PNNL-27208

Pacific Northwest
NATIONAL LABORATORY

Proudly Operated by Battelle Since 1965

Analysis of Potential Performance Metrics for Fitness-for-Duty Programs

Technical Letter Report

January 2018

BE Ford
ME Lerchen
PD Whitney
AL Bunn
AC Dalton
KM Branch

Prepared for the U.S. Nuclear Regulatory Commission
under an Interagency Agreement with the U.S. Department of Energy
Contract DE-AC05-76RL01830

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from the
Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062;
ph: (865) 576-8401
fax: (865) 576-5728
email: reports@adonis.osti.gov

Available to the public from the National Technical Information Service
5301 Shawnee Rd., Alexandria, VA 22312
ph: (800) 553-NTIS (6847)
email: orders@ntis.gov <http://www.ntis.gov/about/form.aspx>
Online ordering: http://www.ntis.gov


This document was printed on recycled paper.
(8/2010)

Analysis of Potential Performance Metrics for Fitness-for-Duty Programs

Technical Letter Report

January 2018

BE Ford
ME Lerchen
PD Whitney
AL Bunn[1]
AC Dalton
KM Branch

Prepared for the U.S. Nuclear Regulatory Commission
under an Interagency Agreement with the
U.S. Department of Energy
Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory
Richland, Washington 99352

[1] Corresponding author

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

EXECUTIVE SUMMARY

This report evaluates the feasibility of using statistical analysis of drug and alcohol test results to develop reliable and useful performance metrics for licensees' and other entities' fitness-for-duty (FFD) programs being implemented under Title 10 of the *Code of Federal Regulations* (CFR) Part 26. The project team developed a methodological approach and a set of nonparametric statistical tests appropriate for the characteristics of the FFD testing data to detect statistically significant differences in estimates of positive result rates (PRRs) across different sites and different points in time. The team specifically identified statistically significant trends in industrywide FFD performance data (i.e., across all sites subject to 10 CFR Part 26 drug and alcohol testing requirements) and statistically significant performance deviations from both industrywide performance trends and historical performance trends at licensee sites. Although a number of sites' PRRs were significantly different from another baseline rate (e.g., industrywide average, year-on-year), however, because the results do not explain *why* a site's PRR is significantly different from the baseline rate, their interpretation should be approached with caution. Although review and analysis of test PRRs provides information about workplace conditions, PRRs, both overall and substance-specific, are of limited value for evaluating the performance of a licensee's FFD program. Additional data, analysis, and modeling would be required to understand what good performance means in the context of FFD drug and alcohol testing, and to identify potential explanatory variables that describe FFD program performance. Based on the findings from this effort, the project team identified a study of other potential metrics for future exploration into FFD program performance.

ACKNOWLEDGMENTS

The PNNL project team gratefully acknowledges the guidance and assistance provided by DaBin Ki, Lawrence Criscione, and Valerie Barnes of the U.S. Nuclear Regulatory Commission (NRC) Office of Nuclear Regulatory Research and by Paul Harris, of the Office of Nuclear Security and Incident Response. We also thank Mr. Brian Zaleski for providing the data for analysis and Ms. Fleur de Peralta and Mr. Dave Engel at PNNL for their peer review.

The accuracy of the information and the views presented in this report are the responsibility of the authors and do not necessarily represent the opinion of the NRC or of any particular individuals or licensees.

ABBREVIATIONS

ADAMS	Agencywide Documents Access and Management System
ARF	Annual Reporting Form
CFR	Code of Federal Regulations
C/V	contractor/vendor
FFD	Fitness-for-Duty
HHS	U.S. Department of Health and Human Services
NRC	U.S. Nuclear Regulatory Commission
PNNL	Pacific Northwest National Laboratory
PRR	positive result rate
SPTF	Single Positive Test Form

1.0 INTRODUCTION

In Title 10 of the *Code of Federal Regulations* (CFR) Part 26 (hereafter, Part 26), the U.S. Nuclear Regulatory Commission (NRC) requires licensees of nuclear facilities and other entities to implement a fitness-for-duty (FFD) program that includes a drug and alcohol testing program. As specified in §26.31(c), these licensees and other entities are required to administer drug and alcohol tests to subject individuals under the following conditions: pre-access, for cause, post-event, follow-up, and random. Subject individuals may be licensee employees or contractor/vendors (C/Vs).

The NRC directed the Pacific Northwest National Laboratory (PNNL) to evaluate the feasibility of using test result data from FFD programs by licensees and other entities subject to Part 26 testing requirements to develop performance metrics that could be used to evaluate the performance of the FFD programs. The NRC requested that PNNL propose a methodology for detecting statistically significant trends in historical drug and alcohol testing program performance data that could be used by NRC inspectors to identify programs that might warrant additional attention. This methodology should complement the analysis provided in the Summary of Fitness for Duty Program Performance Reports that the NRC prepares annually. The NRC is specifically interested in the following:

1. Statistically significant trends in positive result rates (PRRs) reported in the drug and alcohol testing program performance data across all sites subject to the Part 26 drug and alcohol testing requirements (hereafter, industrywide);
2. Statistically significant deviations from industrywide performance trends and historical performance trends at individual sites.

This report describes a methodological approach and a set of statistical tests to detect significant differences in estimates of PRRs or drug and alcohol-use patterns across different sites and different points in time. When applied to historical program data, these tests can provide statistical rigor to inferences regarding industry- and site-specific performance trends. They can also be applied to detect outliers in real time, as new FFD testing data become available. These tools could support discovery of meaningful trends and deviations from the norm for PRRs and could potentially be used for identifying programs that warrant additional attention.

2.0 FITNESS-FOR-DUTY PROGRAM REPORTING REQUIREMENTS

Licensees and other entities subject to the requirements of Part 26 are required to collect and submit performance information about their FFD programs to the NRC. The reporting requirements focus on drug and alcohol testing results, as well as significant FFD policy violations or programmatic failures.

Annual reporting information must be submitted to the NRC before March 1 for the previous calendar year (§26.717(e)). Licensees are required to submit annual reporting data in a consolidated report, as long as the report presents the data separately for each site (§26.717(f)). Reactor construction sites must report the same information under §26.417(b)(2). Licensees voluntarily submit this annual reporting information on NRC Form 891, "Annual Reporting Form for Drug and Alcohol Tests" (hereafter, ARF).[1] Individual drug and alcohol violations are reported on NRC Form 890, "Single Positive Test Form" (SPTF).[2] NRC makes submitted forms available through the web-based Agencywide Documents Access and Management System (ADAMS) information portal.[3] The NRC compiles the annual reportable events and 24-hour and 30-day reportable events for all regulated licensees and other entities in an annual Summary of Fitness for Duty Program Performance Reports.[4]

2.1 Required Drug and Alcohol Performance Data

As specified in §26.717(b), annual drug and alcohol testing reporting require the following information:

(1) The random testing rate;
(2) Drugs for which testing is conducted and cutoff levels, including results of tests using lower cutoff levels, tests for drugs not included in the HHS panel, and any special analyses of dilute specimens permitted under § 26.163(a)(2) [special analyses of dilute specimens];
(3) Populations tested (i.e., individuals in applicant status, permanent licensee employees, C/Vs);
(4) Number of tests administered and results of those tests sorted by population tested (i.e., individuals in applicant status, permanent licensee employees, C/Vs);
(5) Conditions under which the tests were performed, as defined in § 26.31(c);
(6) Substances identified;
(7) Number of subversion attempts by type; and
(8) Summary of management actions.

