noted over the past century. The Holzknecht and Pordes perspective is highly consistent with subsequent reports of Hektoen[192] and

Bloom and Jacobson[271] who, also studying X-ray effects on biological systems, concluded that the 'stimulation was an example of reparative over-compensation after initial damage.'

Discussion

This review has demonstrated that the hypothesis that is today called radiation hormesis has been evaluated by numerous investigators, using highly diverse plant and animal models over the initial decades of the 20th century. Particularly noteworthy were the highly consistent findings of a low dose stimulation, high dose inhibition for an exceptionally wide range of plant species. Likewise, convincing evidence of hormetic response were seen in the research on various fungal species, protozoans, algae and insects. While some of the findings would be considered inadequate or even poor by current standards, many other supportive experimental findings would be considered quite impressive even today. As in the case with that observed with historical features of chemical hormesis,[12] these observations of low dose stimulation were usually quite unexpected. For example, the observations of Davey[168,169] that low doses of X-rays enhanced longevity in the confused flour beetle were at first totally unexpected, but then highly reproducible in subsequent confirmatory experimentation. In fact, this type of process of initially observing an unexpected stimulatory response with follow-up confirmation and extension of the hormetic finding is a general feature of the database of the early decades of the 20th century. This combination of unexpected initial observation and reproducibility are important factors enhancing the credibility of the hormetic hypothesis, since they speak both to a lack of bias on behalf of such investigation and to the consistency of the initial observations.

The assessment also reveals that a large number of reports of hormetic-like findings were conducted by highly prestigious investigators, residing at some of the most outstanding research institutions in Europe, the United States and Japan, and published in the leading journals of that period, such as the *Journal of the American Medical Association*, the *New England Journal of Medicine*, and the *Journal of Immunology*.

Of particular importance is that the stimulatory responses were remarkably similar across the various biological models evaluated following exposure to various types of radiation with respect to stimulatory dose range, maximum stimulatory response, and distance of maximum stimulatory response to the threshold for toxicity (NOAEL). In fact, such responses were also highly consistent with that observed with the developing chemical hormesis database, as well. Further, the stimulatory response was often seen after an initial inhibitory response, thereby suggesting an overcompensation response to an initial disruption in homeostasis.

Despite the extensive earlier findings of a low dose stimulation, high dose inhibition to radiation exposure in numerous models, including humans, the belief that radiation hormesis was a general biological phenomenon came to be severely questioned in the mid 1930's and eventually became a marginalized hypothesis at best, and often the source of ridicule. Given the substantial initial scientific foundations of the hormetic hypothesis in the biological and medical sciences, it is important to consider how the concept of radiation hormesis evolved into a nearly forgotten concept, being ignored by leading radiological and toxicological texts, never the subject of technical sessions at society conferences, and with no place in the curriculum of toxicologists and biomedical scientists. The following article will explore the basis of the remarkable fall of the hormetic hypothesis from that of mainstream theory to an historical footnote and whether this was a justified demotion or whether a bona fide biological hypothesis with potentially profound toxicological and societal implications was inappropriately marginalized.

Acknowledgements

We would like to thank Richard S Szczygiel, Jr. for his dilligent efforts as a library sleuth. This work has been sponsored in part by a grant to the University of Massachusetts (Edward J Calabrese, Principal Investigator) by the US Nuclear Regulatory Commission.

References

1 Luckey TD. Hormesis With Ionizing Radiation. CRC Press, Inc.: Boca Raton, FL, 1980.
2 Luckey TD. Radiation Hormesis. CRC Press, Inc.: Boca Raton, FL, 1991.
3 Calabrese EJ (editor). Biological Effects of Low Level Exposures to Chemicals and Radiation. Proceedings of the First BELLE Conference, April-May, 1991, Amherst, MA. Lewis Publishers: Chelsea, MI, 1992.

4 Calabrese EJ (editor). Biological Effects of Low Level Exposures–Dose Response Relationships. Proceedings of the Second BELLE Conference, April 1993, Crystal City, VA. CRC Press: Boca Raton, FL, 1994.
5 Calabrese EJ (editor). Toxicological Defense Mechanisms and the Shape of Dose-Response Relationships. Proceedings of the Third BELLE Conference, November 1996, Research Triangle Park, NC. *Environmental Health Perspectives* 1998; **106**(Suppl.1): 275–394.
6 Liu SZ (editor). Proceedings of the International Symposium on Biological Effects of Low Level Exposures to Radiation and Relataed Agents. Cangchun, China. Princeton Scientific Publishers, Princeton, NJ, 1994.
7 Sugahara T, Sagan LA, Aoyama T (editors). Low Dose Irradiation and Biological Defense Mechanisms. Proceedings of the International Conference on Low Dose Irradiation and Biological Defense Mechanisms. July 12–16, Kyoto, Japann, 1992.
8 Calabrese EJ, Baldwin LA. Chemical hormesis: Its historical foundations as a biological hypothesis. *Toxicologic Pathology* 1999; **27:** 195–216.
9 Calabrese EJ, Baldwin LA. The marginalization of hormesis. *Toxicologic Pathology* 1999; **27:** 187–194.
10 Calabrese EJ, Baldwin LA. The dose determines the stimulation (and poison): development of a chemical hormesis database. *International Journal of Toxicology* 1997; **16:** 545–559.
11 Calabrese EJ, Baldwin LA. A quantitatively-based methodology for the evaluation of chemical hormesis. *Human and Ecological Risk Assessment* 1997; **3:** 545–554.
12 Calabrese EJ, Baldwin LA. A general classification of U-shaped dose-response relationships in toxicology and their mechanistic foundations. *Human and Experimental Toxicology* 1998; **17:** 353–364.
13 Southam CM, Erlich J. Effects of extracts of western red-cedar heartwood on certain wood-decaying fungi in culture. *Phytopathology* 1943; **33:** 517–524.
14 Schulz H. Zuhre Lehre von der Arzneiwirdung. *Virchows Archiv fur Pathologische Anatomie und Physiologie fur Klinische Medizin* 1887; **108:** 423–445.
15 Schulz H. Uber Hefegifte. *Pflugers Archiv fur die gesamte Physiologie des Menschen und der Tiere* 1888; **42:** 517–541.
16 Hueppe F. Principles of Bacteriology. Translated from the German by E.O. Jordan, The Open Court Publishing Company: Chicago, 1896.
17 Schober A. Ein Versuch mit Rontgen'schen Strahlen auf Keimpflanzen. *Berichte der Deutschen Botanishen Gesellschaft* 14, 1896; **No.7:** 108–110.
18 Maldiney E, Thouvenin S. De l'influence des rayons X sur la germination. *Revue generale de Botanique* 1898; **10:** 81–86.
19 Perthes G. Versuche ueber den einfluss der Rontgen- und Radiumstrahlen auf Zellteilung. *Deutsche Med. Wochenschrift Jahr* 1904; **30:** 632.
20 Euler. Uber die heilende Wirkung der Tontgenstrahlen bei abgegrenzten Eiterungen. *Veroffentlichungen Gebiete Des Militar-Sanitatswesen. Berlin* 1906.
21 Koernicke M. Uber die Wirkung von Rontgenstrahlen auf die Keimung und das Wachstum. *Berichte der Deutschen Botanishen Gesellschaft* 1904; **22:** 148–155.
22 Koernicke M. Weitere Untersuchungen uber die Wirkung von Rontgenund Radiumstrahlen auf die Pflanzen. *Berichte der Deutschen Botanishen Gesellschaft* 1905; **23:** 324–333.
23 Guilleminot H. Effets compares des rayons X et du radium sur la cellule [vegetale]. Valeur de l'unite M en physiologie vegetale. *Comptes Rendus Academie des Science (Paris)* 1907; **145:** 798–800.
24 Schmidt HE. Experimentelle Untersuchungen uber die Wirkung kleinerer und grosserer Strahlenmengen auf junge Zellen. *Berliner Klin. Wochenschrift* 1910; **47,** No. 21: 972–974.
25 Wetterer J. Beitrag zur Kenntnis der biologischen Wirkung der Rontgenstrahlen auf das Wachstum der Pflanzen. *Fortschritte auf dem Gebiete Roentgenstrahlen* 1912–1913; **14:** 172.
26 Promsy G, Drevon M. Influence des rayons X sur la germination. *Revue generale de Botanique* 1912; **24:** 177 197.
27 Schwarz E. Der Wachstumsreiz der Rontgenstrahlen auf pflanzliche [und tierische] Gewebe. *Munch. med. Wochenschr. 60,* 1913; **39:** 2165–2169.
28 Miege E, Coupe H. De l'influence des rayons X sur la vegetation. *Comptes Rendus Academie des Science (Paris)* 159, 1914; **4:** 338–340.
29 Koernicke M. Uver die Wirkung verschieden starker Rontgenstrahlen auf Keimung und Wachstum bei den hoheren Pflanzen. *Jahrbucher fur wissenschaftliche Botanik* 1915; **56:** 416–430.
30 Koernicke M. Wirkung der Tontgenstrahlen auf Pflanzen. *Fortschritte auf dem Gebiete Roentgenstrahlen* 27, 1920; **1:** 661.
31 Yamada M. On the effect of Roentgen rays upon seeds of *Oryza sativa. Journal of Physical Therapy* No. 6, referenced in *Journal of the College of Agriculture. Tokyo Imperial University* 1917; **8:** 2.
32 Nakamura S. Comparative experiments of the effect of X rays. Proceedings of Koto-Kwai No. 111, Japan, 1918; referenced in *Journal of the College of Agriculture. Tokyo Imperial University* 1923; **8:** 2.
33 Sierp A, Robbers F. Uber die Wirkung der Rontgenstrahlen auf das Wachsum der Pflanzen. *Strahlentherapie* 14, 1923; **3:** 538–557.
34 Lallemand S. Etude de l'action des rayons X sur le developpement des plantes. *Archiv D'anatomie, d'histologie et d'embryologie* 1929; **10:** 1–233.
35 Weber F. Fruhtreiben ruhender Plflanzen durch Rontgenstrahlen. *Biochemische Zeitschrift* 1922; **128:** 495–507.
36 Altmann V, Rokhlin D, Gleikhgevikht E. Uber [den] entwicklungsbeschleunigenden und entwicklungshemmenden Einfluss der Rontgenstrahlen. *Fortschritte auf dem Gebiete Roentgenstrahlen* 1923; **31:** 51–62.
37 Komuro H. Studies in the effect of Rontgen rays upon the development of *Vicia faba. Journal of the College of Agriculture, Imperial University, Tokyo* 1923; **8:** 253–292.

