MASSACHUSETTS INSTITUTE OF TECHNOLOGY

CAMBRIDGE, MASSACHUSETTS

March 2, 1978

Jacob Kastner, Chief
Environmental Standards Branch
Office of Standards Development
U.S. Nuclear Regulatory Comm.
Washington, D. C. 20555

Dear Dr. Kastner:

I am enclosing two copies of a critique of the Mancuso,
Stewart, Kneale report requested by you.

My overall conclusion is that it is not an appropriate
analysis of the data. Of the various comments in the enclosed
critique, the most important are:

1. The goals sought in this paper are unreasonable and
   misguided.

2. None of the stated conclusions are justified from the
   given data analysis.

3. No light is shed in the report on the important
   questions involved, namely: What can be said about the
   effect of low level radiation in causing cancer.

4. No effort is made to consider the effects of other
   potentially carcenogentc factors; without such effort
   conclusions of the kind made here are not justified.

5. The methodology used is non-standard as statistics
   not appropriate for the apparent purposes here.

6. A majority of the conclusions in the report are based
   on eight cases of myeloma. While these eight cases
   deserve investigation in the context of what is known
   about epidemology of myeloma, they are insufficient
   to support the conclusions here.

7. Doubling doses obtained in this report imply a threshold
   model for radiation effect in causing cancer, rather
   than the linear model used by the authors. Their

678135

7908150487

deriving them from a linear model would seem to contradict the linear model or their own analysis.

8. Use of Hanford deaths from other causes rather than survivor data is unjustified. Construction workers not receiving radiation hired in 1944 should have been removed from the data set.

9. Treatment of internal radiation is arbitrary and unjustified.

10. The variables used are not really appropriate for this problem.

11. No attempts are made to discuss background, or accuracy of data base.

12. No mention is made of the state of knowledge of the effect of radiation as a cause of cancer.

13. No mention of other efforts to deal with Hanford data are made.

14. Data is grouped and categorized in arbitrary and unjustified ways.

15. No effort is made to assess validity of the linear models used or to obtain confidence intervals on the doubling doses.

In light of these comments and conclusions, I thoroughly concur with the decision that further analysis of this data be handled by a different contractor. I believe such analysis should be made. I do not believe that this report has regulatory implications.

Sincerely yours,

Daniel J. Kleitman

CRITIQUE OF MSK REPORT

By Daniel J. Kleitman

This review consists first of a description of the Mancuso et al. report and the claims, arguments, evidence, etc. that are in it. Second a discussion of questions raised by the report and finally some conclusions.

It is divided into the following sections.

1. DATA SET

We here consider several questions that relate to the validity of the data set. Potential questions are:

1. Is the radiation data accurate?

2. Is the cause of death data accurate?

There are serious questions about accuracy of radiation data that are implied by the existence of background radiation, unmonitored individuals receiving radiation, man made radiation not monitored, and medical uses of radiation.

The existence of background implies that badges never had a zero reading, so that radiation doses must be estimated by subtracting background. This fact means that total lifetime dose cannot be accurately assessed to within a small fraction of background. Individuals who left their job within a year of employment may very well receive job related radiation within the next twenty years, more so than an individual who receives a total on the job dose of less than 1/10 th of the background. Flying in airplanes and visiting certain exposed areas increases off the job radiation. Finally, an individual may receive

radiation therapy for cancer or other diagnostic radiation which if monitored might distort the figures seriously.

The existence of a variety of radiation exposures does not preclude testing for the effects of on the job radiation. It only means that any actual effect of radiation will be smeared out by the scattering of the data caused by non-monitored radiation. It does, however, imply that it is probably unwise to group individuals by radiation dose into groups that are small compared to the scatter of natural and unmonitored radiation dose. Thus the distinction between no radiation and <.2 rad over a lifetime is probably meaningless. The category 'no radiation' in fact may select primarily for certain occupations, or short employment periods.

There is a problem with cause of death data in that the ultimate and proximate causes of death are not necessarily the same. An individual informed of non-curable cancer may commit suicide or drive into a tree in an apparent accident. One kind of cancer can metastasize into another fatal cancer--which variety is the cause of death? A patient may fail to recover from anaesthesia from an operation -- what is the cause of death? It is probably not the business of the authors here to be involved with this question, but an appreciation of varieties and accuracy of certification practices is important if one

is to distinguish among kinds of cancer. Absence of
certificate cause of death can also conceivably bias data.

2. VARIABLES

The authors make use of several major variables
in describing the death data in this study.

These are:

(1) The proportion of those dying of cause X
that were irradiated.

(2) The mean cumulative radiation dose of those
dying of cancer X (who were radiated).

(3) The cumulative mean radiation to some point in
time (or some time before death, or after employment)
among those employed at that time (or at that length of
time before death, etc.).

The proportion of deaths attributed to cause X.
There is no attempt to compare individuals dying of X
with survivors, or to consider absolute death rates.

Use of the statistic: "The proportion dying
of X that were radiated" suffers from the following
problems:

1. The distinguishing feature,   "not being
radiated at all", is a misnomer -- the actual radiation
received by the unradiated group is not much different from
that of the lightly radiated group due to background.

2. To a large extent the "unradiated" represent
individuals hired in 1944 and 1945 who stayed in the job
for less than a year as construction workers, etc. This
group is quite different in character from the rest of
the work force, and can be expected to show different
death patterns. (Over half and perhaps three quarters
of the non-radiated are in this category). These were
war time employees at a time when virtually all able
bodied men without specific skills were in the armed
forces. It is not unlikely that a high percentage of
these employees were 4-F and may have suffered from ail-
ments that increased their risk of non-cancer deaths.
It is evident from Table 1 that they have died much more
frequently than the 1946 or later cohorts, and this re-
flects a considerable difference between these cohorts
and the others which can be expected to effect death
patterns, increasing the proportion of non-cancer deaths
among them.