Under §26.717(d), any licensee or other entity who terminates an individual's authorization or takes administrative action on the basis of the results of a positive initial drug test for marijuana

[1] https://www.nrc.gov/reading-rm/doc-collections/forms/
[2] https://www.nrc.gov/reading-rm/doc-collections/forms/
[3] https://adams.nrc.gov/wba/
[4] http://www.nrc.gov/reactors/operating/ops-experience/fitness-for-duty-programs/performance-reports.html

or cocaine shall also report these test results in the annual summary by processing stage (i.e., initial testing at the licensee testing facility, testing at the HHS-certified laboratory, and MRO determinations). The report must also include the number of terminations and administrative actions taken against individuals for the reporting period.

2.2 Testing Conditions

The licensee is required to administer drug and alcohol tests to subject individuals under the following testing conditions (§26.31(c)):

(1) *Pre-access*. In order to grant initial, updated, or reinstated authorization to an individual, as specified in Subpart C of this part;
(2) *For cause*. In response to an individual's observed behavior or physical condition indicating possible substance abuse or after receiving credible information that an individual is engaging in substance abuse, as defined in § 26.5;
(3) *Post-event*. As soon as practical after an event involving a human error that was committed by an individual who is subject to this subpart, where the human error may have caused or contributed to the event. The licensee or other entity shall test the individual(s) who committed the error(s), and need not test individuals who were affected by the event whose actions likely did not cause or contribute to the event. The individual(s) who committed the human error(s) shall be tested if the event resulted in—
    (i) A significant illness or personal injury to the individual to be tested or another individual, which within 4 hours after the event is recordable under the Department of Labor standards contained in 29 CFR 1904.7, "General Recording Criteria," and subsequent amendments thereto, and results in death, days away from work, restricted work, transfer to another job, medical treatment beyond first aid, loss of consciousness, or other significant illness or injury as diagnosed by a physician or other licensed health care professional, even if it does not result in death, days away from work, restricted work or job transfer, medical treatment beyond first aid, or loss of consciousness;
    (ii) A radiation exposure or release of radioactivity in excess of regulatory limits; or
    (iii) Actual or potential substantial degradations of the level of safety of the plant;
(4) *Follow-up*. As part of a follow-up plan to verify an individual’s continued abstinence from substance abuse; and
(5) *Random*. On a statistically random and unannounced basis, so that all individuals in the population subject to testing have an equal probability of being selected and tested.

2.3 Reportable Events

Significant violations of FFD policy, significant FFD program failures, and errors in drug and alcohol testing are reportable events under the FFD program rather than under the provisions of 10 CFR 73.71, Reporting of safeguards events. These reporting requirements are specified in §26.719.

2.3.1 Policy Violations or Programmatic Failures

Significant FFD policy violations or programmatic failures under §26.719(b) must be reported to the NRC Operations Center by telephone within 24 hours after the licensee or other entity discovers the violation; these are:

(1) The use, sale, distribution, possession, or presence of illegal drugs, or the consumption or presence of alcohol within a protected area;
(2) Any acts by any person licensed under 10 CFR part 55 to operate a power reactor, as well as any acts by SSNM transporters, FFD program personnel, or any supervisory personnel who are authorized under this part, if such acts—
    (i) Involve the use, sale, or possession of a controlled substance;
    (ii) Result in a determination that the individual has violated the licensee's or other entity's FFD policy (including subversion as defined in § 26.5); or
    (iii) Involve the consumption of alcohol within a protected area or while performing the duties that require the individual to be subject to the FFD program;
(3) Any intentional act that casts doubt on the integrity of the FFD program; and
(4) Any programmatic failure, degradation, or discovered vulnerability of the FFD program that may permit undetected drug or alcohol use or abuse by individuals within a protected area, or by individuals who are assigned to perform duties that require them to be subject to the FFD program.

2.3.2 Drug and Alcohol Testing Errors

Required reporting for drug and alcohol testing errors under §26.719(c) are:

(1) Within 30 days of completing an investigation of any testing errors or unsatisfactory performance discovered in performance testing at either a licensee testing facility or an HHS-certified laboratory, in the testing of quality control or actual specimens, or through the processing of reviews under § 26.39 and MRO reviews under § 26.185, as well as any other errors or matters that could adversely reflect on the integrity of the random selection or testing process, the licensee or other entity shall submit to the NRC a report of the incident and corrective actions taken or planned. If the error involves an HHS-certified laboratory, the NRC shall ensure that HHS is notified of the finding.
(2) If a false positive error occurs on a blind performance test sample submitted to an HHS-certified laboratory, the licensee or other entity shall notify the NRC within 24 hours after discovery of the error.
(3) If a false negative error occurs on a quality assurance check of validity screening tests, as required in § 26.137(b), the licensee or other entity shall notify the NRC within 24 hours after discovery of the error.

3.0 DATA DESCRIPTION

NRC provided PNNL with the data analyzed in this report. Two datasets were evaluated: (1) aggregated annual performance data and (2) aggregated single positive test data. These are described in the following subsections.

3.1 Aggregated Annual Performance Data

The aggregated performance data includes tests performed under the drug and alcohol testing programs for calendar years (CYs) 2005–2013. The dataset covers two measures:

1. total number of tests performed
2. number of positive results, which includes positive drug and alcohol tests, results reported as positive due to adulteration or substitution, and refusals to test.

For each year, licensee, and facility, these measures are reported by test condition and employee type. Table 3.1 provides an overview of the data. Each licensee operates one or more sites.

Table 3.1. Data Overview

Variable Name	Data Type	Range
Number of tests performed per site per year	Count	0, 1, 2, 3,…n
Number of positive results per site per year	Count	0, 1, 2, 3,…n
Year	Date	2005–2013
Licensee	Text	company names
Site	Text	site names
Testing condition	Categorical	• pre-access • for cause • post-accident • follow-up • random
Employee type	Categorical	• licensee • C/Vs

Testing data are not available in the dataset for all sites for all years. During the CY2005-2013 reporting period, a subset of licensee reports was not available at the time of annual data collection and the data were subsequently excluded from the dataset for that year. Some sites' FFD programs were operational only during certain portions of the CY2005-2013 reporting period. Nuclear Fuel Services also submitted reports under the name "Erwin," and these are all identified as "Nuclear Fuel Services". Note that missing data for sites with operational FFD programs may skew statistical inferences about industry trends. Table 3.2 summarizes the missing data.