38 Komuro H. Studies on the effect of Roentgen rays upon the germination of *Oryza sativa. Botanical Magazine (Tokyo)* 1924; **38:** 1–21.
39 Czepa A. Das Problem der wachstumsfordernden und funktionssteigernden Rontgen-Radium-Wirkung. *Strahlentherapie* 1924; **16:** 913.
40 Martius H. Bohnenversuche an Rontgenstrahlen. *Fortschritte auf dem Gebiete Roentgenstrahlen* 1924; **32:** 361–365.
41 Geller FC. Die Wirkung der Tontgenstrahlen auf jugendliche Organismen. *Klin. Wochenschr.* 1924; **3:** 561–566.
42 Gambarov GG. The problem of the so-called 'irritating' effect of X rays. *Vestnik Rentgenol. i Radiol.* 1925; 3, **6:** 311–323.
43 Ancel S. Action de faibles doses de rayons X sur des graines seches. *Comptes Rendus Societe Biologie* 1924; **91:** 1435–1436.
44 Breslavets LB. Plants and X-rays. Translation by A. Elbl. A.H. Sparrow (editor). The American Institute of Biological Sciences: Washington, 1946.
45 Iven H. Neuere Untersuchungen uber die Wirkung der Tontgenstrahlen auf Pflanzen. *Strahlentherapie* 19, 1925; **3:** 413–461.
46 Ancel S. Les rayons X appliques sur des graines seches n'ont aucune influence sur l'epoque d'apparition du germe. *Bulletin Societe Botanique de France* 1925; **72:** 195–197.
47 Ancel S. Sur un phenomene de pseudo-excitation determine par les rayons X sur les bourgeons dormants cotyledonaires de la Lentille. *Bulletin Societe Botanique de France* 1925; **72:** 1084–1088.
48 Kol'tsov AV, Kol'tsov LI. The influence of radium emanations and X-rays on the growth and development of plants. *Zapiski Leningr, Sel'skokhoz. Institute* 1925; **2:** 205–222.
49 Ancel S. Recherche du meilleur test de la radio-reaction des graines des Legumineuses. *Bulletin Societe Botanique de France* 1926; **73:** 71–73.
50 Ancel S. De l'influence accelatrice des rayons X sur le developpement des plantes. *Archives de Physiologiques et Biologie* 1926; **5:** 106–118.
51 Bersa E. Strahlenbiologische Untersuchungen. I. Zur Frage der Rontgenreizwirkung bei Keimlingen. *Sitzungsbericgte Akademie der Wissenschaften In Wien, Mathematische–Naturwissenschaftliche Klasse* Abt. 1, 135, 1926; **1:** 425–451.
52 Johnson E. Effect of X rays on growth, development and oxidizing enzymes of *Helianthus annuus. Botanical Gazette* 1926; **82:** 373–402.
53 Johnson E. Growth and germination of sunflowers as influenced by X rays. *American Journal of Botany* 1928; **15:** 65–76.
54 Doroshenko AV. The influence of X radiation on the length of the vegetative period in plants. *Tr. Prikl. Bot., Genet. i Sel.* 23, 1929–1930; **2:** 511–535.
55 Sprague H, Lenz M. The effect of X rays on potato tubers for 'seed.' Science 69:606.Weber, F. 1922. Fruhtreiben ruhender Pflanzen durch Rontgenstrahlen. *Biochemische Zeitschrift* 1929; **128:** 495–507.
56 Patten R, Wigoder S. Effect of X rays on seeds. *Nature* 1929; **123:** 606.
57 Cattell W. The effect of X rays on the growth of wheat seedlings. *Science* 1931; **73:** 531–533.
58 Johnson E. On the alleged stimulating action of X rays upon plants. *American Journal of Botany* 1931; **18:** 603–614.
59 Johnson E. Effect of X radiation upon growth and reproduction of tomato. *Plant Physiology* 1931; **6:** 685–694.
60 Johnson E. The influence of X radiation on *Atriplex hortensis* L. *New Phytologist* 1933; **32:** 297–307.
61 Chekhov VP. The influence of X rays on plants. *Tr. Tomsk. Gosudarst. Univ. im. V. V. Kuibysheva* 1932; **85:** 67–135.
62 Shull C, Mitchell J. Stimulative effects of X rays on plant growth. *Plant Physiology* 1933; **8:** 287–296.
63 Benedict HM, Kersten H. Effect of soft X rays on germination of wheat seeds. *Plant Physiology* 1934; **9:** 173–178.
64 Francis DS. The effects of X-rays on growth and respiration of wheat seedlings. *Bulletin of the Torrey Botanical Club* 1934; **61:** 119–153.
65 Breslavets LP, Afanas'eva AS. Increase of yield under the influence of X rays. Rye. II. Irradiation of seeds. *Tr. Vsesosoyuznyi Institut Udobrenii Agropochvovedeniya I Agrotekhniki* 1935; **8:** 245–253.
66 Breslavets LP, Afanas'eva AS. The action of X rays on rye. II, Irradiation of seeds. *Vestnik Rentgenol. I Radiol.* 1935; **14:** 302–324.
67 Long T, Kersten, H. Stimulation of growth of soy bean seeds by soft X rays. *Plant Physiology* 1936; **11:** 615–621.
68 Frolov G. The action of X rays and ultraviolet rays on the growth of plants. *Tr. Sel'skokhoz. Akad. Im. Timiryazeva* 1, 1936; **2:** 189–206.
69 Johnson E. Effects of X rays upon green plants. In: *Biological Effects of Radiation.* BM Duggar (ed). **2.** pp. 961–985. McGraw-Hill: New York, 1936.
70 Saeki H. Studies on the effects of X-ray radiation upon germination, growth and yield of rice plants. *Journal of the Society of Tropical Agriculture [Taiwan]* 1936; **8:** 28–38.
71 Zankevich E, Brunst V. The influence of X rays on the individual development of tobacco, poppies, flax, and rhubarb. *Inst. Bot. AN UkSSR,* 1937; **10/18:** 77–98.
72 Bless AA. Effects of X-rays on seeds. *Plant Physiology* 1938; **13:** 209–211.
73 Zaurov EI. An experiment on the effect of X rays on Indian hemp. *Biologii Zhurnal* 1937; **6:** 479–486.
74 Breslavets LB. The present status of X-ray biology of plants. Byull. Moskov. Obshchestva Ispytatelei Prirody. *Otdel biol.* 1937; **46:** 359–369.
75 Wort DJ. X-ray effects on the growth and reproduction of wheat. *Plant Physiology* 1941; **16:** 373–383.
76 Porter TM. The Rise of Statistical Testing, 1820–1900. Princeton University Press: Princeton, NJ, 1986.
77 Hudson JC. Roentgen-Ray Dosimetry. In: O. Glasser, The Science of Radiology. Charles C. Thomas: Springfield, IL. Pp. 120–139, 1933.
78 Taliaferro WH, Taliaferro LG. 1951. Effect of X-rays on immunity: A review. *Journal of Immunology* 1951; **66:** 181–212.

79 Pfeiffer T, Simmermacher W. The influence of Rontgen rays on the seeds of *Vicia faba* as shown in the development of the plants. *Landwirtschaftlichen Versuchs-Stationen* 1915; **86:** 35–43.
80 Jungling O. Die praktische Verwendbarkeit der Wurzelreaktion von *Vicia faba equina* zur Bestimmung der biologischen Wertigkeit [der Rontgenstrahlen]. *Munch. med. Wochenschr.* 1920; **40:** 1141–1144.
81 Johnson E. Susceptibility of seventy species of flowering plants to X radiation. *Plant Physiology* 1936; **11:** 319–342.
82 Stebbing ARD. A theory for growth hormesis. *BELLE Newsletter* 1997; **6:** 1–11.
83 Smith EC. Effects of ultra-violet radiation and temperature on Fusarium. II. Stimulation. *Bulletin of the Torrey Botanical Club* 1935; **62:** 151–164.
84 Smith EC. The effects of radiation on fungi. In: *Biological Effects of Radiation, Vol. II.* Duggar BM (ed) pp. 889–918. McGraw-Hill Book Co., Inc.: NewYork. 1936.
85 Seide J. 1929.
86 Failla G, Henshaw PS. The relative biological effectiveness of X-rays and gamma rays. *Radiology* 1931; **17:** 1–43.
87 Sax K. The effect of ionizing radiation on plant growth. *American Journal of Botany* 1955; **42:** 360–364.
88 Sax K. The stimulation of plant growth by ionizing radiation. *Radiation Botany* 1963; **3:** 179–186.
89 Packard C. Roentgen radiations in biological research. *Radiology* 1945; **45:** 522–533.
90 Clark AJ. Handbook of Experimental Pharmacology. Verlig Von Julius Springer: Berlin, 1937.
91 Gager CS. Effects of the rays of radium on plants. *Memoires of the New York Botanical Garden* 1908; **4:** viii+278.
92 Gager CS. Present status of the problem of the effect of radium rays on plant life. *Memoires of the New York Botanical Garden* 1915; **6:** 153–160.
93 Gager CS. The effects of the rays of radium on plants. In: *Biological Effects of Radiation, Vol. II.* BM. Duggar (ed) pp. 987–1013. McGraw-Hill Book Co.: New York, 1936.
94 Ewart AJ. *Journal of the Department of Agriculture, Victoria* 1912; **10:** 417–421.
95 Sutton HF. Radium and plant growth. *Gardeners' Chronicle* 3rd Ser. 1915; **58:** 102.
96 Ross WH. The use of radioactive substances as fertilizers. U.S. Department of Agriculture Bulletin **149,** 1914.
97 Hopkins CG, Sachs WH. Radium fertilizer in field tests. *Science* 1915; **41:** 732–735.
98 Ramsey RR. Radium fertilizer. *Science* 1915; **42:** 219.
99 Stoklasa J. Influence du selenium et du radium sur la germination des grains. *Comptes Rendus Academie des Science (Paris)* 1922; **174:** 1075–1077.
100 Stoklasa J *et al.* Die Dynamik der physiologischen Wirkung der Radioaktivitat auf das Protoplasma. *Beitrage zur Biologie der Pflanzen* 1930; **18:** 185–224.
101 Stoklasa J, Penkava J. Biologie des Radiums und Uraniums. Paul Parey: Berlin, 1932.
102 Stoklasa J. Bedeutung der Radioaktivitat in der Physiologie. *Chemische Zeitschrift* 1913; **37:** 1176.
103 Stoklasa J. Bedeutung der Radioaktivitat in der Physiologie. *Centralblatt fur Bakteriologie* 2. Abt. 1914; **40:** 266–280.
104 Kotzareff A, Chodat F. De l'action exercee par l'emanation du radium sur les levures. *Compte Rendu Societe de Physique et d'Histoire Naturelle de Geneve* 1923; **40:** 36–39.
105 Doumer E. Radiumenanation und die Keimung der Samen. Beiblatter, VI. Internat. Kongress Allgem. Und Arztl. Elektrol. und Radiol. Prog., 1912.
106 Agulhon H, Robert T. The action of radium and radioactivity on germination in the higher plants. *Annales de l'Institut Pasteur, Paris* 1915; **29:** 261–273.
107 Montet D. Action des faibles radioactivities sur la germination des graines. *Comptes Rendus Hebdomadaires des Seances de l'Academie des Sciences* 1932; **194:** 304–306.
108 Montet D. L'action de la radioactivite dur la germination des bulbes. *Comptes Rendus Hebdomadaires des Seances de l'Academie des Sciences* 1932; **194:** 1093–1095.
109 Montet D. De l'influence des faibles radioactivities sur la germination. *Comptes Rendus Hebdomadaires des Seances de la Societe Biologie* 1932; **109:** 678–680.
110 Alexander LT. Radioactive materials as plant stimulants–field results. *Agronomy Journal* 1950; **42:** 252–255.
111 Lemmermann OG, Schmidt MOC. Die Wirkung von radiumhaltigen Dungemitteln. *Zeitschrift fur Pflanzenernahrung, Dungg und Bodenkunde* 1947; **38:** 185–195.
112 Burstenbinder R. Radioaktive Dungemittel. *Zeitschrift fur Pflanzenernahrung, Dungg und Bodenkunde* 1947; **38:** 200–201.
113 Marx TH, Salm U. Uber die Radioaktivitat und den Dungewert von Plantoradon. *Zeitschrift fur Pflanzenernahrung, Dungg und Bodenkunde* 1947; **38:** 195–200.
114 Gericke S. Dungungsversuche mit Planteondon. *Zeitschrift fur Pflanzenernahrung, Dungg und Bodenkunde* 1948; **42:** 143–147.
115 Baetge HH, Begemann E. Dungungsversuche mit Planteradon. *Zeitschrift fur Pflanzenernahrung, Dungg und Bodenkunde* 1949; **44:** 198–206.
116 Kaindl K, Linser H. Radiation in agricultural research and practice. Developments in the Peaceful Applications of Nuclear Energy No. 10, International Atomic Energy Agency, Karntner Ring, Vienna, Austria, 1961.
117 Linser H, Pelikan W. Gefassversuche mit radioaktiver Dungung. *Bodenkultur* 1951; **5:** 417–424.
118 Kaindl D. Verhalten von Radium in Boden und Pflanze. *Bodenkultur* 1951; **5:** 425–436.
119 Nadson GA, Philippov GS. Action excitante des rayons ultra-violets sur le developpement des levures et des moisissures. *Comptes Rendus Societe de Biologie [Paris]* 1928; **98:** 366–368.
120 Luyet BJ. The effects of ultra violet, X-rays and cathode rays on the spores of Mucoraceae. *Radiology* 1932; **18:** 1019–1022.