(3) The remainder of the non-radiated are dis-
tinguished occupationally from the "radiated". They,
therefore, would be expected to exhibit differences in
death patterns associated with occupation, level of
physical activity, etc., that seem apriori far more
significant than a trace of monitored radiation over
background.

The use of mean radiation dose suffers from the fact that the dose is not nicely distributed. The number of individuals who have received doses in the various categories are: above 10 rad 100, above 5 rad 150, above 1 rad 600, above 0 rad 2200, exactly 0 1300. The average dose is not in any sense a representative dose.

If one believes in a linear model, the average dose is, of course, exactly what measures the radiation effect, so that there is some justification for using it. However, the nature of the distribution here implies that the average is dominated by the presence or absence of relatively few high dose individuals.

Thus consider the RES Neoplasm chart on page 50. One form of RES Neoplasm is myeloma in which the eight unradiated cases received _average_ dose of over 10 rad each. Just one such case added or subtracted to the RES neoplasm column on page 50 would drastically alter its appearance -- putting the entries in it down with these in the other columns or lifting them significantly higher (by 22 in the last entry). Thus we see a whole column of figures whose significance all dwells on the diagnosis of one single case. When one case can drastically affect all conclusions the strain on the accuracy of the data base can become excessive.

When comparisons are made with national statistics, these are always done as proportions of deaths by various

causes, proportions of cancer deaths, etc.

When only death date is available this procedure
is unavoidable. When survivor data is available, however,
it is generally far superior to consider death rate as a
function of population at risk. Why should one make one's
statistic vulnerable to fluctuations in other causes of
death if one can avoid doing so?   With the statistics
considered, there is no way to distinguish a positive cor-
relation with cancer from a negative correlation with
heart disease, or any other cause of death. An outbreak
of cholera among radiated at Hanford could lead to the con-
clusion that radiation prevented cancer -- quite erroneously --
with the present data set-up.

3. CONTROL STRATEGY

        The general methodology used to supply a control
to the cancer death data is to use the death statistics
from non-cancer victims.

        There are a number of problems with this approach:

        1. These controls have a significant contribution
from the 1944 cohort of construction workers, who are
basically irrelevant to the study.

        2. As already noted, it permits no distinction
between upward fluctuation in cancer and downward fluctuation
in non-cancer.

3. It was chosen after the data was seen. That
is, when data is taken according to a fixed predetermined
pattern, and one applies predetermined tests to it,
traditional statistical techniques describe how to deter-
mine "significance" of data. When a method of analysis
that yields a "significant" effect is chosen after seeing
the data and noting that some other equally reasonable
procedure fails to yield such effect, is not so easily
analyzed. Certainly attributions of significance must be
sharply reduced under these circumstances. Extra pains
must be taken to show that the conclusions drawn are not
artifacts of the method.

It seems to me that age matched survivor data
is much more appropriate as a control. If conclusions
depend on the non-cancer death rather than survivor as
control they are highly suspect.

(It is the authors themselves who point out that
analysis had been held back for some time by attempting to
use survivor data and who suggest that they were only able
to find effects.with the present comparisons).

4. INHOMOGENEITY

The data is not homogeneous. With regard to lung
cancer, there are smokers and non-smokers; there are those
who have evidence of internal radiation and not, there are

a number of dissimilar occupational groups. Included in
the sample are many who had contact with job related
radiation at the Hanford facility only for very short
periods. Moreover, external radiation is not randomly
distributed with regard to most of these factors. There
are very significant correlations with internal radiation
and occupation.

It is, therefore, impossible to draw conclusions
of the type sought in this paper without confronting the
following questions:

1. Are any of these other factors in themselves
correlated to the onset of cancer of various kinds?

2. How are these factors correlated with
radiation rate?

3. Do these factors provide alternate explanations
of the data?

The standard way to handle factors like these is a
multivariate regression as opposed to the single variable
regression used here.

In the absence of any attempt to consider such
variables one can draw no hard conclusions from the data.

Of the various factors, smoking correlates with
lung cancer, internal radiation may correlate with lung
cancer, occupations like plumber and chemical worker cor-
relate with cancer and are inhomogeneous in this data

(as noted by Gilbert). 1944-1945 unskilled male employees may correlate with 4-F's and enhanced risk from non-cancerous dise....

5. TEST METHODOLOGY

The only tests mentioned in the paper are:

(1) Rank tests on age-distributed dose-grouped overall-cancer-death percentages.

(2) t-tests on average cumulative doses at specific ages before death for specific cancer varieties (vs) non-cancer deaths in the same range.

The rank tests are not unreasonable although they are not the most natural tests to apply to determine correlation between cancer death and radiation. Death rate would be more natural as a variable. No attempt is made to correct for occupational or other inhomogeneities. The choice of categories is not particularly natural. (That is, the divisions should probably be

<1 Rem   1-5Rem   5-10 Rem   over 10 Rem

rather than those used.

The results would be quite different with this division and probably would lose significance. As a test of the effect of radiation, this test is weak and not very convincing. It does not   justify the almost absolute

certainty with which the authors state their conclusions.

The t-tests,as noted by the authors in Appendix 3, should not be compared with standard t tables, since the distribution is far from normal. They discover empirically that 10^{-4} should be 6×10^{-3}, and 10^{-3} is more like 10^{-2}; if these results are general, t=2 is not significant at all for this test, and t=2.5 is probably not significant. This calls into question many of the significance conclusions deduced by their arguments.

There is something else methodologically questionable about this t-test procedure. The authors examine a large data set, seek out the most extreme t values in it, and apply tests to the significance of these points. Now, supposedly, if one chooses 20 independent data points, one of them can be expected to deviate significantly at a 5% level. If one examines 20 types of cancer by any fair approach, likewise one should expect one to deviate significantly at a 5% level. It is not appropriate really to seize upon such deviations and take them as an in-dication of a cause of cancer. How can one be sure they are not fluctuations in data?

The testing procedure is deficient also in putting a blind eye toward factors other than radiation that might correlate positively with cancer and which are not uniform-ly distributed among radiation doses. Where such features

are known to exist (here occupations, internal radiation)
it is inappropriate to analyze data as if they did not.