Table 3.2. Summary of Missing Data

Licensee	Site	Years Missing
Energy Northwest	Columbia	2006 (first half)
Exelon/Energy Solutions	Zion	2007–2010
Dominion Generation	Kewaunee	2005 (second half)
Dominion Generation	Millstone	2005 (second half)
Dominion Generation	North Anna	2005 (second half)
Dominion Generation	Surry	2005 (second half)
Nuclear Fuel Services	Erwin	2005 (first half), 2006
Pacific Gas & Electric	Humboldt Bay	2005 (first half), 2007–2013
South Carolina Electric & Gas	V.C. Summer Units 2 & 3 (construction)	2005-2010
Southern Nuclear	Vogtle Units 3 & 4 (construction)	2005-2008
Westinghouse	Westinghouse	2005–2006, 2010–2013

In total, the aggregated performance data includes 1,413,908 tests conducted and 8,975 positive results. The term "positive result" encompasses all violation types, including positive drug and alcohol tests and results reported as an FFD policy violation due to adulteration, substitution, or refusal to test. Note that this definition differs from that for positive result contained in the text of the regulation itself (§26.5 Definitions), which states the following:

> Positive result means, for drug testing, the result reported by a licensee testing facility or HHS-certified laboratory when a specimen contains a drug or drug metabolite equal to or greater than the cutoff concentration. A result reported by an HHS-certified laboratory that a specimen contains a drug or drug metabolite below the cutoff concentration is also a positive result when the laboratory has conducted the special analysis permitted in § 26.163(a)(2). For alcohol testing, a positive result means the result reported by a collection site when the BAC indicated by testing a specimen exceeds the cutoff concentrations established in this part.[1]

This report uses the more expansive definition in accordance with guidance from the NRC and thus FFD policy violations for adulteration, substitution, or refusal to test are included in the counts of positive results for this analysis.[2]

3.2 Additional Performance Data

Following submission of the initial draft report, the NRC provided a second workbook containing two additional datasets covering the period from 2009 through 2013. The first dataset, derived from licensees' voluntary SPTF (Form 890) submissions, provided more information on the

[1] 10 CFR §26.5 (HHS means U.S. Department of Health and Human Services; BAC means blood alcohol concentration).
[2] P. Harris (e-mail communication, August 14, 2014).

subset of tests with a positive result. Depending upon the type of test conducted (i.e., drug or alcohol), data are included on the employee's labor category (i.e., whether or not the test was conducted in connection with a 24-hour reportable event, substance(s) detected, blood alcohol content (if applicable), validity test results, subversion attempts (if applicable), and sanctions imposed as a consequence of the positive result). That dataset, which was analyzed in this report, contains 3,868 records with each record accounting for a positive result.

The other dataset in the workbook was derived from licensees' annual reports submitted voluntarily on ARF (Form 891) submissions and provides count data for drug and alcohol tests from CY2009–2013. For the most part, the data structure is similar to the aggregated performance data, including total test counts and positive result counts, although the data themselves cover a shorter time period. In addition, the ARF dataset includes random testing pool size; whether or not a licensee testing facility was used; whether limit-of-detection testing was used; and, in some cases, the number of dilute specimens. This dataset contains 322 records, encompassing 734,692 tests and 4,490 positive results. The dataset is not complete; data for all years are not included for all sites. This data covered similar information to the aggregated performance data but for fewer years; this dataset was not analyzed in this document.

3.3 Data Limitations

Although it would be of interest to calculate PRRs for individual substances and the rates for subversion attempts based on these additional data, such detailed analyses were not feasible due to differences between the original aggregated performance data and the ARF dataset received in the second workbook. As noted above, the 2009–2013 ARF data counts 4,490 total positive results, while there are only 3,868 records in the SPTF dataset, a difference of 622 positive results. Further, detailed data on substances and subversion attempts are contained only in the SPTF dataset, but the SPTF dataset accounts for less than 90 percent of positive results reported in the ARF dataset. The approach selected for this initial evaluation would require the records in the SPTF to be a complete subset of the positive results reported in the ARF dataset. Consequently, PNNL did not perform statistical tests at the level of individual substance or subversion type.

4.0 METHODOLOGY

The FFD performance study considered two types of analysis: 1) exploratory analysis to understand the characteristics and latent structures within the data, and 2) performance metrics analysis to identify historical trends and outliers in the data and to identify significant relationships between the PRR and licensee or test subject characteristics.

4.1 Exploratory Analysis

The exploratory analysis characterizes the data, identifies patterns in and relationship between the variables listed in Table 3.1, and provides insight into the underlying structure of the FFD drug and alcohol testing program performance data.

The exploratory analysis of the FFD program data used three statistical techniques:

- The relationships between the PRR and the testing condition (e.g., pre-access or for cause) as well as the PRR and the population tested (e.g., licensee and C/V employees) were evaluated using **descriptive statistics and data visualization**. These techniques were also used to investigate positive result counts by substance and subversion attempt type.
- The drug and alcohol testing results were further investigated by determining if a relationship existed between employee type and the PRR. A **chi-square test of independence** was used to determine if the relationships arose by random chance. This test is used for categorical data to detect whether there are statistically significant differences between the categorical variables.
- The data was also evaluated to determine whether any interpretable latent structure underlies the data that could be used as a metric for the FFD program in the future. **Factor analysis** was used to reduce the dimensionality of a dataset, making it possible to discern potentially meaningful latent structures.

4.2 Performance Metrics Analysis

The statistical tests are designed to identify PRRs that are outliers relative to an industrywide or year-on-year baseline. These statistical tests are intended to answer the following specific questions:

- Industrywide, are PRRs significantly different this year from PRRs for the previous year? Break out by testing condition and employee type.
- Is a site's PRR significantly different this year from the site's PRR for the previous year? Break out by testing condition (and employee type.
- Is a site's pre-access PRR significantly different from that site's random PRR? Break out by employee type.
- Is a site's random PRR significantly different from the combined PRR for post-event and for cause tests for that site? Break out by employee type.
- Is a site's PRR significantly different from the industry average PRR? Break out by testing condition and employee type.

Compared to the techniques used in the exploratory analyses, a different statistical approach was used to address these questions. Whereas in the exploratory analyses, the data was assumed to have a parameterized, underlying distribution (e.g., normal or chi-squared), for these questions addressing performance metrics, the approach was to use the FFD program data itself to construct nonparametric, empirical distributions to evaluate differences between the variables associated with the question. Nonparametric bootstrap resampling was used to construct a confidence interval for each point estimate of the difference between two PRRs (Fleiss et al. 2003; DiCiccio and Efron 1996; Stephenson et al. 2010). This confidence interval was used to determine if a statistically significant difference exists between estimates of those of the PRRs. Appendix A provides additional detail on the statistical methodologies.

5.0 RESULTS

The results of the exploratory analysis and analysis to identify potential metrics for the FFD drug and alcohol program are discussed in this section. Illustrations of confidence intervals from these analyses are presented in Appendix B.

5.1 Exploratory Analysis

The exploratory analysis evaluated the aggregated performance data and the SPTF data using the three statistical approaches discussed in Section 4.0. The results of each of those approaches are discussed here.

5.1.1 Results of Descriptive Statistics and Data Visualization

The descriptive summary presented here explores a few key dimensions of the drug and alcohol testing program performance data. The Summary of Fitness for Duty Program Performance Reports that the NRC prepares annually provide a far more detailed and comprehensive descriptive analysis of the performance data. The statistical analysis presented in the following section should complement the descriptive analysis contained in the NRC's summary reports.

Individual evaluations of the aggregated performance and SPTF datasets are described in the following sections.

5.1.1.1 Aggregated Performance Dataset Evaluation

In total, the aggregated performance dataset accounts for 1,413,908 tests conducted and 8,975 positive results from CY2005–2013. Table 5.1 provides the positive count, test count, and calculated PRR for all years and licensees in the aggregated performance data. For all testing conditions except random tests, the test counts are higher for C/Vs than for licensee employees. The for cause employee type PRRs are much higher than the PRRs for any other testing condition. For both licensee employees and C/Vs, the PRRs for the random testing are the lowest of all the testing conditions.