121 Schreiber H. Strahlenbiologische Untersuchungen besonders im ultravioletten Spektralbezirk an Saccharomyces turbidans Hansen. *Strahlentherapie* 1934; **49:** 541–595.
122 Chavarria AP, Clark JH. The reaction of pathogenic fungi to ultra-violet light and the role played by pigment in this reaction. *American Journal of Hygiene* 1924; **4:** 639–649.
123 Sperti GS, Loofbourow JR, Dwyer CM. Proliferation-promoting factors from ultraviolet injured cells. *Ins. Divi. Thomae* 1937; **1:** 163–191.
124 Sperti GS, Loofbourow JR, Lane MM. Effects on tissue cultures of intercellular hormones from injured cells. *Science* 1937; **86:** 611.
125 Loofbourow JR. Intercellular hormones. 7. Release of aminoacids by damaged living yeast cells. *Biochemical Journal* 1947; **41:** 119–122.
126 Loofbourow JR, Dwyer CM. Production of intercellular hormones. *Nature (London)* 1939; **143:** 725–726.
127 Loofbourow JR, Morgan MN. Intercellular wound hormones from ultraviolet injured cells. *Stud. Inst. Divi Thomae* 1938; **1:** 137–153.
128 Loofbourow JR., Cook ES, Stimson MM. Chemical nature of proliferation-promoting factors from injured cells. *Nature (London)* 1938; **142:** 573.
129 Loofbourow JR, Cronin AG, Lane MM. Proliferation producing intercellular hormones. I. Quantitative study of factors produced by injured animal tissue cells. *Biochemical Journal* 1940; **34:** 432–441.
130 Loofbourow JR, Englert ME, Dwyer CM. Increased yield of nucleic acid-like substances from irradiated yeast. *Nature (London)* 1941; **148:** 113.
131 Loofbourow JR, Oppenheim-Errera S, Loofbourow DG. Intercellular hormones. 8. Release of nucleotides and nucleosides by damaged living cells. *Biochemical Journal* 1947; **41:** 122–129.
132 Loofbourow JR, Webb AM, Abramowitz RK. Relation of aeration to the activity of proliferation promoting action of injured cells. *Nature (Lond.)* 1942; **149:** 272.
133 Giese AC. Radiations and cell division. *Quarterly Review of Biology* 1947; **22:** 253–282.
134 Stevens FL. Effects of ultra-violet radiation on various fungi. *Botanical Gazette* 1928; **86:** 210–225.
135 Stevens FL. The sexual stage of fungi induced by ultra-violet rays. *Science* 1928; **67:** 514–515.
136 Stevens FL. The effects of ultra-violet irradiation on various Ascomycetes, Sphaeropsidales and Hyphomycetes. *Centralblatt fur Bakteriologie (Abt. 2)* 1930; **82:** 161.
137 Stevens FL. The ascigerous stage of *Colletotrichum lagenarium* induced by ultra-violet irradiation. *Mycologia* 1931; **23:** 134–139.
138 Purvis JE, Warwick GR. The influence of spectral colors on the sporulation of Saccharomyces. *Proceedings of the Cambridge Philosophical Society* 1907; **14:** 30–40.
139 Dillon-Weston WAR. III. Studies on the reaction of disease organisms to light. The reaction of disease organisms to certain wave-lengths in the visible and invisible spectrum. *Scientific Agronomy* 1932; **12:** 352–356.
140 Hutchinson AH, Ashton MR. The effects of radiant energy on growth and sporulation in *Colletotrichum phomoides*. *Canadian Journal of Research* 1930; **3:** 187–198.
141 Ramsey GB, Bailey AA. Effect of ultra-violet radiation on sporulation in *Macrosporium* and *Fusarium*. *Botanical Gazette* 1930; **89:** 113–136.
142 Bailey AA. Effects of ultraviolet radiation upon representative species of Fusarium. *Botanical Gazette* 1932; **94:** 225–271.
143 Porter CL, Bockstahler HW. Concerning the reaction of certain fungi to various wave lenghts of light. *Proceedings of the Indiana Academy of Sciences* 1928; **38:** 133–135.
144 Lacassagne A, Holweck F. Sur la radiosensibilite de la levure *Saccharomyces ellipsoideus*. *Comptes Rendus Societe de Biologie [Paris]* 1930; **104:** 1221–1223.
145 Wycoff RWG, Luyet BJ. The effects of X-rays, cathode and ultra-violet rays on yeast. *Radiology* 1931; **17:** 1171–1175.
146 Zeller H. Wirkung von Arzneimitteln und Strahlen auf Hefe. III. Wirkung von Rontgenstrahlen auf Hefe. *Strahlentherapie* 1926; **23:** 336–353.
147 Fabre G. Effets de l'Activation de l'Atmosphere par l'Emanation de Radium sur la Germination et la Poussee de Divers Organisms Vegetaux. *Comptes Rendus Societe de Biologie [Paris]* 1911; **70:** 187–188.
148 Ingber E. Uber die Radiumsensibilitat des Actinomycespilzes. *Strahlentherapie* 1929; **28:** 581–588.
149 Kayser E, Delaval H. Radioactivite, fixateurs d'azote et levures alcooliques. *Comptes Rendus Academie des Science (Paris)* 1925; **181:** 151–153.
150 Colley MW. Stimulation phenomena in the growth of bacteria as determined by nephelometry. *American Journal of Botany* 1931; **18:** 266–287.
151 Townsend CO. The correlation of growth under the influence of injuries. *Annals of Botany* 1897; **11:** 509–532.
152 Meier FE. Lethal action of ultra-violet light on a unicellular green alga. *Smithsonian Miscellaneous Collections* 1932; **87:** 1–11.
153 Meier FE. Lethal response of the alga *Chlorella vulgaris* to ultraviolet rays. *Smithsonian Miscellaneous Collections* 1934; **92:** 1–12.
154 Meier FE. Lethal effect of short wave lengths of the ultraviolet on the alga *Chlorella vulgaris*. *Smithsonian Miscellaneous Collections* 1936; **95:** 1–19.
155 Meier FE. Stimulative effect of short wave lengths of the ultraviolet on the alga *Stichococcus bacillaris*. *Smithsonian Miscellaneous Collections* 1939; **98:** 1–19.
156 Meier-Chase FM. Increased stimulation of the alga *Stichococcus bacillaris* by successeive exposures to short wave lengths of the ultraviolet. *Smithsonian Miscellaneous Collections* 1941; **99:** 1–16.
157 Crowther JA. The action of X-rays on *Colpidium colpoda*. *Proceedings of the Royal Society of London,* 1926; **B100:** 390–404.

158 Markowits E. Cytologische Veranderungen der Einzeller Paramoecium nach Bestrahlung mit Mesothorium. *Arch. Zellforsch.* 1922; **16:** 238.
159 Bovie WT, Hughes DM. The effects of quartz ultra-violet light on the rate of division of *Paramecium caudatum. Journal of Medical Research* 1918; **39:** 223.
160 Hinrichs MA. Ultraviolet radiation and division in *Paramecium caudatum. Physiological Zoology* 1928; **1:** 394–415.
161 MacDougall MS. Modification in *Chilodon uncinatus* produced by ultraviolet radiation. *Journal of Experimental Zoology* 1929; **54:** 95–109.
162 MacDougall MS. Another mutation of *Chilodon uncinatus* produced by ultraviolet radiation with a description of its maturation processes. *Journal of Experimental Zoology* 1931; **58:** 229–236.
163 Roskin G, Romanowa K. Arzneimittel und ultraviolette Strahlen; die Chininwirkung auf die Zelle bei gleichzeitiger Bestrahlung derselben mit ultravioletten Strahlen. *Zeitschrift fur Immunitatforschrung* 1929; **62:** 147–158.
164 Alpatov NV, Nastiukova DK. The influence of different quantities of ultraviolet radiation on the division rate of *Paramecium. C. r. Acad. Sci. USSR* 1933; **1:** 290–291.
165 Kimball RF. The effects of radiation on protozoa and the eggs of invertebrates other than insects. In: Radiation Biology Vol III: Ultraviolet and Related Radiations. Hollander A (ed). McGraw-Hill Book Company, New York. Pp. 285–331, 1955.
166 Chase HY. The effect of ultraviolet light upon early development in the egg of *Urechis caupo. Biological Bulletin* 1937; **72:** 377–383.
167 Chase HY. The effect of ultraviolet light upon cleavage in certain marine eggs. *Biological Bulletin* 1938; **75:** 134–144.
168 Davey WP. The effect of X-rays on the length of life of *Tribolium confusum. Journal of Experimental Zoology* 1917; **22:** 573–592.
169 Davey WP. Prolongation of life of *Tribolium confusum* apparently due to small doses of X-rays. *Journal of Experimental Zoology* 1919; **28:** 447–458.
170 Cork JM. Gamma radiation and longevity of the flour beetle. *Radiation Research* 1957; **7:** 551–557.
171 Gilman PK, Baetjer FH. Some effects of the rontgen rays on the development of embryos. *American Journal of Physiology* 1904; **10:** 222–224.
172 Essenberg JM. Effect of X-rays on the incubation period, sexual development, and egg-laying in white and brown leghorn chickens. *Poultry Science* 1935; **14:** 284, 293, 317.
173 Bless AA, Romanoff AL. The effect of X-ray stimulation on the bioelectric potentials of the avian egg. *Proceedings of the National Academy of Science* 1942; **28:** 306–311.
174 Bless AA, Romanoff AL. Stimulation of the development of the avian embryo by X-rays. *Journal of Cellular and Comparative Physiology* 1943; **21:** 117–121.
175 Brunst VV. Influence of local X-ray treatment on the development of extremities of the young axolotl (*Siredon mexicanum*). *Journal of Experimental Zoology* 1950; **114:** 1–49.
176 Brunst VV. Histopathology of roentgen death of young axolotl (*Siredon mexicanum*). *American Journal of Roentgenology* 1957; **78:** 518–545.
177 Brunst VV. Effects of ionizing radiation on the development of amphibians. *The Quarterly Review of Biology* 1965; **40:** 1–66.
178 Caspari W. Physiologie der Rontgen und Radiumstrahlen. *Handbuch der norm. und pathol. Physiologie* 1926; **17:** 343–390.
179 Strelin GS. Rontgenologische Untersuchungen an Hydren II. Die histologischen Veranderungen im Korperbau vom Pelmatohydra oligactis unter der Wirkung der Rontgenstrahlen und ihre Bedeutung fur die Regeneration und Vermehrung. *Wilhelm Roux' Archiv fur Entwicklungsmechanik* 1929; **115:** 27–57.
180 Zawarzin AA. Rontgenologische Untersuchungen an Hydren I. Die Wirkung der Rontgenstrahlen auf die Vermehrung und Regeneration bei Pelmatrohydra oligactis. *Wilhelm Roux' Archiv fur Entwicklungsmechanik* 1929; **115:** 1–26.
181 Scheremetjewa EA, Brunst VV. Untersuchung des Einflusses von Rontgenstrahlen auf die Regeneration des Schwanzes bei den Kaulquappen von Pelobates fuscus. *Wilhelm Roux' Archiv fur Entwicklungsmechanik* 1933; **130:** 771–791.
182 Brunst VV, Scheremetjewa EA. Untersuchung des Einflusses von Rontgenstrahlen auf die Regeneration der Extremitaten beim Triton. *Wilhelm Roux' Archiv fur Entwicklungsmechanik* 1933; **128:** 181–215.
183 Gordon MB. The stimulative effect of roentgen rays upon the glands of internal secretion. *Endocrinology* 1930; **14:** 411–437.
184 Gordon MB. The stimulative effect of roentgen rays upon the glands of internal secretion. *Radiology* 1931; **17:** 1309–1311.
185 Freund L. Elements of General Radiotherapy (English translation). New York, Rebman Co.: New York, 1904.
186 Borak J. Theories on the effectiveness of roentgen therapy in inflammatory conditions. *Radiology* 1944; **42:** 249–254.
187 Chen W. The laboratory as business: Sir Almroth Wright's vaccine programme and the construction of penicillin. In: The Laboratory Revolution in Medicine. Cunningham, A. and Williams, P. (eds) Cambridge University Press: London, 1992.
188 Heidenhain L, Fried C. Rontgenstrahlen und Entzundung. *Klin. Wchnschr.* 1924; **3:** 1121–1122.
189 Fried C. Die Rontgenbehandlung der akuten Entzundungen. *Strahlentherapie* 1927; **26:** 484–506.
190 Hartley P. Effect of radiation on production of specific antibodies. *Journal of Experimental Pathology* 1924; **5:** 306–313.
191 Hektoen L. Influence of the X-ray on production of antibodies. *Journal of Infectious Diseases* 1915; **17:** 415–422.
192 Hektoen L. Further studies on the effects of roentgen ray on antibody-production. *Journal of Infectious Diseases* 1918; **22:** 28–33.
193 Westman A. Studies on influence of roentgen and radium rays on phagocytosis. *Acta Radiology* 1923; **2:** 57–69.