In particular, internal radiation appears to be
so heavily correlated with higher external radiation here
that it is obviously impossible to separate their con-
sequences. It is inappropriate to ignore such questions
when testing for dependence in closely correlated variables.
It is also inappropriate to ignore other forms of inhomo-
geneity in the data set such as occupation.

To summarize, the testing methodology here is by
no means the standard approach to problems of this kind
in the following respects:

1. No effort is made to consider other variables
in the problem (occupation, internal radiation) that are
probably important.

2. t tests are applied to many different
variables (hundreds, though not all are independent) and
significance attributed to all with values above t = 2
(despite the fact, noted in Appendix 3, that
t = 2 may not be an appropriate criterion for significance
even for one test in certain cases).

3. The choice of categories for the rank test
is arbitrary and unnatural.

4. The method of choosing "critical" year before

death and age sensitivity involves searching among columns
of numbers for the one with the largest t value. The
test then applied to that statistic appears to be the t
test advocated by the authors here. It is, of course,
wrong to use any such values in any test. This procedure
may be fine as a search technique. It is absolutely
unacceptable as a test technique.

Each one of these considerations call into question
the conclusions of these tests.

6. DOUBLING DOSE ESTIMATES

Doubling doses are estimated by the standard one
variable linear regression formula, which gives doubling
dose for linear dependence of (probability of disease)
vs. (radiation) as a function of number of victims of a
given disease x, t value, and parameters of the total
population.

The formula used is $D = \frac{\sqrt{n}V}{t} - R$ (for n<<N) where
n is the number of disease x victims, t is the variable
used in the previous test, and V and R are parameters of
the entire population.

This formula gives a "best" linear fit to the
data in the sense that the sum of the squares of the de-
viations of the data points from the linear relation is
minimized.

Use of this procedure deviates from standard in the following respects.

1. It is customary, when using such an approach, to test the "goodness" of fit of the linear model by computing the "residuals". The ratio between variance in the data and the sum of residuals gives some notion of the utility of the linear model in explaining the data. No such computations are made here, and there is, therefore, no mention of how well the linear model succeeds in explaining this data. The estimates given are mere estimates; confidence intervals in doubling doses would be much more appropriate.

2. When other variables (such as occupation, internal radiation reading) are present (as in this case) multivariable regression is the standard approach. The extent to which use of radiation as a variable reduces the residuals is the standard measure of significance of radiation here. Nothing like this analysis is applied here - other variables are simply ignored.

3. The procedure used for computing doubling doses used here involves first selecting those cancer types for which t values above 2 are encountered in cumulative radiation dose. Since D is determined by t(and n) what we are doing here really is:

considering a number of cancer types, with corresponding t and n values, selecting those

with highest t values, and giving the D value $\frac{m}{t}$ V-R as if it were not a fluctuation but rather was due to a linear model.

This procedure is dangerous, since with random causes of death one expects to find one with a t- value significant at a 5% level. Such a value will automatically lead to a doubling dose $\frac{\sqrt{n}V}{t}$ - R which for small n will probably be rather small - almost independent of the phenomenon under consideration. In other words, this procedure can take probable fluctuations and make them into rather small doubling doses (particular for diseases with relatively few victims (small n). It would be much safer here to divide the data arbitrarily into two pieces, using one to select for high t, the other to compute doubling dose from observed t value.

4. The procedure is rendered even more unusual in that for each cause of death the statistic is chosen that maximizes t and hence minizes D among a number of possible models. That is, the variable considered is "cumulative radiation dose up to k years before death". t values are determined for each k, and the doubling dose is apparently calculated for each disease with the k value that gives largest t or smallest doubling dose D.

A standard approach here might involve testing
a model with a k value given ab initio. To use the same
data to determine both the k value with highest t and the
doubling dose determined by that t runs the risk of
drawing a long chain of consequences from one single
fluctuation.

It would be much safer here to divide the data
randomly into groups, compute k from one and doubling
dose for that k from the other, so that one cannot
unduly compound the effects of single pieces of data or
fluctuations.

5. The model of cancer probability used here is
linear, $p = a(D+x)$, where D is a measure of the
susceptibility to cancer (on a scale determined by the scale
of x) without job related monitored radiation. It is
obvious that D must be positive, or p would become
negative for small x, an absurdity. Moreover, in a linear
model (as opposed to a threshold model which the authors
reject) D must exceed the average background radiation
level B as well, since without background p must still
be positive.

A conclusion that $D < \frac{1}{40} B$ (as obtained for bone
marrow cancers in higher age brackets) is therefore completely
inconsistent with a linear model, as are a number of other
doubling dose results in this paper. The true conclusion
from such doubling doses would have to be that they are

incompatible with the linear model that spawned them.

It is possible that the authors have proven the invalidity of a purely linear model for some of this data, granting the validity of their procedures. (It is likely that the implications of a (cumulative lifetime) doubling dose of .1 rad at 71 years for bone marrow cancer would have observable consequences in high radiation areas even with the threshold model presumably assumed by the authors here.)

6. As a justification of their doubling dose conclusion, the authors note that the increase in proportion of deaths predicted by the doubling doses is less than the excess proportion of such deaths observed (over U.S. statistics).

Contrary to the premises of that argument, the doubling dose can be both too low and still can predict too few excess deaths.

This will occur when the death rate data is in fact not particularly linear in cancer death vs. radiation dose, namely when small amounts of radiation have cancer rates that are relatively high compared to those with high amounts of radiation. In the varieties of cancer here (except myeloma) there are many more of the former cases than the

latter and this phenomenon occurs in the sense that the best linear fit to the 0 and low radiation data has a larger slope than the best fit to the 0 and higher radiation data. The linear model gives a poor fit to the data.

Since there are many more low radiation data points, the best fit can tend to underestimate the cancer death rate for many more points than it overestimates them, even while exaggerating the slope.