Table 5.1. Test Counts by Testing Condition and Employee Type

Testing Condition	Employee Type	Positive Count	Test Count	PRR
Follow-up	C/Vs	291	29,588	0.98%
	Licensee employees	142	25,827	0.55%
For cause	C/Vs	603	3,754	16.06%
	Licensee employees	179	2,453	7.30%
Post-event	C/Vs	41	4,397	0.93%
	Licensee employees	8	3,059	0.26%
Pre-access	C/Vs	5,981	724,120	0.83%
	Licensee employees	261	91,159	0.29%
Random	C/Vs	932	190,263	0.49%
	Licensee employees	537	339,288	0.16%

Figure 5.1 illustrates that the distribution of PRRs, aggregated by site-year, is bimodal, with modes at zero and 0.6 percent, and a median of 0.54 percent.[1] In just under 8 percent of site-years, there were no positive results reported. At the higher end of the distribution shown in Figure 5.1, two sites reported PRRs across all testing conditions, employee types at or above 2 percent: Vogtle Units 3 & 4 in 2011 (2.0 percent) and V. C. Summer Units 2 & 3 in 2013 (2.3 percent). In its annual *Summary of Fitness for Duty Program Performance for Calendar Year 2013* (NRC 2014), NRC notes:

> The two reactor construction sites conducted 5.8 percent of the industry tests in CY2013, but accounted for 17.7 percent of the positive test results and testing refusals. V.C. Summer (Units 2 and 3) and Vogtle (Units 3 and 4) reported performing 9,394 tests with 178 positive drug and alcohol tests results and testing refusals.

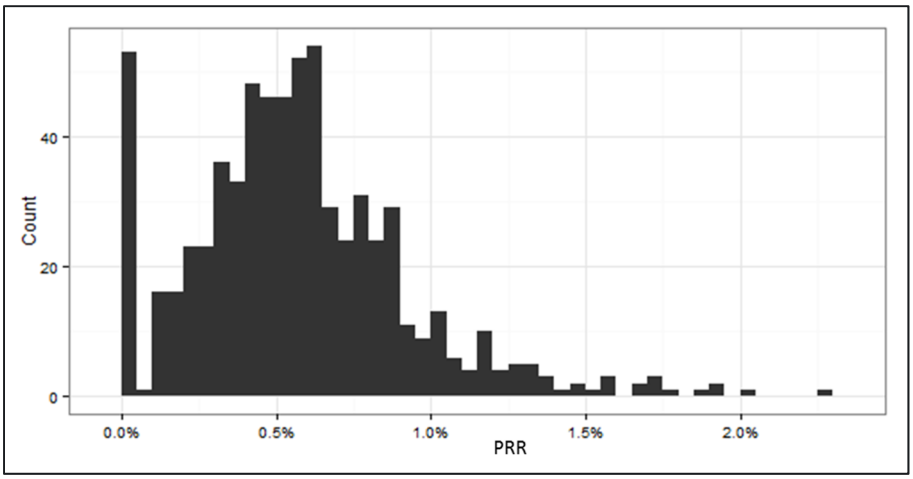


Figure 5.1. Distribution of PRRs, Aggregated by Site-Year

Figure 5.2 shows the commercial nuclear power sites across the United States, and identifies the six sites where the PRR across all testing conditions, employee types, and years is greater than 0.9 percent. Among these, the PRRs for licensees South Carolina Electric & Gas and Southern Nuclear exceed 1 percent, and these two licensees operate four of these six sites.

Figure 5.3 shows the PRR by year aggregated across the industry for each testing condition and employee type. For all testing conditions, except for cause, annual PRRs vary between 0 and 1 percent, with little variation across years for either C/Vs or licensee employees. For cause PRRs, however, vary between 10 and 25 percent for C/Vs, and between 5 and 12 percent for licensee employees. This is to be expected, as the testing population for for cause tests consists of individuals whose behavior or physical condition indicates possible substance abuse or about whom credible information has been received indicating that the individual is engaging in substance abuse.[2] Some temporal correlation appears to exist between the for cause PRRs for C/Vs and licensee employees; however, this is likely due to random variation

[1] "Site-year" refers to the individual data at a site for one year. There should be nine site-years of data for each site (CY2005-2013), although as noted in Section **Error! Reference source not found.**, data are not available for some site-years (Table 3.2).
[2] 10 CFR §26.31(c)(2)

rather than a structural pattern because sites conduct so few for cause tests that the PRR can vary dramatically from year to year. It is notable that in 2009, when the for cause PRR appears to spike for both employee types, a greater number of sites than normal reported at least one for cause positive result.