194 Castellino PG. L'azione dell'irradiazioni roentgen ed ultraviolette sul sistema reticulo, istiocitario cutaneo. *Arch. di radiol.* 1930; **6:** 681–689.
195 Tannenberg J, Bayer L. Der heilungavorgang von entzundlichen Veranderungen unter dem Einfluss von Rontgenstrahlen. *Strahlentherapie* 1933; **47:** 408–425.
196 Pendergrass EP, Hodes PJ. Roentgen irradiation in the treatment of inflammations. *American Journal of Roentgenology and Radium Therapy* 1941; **45:** 74–106.
197 Desjardins AU. Radiotherapy for inflammatory conditions. *Journal of the American Medical Association* 1931; **98:** 401–408.
198 Desjardins AU. The action of roentgen rays or radium on inflammatory processes. *Radiology* 1937; **29:** 436–445.
199 Desjardins AU. Dosage and method of roentgen therapy for inflammatory conditions. *Radiology* 1939; **32:** 699–707.
200 Desjardins AU. Radiotherapy for inflammatory conditions. *New England Journal of Medicine* 1939a; **221:** 801–809.
201 Desjardins AU. The action of roentgen rays on inflammatory conditions. *Radiology* 1942; **38:** 274–280.
202 Heidenhain L. Rontgenbestrahlung und Entzundung. *Strahlentherapie* 1926; **24:** 37–51.
203 Musser JH, Edsall DL. A study of metabolism in leukemia, under the influence of the X-ray. *Transactions of the Association of American Physicians* 1905; **20:** 294–323.
204 Edsall DL, Pemberton R. The use of X-rays in unresolved pneumonia. *American Journal of Medical Science* 1907; **133:** 286–297.
205 Quimby AJ, Quimby WA. Unresolved pneumonia: Successful treatment by roentgen ray. *New York Medical Journal* 1916; **103:** 681–683.
206 Mayou MS. A case of trachoma treated by X-rays. *Transactions of the Ophthamology Society of the United Kingdom* 1902; **22:** 95–95.
207 Mayou MS. The treatment of trachoma by X-rays. *Transactions of the Ophthamology Society of the United Kingdom* 1903; **23:** 10–23.
208 Coyle RR. Odds and ends of X-ray work, including some cases of carbuncle. *M. Electrol. and Radiology* 1906; **7:** 139–142.
209 Dunham K. Treatment of carbuncles by the roentgen ray. *American Journal of Roentgenology* 1916; **3:** 259–260.
210 Ross GG. X-ray treatment of carbuncle of face. *Annals of Surgery* 1917; **66:** 99–100.
211 Richards CE. Some of the less common uses of X-ray therapy. *Journal of Radiology* 1922; **3:** 271–273.
212 Lewis RW. A conservative treatment of carbuncles. *Annals of Surgery* 1923; **78:** 649–659.
213 Hodges FM. The roentgen ray in the treatment of carbuncles and other infections. *American Journal of Roentgenology* 1924; **11:** 442–445.
214 Hodges FM. Roentgen ray in treatment of local inflammations. Cellulitis and carbuncles. *Journal of the American Medical Association* 1925; **85:** 1292–1294.
215 Pordes F. Uber die Natur der Wirkung der Rontgenstrahlen, speziell uber das Verschwinden von Anurie nach Nierenbestrahlung. *Wien. Klin. Wchnschr.* 1923; **36:** 656–637.
216 Pordes F. Der Mechanismus der Rontgenwirkung: Ein Erklarungsversuch. *Fortschritte auf dem Gebiete Roentgenstrahlen* 1923–1924; **31:** 287–297.
217 Pordes F. Ueber Rontgenbehandlung entzundlicher Erkrankungen: Allgemeines und Spezielles. *Strahlentherapie* 1926; **24:** 73–86.
218 Pordes F. Die Verlaufsanderung akuter Entzundungen nach Rontgenbestrahklung. *Strahlentherapie* 1929; **33:** 147–151.
219 Holzknecht G. Roentgen treatment of spontaneous post-traumatic and postoperative coccus infections and suppurations. *American Journal of Roentgenology* 1926; **15:** 332–336.
220 Gerber I. Roentgen-ray treatment of superficial pyogenic infectioms. *Rhode Island Medical Journal* 1926; **9:** 33–38.
221 Fraenkel SR, Nissnjewitsch LM. Ueber die Rontgenbehandlung der chirirgischen entzundlichen Vorgange. *Strahlentherapie* 1926; **24:** 87–100.
222 Solomon S, Blondeau A. La roentgentherapie dans les affections inflammatoires. *J. de radiol. et d'electrol.* 1927; **11:** 465–469.
223 Carp L. Treatment of carbuncles. *Annals of Surgery* 1927; **86:** 702–706.
224 Light RU, Sosman MC. Treatment of carbuncles by the roentgen ray. *New England Journal of Medicine* 1930; **203:** 549–555.
225 King CO. Radiation therapy of carbuncles. *Southern Medical Journal* 1937; **30:** 903–906.
226 Krost GN. Unresolved pneumonia in choldren; treatment with roentgen ray. *American Journal of Diseases of Children* 1925; **30:** 57–71.
227 Torrey RG. Roentgentherapy in disorders of the respiratory tract, particularly those associated with enlargement or persistence of the thymus gland, and those associated with unresolved pneumonic exudates. *S. Clin. N. America* 1927; **7:** 221–235.
228 Heidenhain L. UeberKopfverletzungen durch stumpfe Gewalteinwirkung. *Munchen. Med. Wchnschr.* 1917; **64:** 600–602.
229 Kaess FW. Rontgenbestrahlung bei postoperativer Pneumonie. *Mitt. a. d. Grenzgeb. d. Med. u. Chir.* 1925; **38:** 509–515.
230 Fried C. Rontgentherapie der post-operative Pneumonie. *Klin. Wchnschr.* 1926; **5:** 15–18.
231 Gadjanski B. Ueber die Radiotherapie der akuten Entzundungen. *Serb. Arch. f. d. ges. Med.* 1927; **4:** 191.
232 Glas K. Rontgenbehandlung von Lungenstorungen nach Operationen. *Wien. klin. Wehnschr.* 1927; **40:** 1054–1055.
233 Holtz L. Ueber die Behandlung der Pneumonie mit Rontgenstrahlen. *Ztschr. f. klin. Med.* 1929; **109:** 698–713.
234 Merritt EA, McPeak EM. Roentgen irradiation in unresolved pneumonia. *American Journal of Roentgenology* 1930; **23:** 45–48.
235 McIntire FT, Smith JH. X-ray therapy in treatment of pneumonia. *Texas State Journal of Medicine* 1937; **33:** 422–426.
236 Powell EV. Roentgen therapy of lobar pneumonia. *Journal of the American Medical Association* 1938; **110:** 19–22.
237 Powell EV. Treatment if acute pneumonias with roentgen rays. *American Journal of Roentgenology and Radium Therapy* 1939; **41:** 404–414.

238 Stephenson S, Walsh D. Short note on the cure of trachoma by X-ray tube exposure and by high-frequency brush discharges. *Lancet* 1903; **1:** 237.
239 Bettremieux. Rayons X en therapeutiqueoculaire. *Clin. Opht.* 1903; **9:** 225–227.
240 Cassidy HF, Rayne FC. Trachome chronique tres ameliore par les rayons X. *Tr. Ann. d'electrobiol.* 1903; **6:** 617–619.
241 Geyser AC. The successful treatment of eighteen cases of granular lids by the X-ray and high-frequency vacuum electrodes. *American Therapist* 1903–1904; **12:** 41–43.
242 Geyser AC. The successful treatment of eighteen cases of granular lids by the X-ray and high-frequency vacuum electrodes. *Journal of Advanced Therapeutics* 1904; **22:** 299–302.
243 Pardo R. Due casi di tracoma trattati coi raggi di Rontgen. *Gazz. d. osp.* 1904; **25:** 459–460.
244 Horniker E, Romanin V. Ueber einen Hulfsapparat zur Behandlung des Trachoms mit Rontgenstrahlen. *Ztschr. f. Augenh.* 1905; **14:** 569–575.
245 Stargardt K. Ueber die Wirkung der Rontgenstrahlen auf Trachomfollikel. *Ztschr. f. Augenh., Ber.* 1905; **14:** 251–258.
246 Thielemann R. Zur Wirkungsweise des Radiumbestrahlung auf die trachomatose Bindehaut. *Ztschr. f. Augenh.* 1905; **14:** 559–569.
247 Newcomet WS. Radium and radioactive salts compared with X-rays. *Journal of Advanced Therapeutics* 1912; **30:** 72–77.
248 Jacqueau, Lemoine, Arcelin. Trachome et radiotherapie. *Lyon Med.* 1920; **139:** 869–870.
249 Rollet M, Bussy D. Conjunctive granuleuse grave: Radiotherapie amelioration. *Lyon med.* 1927; **129:** 458.
250 Cochard M. La radiotherapie du trachome. Thesis 128. Lyons, France, 1921.
251 Meldolesi G, Sabbadini D. Sulla terapia del con le radiazioni secondarie ottenute col metodo del Ghilarducci. *Radiol. med.* 1923; **10:** 222.
252 Meldolesi G. Sulla terapia del tracoma con le radiazioni secondarie (secondo il metodo del Ghilarducci). *Actinoterapia* 1924; **4:** 97–162.
253 Lane LA. Radium in ophthalmology, with special reference to its use in benign affections. *Journal of the American Medical Association* 1924; **83:** 1838–1845.
254 Sabbadini D. Le modificazioni cliniche ed anatomo-pathologiche della conguintiva tracomatosa in seguito all'azione delle radiazioni secondarie. *Arch. di Radio.* 1926; **2:** 584–688.
255 Kelly JF. The X-ray as an aid in treatment of gas gangrene. *Radiology* 1933; **20:** 296–304.
256 Kelly JF. Present status of the x-ray as an aid in treatment of gas gangrene. *Radiology* 1936; **26:** 41–44.
257 Hubeny MJ, McNattin RF. X-ray therapy: Evaluation of its use in treatment of diseases of non-malignant neoplastic nature. *Urol. and Cutan. Rev.* **42:** 436–441.
258 Kelly JF, Dowell DA. Present status of X-rays as aid in treatment of gas gangrene. *J.A.M.A.* 1936; **107:** 1114–1118.
259 Kelly JF, Dowell DA. Twelve-year review of X-ray therapy of gas gangrene. *Radiology* 1941; **37:** 421–439.
260 Altemeier WA, Jones HC. Experimental peritonitis: Its prevention by X-ray irradiation. *Journal of the American Medical Association* 1940; **114:** 27–29.
261 Bates MT. Gas gangrene; review of 32 cases with special reference to the use of serum, both prophylactic and therapeutic. *Annals of Surgery* 1937; **105:** 257–264.
262 Faust JJ. Radiation therapy of gas bacillus infecton. *Illinois Medical Journal* 1934; **66:** 547–551.
263 Faust JJ. Report on x-ray treatments in gas gangrene cases. *Radiology* 1934; **22:** 105–106.
264 Kelly JF, Dowell DA, Russum BC and Colien FE. Practical and experimental aspects of roentgen treatment of *Bacillus welchii* (gas gangrene) and other gas-forming infections. *Radiology* 1938; **31:** 608–619.
265 Kelly JF, Dowell DA. Roentgen Treatment of Infections. The Yearbook Publishers, Chicago, IL. 432 pp. 1942.
266 Dowdy AH, Sewell RL. Roentgen radiation in experimental *Clostridium welchii* ingection (gas gangrene) in dogs: Preliminary report. *Radiology* 1941; **37:** 440–442.
267 Merritt EA, Den AJ, Wilcox UV. The effects of radiation therapy in Clostridium infection in sheep. *Radiology* 1944; **43:** 325–332.
268 Cantril ST, Buschke F. Roentgen therapy in gas bacillus infection. *Radiology* 1944; **43:** 333–345.
269 Glenn Jr JC. Studies on the effects of X-rays of phagocytic indices of healthy rabbits. *J. Immunology* 1946; **52:** 65–69.
270 Glenn Jr JC. Further studies on the influence of X-rays of phagocytic indices of healthy rabbits. *J. Immunology* 1946a; **53:** 95–100.
271 Bloom W, Jacobson LO. Some hematologic effects of irradiation. *Blood* 1948; **3:** 586–592.

TOXICOLOGICAL SCIENCES **94(2)**, 368–378 (2006)
doi:10.1093/toxsci/kfl098
Advance Access publication September 1, 2006

Hormesis Outperforms Threshold Model in National Cancer Institute Antitumor Drug Screening Database

Edward J. Calabrese,*,[1] John W. Staudenmayer,† Edward J. Stanek III,‡ and George R. Hoffmann§

*Department of Public Health, Environmental Health Sciences Program, Morrill I, N344, University of Massachusetts, Amherst, Massachusetts 01003; †Department of Mathematics and Statistics, Lederle Graduate Research Tower, University of Massachusetts, Amherst, Massachusetts 01003; ‡Department of Public Health, Biostatistics and Epidemiology Program, School of Public Health and Health Sciences, Arnold House, University of Massachusetts, Amherst, Massachusetts 01003; and §Department of Biology, College of the Holy Cross, O'Neil Hall, 107, Worcester, Massachusetts 01610-2395

Received March 31, 2006; accepted August 30, 2006

Which dose-response model best explains low-dose responses is a critical issue in toxicology, pharmacology, and risk assessment. The present paper utilized the U.S. National Cancer Institute yeast screening database that contains 56,914 dose-response studies representing the replicated effects of 2189 chemically diverse possible antitumor drugs on cell proliferation in 13 different yeast strains. Multiple evaluation methods indicated that the observed data are inconsistent with the threshold model while supporting the hormetic model. Hormetic response patterns were observed approximately four times more often than would be expected by chance alone. The data call for the rejection of the threshold model for low-dose prediction, and they support the hormetic model as the default model for scientific interpretation of low-dose toxicological responses.

Key Words: **hormesis; threshold; dose-response; yeast; NCI; U-shaped; J-shaped; bell-shaped; risk assessment; carcinogens; chemotherapeutics; cell proliferation; *Saccharomyces*.**

The threshold dose-response model has long been recognized as the dominant dose-response model in the biological sciences, including pharmacology and toxicology (Clark, 1926, 1933, 1937). The threshold model dominates discussion in the leading pharmacological (Hardman and Limbird, 2001) and toxicological textbooks (Eaton and Klaassen, 2001; Hayes, 2001), development of study designs that drive hazard assessment procedures for pharmaceutical and chemical agents, and risk assessment processes used by regulatory and public health agencies worldwide. Despite this fact and a history of broad acceptance in many biological disciplines, the assumption that the threshold model should be used has been recently challenged. An alternative model, the hormesis model, has been proposed based on evidence of its generalizability by biological system, endpoint measured, chemical class tested (Calabrese, 2004, 2005; Calabrese and Baldwin, 2001a,b,c, 2003; Calabrese and Blain, 2005; Calabrese *et al.*, 1999), and high frequency in the toxicological literature (Calabrese and Baldwin, 2001b). Using *a priori* entry and evaluative criteria, Calabrese and Baldwin (2003) reported that the hormesis model far outperformed the threshold model in a toxicological assessment using approximately 800 dose-response relationships that were broadly representative of commonly employed biological models, endpoints, and chemical agents. The present paper extends these findings by a systematic in-depth analysis of 56,914 dose-response studies in yeast. The analyses demonstrate that the hormesis dose-response model strongly outperforms the threshold model when applied to the extensive and highly standardized National Cancer Institute (NCI) tumor drug screening database using yeast as the test organism.