In consequence of this phenomenon, the justification of doubling dose results given in the paper is without merit.

7. BACKGROUND AND DOUBLING DOSE (see Appendix II of MSK)

A linear radiation model would have the probability of cancer proportional to radiation dose. The authors use an expression

$$p = a(D+x)$$

If an individual receives background or unmonitored radiation b in addition to the monitored radiation x, in this model one has $p = a(D+b+x)$. If one assumes that b and x are uncorrelated, we should have in our model $p = a(D+\overline{b} + x)$ and with our same definition of doubling

dose, what the authors call doubling dose is really $D+\bar{b}$, the doubling dose plus the average background dose. Since $\bar{b}$ on a 71 year old is doubtless at least four rads, the true doubling dose is actually what the authors compute less at least four rads. This is in many cases negative. A negative value D means that the model is nonsense - that more than all the cancer is radiation induced, or that one must choose a meaningless linear model, or resort to a threshold model.

If background is four rads and there are typical fluctuations of 1 rad in background, an observed doubling dose of .1 must, furthermore, mean that the threshold must be well over 5 rads or else many individuals would cross over it from background alone.

Thus, predicted doubling doses that are small fractions of background, if taken literally, imply that there is a threshold and that the threshold lies well above average background level.

Unfortunately, the authors doubling dose arguments are not sufficiently justified by confidence intervals, etc. to be taken as proof of the threshold theory. If taken alone, however, they do support it.

corrected page

(Qualitatively, the only way doubling dose can be much smaller than background is if it is the last straw that produces cancer in the victims. This must mean that most background doses, including those that fluctuate above average background, do not have this last straw, and hence life below the threshold.)

8. RES NEOPLASM

Many of the strongest conclusions of the paper concern bone marrow cancer and RES Neoplasm, and leukemias in general. These raise the following questions:

1. What is the present data for these diseases?
2. For which are these indications of a linkage with radiation?
3. What is known or suspected from other sources about causal relations and patterns of onset or general epidemology of these diseases?
4. What are potential explanations of the data?
5. What conclusions may be drawn from the data?
6. Are the conclusions drawn by the authors correct?

The types of RES Neoplasm are classified into five groups in this study. There are:

678156

corrected page

	No. cases	No. radiated	No. expected cases	avg. dose
Lymphoma (200-2)	34	28	27.7	1.45
Lymphatic Leukemia (204)	3	2	9.4	.29
Myeloma (203)	11	8	7.6	10.66
Myeloid Leukemia (205)	11	6	5.8	2.21
Other Leukemia (206-9)	5	3	20.3	.19
All non cancer	2850			1.62

These columns represent the number of cases, the number of these with radiation records, the number of cases expected if 670 cancer victims were distributed according to national statistics and the average radiation dose of those receiving radiation.

In addition, according to tables 19 and 20 there were 17 victims with internal radiation and 47 without, as opposed to 619 non-cancers receiving such and 2159 (not). In a crude sense the caseswith IR could be considered as these receiving external radiation as opposed to the essentially unradiated others. The figures then show little or no death excess in radiated over unradiated here.

If one examines each type of RES Neoplasm separately one finds the following:

678157

Lymphoma : relatively few zero radiation cases, not many
"radiated cases". Large number of cases with low level of
radiation; t value with respect to non-cancers is small.
No significant indication of radiational effect. If
classified by radiated vs. hardly radiated, the t value
is negative. Not even qualitative indication of radiation
effect.

Lymphatic Leukemia and Other Leukemias: -- There is a
relative deficiency of cases of these diseases in this data.
In fact with 29.7 cases expected, there were only 8, and
none of these was in the "radiated" group (for which there
was internal radiation). By contrast roughly 1/3 of the
control group (non cancers) was in this category. If these
cancer types are joined together, this lack of radiated cases
is significant at the 5% level.

   Considering that the number of cases of these diseases
is relatively low by national standards (by a factor of 2)
even in the control group, the significance of this lack
could be made much higher by alternate analysis.

Myeloid Leukemia: -- There were 11 cases, 5 with no monitored
radiation; of the other 6 cases, 3 or 4 were in the "irradiated"
group. This data is not significant, and shows no particular
trend, it gives no evidence even suggesting a radiation cause.

<u>Myeloma (203)</u>:-- Here there were 11 cases, 3 of whom had no radiation records. The remaining 8 were almost all "irradiated", with an average dose of 10.66 rads.

This data shows a significant correlation of radiation with deaths at the .5% level using the authors' tests.

This myeloma data, these 8 cases, form the only leukemia data that has any potential significance as a link between this radiation and cancer.

To summarize the data:

A. For myeloma (203) there are 7 or 8 irradiated cases with average dosage of 10.66 and 3 unradiated cases. 0 or 1 case with very low level radiation   (There are 6 expected cases.)

B. For (204) + (206-9) there are no cases with more than   trace amounts of radiation, 4 cases with no reading, 5 with very low level of radiation (25 expected cases).

C. Nothing else is even potentially significant The possible explanations of this data are:

1.   Existence of causal links between leukemia and cancer

2.   Fluctuation

3.   Mislabeling, misidentification, non standard

identification of these diseases

4. Clerical or bookkeeping error in tabulating or recording data.

In attempting to distinguish among these causes, it is appropriate to consider what is known and suspected about the epidemology of these diseases and the effect of radiation on them. In particular, the answer to the following questions are important:

1. Does or did myeloma occur frequently in those undergoing large radiation exposures, bomb victims, those undergoing radiation therapy, etc.?

2. Do diseases 204 and (206-9) occur at all in these groups?

3. Do these diseases tend to occur in clusters? What is known about their pattern of occurrence?

I do not know the answers to these questions. If radiation has been observed to cause myeloma, this data may well be in part or in whole a manifestation of it. If it does not, it is very unlikely to be the cause of these cases; low level radiation data is not the place to look for new evidence of diseases caused by radiation.

If myeloma and the other diseases tend to occur in clusters that are often not radiation related the probability that this data is a fluctuation is greatly enhanced.