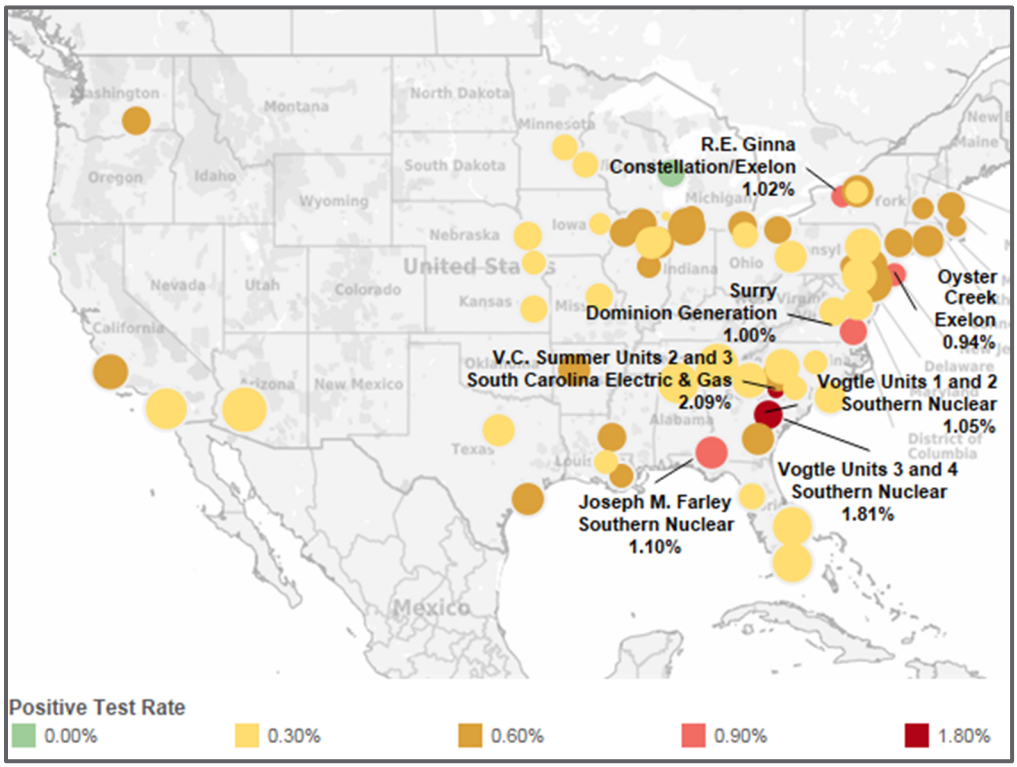


Figure 5.2. Map of Sites with PRR over All Testing Conditions and Years

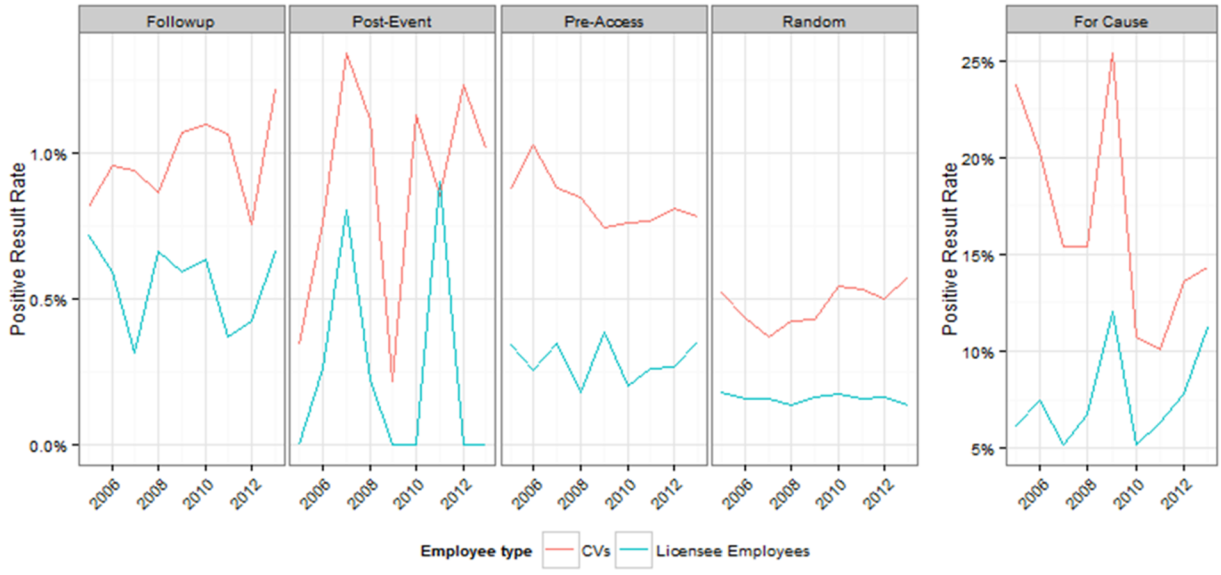


Figure 5.3. Industrywide PRRs by Year, Testing Condition, and Employee Type (Note that two different scales are used to present the data)

5.1.1.2 Single Positive Test Form Dataset Evaluation

The SPTF dataset included 3,868 records from CY2009–2013, as shown in Figure 5.4. Figure 5.4 shows that marijuana is the substance most commonly detected for positive test results, followed by alcohol and cocaine. Amphetamines and methamphetamines slightly increased their respective shares of total positive results from 2009 to 2013. All other drugs constitute less than 2 percent of all positive results.

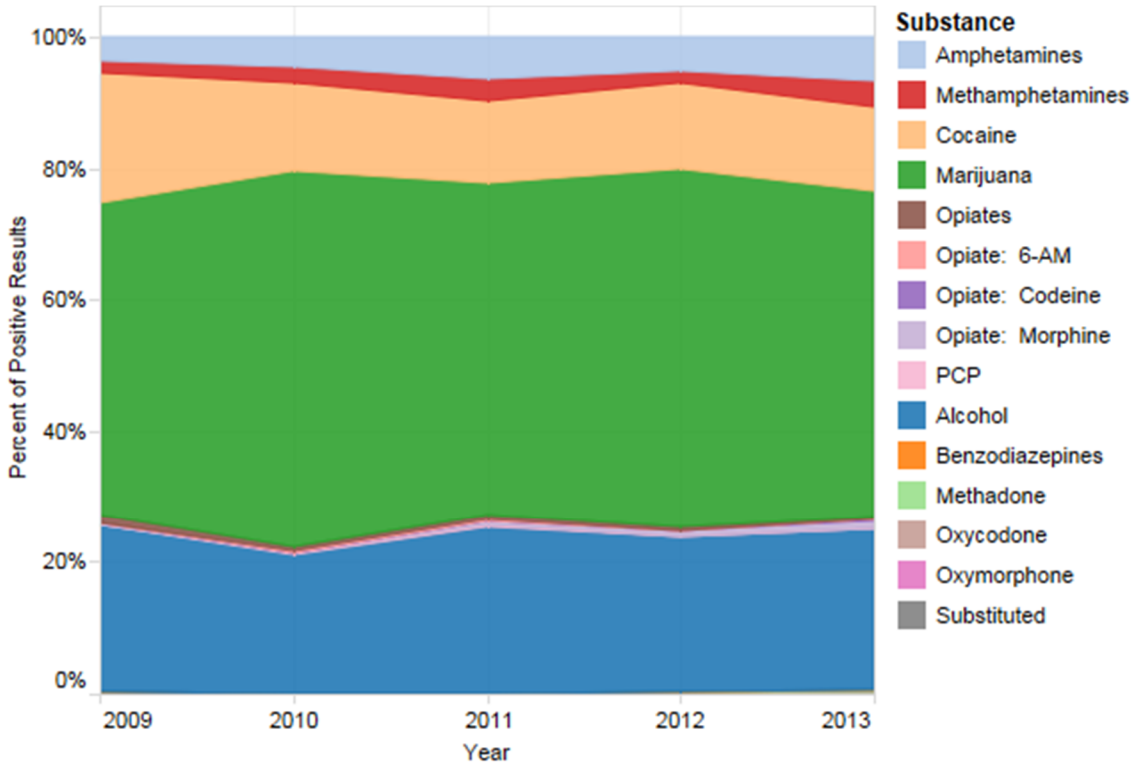


Figure 5.4. Percent of Positive Results by Year and Substance

Table 5.2 shows positive result counts for the most common substances at the sites with the top ten highest positive result counts for the CY2009–2013 period. Consistent with the high PRRs shown in Figure 5.2, Vogtle Units 1–4, Joseph M. Farley, and V.C. Summer Units 2 & 3 all top the list for most total positive results. Results for Vogtle Units 3 & 4 are particularly noteworthy, with more than double the number of total positive results reported for the site with the next highest count, Joseph M. Farley. In general, the rank order of substances at individual sites follows that of the larger industry, with marijuana-related positive results ahead of all others at every site.

Table 5.2. Count of Positive Results by Most Common Substance for Sites with Top Ten Highest Positive Result Counts

Site	Total	Alcohol	Marijuana	Cocaine	Amphetamines	Methamphetamines
Vogtle Units 3 & 4	336	46	138	55	25	15
Joseph M. Farley	166	20	112	23	10	10
Vogtle Units 1 & 2	163	15	98	28	15	11
V.C. Summer Units 2 & 3	136	30	64	21	9	5
Salem/Hope Creek	133	42	62	17	0	0
E.I. Hatch	120	20	77	24	1	1
Nine Mile Point	104	30	47	18	1	0
Arkansas Nuclear One	87	14	44	3	15	0
Limerick	86	26	29	12	7	2
Palo Verde	82	21	43	10	7	6

As shown in previous figures, certain sites reported a disproportionately high number of positive results over the CY2009–2013 period. Table 5.3 shows counts of total subversion attempts and refusals to test for sites with the top ten highest subversion attempt counts. Again, Vogtle Units 1–4, V.C. Summer Units 2 & 3, and Joseph M. Farley are at or near the top of the list. Vogtle Units 3 & 4 report more than four times more subversion attempts than Perry, the site with the next highest count on the list. Note that in some cases, a single positive result might be characterized with one or more subversion descriptors.[3] For example, if a donor submitted an initial specimen that was out of the acceptable temperature range, and during the subsequent collection under direct observation, a plastic bottle was discovered in the donor's underwear, this subversion attempt could be characterized with two subversion descriptors (i.e., temperature and paraphernalia) in the dataset.