[1] To whom correspondence should be addressed. Fax: 413-545-4692, E-mail: edwardc@schoolph.umass.edu.

MATERIALS AND METHODS

This study utilized data from the NCI yeast anticancer drug screen, described in detail by Holbeck (2004) and at the NCI Web site (http://www.dtp.nci.nih.gov/yacds/index.html). Briefly, data from stage 2, which contains the most promising compounds based on preliminary testing, were selected for evaluation; the agents were tested at five concentrations (1.2, 3.7, 11, 33, and 100μM) in 13 yeast strains. The yeast comprises a panel of *Saccharomyces cerevisiae* strains altered in DNA damage repair or cell cycle control genes, along with the wild-type (wt) strain without such genetic alterations. The NCI Web site contains a description of the genotype of each yeast strain used. The responses reported are derived from the fraction of growth of the yeast strain exposed to the compound relative to the growth of the same yeast strain treated with solvent (i.e., DMSO) control. Yeast cells in the exponential phase of growth were inoculated into synthetic complete medium containing 2% glucose and the test chemical. The starting cell density was 10^4 cells per well containing 200 μl of medium. (Julian Simon, personal communication).

This study, like any analysis of preexisting data, has limitations based on the data that are available. Factors limiting the range of questions that we could analyze were the fact that the NCI database provides the average of two responses and the difference between them but not original optical densities or

raw data permitting us to match chemicals to plates. Possible sources of variation were differences among columns or rows in 96-well plates and the lack of randomization of treatments within plates, plates within stacks, and stacks within the incubator. In the main, however, the large size, apparent quality control, and internal consistency of the NCI database minimize risks associated with experimental protocols, and we restricted our analysis to questions where experimental variation can be properly analyzed. Moreover, in each instance where such factors may be relevant to our analysis, we specifically point it out. Whenever a choice of assumptions was possible, we made assumptions so as to ensure that our analysis was conservative, in the sense that the bias, if any, would reduce the chance of observing hormesis.

Replication Procedure

Each chemical was tested four times at the same five concentrations in each of the 13 yeast strains (Table 1). Ninety-six–well plates were used, with 80 chemicals being tested at the same concentration (1.2, 3.7, 10, 33, or 100μM) on one plate, leaving the 16 peripheral wells for controls. Each concentration for that drug was incubated over the same 12-h period on a different plate such that there were five plates run on the same chemical at the same time. Of the 16 control wells, four were allocated to unexposed controls, eight to solvent controls, and four to controls using cycloheximide. All strains were cycloheximide sensitive, and the assay was deemed invalid if there was growth in the presence of cycloheximide. The order of the plates in the growth chamber with respect to chemical tested was systematic and not randomly allocated. The relative position of the controls on each plate was constant. The variability in response was designed to be maximized by using a different source of chemical on a different day and different yeast cultures in each different test (Julian Simon, personal communication). Slightly greater evaporation from the peripheral wells containing the controls may have caused control values to be several percentage points above a normal background. This potential bias was systematic for all plates and, if present, would introduce a negative bias with respect to discerning possible stimulatory responses (Faessel *et al.*, 1999). To ensure that our analysis was conservative with respect to the detection of hormesis, no correction was made for this factor. Factors other than evaporation from peripheral wells may contribute to variation in data from 96-well plates. An analysis of such factors by Faessel *et al.* (1999) found that differences associated with positions of plates in stacks and of stacks in an incubator tended to be smaller than differences between the middle and edges of plates. Moreover, the importance of these sources of variation is minimal for our purposes as each plate contained its own set of controls.

The response data consisted of a ratio of the optical density (OD) of the response well for the treatment divided by the mean of the OD readings of the eight solvent control wells for each concentration. OD readings were at 600 nm. This process was repeated on the second day, and the ratios from the 2 days were averaged. We refer to the average response as the replication response. Two replication responses were produced for each concentration and for each strain/experiment. Data on the NCI Web site provide the average of the two response values and the difference between the two values, but the original OD values are not available. Data were also not available to match chemicals to plates to take advantage of plate effects. It is known that even if the control and treatment responses have the same mean, then the mean of their ratio will be greater than 100%, and the mean of the ratio approaches 100% as the variance of controls gets smaller (Casella and Berger, 2002). This creates a slight bias in favor of a hormetic model, but since the design used eight controls for each response, the effect is slight. Assuming a lognormal model, if the coefficient of variation is 10%, then the mean of the ratio of a control and treatment response with the same mean is 100.1%.

Evaluation Strategy

Since the goal of this research is to evaluate whether there is nonrandom biological activity as measured by cell proliferation below the toxicological threshold, it is necessary to evaluate individual dose-response relationships. In the five-concentration protocol of NCI, the dose-response study should ideally have at least one concentration in the toxic (i.e., above threshold) domain, a concentration with a response that approximates the control response (i.e., the so-called no observed effect level [NOEL] or the highest dose that does not differ in a significant manner from the control) and several lower concentrations that would be evaluated for biological activity below this NOEL or threshold-like value. The NCI yeast database is unique and useful for this purpose but not ideal, in that some experiments show toxicity but insufficient doses below the toxicologic threshold. Our *a priori* entry criterion required at least one measurement at a concentration below that that was used to estimate the

TABLE 1
NCI Yeast Screening Replication Methodology

	Replication 1	Replication 2
Day 1	• Five concentrations, each single concentration on a different 96-well plate • Eight solvent control wells per 96-well plate • The single treatment value for each concentration is divided by the average of the eight solvent control values for each well plate	
Day 2	• Five concentrations, each single concentration on a different 96-well plate • Eight solvent control wells per 96-well plate • The single treatment value for each concentration is divided by the average of the eight solvent control values for each well plate The average of day 1 and 2 values creates what is designated as replication 1	
Days 3 and 4		• The procedures on days 3 and 4 were identical to those on days 1 and 2 The average of day 3 and 4 values creates what is designated as replication 2

toxicological threshold. A response in the toxic zone at the highest concentration was considered desirable but not essential since lack of toxicity at the highest concentration in a screening bioassay may simply reflect an inadequate concentration range.

A two-stage approach was used to assess possible below–toxic threshold biological activity. The first involved estimation of the toxic threshold, while the second assessed the distribution of responses at concentrations lower than the estimated threshold concentration.

Threshold Estimation Strategies

Two strategies were used to estimate the toxicological threshold: the Benchmark Dose (BMD) and the NOEL. Since the results were similar, we only present the BMD results here. The Benchmark Dose 10 (BMD(10)) is the dose at which the response is estimated to have decreased 10% below control value (Crump, 1984). The BMD(10) was selected since 10% bounded the variability of the most variable yeast strain (i.e., response SD ranged from 3.0 to 7.5% for the 13 strains). Use of a BMD based on less than 10% (e.g., 2.5–7.5%) would yield progressively higher estimates of the frequency of hormesis for all parameters estimated (see ''Results'' section, Fig. 3 and ''Supplementary Data'' section), suggesting that the current approach (i.e., BMD(10) method) would lead to an overall underestimation of hormesis frequency. An identical analysis using the BMD(2.5) for all parameters reported for the BMD(10) analysis is given within the ''Supplementary Data'' section. Our method of estimating a BMD(10) is explained below (Fig. 1). Since our goal is to classify toxicity, we did not calculate the lower bound of a confidence interval for that dose. The BMD(10) approximates the control but probably entails a low degree of toxicity. It corresponds to a dose that is slightly higher than the toxicological threshold. This suggests that a dose immediately below and very close to the BMD(10) may itself be within the toxic zone (i.e., a slightly higher concentration than the actual toxicological threshold). This would become less likely with increasing distance between the BMD(10) and the concentration below the BMD(10). For example, for agents with a BMD(10) near 3.7μM, the 1.2μM dose would be close to the toxicity threshold. In contrast, for agents with a BMD(10) approaching 100μM, the 1.2μM dose would be nearly two orders of magnitude below the toxicity threshold.

A BMD(10) was calculated for each of the 28,457 (2189 chemicals and 13 strains) dose-response experiments using the average of two replications as response. The BMD(10) was estimated through the following procedure.

1. The largest concentration with an average response below 90% is identified. Let this concentration be C_{below}, and let the associated response be R_{below}.

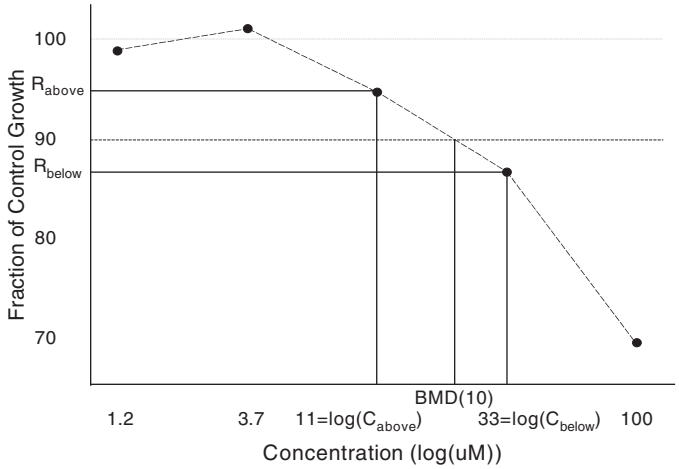


FIG. 1. General scheme used for the derivation of the BMD(10) used in the present paper.

2. If the average response at the next smallest concentration is at least 90%, then let this concentration be C_{above}, and let the associated response be R_{above}. The BMD(10) is estimated by linear interpolation on the log concentration scale:

$$\text{BMD}(10) = \exp[\log(C_{\text{above}}) + (0.90 - R_{\text{above}})(\log(C_{\text{below}}) - \log(C_{\text{above}}))/(R_{\text{below}} - R_{\text{above}})].$$

3. If the average response at the next lowest concentration below C_{below} is less than 90%, then let this concentration be C_{below} with response R_{below} and return to step 2.

If all responses for a particular chemical-strain experiment were above or below 90% then ''greater than 100μM'' or ''less than 1.2μM,'' respectively, was reported.

In contrast to a linear or nonlinear regression approach to calculating BMD, the procedure described above is ''local'' in the sense that the BMD(10) is only calculated using the responses at concentrations that are adjacent to the BMD(10). Further, when this approach is used, the two concentration-response pairs that surround the BMD(10) are chosen using only concentrations above the BMD(10) and one concentration below the BMD(10). Responses at concentrations used to estimate below-threshold responses were not used in the estimation of the BMD.

Assessing the Distribution of Responses below the Toxic Threshold

We used two approaches to assess evidence of stimulated biological activity at concentrations below the estimated threshold of toxic response. The first is a pattern analysis that counts how often both replicates in each experiment were above and below (or equal to) 100% and compares those counts to expected values, assuming a threshold model. The second approach compares the frequency of responses at various levels above and below 100%.

Pattern analysis. This approach categorized each concentration-response replication for all chemical-strain combinations and analyzed the patterns of responses at each concentration that were above and below (or equal to) 100%. In this approach, each chemical-strain repetition can express one of two different responses: H (response above 100%) and L (response less than or equal to 100%). A simple ''fair coin'' model was posited for the responses below the BMD(10) where each single replication would have a 50% chance of being above or below (or equal to) 100% and there is statistical independence across responses. This model assumes that the responses at concentrations below the BMD(10) have a median of 100%. It therefore describes a threshold model with minimal distributional assumptions. The fidelity between the observed data and this hypothesized model was tested.

Comparison of above/below-control values in the subtoxic zone of the dose-response. The threshold dose-response model predicts that responses below the toxicological threshold should randomly vary on either side (i.e., above or below) of control group values (100% response). The hormetic model predicts that there should be a nonrandom stimulatory response (i.e., responses greater than 100%) below the toxic threshold. In order to test which model best accounts for the observed data, above-control (> 100% response) to below-control (≤ 100% response) ratios were detailed for all yeast strains in the various BMD(10) classifications. The nonrandom distribution predicted by the hormesis model would be reflected in a greater frequency of responses above than below the control and in the magnitude of the deviation from the control.

Comparisons were made to responses above 100, 105, 110, 115, and 120% and then to below-the-appropriate-control group response using the formula:

$$\frac{\text{Control}}{\text{Above Response Level}} = \text{Below Response Comparison}$$

$$\left(\text{e.g.,}\frac{100\%}{120\%} = 83.33\%\right).$$

This methodology is based on the observation that the 100% control value is 83.33 of 120%. This model indicates that a 20% increase in response over the

100% value is equivalent to a 16.7% decrease from the control. Using this approach, ratios of counts comparing the following levels were made > 100%/≤ 100%, > 105%/≤ 95.24%, > 110%/≤ 90.91%, > 115%/≤ 86.96%, and > 120%/≤ 83.33%. This methodology was used to take into account the possibility of an unrestricted stimulatory response while the maximum inhibitory response was fixed at zero.