(There is some evidence of this clustering.)

The possibility of error or non standard diagnostic classification are worth checking, since the excess of 203 cases almost neatly balances the deficit in 204 and 206-9) and this might tend to explain this data.

Ideally, Hanford data could be used to distinguish between linear and threshold models for low level radiation. If there is any data here that would support the existence of effects of low level radiation and, therefore, support a linear theory, it is this myeloma data.

It is a problem that there are only eight cases. A serious study of these cases, in the context of what is known about myeloma would be interesting and is in my opinion worth doing.

The non existence of 204, 206-9 cases among radiated is mysterious as well. Could this be a clustering phenomena that has avoided Hanford? It is very hard to believe that the radiation present here would prevent these diseases.

One purpose of the discussion above was to point out that the 8 myeloma cases represent practically the only feature of RES neoplasma data that correlates cancer positively with radiation.

On the other hand, these 8 cases averaged 10.66 rads
of radiation each, and this dosage is enormous compared to
that of most of the others in this data set (though not
quite so much by comparison with survivors).(This is in
part because of the temporary workers of 1944 in the present
data set.)

The authors do amazing things with these 8 cases:
I list some of these.

— They group them with the 11 myeloid leukemia cases
into a group called "bone marrow cancers". This grouping
gives rise to the most significant t value; in the entire
analysis. (See Table 14). It also leads to a conclusion
that 9 years before death is the crucial time for measuring
dosage (p. 17), a body of information as to age similarity
doubling dose (p. 21), an estimate of "actual" doubling dose
(table 16) of .8 rad, and an estimate that 9.3 bone marrow
deaths were radiation induced! All of these conclusions and
significances are not indicated at all by the myeloid leukemia
data. It all comes from the 8 cases of myeloma. (Incidentially,
the conclusions would have been even more extreme for myeloma
itself, except of course, for the number of radiation
induced deaths.)

Next the authors obtain all sorts of conclusions for

"all RES neoplasmas". Again, the 8 cases dominate the data, giving large t values and low doubling doses. Again a raft of conclusions about age and critical years follow. Again all of this vanishes for any combination of RES neoplasmas other than myeloma. The myeloma data, the 8 myeloma cases, are responsible for every single conclusion involving RES neoplasmas.

Next the authors include these cases, as they obviously belong, in the category "all cancers", again all significant conclusions ($t \geq 2$) disappear if these (11) cases are omitted. The level of radiation 10.66 rads for the 8 exposed here is so much higher than that of most of the rest of the cases in the study that these 8 cases out of 670 are responsible for 10-20% of the observed t values (which depend on year before death) for the "all cancers" category.

There are numerous conclusions in each of these categories about age sensitivity, critical year before death, doubling dose, significance of t value, all of which are traceable to the same 8 cases of myeloma.

Fascinating and suggestive as these 8 cases may be, they cannot support this weight of conclusion, or any con- clusion about any disease other than myeloma.

9. INTERNAL RADIATION

Data was taken from urine-analysis of the presence
of ingested radioactive matter. Such matter could come
through skin, through lungs or through the digestive tract.
Presumeably this radiation came primarily through the lungs.
The presence of a quantity of radiactive matter deposited
within the lung, remaining perhaps for long periods and of
sufficient magnitude to show up in urine, could well be a
causative factor for lung cancer. Radioactive dust has long
been known to contribute to lung cancer.

That such radiation was observed in a large portion
of those deceased who got more than trace amounts of external
radiation calls into question any conclusions that excess of
lung cancer deaths among these was produced entirely by
external radiation.

It may be that this internal radiation was minor or
trivial and could not be or should not be considered a possible
cause of cancer here.

Without any such assurance it is difficult to under-
stand how the authors can treat internal radiation the way
they do. With regard to lung and pancreatic cancer, it is
obvious from tables 19 and 20 that there is a significant

correlation between these types of cancer and internal
radiation. The relevant data is:

Non-Cancer: Number of cases, (external average doses)

with IR	691	(3.8 rads)
without	2159	(.4 rads)
ratio	=	.32

RES Neoplasmas (including myeloma)

with IR	17	(4.2 rads)
without	47	(.5 rads)
ratio	=	.36

Lung and Pancreas cancer (mostly lung)

with IR	86	(4.6 rads)
without	155	(.36 rads)
ratio	=	.555

Other solid tumors

with IR	91	(3.8 rads)
without	274	(.4 rads)
ratio	=	.33

It may be that the lung and pancreas data represents a fluctuation; it may be that it is due to external radiation but it may be that it has something to do with internal radiation (and smoking as well and occupation also). It seems incredible that the authors ignore this possibility without any explanation.

The use of internal radiation here with regard to survivor data is even harder to understand.

The authors were confronted with the embarrassing fact that survivors had even higher radiation doses than cancer victims; and as they had chosen to ignore survivor data, they attempted to explain this fact. Their explanation is on page 26. They chose not to include survivors in their rank test for reasons unknown to me. They did not attempt to use age as an explanation for survivors' radiation patterns, so I presume age was unable to explain them. Instead they sought to understand the survivor radiation data through three factors: cohort (year of hire), exposure year and internal radiation.

They note (see table 1 and table 26) that many fewer survivors were hired in 1944-5 than the proportion of non-survivors, and relatively more were hired comparatively

recently. They, therefore, gave weight to each individual according to his "cohort" to make up for this factor. To some extent this is an age correction; it also gives higher weight to surviving irradiated construction workers and, therefore, lowers the average radiation levels of the "normalized survivors". But the effect is not nearly enough to reach the comparatively low level of cancer victims.

They then note that survivors have tended to have more of their radiation recently, i.e., before it could take effect. Counting radiation more and less for survivors to obtain the same average time pattern for survivors and non-survivors again has an effect in reducing average survivor radiation levels; but still not enough even to match the level of cancer victims.