Table 5.3. Count of Total Subversion Attempts and Refusals to Test for Sites with Top Ten Highest Subversion Attempt Counts

Site	Total Subversion Attempts	Refusals to Test
Vogtle Units 3 & 4	126	78
Perry	37	12
V.C. Summer Units 2 & 3	36	16
Vogtle Units 1 & 2	36	15
Arkansas Nuclear One	20	9
Joseph M. Farley	20	7
Salem/Hope Creek	18	11
Limerick	17	14
South Texas Project	16	7
River Bend	15	4

[3] In the SPTF dataset, subversion descriptors include: Subversion-No Specimen 1, Subversion-No Specimen 2, Subversion-Temperature, Subversion-Characteristics, Subversion-Paraphernalia, Subversion-Invalid Test, Subversion-Refused Direction, and Subversion-Other.

5.1.2 Evaluation of PRRs across Employee Types Using Chi-Square Test of Independence

The industrywide descriptive statistics suggest a distinction in PRRs across employee types. As shown in Table 5.1, for each testing condition, the PRR for C/Vs is roughly two or more times as high as the PRR for licensee employees. One way to test the hypothesis that there is a relationship between PRR and employee type is to conduct a chi-square test of independence. The relation between these variables, PRR and employee type, was significant for all testing conditions.[4] This result supports the conclusion that there is a statistically significant difference in PRRs between C/Vs and licensee employees.

The industrywide statistics also suggest a distinction in PRRs between construction and non-construction sites. As shown in Figure 5.2, the two sites with the highest overall PRRs are V.C. Summer Units 2 & 3 and Vogtle Units 3 & 4. These sites were both under construction during some of the years evaluated. A chi-square test of independence was performed to investigate the relationship between PRR and construction site status. The relation between these variables was significant for all testing conditions.[5] This result supports the conclusion that there is a statistically significant difference in PRRs between construction and non-construction sites.

5.1.3 Results of Factor Analysis

To further explore the presence of any latent patterns of variation or underlying structure in the aggregated performance data (CY2005–2013), an exploratory factor analysis was performed. Exploratory factor analysis is a mathematical technique that groups data and variables in such a way that potentially meaningful patterns of relationships among the variables might emerge (Rummel 1970; Yong and Pearce 2013). (See Appendix A for a technical note on factor analysis.) The goal of the factor analysis was to provide an additional means of examining the aggregated performance data and to generate added insight beyond the descriptive statistics and data visualization analyses discussed in Section 4.0 and other subsections in Section 5.0. However, the validity of factor analysis is affected by the quality of the dataset (e.g., missing data, as is the case for this dataset (Table 3.2)) and limited by the scope of information covered. Thus, caution should be taken when interpreting the results, particularly when extending beyond the scope substantiated by the dataset.

5.1.3.1 Data Transformation

To conduct an exploratory factor analysis, the aggregated performance data (CY2005–2013) were transformed to produce a PRR for each site, testing condition, and employee-type combination. The unit of analysis of the transformed data is the individual site, which allows for

---

[4] For random tests, $\chi^2(1, N = 529{,}551) = 484.5$, p=0.0000; for for cause tests, $\chi^2(1, N = 6{,}207) = 103.5$, p =0.0000; for post-event tests, $\chi^2(1, N = 7{,}456) = 12.5$, p= 0.0004; for follow-up tests, $\chi^2(1, N = 55{,}415) = 33.5$, p=0.0000; for pre-access tests, $\chi^2(1, N = 815{,}279) = 310.4$, p=0.0000.

[5] For random tests, $\chi^2(1, N = 529{,}551) = 219.1$, p=0.0001; for for cause tests, $\chi^2(1, N = 6{,}207) = 11.6$, p =0.0007; for post-event tests, $\chi^2(1, N = 7{,}456) = 3.0$, p= 0.0810 [note: significant at 90% confidence level]; for follow-up tests, $\chi^2(1, N = 55{,}415) = 16.3$, p=0.0001; for pre-access tests, $\chi^2(1, N = 815{,}279) = 359.5$, p=0.0000.

a direct comparison of PRRs across sites, including those sites with missing years of data. The resulting dataset contains one observation for each site (N = 81). The observation consists of ten independent variables. Each variable is a combination of one testing condition (i.e., pre-access, random, for cause, post-event, or follow-up) and an employee type (i.e., licensee employee or C/V).

5.1.3.2 Factor Analysis Results

Factor analysis models can produce one or more factors, with each factor potentially representing a latent structure among the variables. For data-exploratory purposes, the transformed data produced a two-factor model.[6] The results of the two-factor model are shown in Table 5.4. Factor loadings for all variables are reported in columns "Factor 1" and "Factor 2." Communality and uniqueness measures are also reported for each variable.

Factor loadings measure the correlation between variables and the factor(s). That is, factor loadings give a quantitative indication of the extent to which variables play a role in characterizing a latent structure in the data (i.e., a factor). For example, the high loadings for pre-access C/V (0.90) and random C/V (0.62) on Factor 1 suggest that two testing conditions: pre-access and random, and one employee type (i.e., C/V) play a large role in characterizing a latent structure in the data (i.e., Factor 1). For Factor 2, the high loadings for pre-access licensee employee (0.70) and random licensee employee (0.65) suggest these two testing conditions (i.e., pre-access and random) and one employee type (licensee employee) play a large role in characterizing another latent structure (i.e., Factor 2).

Table 5.4. Factor Loadings for Two-Factor Model

Variable Name	Factor 1	Factor 2	Communality	Uniqueness
Pre-access Licensee employee	0.21	0.70	0.54	0.46
Pre-access C/V	0.90	-0.09	0.82	0.18
For cause Licensee employee	-0.01	0.14	0.02	0.98
For cause C/V	0.13	0.08	0.02	0.98
Post-event Licensee employee	-0.05	0.21	0.05	0.96
Post-event C/V	0.23	0.07	0.06	0.94
Follow-up Licensee employee	0.11	0.15	0.03	0.97
Follow-up C/V	0.58	0.04	0.34	0.66
Random Licensee employee	0.05	0.65	0.42	0.58
Random C/V	0.62	-0.16	0.41	0.59

[6] The analysis was run for one-, two-, three-, four-, and five-factor models. Factor selection was based on factor loading values and the screen test (D'agostino and Russell 2005; Cattell 1966), a graphical method to determine which factors in the model should be retained. The three-, four-, and five-factor models resulted in low or zero loadings on each factor above two, adding no additional explanatory value to the model, and are, therefore, not discussed here. In addition, a separate data transformation was done for a one-factor model, which yielded results no more informative than the two-factor model discussed here.

Communality measures the amount of variation in a variable explained by a factor or factors. For each variable, communality is calculated as the sum of the squared factor loading values across all the factors. For instance, the communality value for the variable, pre-access licensee employee, is 0.54, indicating that about 54 percent of the variation in the variable is explained by Factor 1 and Factor 2 in this two-factor model.

Uniqueness measures the portion of variation in the variable not explained by a factor or factors. It is computed as the difference between 1 and the communality value (i.e., uniqueness value = 1 – communality value). For example, pre-access C/V has a uniqueness value of 0.18, suggesting 18 percent of the variation in pre-access C/V is explained neither by Factor 1 nor Factor 2.

In summary, the two-factor model detailed in Table 5.4 (RMSR = 0.07)[7] shows high factor loadings on the pre-access and random testing condition variables for each factor, but those high loadings are separated by employee type across the two factors, with Factor 1 being highly loaded for the C/V variables (i.e., pre-access C/V and random C/V) and Factor 2 highly loaded for the licensee employee variables (i.e., pre-access licensee employee and random licensee employee). Factor 1 also has a fairly high loading on the follow-up testing condition variable for C/Vs (i.e., follow-up C/V), although not as high as for pre-access or random testing conditions. These results suggest that across all sites, there is a substantive difference between PRRs for C/Vs and licensee employees for pre-access and random tests for drugs and alcohol.