RESULTS

Figure 2 describes the concentration-response relationships of the 13 yeast strains to the 2189 chemical agents tested. While there was little change on average from the control response at the lowest concentration (1.2μM), indications of average toxicity start to become evident at 3.7μM, progressing in dose-dependent fashion over the next three concentrations (11, 33, and 100μM). Table 2 shows the number of chemicals with BMD(10) values within each of six BMD(10) classification ranges for each of the 13 strains. The more toxic chemicals are included in the low BMD(10) range (e.g., < 1.2μM), while the chemicals with the lowest toxic potential comprise the highest BMD(10) categories. The data indicate that the wild type and SPY50780 yeast strains had the lowest number (139/2189 and 143/2189) of concentration-responses with BMD(10)s < 1.2μM, indicating that they were the least susceptible strains, a perspective that is supported by plotting of the overall data in Figure 2. In contrast, strains carrying *rad50*, *rad50EPP+*, *rad18*, *rad52*, and *sgs1* were the most susceptible (Fig. 2, Table 2).

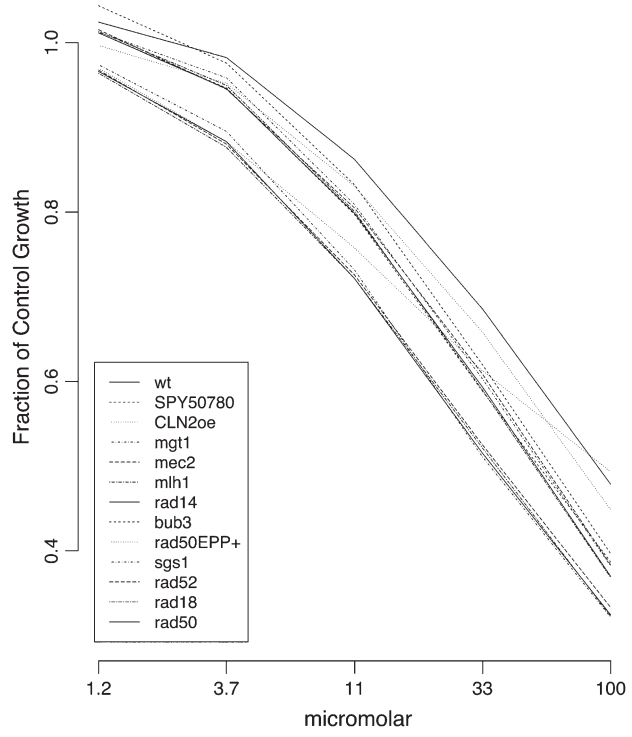


FIG. 2. Average concentration-response of 2189 chemicals on the 13 yeast strains.

Table 3 is a summary of the below-BMD(10) mean responses and SD for each of the 13 yeast strains. The numbers of chemical-concentration relationships satisfying *a priori* entry criteria using the BMD(10) methodology are shown in Table 2. Of 28,457 concentration-responses to the 2189 chemicals, 16.7% (4763) gave no evidence of toxicity, having BMD(10) values ≥ 100μM. Assessments were performed on this subgroup of responses under the assumption that at higher concentrations, a toxic response would have occurred. When the BMD(10) is less than 3.7μM, there is no concentration that can be assessed for biological activity below the BMD(10), thereby not satisfying our entry criteria. There were 7558 such responses (3798 with BMD(10) < 1.2μM and 3760 with 1.2 ≤ BMD(10) < 3.7), accounting for 26.6% of the total responses. Therefore, 73.4% of the total dose-responses were evaluated. Similar findings were observed with the NOEL methodology (data not shown).

BMD(10) Response Evaluation

We averaged responses at concentrations when the concentration was below the BMD(10) for each strain. This resulted in averages of between 196 and 572 responses, with the mean and SD (Table 3). The below-BMD(10) mean values (Table 3) are generally consistent across each of the 13 yeast strains within a specific BMD(10) classification as well as across BMD(10) classifications. However, the mean values are modestly lower in the 3.7 ≤ BMD(10) ≤ 11μM group than in the other groups ($p < 0.001$), which do not differ significantly from each other. These trends are consistent with median values as well. Consequently, all 13 strains with each of the four BMD(10) classifications had average responses significantly greater than the control ($p < 0.001$ for each of the four columns in Table 3). These findings are consistent with a nonrandom distribution of responses in the direction of the hormetic dose-response. Findings with the NOEL methodology were similar (data not shown) except that the responses were usually several percentage points higher per strain than for the BMD(10) methodology. Similarly, if a smaller BMD(2.5–7.5) cutoff point were used instead of the BMD(10), the mean responses become progressively higher as the BMD value decreases (Fig. 3 and supplementary data, Table S1).

According to the threshold dose-response model, the distribution of responses below the estimated threshold (e.g., BMD(10)) should approach a 1:1 ratio for above- and below-control values. This was assessed for each BMD(10) classification group for responses > 100/≤ 100%, > 105/≤ 95.24%, > 110/≤ 90.91%, > 115/≤ 86.96%, and > 120/≤ 83.33%. Alternatively, one could use a different model and assume the equivalency of a symmetrical response (e.g., > 120%/≤ 80% rather than > 120%/≤ 83.3%), but this paper used the prior and more conservative approach. This approach was

TABLE 2
Number of Chemicals Tested Per Yeast Strain Classified on the Basis of BMD(10)

Yeast strains	BMD(10) < 1.2	1.2 ≤ BMD(10) < 3.7	3.7 ≤ BMD(10) < 11	11 ≤ BMD(10) < 33	33 ≤ BMD(10) < 100	BMD(10) ≥ 100	Totals
Wild type	139	249	365	443	462	531	2189
SPY50780	143	253	411	536	430	416	2189
CLN2oe	236	246	379	456	428	444	2189
mgt1	218	256	408	551	380	376	2189
mec2	259	269	399	462	446	354	2189
mlh1	227	265	417	572	363	345	2189
rad14	227	285	423	550	361	343	2189
bub3	244	274	453	488	363	367	2189
rad50EPP+	414	367	405	330	196	477	2189
sgs1	435	291	454	498	241	270	2189
rad52	424	334	398	452	289	292	2189
rad18	419	321	403	464	302	280	2189
rad50	413	350	411	464	283	268	2189
Totals	3798	3760	5326	6266	4544	4763	28457

selected in order to make hormesis more difficult to detect. Table 4, A–D, indicates the distribution of responses below the BMD(10) for the chemicals in the various BMD(10) ranges. For example, the distribution of responses in the 33 ≤ BMD(10) < 100μM classification range (Table 4, C) for the 13 yeast strains is nonrandomly distributed in the direction of a hormetic response regardless of the degree of variability in the data. The findings are inconsistent with the threshold model, which predicts a ratio closely approximating 1:1. A comparison of the respective BMD(10) classification groups reveals that in the large variation comparisons (i.e., > 110%/≤ 90.91%), the proportion of stimulatory responses exceeds those on the "below" side by threefold to over 10-fold. The comparisons are generally similar among the 11 ≤ BMD(10) < 33μM, 33 ≤ BMD(10) < 100μM, and BMD(10) ≥ 100 classifications. While the 3.7 ≤ BMD(10) < 11μM classification (Table 4, A) also shows an excess of above-control values, the magnitude of the above/below differential is notably less. The most likely explanation for the reduced response in the 3.7 ≤ BMD(10) < 11μM classification is that responses at the 1.2μM concentration in these experiments may have displayed toxicity for some of the chemicals since 1.2μM is very close to

TABLE 3
Below-BMD(10) Mean Responses (%) by Yeast Strain and BMD(10) Grouping[a]

	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Yeast strains	3.7 ≤ BMD(10) < 11		11 ≤ BMD(10) < 33		33 ≤ BMD(10) < 100		BMD(10) ≥ 100	
Wild type	102.6	(12.7)	107.2	(15.1)	105.8	(12.9)	105.1	(11.1)
SPY50780	106.1	(13.6)	108.3	(15.0)	108.8	(14.7)	105.5	(11.8)
CLN2oe	101.7	(10.3)	103.7	(11.4)	104.6	(10.8)	104.8	(10.2)
mgt1	102.7	(13.1)	106.6	(14.6)	106.5	(13.9)	105.0	(11.2)
mec2	105.4	(16.8)	107.3	(16.4)	105.8	(14.4)	106.0	(14.7)
mlh1	103.8	(15.2)	107.2	(15.5)	105.9	(14.7)	104.5	(11.0)
rad14	103.9	(12.9)	107.4	(13.7)	106.5	(13.3)	106.4	(12.5)
bub3	104.8	(13.0)	106.0	(12.5)	106.8	(12.2)	106.0	(10.6)
rad50EPP+	102.2	(10.5)	105.3	(16.1)	106.3	(14.4)	107.7	(15.4)
sgs1	103.3	(11.0)	106.7	(14.4)	106.8	(14.9)	104.8	(11.4)
rad52	103.6	(12.6)	106.8	(15.3)	105.9	(13.6)	104.0	(10.8)
rad18	103.9	(12.5)	106.2	(14.2)	106.5	(14.2)	106.6	(12.2)
rad50	102.9	(12.7)	105.5	(13.8)	104.4	(14.9)	104.7	(11.4)
Overall	103.6	(13.0)	106.6	(14.5)	106.2	(13.7)	105.5	(12.1)

[a]The number of concentration-responses for each mean value is given in Table 2. The average number of concentration-responses on which a single mean value is based is 402 (196–572 range). The 3.7 ≤ BMD(10) < 11 column is based on responses at 1.2μM, the 11 ≤ BMD(10) < 33 column is based on responses at 1.2 and 3.7μM, the 33 ≤ BMD(10) < 100 column is based on responses at 1.2, 3.7, and 11μM, and the BMD(10)≥100 column is based on responses at 1.2, 3.7, 11, and 33μM.

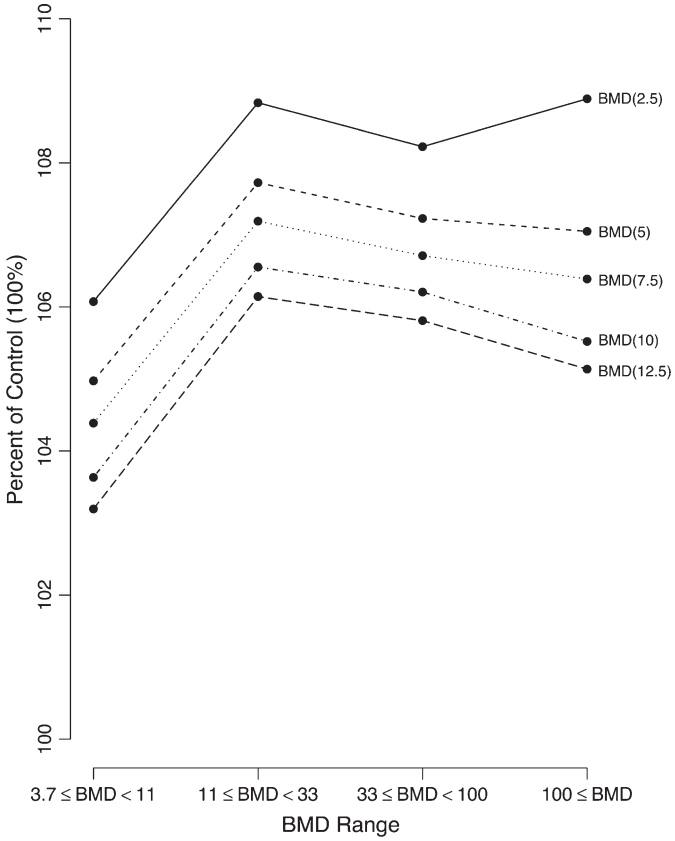


FIG. 3. BMD cutoff point influence on mean yeast growth for 13 yeast strains. (Data table for all yeast strains is found in supplementary data, Table S1).

the BMD(10) value. Figure 4 provides a simplifying summary of the information in Table 4. The comparisons in Table 4 were also used to compute weighted average estimates of overall ratios of above-control responses to below-control responses. For experiments in the $3.7 \leq \text{BMD}(10) < 11\mu\text{M}$ classification range, above-control responses were seen 1.96 times as often as below-control responses. For the $11 \leq \text{BMD}(10) < 33\mu\text{M}$, $33 \leq \text{BMD}(10) < 100\mu\text{M}$, and $\text{BMD}(10) \geq 100\mu\text{M}$ classifications, above-control responses were seen 4.64, 4.22, and 6.94 times as often as below-control responses, respectively. A weighted average calculated across all four BMD(10) classifications revealed that the above-control responses predicted by the hormesis model are 4.39 times as frequent as below-control responses. A similar assessment was performed using the BMD(2.5) (see supplementary data—Table S2A–D and Figure S1) with findings consistent with the BMD(10) analysis but even more supportive of the hormetic model.