They find that the largest reduction in survivor radiation of all is obtained by giving added weight to those without record of internal radiation. It seems that survivors had even more internal radiation than cancer victims. The first two corrections can be rationalized - to some extent - as attempts to take into account age and latency period factors. The last correction is of the "fudge" variety -- unless there is some reason that an internal radiation reading should correlate negatively

or prevent the cancer that would otherwise be caused by external radiation, there is no sense to this correction.

The treatment of internal radiation and of survivors certainly casts doubt on all the positive conclusions concerning lung cancer described in the paper, and probably casts doubt on all the conclusions except those concerning the 8 myeloma cas .

Certainly survivor data and internal radiation as an additional variable should have been in the data set in all analyses.

10. IMPLICATIONS OF THE RESULTS

Among the questions that should be raised here are

1. Does the data and analysis here support the conclusions stated?

2. What conclusions are implied by these analyses?

It is appropriate to review the stated conclusions: (Summary) "The study shows that there is a definite relationship between low level ionizing radiation and the development of cancer". "Sensitivity to the cancer-induction effects of radiation is at a low ebb between 25 and 45 years of age." "There is a hazard associated with bone marrow cancers more than other neoplasmas and cancers of the pancreas and lung more than other solid tumors".

"Further analyses will be needed to rule out the now remote possibility that the positive findings were merely the result of the radiation exposures having associations with other cancer related factors."

The tests applied to the data here were:

A.   Rank tests on five grouped radiation level data for each of five age groups.

B.   t tests on average radiation level (with age and for various cancer types vs. non-cancer).

First, since no effort of any kind appears to have been   .to examining other cancer related factors among the data (in particular, occupational, smoking habits, and internal radiation), no conclusions of the type indicated can be made from this analysis.

Secondly, even ignoring the first objection, the fact that a large body of data relating to survivors was left out of the data base, and that such data showed relatively high radiation levels among survivors, reduces the credibility of the conclusions, and the significance of any findings.

Thirdly, even ig ring the objections above, the rank test (A) result is insufficie to justify the sweeping and forceful conclusions. The significance level in the

test result is 5%, and even this is based upon splitting the
low radiation group into three parts (0,0-. 2rad, .2-1 rad)
and would lose significance if these were merged.

In looking at the individual cancer types (B) one
finds that there were 11 cases of myeloma, 8 of which
received radiation at an average level of 10 rads each.
This single group of 8 cases is alone responsible for almost
all the positive findings (t values $\geq$ 2) in the report.

This includes all positive findings involving the
categories:

bone marrow cancer

RES neoplasmas

all cancer.

That is, if one considers instead the categories:
bone-marrow-cancer-other-than-myeloma

RES neoplasmas other than myeloma

all-cancer-other-than-myeloma
all positive findings disappear (in particular all the starred
entries in Tables 13, 14) except those for lung and pancreatic
cancer.

A fuller discussion of RES neoplasmas data is contained
in another section of this critique.

However, neither these 8 myeloma cases, nor the lung
and pancreatic evidence are sufficient to support the con-
clusions stated.

What conclusions can be drawn from this data and analysis?

1. There is a cluster of 8 myeloma cases that deserves further investigation. This investigation should be carried out in conjunction with whatever is known from experience in the outside world about causes, about radiation effects on, and about clustering of myeloma cases.

2. The lung and pancreas data should be examined in terms of other related factors (occupation, smoking, internal radiation). It could ultimately be significant evidence of the kind sought here. No conclusion about them is warranted at this stage.

3. If the doubling dose calculations in the paper are correct, the authors have disproven the linear model in favor of a threshold model. The implications of this proof would be very serious; since no confidence limits on doubling doses are obtained this conclusion is dubious.

It might, however, be possible to draw conclusions about this important question, though not from the analyses presented here. No single conclusion claimed by the authors here is sustained by the analysis.

11.   DISCUSSION AND CONCLUSION

There are two kinds of papers in applied statistics in areas like these; those that attempt to extract conclusions from data, and those that are exploratory, that attempt to find new phenomena for later testing.

Traditional statistics concentrates on the first of these kinds; there are rules that must be followed if one is to write such a paper. One must not suppress or falsify or even go around inconvenient data. One must not neglect factors. One must construct the models before seeing the data, not after. One must not merge or group data for effect. And so on.

There is I think a place for another, an exploratory kind of paper. This is one that seeks hidden patterns or phenomenas, asks, what combination of data has the highest t value? rather than what t value is appropriate in a predetermined model?

This kind of paper is not one to draw conclusions from, but rather one to suggest new hypotheses, possibly new phenomena. It is a form of pattern recognition. Traditional statistics frowns on all this, but there is a place for it. If the hypotheses and phenomena are spurious at least they will stimulate data that proves it and may find the right patterns.

The present paper claims to be of the first kind and
is of the second. It claims to prove radiation as a cause,
but merely extracts statistics that have large t values.
It does so in a somewhat haphazard way, and is by no means
a model of its kind. It suffers from the false claims
made for itself, from willfully ignoring inconvenient data,
but it appears to be an honest attempt to use the data base
to explore for the unlikely (larger t) phenomena in it.

Perhaps the most serious question one can raise about
the paper is: is it appropriate here to be searching
for radiation effects on types of cancer? The authors here
procede as if the rest of the world did not exist, nothing
was known about the problem, and they were to take this data
base and look for indications that radiation causes cancer.

But is this the place to look for such indications?
It seems to me that the fact that radiation causes cancer is
so well established that it is pointless to attempt to establish
it here. It is not pointless to look for new
kinds of cancer caused by radiation. It _is_ pointless
to look for them in this data base. Let's be specific; if
a doubling dose of .1 rad exists for some variety of cancer
at some age, how could this fail to show among bomb survivors

or others who received doses thousands of times greater than this
dose? If one believes in a linear theory, individuals with
heavy doses should show their effects far more than those with
small doses. The authors can use their approach on Hiroshima
survivors -- they might then find something new with their
exploratory methods.