5.1.3.3 Observations of Factor Analysis Results

The two-factor model results show that the type of employee (licensee employee or C/V) is correlated with PRRs for two testing conditions: pre-access and random, across all sites included in the aggregated performance data. These findings highlight several attributes of the PRRs across testing conditions and employee types. First, the factor analysis results offer further evidence that there is a substantive difference in the positive rates for C/Vs and licensee employees. In combination with the descriptive statistics that show higher rates for C/Vs and the results of the chi-square test, the differentiation between the two employee types supports the conclusion that C/Vs tend to have higher PRRs than licensee employees. In addition, because the high loadings were found on the pre-access and random testing condition variables for each factor in the model, this indicates that the PRRs for those testing conditions are strongly correlated. By comparison, the low loadings on the other testing condition variables suggests a lack of correlation among for cause, post-event, and follow-up testing. It is worth noting that while industrywide random PRRs for both employee types is lower than the pre-access PRR, those sites with higher pre-access PRRs tend to also have higher random PRRs. However, since the populations that were subject to pre-access and random drug and alcohol testing were different, it remains unclear whether the correlation between these sites' pre-access PRRs and random PRRs is meaningful for evaluating their FFD programs' performance.

[7] RMSR (root mean square residual) is a measure for goodness of model fit. It indicates the average matrix correlation among the variables that is explained by the latent pattern or patterns (a factor or factors) in the data.

5.2 Performance Metric Analysis

The following section describes the results of statistical analyses conducted to answer the set of questions presented in Section 4.2 that are designed to identify PRRs that are outliers relative to an industrywide or year-on-year baseline. Appendix A provides a more detailed discussion on interpreting the results using the nonparametric bootstrapping analysis. Appendix B presents the figures illustrating the results of the statistical tests for all but the first question.

5.2.1 Industrywide, are PRRs for This Year Significantly Different This Year from PRRs for the Previous Year?

Industrywide, there is little significant year-on-year variation in PRRs, by testing condition or employee type, with the exception of for cause tests. In 2009, for cause PRRs increased significantly for both C/Vs (10 percent) and licensee employees (5 percent), then decreased in 2010 to below the 2008 rates. Figure 5.5 shows this analysis (note the difference in scales for the testing conditions). While there are significant differences for the PRRs by year, there is no trend in the testing conditions.

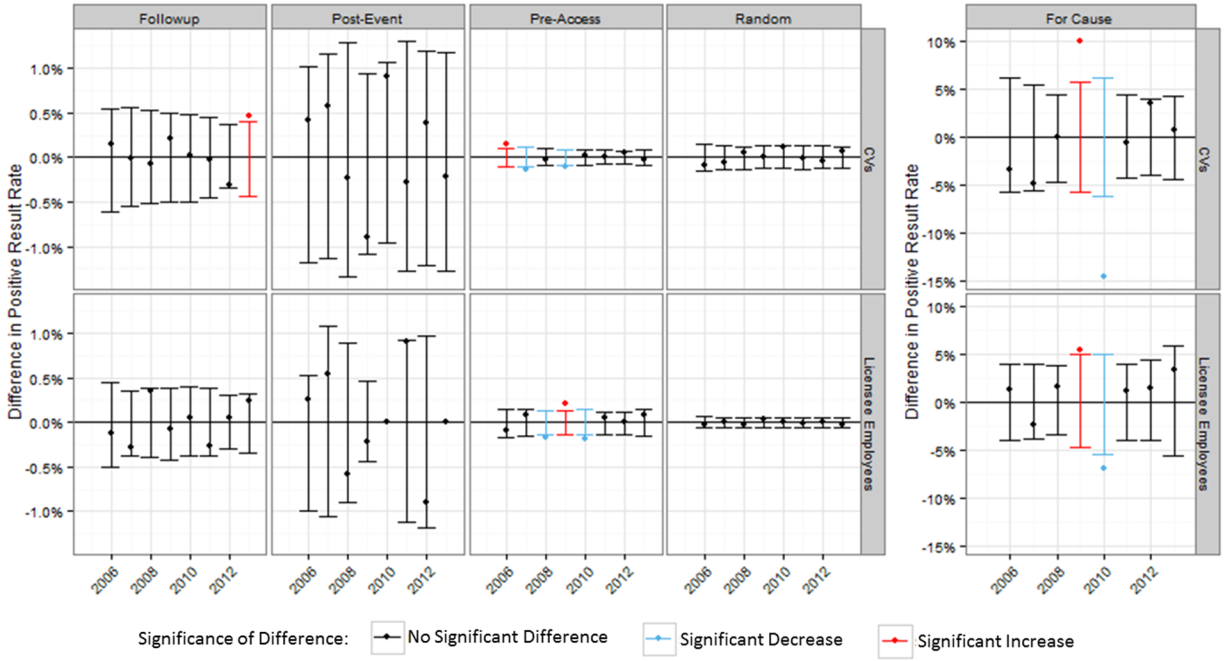


Figure 5.5. Industrywide, Year-On-Year Difference in PRR by Testing Condition and Employee Type, Bootstrapped Confidence Intervals (Note that two different scales are used)

5.2.2 Is a Site's PRR Significantly Different This Year from That Site's PRR for the Previous Year?

5.2.2.1 Site-Specific, Year-on-Year Difference in Pre-Access Positive Rate, C/Vs

Across all sites, pre-access PRRs tend to vary year-to-year by 0 to 3 percent for C/Vs. Visually, the time series appear to show a fair amount of annual variability in PRRs. However, only 23 out of 711 site-years show a statistically significant increase. Among site-years with a significant increase, the average year-to-year difference is 1.08 percent. The largest statistically

significant increase is 3.08 percent at Crystal River from 2006 to 2007, which went from zero positive results in the preceding year to four in the following. The top five largest statistically significant increases are listed in Table 5.5.

Table 5.5. Top Five Largest Site-Specific, Year-On-Year Increases in Pre-Access PRRs for C/Vs

Year	Company	Site	Difference in PRRs
2007	Progress Energy	Crystal River	3.08%
2012	Southern Nuclear	Vogtle Units 1 & 2	2.10%
2013	South Carolina Electric & Gas	V.C. Summer Unit 1	1.72%
2010	Entergy Nuclear	Grand Gulf	1.69%
2006	Detroit Edison	Fermi 2	1.66%

5.2.2.2 Site-Specific, Year-on-Year Difference in Pre-Access Positive Rate, Licensee Employees

With few exceptions, there is far less variability in pre-access PRRs for licensee employees than for C/Vs. The average absolute difference year-to-year is 0.4 percent, and most spikes are attributable to non-zero positive result counts coupled with low numbers of total tests. Of 711 site-years, 550 have zero positive results for licensee employees. Only two sites show a statistically significant increase in PRR: Oyster Creek from 2011 to 2012 and St. Lucie from 2006 to 2007.

5.2.2.3 Site-Specific, Year-on-Year Difference in Random Positive Rate, C/Vs

The time series for random positive rates for C/Vs appear to show some variability within a range of 0 to 3 percent across sites; however, most of these differences are not statistically significant. The apparent variability is due mostly to small changes in very low positive result counts. For example, random positive result counts for C/Vs at FitzPatrick vary between zero and two over the entire period. Only one site-year shows a significant increase: Comanche Peak from 2010 to 2011 (1.1 percent).