Pattern Analysis

Patterns of response below the BMD(10) were compared for six log-spaced ranges of BMD(10)'s. Figure 5 presents these comparisons, based on observed responses that were not used in the calculation of the BMD(10), along with the expected counts under the fair coin threshold model. For instance, for BMD(10)s in the range $3.7 \leq \text{BMD} < 6.3$, the responses at 3.7μM and 11μM were used to calculate the BMD(10)s, and the figure summarizes the pattern of three responses for the replicates at 1.2μM. For the $11 \leq \text{BMD} < 33\mu\text{M}$ range, there were five possible patterns for the four replicates at 1.2μM: 4H 0L, 3H 1L, 2H 2L, 1H 3L, and 0H 4L. Under the fair coin threshold model, we would expect the fractions of responses that fit those patterns to be 1/16, 4/16, 6/16, 4/16, and 1/16, respectively. Similar procedures were used with BMD(10) values ranging from 33 to 100μM.

Strikingly, for each BMD(10) category, the observed responses markedly skew toward patterns that have more "H" (> 100%) responses, and the skewness increases as the BMD(10) increases and toxicity decreases ($p < 0.0001$). Considering the all-H patterns (left most pattern in each panel), the observed patterns are 1.3, 1.7, 5.3, 6.8, 20.7, and 25.1 times more frequent than the expected counts as the BMD(10) increases from the lowest range (3.7–6.3μM) up to the highest (57–100μM), where 57μM is halfway between 33 and 100μM on the log scale. Further, there is a strong general pattern with the BMD(10) ranges with counts tending to decrease monotonically as the number of Ls in the pattern increases. An assessment using the BMD(2.5) reveals similar findings to the BMD(10) (Fig. 5) but even more supportive of the hormetic model (supplementary data, Figure S2).

Figure 6 plots the fraction of H responses for each strain at the lowest concentration (1.2μM) as a function of the log distance below the BMD(10). As the distance of the 1.2μM concentration below the BMD(10) increases, the frequency of having both replicated responses at this first concentration (1.2μM) being greater than 100% increases to almost 60%, while only 25% would have been expected by chance assuming a threshold model. The figure shows that as the 1.2μM concentration reaches a value of ≤ 1/4th of the BMD(10), the probability of responses consistent with the hormesis model become far more common than chance for all strains; for the more highly toxic agents in the lowest BMD(10) category, the response at 1.2μM is often below control values. A similar assessment was performed using the BMD(2.5). It revealed similar findings to the BMD(10) which were even more supportive of the hormetic model (supplementary data, Figure S3).

DISCUSSION

The data indicate that responses to concentrations below the toxicological threshold for each of the 13 yeast strains tested with many hundreds of chemically diverse agents are non-randomly distributed with respect to the control. A variety of complementary methodological evaluations (Tables 3 and 4, Figs. 4–6) support the same interpretation. These findings indicate that the threshold dose-response model inadequately accounts for biological activity below the threshold. However, the results are consistent with predictions of the hormesis

TABLE 4

Evaluation of Below–Concentration Threshold Responses Based on Threshold Dose-Response Model Predictions. The Table Summarizes the Distribution of Responses Below the BMD(10) Level, Comparing the Frequencies of Levels Above and Below the 100% Control Value. The Above and Below Cutoffs that We Consider Are Comparison 1, > 100%/≤ 100%; Comparison 2, > 105%/≤ 95.2%; Comparison 3, > 110%/≤ 90.9%; Comparison 4, > 115%/≤ 87.0%; and Comparison 5, > 120%/≤ 83.3%. The Threshold Model Predicts a Ratio Closely Approximating 1:1 in Each of the Five Comparisons

	Comparison 1 (> 100%/≤ 100%)	Comparison 2 (> 105%/≤ 95.2%)	Comparison 3 (> 110%/≤ 90.9%)	Comparison 4 (> 115%/≤ 87.0%)	Comparison 5 (> 120%/≤ 83.3%)
(A) 3.7 ≤ BMD(10) < 11μM[a]					
Observed above/below–control group ratios	0.995/1	1.70/1	3.08/1	4.78/1	6.13/1
Number of responses above the designated percentile	2656 (49.9%)	1282 (24.1%)	850 (16.0%)	627 (11.8%)	472 (8.9%)
Number of responses below the designated percentile	2670 (50.1%)	754 (14.2%)	276 (5.2%)	131 (2.5%)	77 (1.4%)
(B) 11 ≤ BMD(10) ≤ 33μM[b]					
Observed above/below–control group ratios	1.75/1	3.29/1	6.81/1	11.67/1	15.77/1
Number of responses above the designated percentile	7930 (63.3%)	4072 (32.5%)	2866 (22.9%)	2182 (17.4%)	1687 (13.5%)
Number of responses below the designated percentile	4602 (36.7%)	1239 (9.9%)	421 (3.3%)	187 (1.5%)	107 (0.8%)
(C) 33 ≤ BMD(10) < 100μM[c]					
Observed above/below–control group ratio	2.03/1	3.78/1	6.58/1	9.85/1	11.01/1
Number of responses above the designated percentile	9133 (67.0%)	4217 (30.9%)	2755 (20.2%)	2099 (15.4%)	1596 (11.7%)
Number of responses below the designated percentile	4499 (33.0%)	1116 (8.2%)	419 (3.1%)	213 (1.6%)	145 (1.1%)
(D) BMD(10) ≥ 100μM[d]					
Observed above/below–control group ratio	1.96/1	4.60/1	40.79/1	Ratio can not be calculated	Ratio can not be calculated
Number of responses above the designated percentile	12618 (66.2%)	4961 (26.0%)	3141 (16.5%)	2399 (12.6%)	1849 (9.7%)
Number of responses below the designated percentile	6434 (33.8%)	1078 (5.7%)	77 (0.4%)	0 (0.0%)	0 (0.0%)

[a]Total of 5326 responses at 1.2μM.
[b]Total of 12,532 responses; 6266 responses each at 1.2 and 3.7μM.
[c]Total of 13,632 responses; 4544 responses each at 1.2, 3.7, and 11μM.
[d]Total of 19,052 responses; 4763 responses each at 1.2, 3.7, 11, and 33μM.

dose-response model. The conclusions based on this large database in yeast are similar to those made in earlier reports of Calabrese and Baldwin (2001b, 2003) using toxicological data representative of a broad range of biological models, endpoints, and chemical agents. The evidence of hormesis in yeast, along with the previous studies, is biologically significant and has potentially important implications for essentially all drug and chemical hazard assessment studies and risk assessments worldwide.

The similar overall response patterns below the toxic zone in all 13 yeast strains for the large number of chemicals in this extensively evaluated public database is a novel finding. Previous publications (see Holbeck, 2004, for a review) with this database have focused on the nature of the above-threshold (rather than below-threshold) responses and their underlying toxicological mechanisms since the goal of the NCI has been principally oriented toward identifying possible antitumor drugs rather than assessing the nature of the dose-response in the low-dose zone.

The pattern analysis assessment indicated that the total number of concentration-responses below the BMD(10) are skewed strongly in the direction predicted by the hormesis model (Fig. 5). These findings are consistent with the earlier reports of Calabrese and Baldwin (2001b, 2003) indicating a similar relationship for data derived from the toxicological literature.

The quantitative nature of the dose-response in the hormetic zone in the present study is also consistent with findings reported for hormesis with other biological models, endpoints, and chemical agents. That is, the hormesis response is usually modest, with the maximum response typically being only 30–60% greater than the controls (Calabrese and Blain, 2005). For example, the data for the wild-type yeast strain in the 11 ≤ BMD(10) < 33μM and 33 ≤ BMD(10) < 100μM classifications indicate that the proportion of responses exceeding 120% was 26.3 and 16.5%, respectively. In the case of strain SP47080, the respective proportion of responses >120% were 26.2 and 23.7%, respectively.

The present findings illustrate the importance of study design in the assessment of hormesis. A comparison of the lowest concentration (1.2μM) to the BMD(10) revealed that the chance of a stimulatory response becomes greater as the difference between 1.2μM and the BMD(10) increases. Approximately 60% of the time, both replicate responses at 1.2μM exceeded control values for BMD(10) values between 50 and 100μM, compared to only 16% for BMD(10) values between 1.2 and 3.7μM. These findings are consistent with the

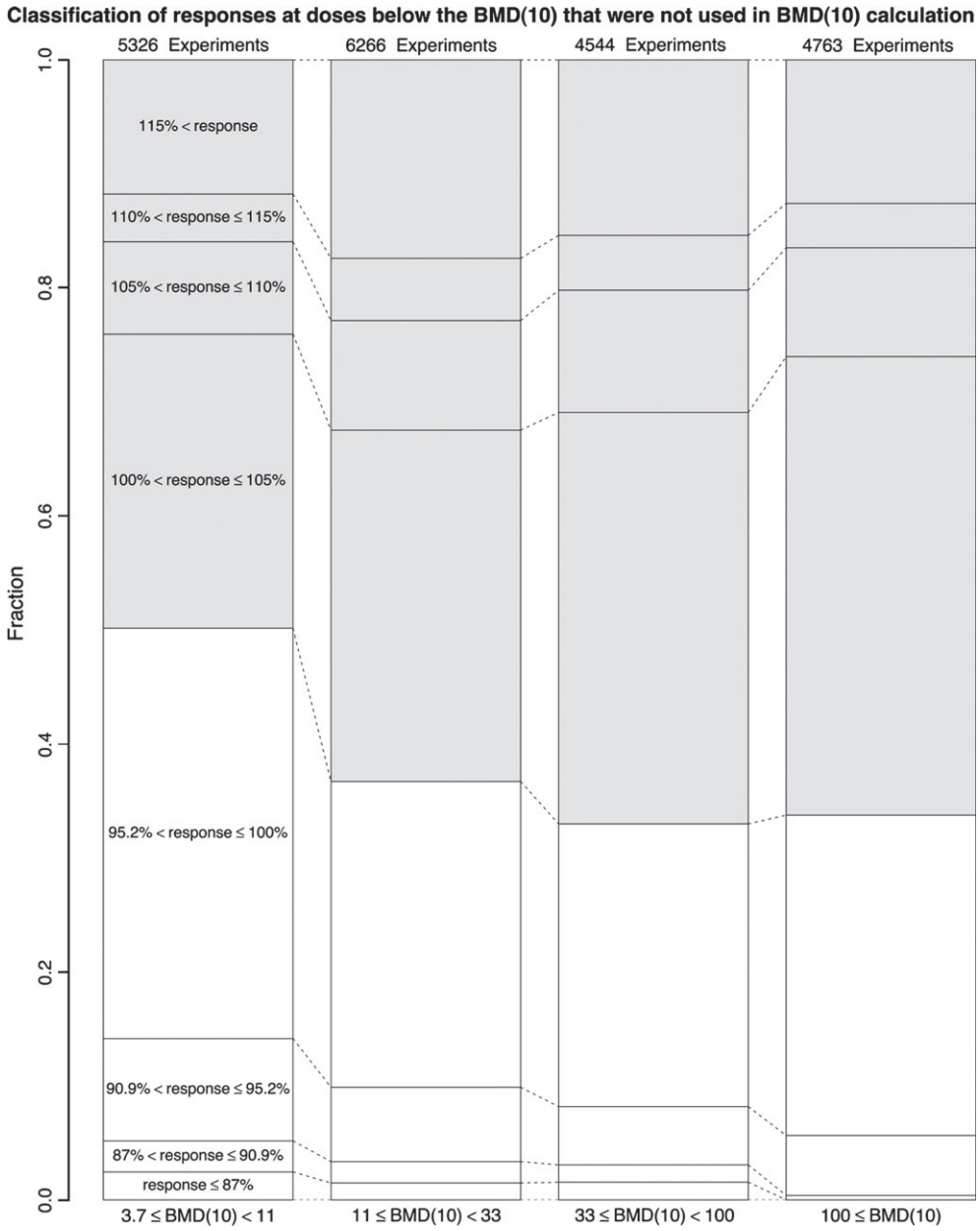


FIG. 4. Distribution of responses below the BMD(10). Shaded panels represent above-control responses, and clear panels are below-control responses.

observations of Calabrese and Baldwin (2001b) that there is an optimal range of hormetic responses starting at about 1/3–1/4 of the estimated toxic threshold. It is likely that the low response at the concentration below the BMD(10) in the 3.7 $\leq$ BMD(10) $<$ 11μM classification (Table 3) was due to there being a substantial proportion of such responses within the toxicity zone. The likelihood of a hormetic response increases as the distance from the BMD(10) increases, at least up to the limits presented in the present database. Since the semilog concentration spacing covered only a 100-fold concentration range, usually including toxicity at the high concentrations, it was not possible to explore the concentration-response range at which a return to control values would be expected. This would have required several concentrations lower than 1.2μM. Based on the hormesis database (Calabrese and Blain, 2005), about 80% of the hormetic responses are within 100-fold of the dose of the toxic threshold. Regardless of the genetic differences among the 13 yeast strains, the overall response to the 2189 chemicals was similar. These findings suggest that the hormetic response is a general one, unrelated to a specific cell cycle regulatory mechanism or DNA repair pathway. Similar quantitative features of the hormetic dose-response occur in models representing broad phylogenetic diversity and various cancer and noncancer-related endpoints (Calabrese and Baldwin,