Now it happens that a .1 rad cumulative doubling
dose contradicts the linear theory, but that is not important.
What is important is that they are looking for the wrong
things. Why are they looking for doubling doses? Why are
they searching for estimates of critical year and doubling
dose for varieties of cancer in this data? The majority of
the people in this data base have very small monitored
radiation doses - the authors are looking for these
phenomena in the worst possible data base for their apparent
purposes.

What should they be doing? The crucial question is
now how radiation affects cancer, but rather, how does low
level radiation affect cancer? The authors attack this
question entirely independently of available information
about the first one. They give the impression of attempting
to follow their data wherever and however it leads them, with-
out any reference to what is known about radiation as a cause

of cancer.

It is crazy to try to obtain a more accurate doubling dose from a small number of low level radiated cases than one has already from high level radiation cases of greater quantity.

The crucial question is: how do the high level results extrapolate to low levels? The most reasonable sounding models are linear, or have a threshold. There are further complications since one can get radiation all at once or over a time period. The authors should stick to specific diseases, take the best information available for the influence of radiation on cancer or cancer types, construct a variety of models extrapolating them to low levels of radiation, and see if the data in this data base is capable of distinguishing among them.

Unfortunately the answer is probably negative; this data is probably incapable of distinguishing among linear and threshold models. But the data base and all the money spent on it probably deserve a try to answer such questions. There is no clue from the present report as to these answers. This is what the authors should have been looking at; they should have tried to distinguish among low level radiation models, and not have computed daffy doubling doses. They should have shown some awareness of what was known about the subject and what their goals should be. It seems to me that the authors display insufficient understanding of goal, of the general problem, and of methodology to get anywhere

in the right direction on this problem. My own assessment
of this paper leads me to believe that sponsors of this research
were fully justified in transferring their support elsewhere,
and that little of use could be expected from these authors
toward resolving the major issues.

(In addition to the flaws mentioned above, the paper
suffers in that it does not refer to other attempts to deal
with parts of the same data base, particularly by Gilbert et al.
Gilbert found that occupational corrections rendered positive
conclusions dubious. One cannot in a scientific paper ignore
the efforts of others, particularly when they grapple with
complications ignored in ones own work.

Other minor problems with the paper are the gross
inconsistency among the tables and within certain tables, and
the lack of any coherent list of t values and predicted
doubling doses.)

APPENDIX I

1. SUMMARY OF MSK PAPER

The summary to the report contains the following
conclusions:

"The study shows that there is a definite re-
lationship between low level ionizing radiation
and the development of cancer. Data from the
Hanford study have shown that sensitivity to the
cancer-induction effects of radiation is at a
low ebb between 25 and 45 years of age. At
younger and older ages there is a cancer hazard
associated with low level radiation which affects
bone marrow cancers more than other neoplasms
and cancers of the pancreas and lung more than
other solid tumors".

Pages 1 and 2 contain description of the data set, which
consists of ERDA records of external and internal radiation,
date, occupation and cause of death of Hanford workers.
Pages 3 to 5 contain a description of preliminary findings,
namely:

1. Cohort of 1944 - hired individuals was largest
and included large number of unmonitored and short term
workers.

2. High proportion of nonexposed workers in 1944
cohort, and relativity low doses before 1954 and for men

with short records of employment, are reasons to expect
non-survivors to have lower radiation doses than survivors;
but taking all certified deaths as sample weighs cancers
and non-cancers equally among 1944 cohort and later cohorts.

3.   Cancer vis a vis other deaths had more positive
dose readings and higher mean cumulative radiate dose.

4.   There is variation of mean dose level among
victims of diseases but the category all cancers' is higher
than all non-malignant disease, and many cancers have high
doses as given in Table 3.

5.   Disease vs. dose level and proportion of victims
registering radiation are listed in Table 2. Of 17 neoplasm
types, 8 showed high radiation among victims with 79 more
deaths than expected by U.S. statistics, while 9 showed lower
radiation among victims with 79 fewer deaths than expected.
(Presumably the authors mean that the same number (670) of
deaths distributed over cancers according to overall U.S.
statistics would rearrange themselves as indicated).

Pages 5 to 12 contained description of controlled
analyses: This section describes tables which indicate how
average dose cancer victims and average dose non-cancer
victims appear when broken up by calendar year (two year
periods); employment year; pre-death year; exposure age.
Distribution of deaths among various cancers is given by
cumulative radiation doses as a function of age.

The following test is described (Table 11):
deaths are divided into five age groups and five radiation
levels: and in each category percent of cancer victims are
ranked within age bracket; a Spearman rank test on this
data is claimed to show significant cancer excess at 5%
level.

Page 12: "Special Tests of Radiation Association".
After a preliminary discussion a "three stage test" is
described:

1. Test for cancers with Definite Radiation
   Association compute

   $$V = \frac{N}{N-1} \ (S-R^2) \quad t = (r-R)/\sqrt{V(\frac{1}{n} \ \frac{1}{N})}.$$

2. If t > 2 consider.

   If null hypothesis rejected by $t \geq 2$ compute
   D, the "Doubling Dose", according to
   $$\frac{R+/D}{1+R/D} = r \ \text{or} \ D = (\frac{S-rR}{r-R}). \quad D = \frac{N-1}{N}\sqrt{V(\frac{1}{n} - \frac{1}{N})}/t-R$$

   $$R^2 + Rt\sqrt{V(\frac{1}{n} - \frac{1}{N})}.$$

3. Perform similar analyses for data from specific
   time periods or ages. The possible implications
   of results are described. Namely, one could
   estimate sensitivity of different tissues,
   intervals between initiation and death,

sensitivity ages, etc.

Pages 16 and 17: Radiosensitivity and critical
pre-death periods:

1. No radiation dose vs age before death,
    implications for non-cancers.

2. Significant results for certain cancer types
    very strong for bone marrow cancers, etc.

3. All cancers, twelve years before detah gave
    t = 2.4, RLS neoplasms t = 2.71, bone marrow
    t = 6.1 etc.

4. Some 26 years before death information at
    t = 1.8.

Page 18: Doubling Doses for Radiosensitive cancers
are noted. They are considerably lower than other estimates.

Excess mortality in some (by %) compared to U.S.
male death distribution (age corrected?) are seen to be less
than that predicted by doubling dose.

Pages 19 and 20: SMRs computed appear conservative.

Internal Radiation:

Data not in form for testing effect of IR. More
cancers occurred for IR group, much more external radiation
(factor of 15) doubling doses similar for two groups for

periods associated with positive findings. 17 RES was
47 (7-15 for bone marrow). Doubling doses obtained from
positive findings.

Page 21: Age sensitivity:

RES Neoplasms and solid tumors are compared with
non-cancers for cumulative radiation dose as a function of
age of death. Very little RES neoplasm data, but suggestions
of an exponential increase in sensitivity: Large findings
for selected solid cancers at certain ages. This is sug-
gestive of greater sensitivity to cancer-induction in early
and later adult life.

Page 23: Females

Less data, smaller proportion radiated; no excess
of cancer deaths among radiated, yet more radiation for can-
certs than non-cancers.

Spearman test gives significant correlation between
radiation and proportion of cancer deaths.

Page 24 contains discussion of estimates of exces
deaths; estimates a gument are given; there were approxi-
mately 25.8 excess deaths due to radiation.

Discussion (quoted in full): "A preliminary
analysis of the records relating to external radiation has
shown that there is sufficient data in the Hanford study
to (i) identify some of the more radiosensitive cancers;

(ii) quantify the radiosensitivity of these neoplasms;
(iii) obtain estimates of characteristic intervals between
initiation and death; and (iv) recognize the ages of
maximum and minimum sensitivity to the cancer-induction
effects of radiation.

Further analyses will be needed to rule out the
now remote possibility that the positive findings were
merely the result of the radiation exposures having
associations with other cancer-related factors. These
analyses will proceed in two directions. First, there
will be joing standardizat:·n for all the factors with
known or suspected radiation or cancer associations (e.g.,
exposure age, interval between hire and exposure, intervals
between exposure and death and depositions of radio-
active substances). Secondly, there will be an extension
of these analyses from non-survivors with certified
causes of death to other members of the monitored population,
or worke·s who are still alive at the time of follow-up.

Meanwhile cursory inspection of the records
relating to men who were still alive in 1973 (Table 1)
has shown that one of the reasons why the doses of ex-
ternal radiation have always been higher for survivors
than non-survivors[3] is because the survivors include a
disproportionately large number of men with positive urine

analyses (Table 25). -- This bias is due to an association
between high risk occupations and young recruits which
has caused the proportion of young recruits to be different
for: (i) singly and doubly monitored occupations;
(ii) men with positive and negative urine analysis and
(iii) survivors and non-survivors.

Since workers with positive urine analysis were
more often and more intensively exposed to external
radiation than other workers (Table 18), it is essential,
when comparing survivors with non-survivors to include
internal radiation among the controlling factors. This
necessity is clearly seen in Table 26 where five sets of
standardized radiation doses are shown for three groups
in Table 1 (survivors, non-cancers and cancers). For
instance even controlling for two factors simultaneously
(i.e., exposure year and cohort), still left the survivors
with a higher dose (127) than the non-cancers (79) or the
cancers (94), but when internal radiation was added to the
other controlling factors the standardized dose was not
only lower for non-cancers (84) than cancers (112), but
also lower for the survivors (101) than cancers.

Nevertheless, the absolute doses were higher for
the men who were still alive in 1973 than for the non-
survivors included in the present investigation, and for
Hanford workers as a whole the trend of radiation doses

(and proportions of exposed workers) is in an upward direction. Therefore, we should be prepared for future analysis of Hanford data to show both a wider range of cancers with definite radiation associations (due to better representation of cancers with long latent periods), and a higher proportion of radiation-induced cancers among the exposed workers."

Page 28: Appendix:

I. A log logistic model is described and claimed to justify the t-test used in the text.

II. Derivation of estimation of doubling dose via a linear model (not log logistic) is described.

III. The question of validity of t-test due to skewness of the distribution was tested by a Monte-Carlo technique. A probability of 6/1000 was found for bone marrow results instead of 10^{-4} from that table. For pancreatic tumors 10^{-2} was found empirically instead of 10^{-3} from t Table.

The remainder of the paper consists of tables (to be discussed later) and descriptions of other data sets.

SUMMARY OF CONCLUSIONS

1. The goals sought in this paper are unreasonable and mis-
   guided.

2. None of the stated conclusions are justified from the given
   data analysis.

3. No light is shed in the report on the important questions
   involved, namely: What can be said about the effect of low
   level radiation on the causing cancer.

4. No effort is made to consider the effects of other potentially
   cancerogenic factors; without such effort conclusions of the
   kind made here are not justified.

5. The methodology used is non-standard as statistics not
   appropriate for the apparent purposes here.

6. A majority of the conclusions in the report are based on eight
   cases of myeloma. While these eight cases deserve investigation
   in the contex' of what is known about epidemology of myeloma,
   they are insufficient to support the conclusions here.

7. Doubling doses obtained in this report imply a threshold model
   for radiation effect on causing cancer, rather than the linear
   model used by the authors. Their deriving them from a linear
   model could seem to contradict the linear model or their
   own analysis.

8. Use of Hanford deaths from other causes rather than survivor
   data is unjustified. Construction workers not receiving
   radiation hired in 1944 should have been removed from the data
   set.

9. Treatment of internal radiation is arbitrary and unjustified.

10. The variables used are not really appropriate for this problem.

11. No attempts are made to discuss background, or accuracy of data base.

12. No mention is made of the state of knowledge of effect of radiation as a cause of cancer.

13. No mention of other efforts to deal with Hanford data are made.

14. Data is grouped and categorized in arbitrary and unjustified ways.

15. No effort is made to access validity of the linear models used or to obtain confidence intervals on the doubling doses.