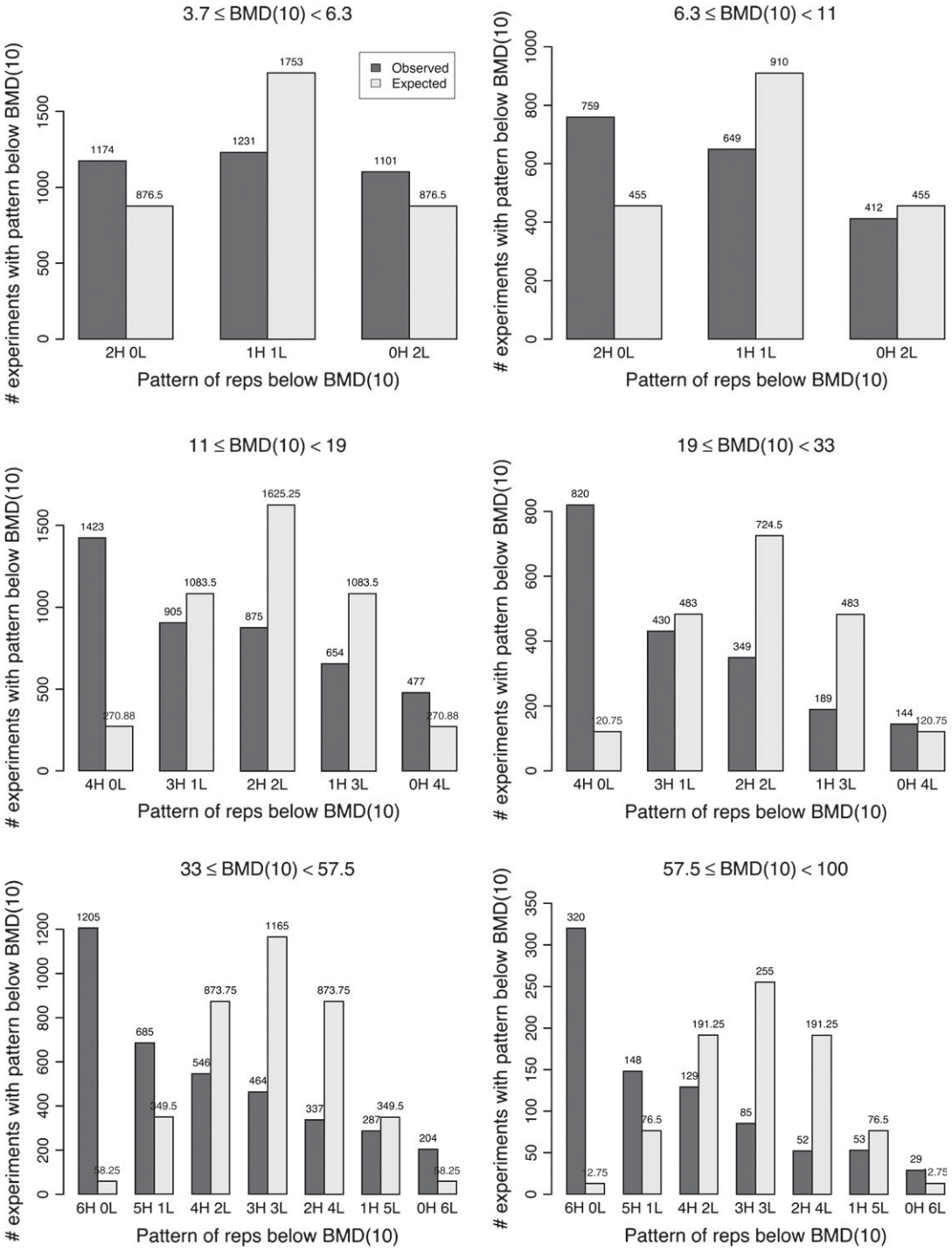


FIG. 5. Pattern analysis of below-BMD(10) responses: a test of threshold and hormetic dose-response model predictions. The levels 3.7, 6.3, 11, 19, 33, 57.5, and 100μM are approximately evenly spaced on the log scale. For example, take the panel labeled $11 \leq \text{BMD}(10) < 19$. There are responses at two concentrations below the BMD(10) (1.2 and 3.7uM) and two replications at each concentration. An experiment falls into the 3H 1L pattern if three replication responses were greater than 100%, and one was less than or equal to 100%. The dark bar is the observed count in that pattern, and the lighter bar is the expected count assuming a threshold model holds where H and L each occur with a probability of 1/2 at each replication below the BMD(10).

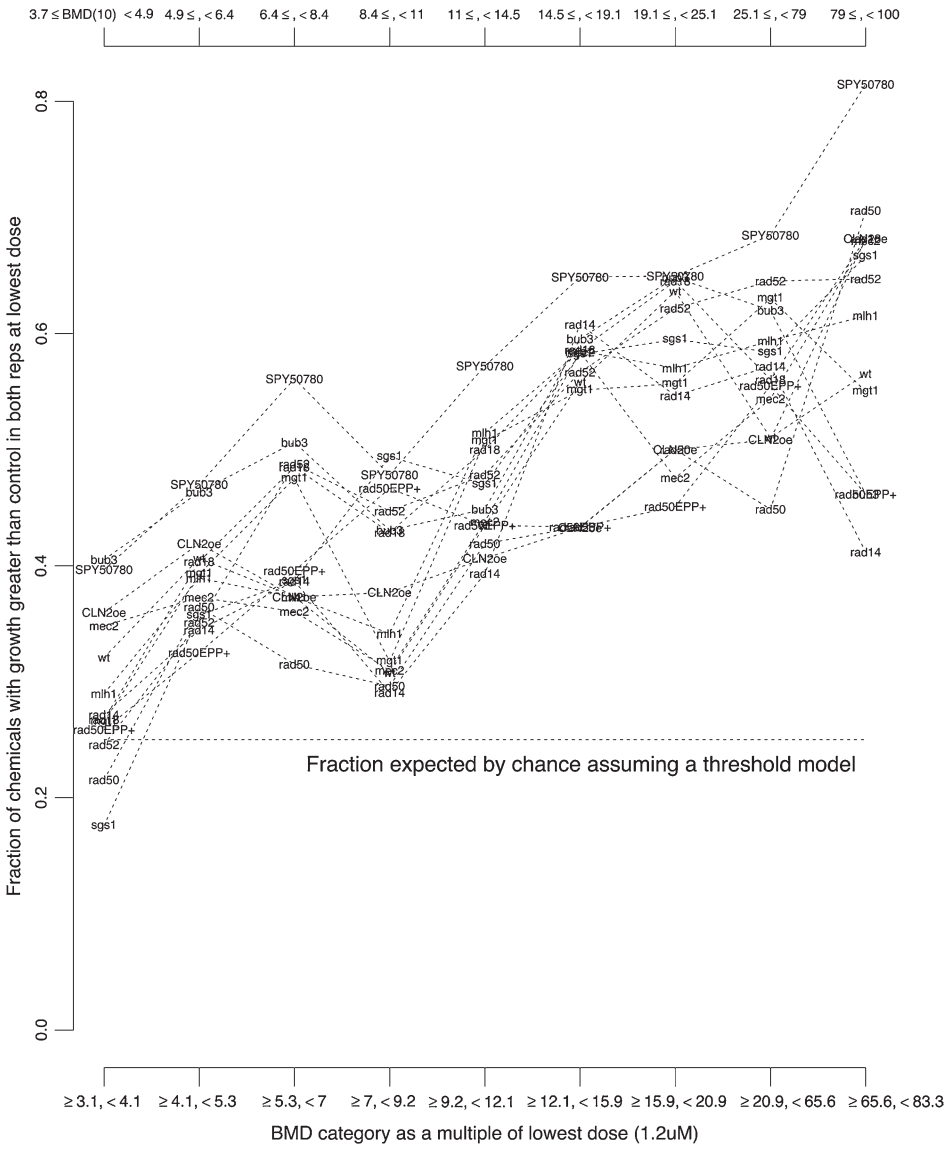


FIG. 6. Frequency of a hormetic response at 1.2μM in relation to the distance from the BMD(10).

2001b, 2003). The findings with yeast are consistent with those seen with NCI cancer drug screening data for 70 human tumor cell lines and up to 55,000 chemicals, involving over 3.3 million dose-responses (Calabrese, Staudenmayer, and Stanek, in preparation) (http://dtp.nci.nih.gov). Our present findings are drawn from the U.S. NCI database for screening of potential antitumor agents, and the consistently observed stimulation of proliferation in the below-threshold zone may have significant implications for the design of new antitumor drugs, drug testing, and the management of patients in clinical settings (Calabrese *et al.*, 2006).

The current findings are particularly important because they demonstrate the inadequacy of the traditional threshold dose-response model in predicting below-threshold responses. They also indicate that the hormetic model is consistent with these subtoxic responses. The findings suggest that the hormetic responses are more fundamental than threshold responses and support recent arguments that the hormesis model should be considered as the default dose-response model for scientific interpretation of toxicological responses (Calabrese, 2004). Several features of the study design and/or methodological evaluation (e.g., peripheral placement of controls on the 96-well plate, use of a nonsymmetrical model to assess above/below 100% responses, use of the BMD(10) instead of the NOEL or BMD's with lower cutoff points [2.5, 5.0, and 7.5]) favored conservative estimates and may have caused an underestimation of the frequency of hormesis. Thus, the inadequacies of the threshold model are probably greater than presented, while the predictive capacity of the hormesis model in the below-threshold zone exceeds that reported. The findings

argue for a paradigm shift in our understanding of the dose-response relationship, the central pillar of pharmacology and toxicology.

SUPPLEMENTARY DATA

Supplementary data are available online at http://toxsci.oxfordjournals.org/.

ACKNOWLEDGMENTS

Effort sponsored by the Air Force Office of Scientific Research, Air Force Material Command, United States Air Force (USAF). The assistance of Dr Julian Simon, Fred C. Hutchinson Cancer Research Center, is gratefully acknowledged.

REFERENCES

Calabrese, E. J. (2004). Hormesis: From marginalization to mainstream. A case for hormesis as the default dose-response model in risk assessment. *Toxicol. Appl. Pharmacol.* **197,** 125–136.

Calabrese, E. J. (2005). Paradigm lost, paradigm found: The re-emergence of hormesis as a fundamental dose response model in the toxicological sciences. *Environ. Pollut.* **138,** 379–412.

Calabrese, E. J., and Baldwin, L. A. (2001a). Hormesis: U-shaped dose-responses and their centrality in toxicology. *Trends Pharmacol. Sci.* **22,** 285–291.

Calabrese, E. J., and Baldwin, L. A., Eds. (2001b). Introduction: Scientific foundations of hormesis. *Crit. Rev. Toxicol.* **31,** 351.

Calabrese, E. J., and Baldwin, L. A. (2001c). The frequency of U-shaped responses in the toxicological literature. *Toxicol. Sci.* **62,** 330–338.

Calabrese, E. J., and Baldwin, L. A. (2003). The hormetic dose response model is more common than the threshold model in toxicology. *Toxicol. Sci.* **71,** 246–250.

Calabrese, E. J., Baldwin, L. A., and Holland, C. (1999). Hormesis: A highly generalizable and reproducible phenomenon with important implications for risk assessment. *Risk Anal.* **19,** 261–281.

Calabrese, E. J., and Blain, R. (2005). The occurrence of hormetic dose responses in the toxicological literature, the hormesis database: An overview. *Toxicol. Appl. Pharmacol.* **202,** 289–301.

Calabrese, E. J., Staudenmayer, J., and Stanek E. J. (2006). Drug development and hormesis. Changing conceptual understanding of the dose response creates new challenges and opportunities for more effective drugs. *Curr. Opin. Drug Discov. Devel.* **9,** 117–123.

Casella, G., and Berger, R. L. (2002). *Statistical Inference*, 2nd ed. p. 190. Duxbury, Pacific Grove, CA.

Clark, A. J. (1926). *Applied Pharmacology.* 1st ed. J & A Churchill, London.

Clark, A. J. (1933). *Mode of Action of Drugs*. Arnold, London.

Clark, A. J. (1937). *Handbook of Experimental Pharmacology.* Verlig Von Julius Springer, Berlin.

Crump, K. S. (1984). A new method for determining allowable daily intakes. *Fundam. Appl. Toxicol.* **4,** 854–871.

Eaton, D. L, and Klaassen, C. D. (2001). Principles of Toxicology. *Casarett and Doull's Toxicology: The Basic Science of Poisons*, 6th ed. (C. D. Klaassen, Ed.). pp. 11–34. McGraw-Hill, New York.

Faessel, H. M, Levasseur, L. M., Slocum, H. K., and Greco, W. R. (1999). Parabolic growth patterns in 96-well plate cell growth experiments. *In Vitro Cell. Dev. Biol.* **35,** 270–278.

Hardman, J. G., Limbird, L. E., and Gilman, A. G. Eds. (2001). *Goodman & Gilman's The Pharmacological Basis of Therapeutics*, 10th ed. McGraw-Hill, New York.

Hayes, A. W., Ed. (2001). *Principles and Methods of Toxicology*, 4th ed. Taylor and Francis, Philadelphia, PA.

Holbeck, S. L. (2004). Update on NCI *in vitro* drug screen utilities. *Eur. J. Cancer* **40,** 785–793.