5.2.2.4 Site-Specific, Year-on-Year Difference in Random Positive Rate, Licensee Employees

Across all sites, little variability exists in random PRRs for licensee employees. At most sites, positive rates vary between 0 and 1 percent and year-on-year differences are not statistically significant. In more than half of all cases, zero positive results were reported. Only one site-year shows a statistically significant increase in PRR: Columbia from 2008 to 2009 (0.8 percent).

5.2.2.5 Site-Specific, Year-on-Year Difference in for Cause Positive Rate, C/Vs

For cause PRRs for C/Vs show considerable variability, ranging from 0 to 100 percent for many sites. The apparently large variability is due in many cases to very low numbers of total tests conducted for a given site-year. With such low total test counts, PRRs can vary substantially; however, confidence intervals constructed from a limited number of observations tend to be very wide relative to the magnitude of the difference estimate, leading to few findings of statistical

significance. Of the five cases of a significant year-on-year increase in the PRR, all but one had fewer than 20 total tests conducted for the site-year. Overall, the average number of for cause tests conducted for C/Vs is just over five per site-year.

5.2.2.6 Site-Specific, Year-on-Year Difference in for Cause Positive Rate, Licensee Employees

Similar to rates for C/Vs, for cause PRRs for licensee employees appear to show a wide range of variability across sites and years, ranging from 0 to 100 percent. Again, this is due primarily to the small numbers of total tests conducted for a given site-year. Of 711 site-years, 593 have five or fewer total for cause tests conducted for licensee employees. There are no statistically significant year-on-year increases in the PRR at any site. The spike in the industrywide for cause PRR observed from 2009 cannot be attributed to a single site or group of sites.

5.2.3 Is a Site's Pre-access PRR Significantly Different from That Site's Random PRR?

5.2.3.1 Site-Specific, Within-Year Difference in Pre-Access and Random Rates, C/Vs

Across all sites, pre-access PRRs tend to be either higher than random PRRs or not significantly different. This could reflect systematic differences in terms of drug or alcohol-use behavior between the pre-access test population (i.e., C/V job applicants and those applying for reinstatement of access authorization) and the incumbent contractor and vendor population (that excludes individuals with positive results from pre-access testing).

5.2.3.2 Site-Specific, Within-Year Difference in Pre-Access and Random Rates, Licensee Employees

Similar to rates for C/Vs, across all sites, pre-access PRRs for licensee employees tend to be either higher or not significantly different than random PRRs. There are no cases where the random positive test rate is found to be significantly higher than the pre-access PRR.

5.2.4 Is a Site's Random PRR Significantly Different from the Combined PRR for Post-event and for Cause Tests for That Site?

5.2.4.1 Site-Specific, Within-Year Difference in Random and (for cause+post-event) Rates, C/Vs

In all cases, random PRRs are significantly lower than, or not significantly different from, combined for cause and post-event rates for C/Vs.

5.2.4.2 Site-Specific, Within-Year Difference in Random and (for cause+post-event) Rates, Licensee Employees

The interpretation of the comparison between random and combined for cause and post-event PRRs for licensee employees is similar to that for C/Vs. There are no cases where the random PRR is significantly higher than the combined for cause and post-event rate.

5.2.5 Is a Site's PRR Significantly Different from the Industry Average PRR?

5.2.5.1 Pre-Access PRRs

Within-Year Difference between Site and Industry Average Pre-Access PRR, C/Vs

There is a fair degree of variability across sites in terms of pre-access testing performance for C/Vs relative to the industry average. Several sites consistently perform worse than the industry average, while a greater number of sites show positive rates significantly lower than or not significantly different from the industry average. Three sites in particular—Vogtle Units 1 & 2, Vogtle Units 3 & 4, and Joseph M. Farley—show more years with PRRs significantly higher than the industry average than years with no significant difference and years with significantly lower rates than the industry average combined. The average PRR industrywide across all years is 0.8 percent.

Within-Year Difference between Site and Industry Average Pre-Access PRR, Licensee Employees

With a single exception, pre-access PRRs for licensee employees do not tend to significantly exceed the industry average at any site. Only Beaver Valley shows more years with PRRs significantly higher than the industry average than years with no significant difference. An important part of the explanation may be that in 77 percent of cases, there are zero positive pre-access tests reported for licensee employees at a site in a given year. No sites showed any years with significantly lower rates than the industry average. The average PRR industrywide across all years is 0.3 percent.

5.2.5.2 Random PRRs

Within-Year Difference between Site and Industry Average Random PRR, C/Vs

Despite the appearance in the time series graphs of a fair amount of variability across all sites relative to the industry average, there are few statistically significant positive differences in random PRRs for C/Vs. V.C. Summer Units 2 & 3 and Vogtle Units 3 & 4 consistently show PRRs significantly higher than the industry average. The large spike at Dominion Generation Corporate in 2006 is due to 1 positive result out of 11 total random tests conducted on C/Vs in that year, which exceeds the industry average of 0.5 percent by 8.5 percent.

Within-Year Difference between Site and Industry Average Random PRR, Licensee Employees

Overall, random testing performance for licensee employees relative to the industry average is similar to random testing performance as reported for C/Vs. There are few cases where a site's random PRR for licensee employees significantly exceeds the industry average. The large, significant spike at Zion in 2005 is due to 1 positive result out of 27 total, which exceeds the industry average of 0.2 percent by 3.5 percent.

5.2.5.3 For Cause PRRs

Within-Year Difference between Site and Industry Average for Cause PRR, C/Vs

Substantial variability exists across sites in terms of for cause testing performance for C/Vs relative to the industry average. Although industrywide there appears to be a fairly steady for cause PRR for C/Vs, at individual sites these rates vary wildly based on the handful of individuals suspected of being in violation of FFD regulations each year. Because total for cause test counts are so low, even if the for cause PRR at a site is significantly different from the industry average, caution should be exercised when interpreting the results of this test.

Within-Year Difference between Site and Industry Average for Cause PRR, Licensee Employees

The comparison of sites' for cause PRRs for licensee employees to the industry average shows the same trend as for C/Vs. Again, because total for cause test counts are so low, even if the for cause PRR at a site is significantly different from the industry average, caution should be taken when interpreting these results.

6.0 DISCUSSION

This study examined the feasibility of using statistical analysis of drug and alcohol program data to develop reliable and useful performance metrics for licensees' and other entities' FFD programs under Part 26. The approach relied upon a set of nonparametric statistical tests to detect significant differences in estimates of PRRs from drug and alcohol testing across different sites and time periods. Results from this approach showed that there are statistically significant trends in industrywide FFD performance data and deviations from both industrywide performance trends and historical performance trends at licensee sites. However, these findings have limited value in performance measurement because they do not provide insight into *why* a site's PRR is significantly different from the baseline rate.

At the close of this study, NRC remains interested in establishing performance metrics for drug testing based on measures derived from program compliance data reported to NRC.[1] Measures related to the program's role in detecting or deterring drug use will be challenging if based solely on the relatively sparse PRRs. Looking beyond PRRs, other measures should be explored such as those related to detection (e.g., number of tests exceeds fifty percent of covered population (§ 26.31(d)(2(vii)), based on goals (e.g., PRR for covered population is less than 0.5%), or possible new measures discovered through other data analytic techniques.[2] Performance metrics may rely on multiple measures to inform different dimensions of a particular aspect of overall program performance.

6.1 Data Limitations and Caveats

The statistical tests described in this report can be used to compare PRRs across different sites and different points in time. They provide a basis for identifying statistically significant outliers relative to a predefined baseline PRR. As such, they could identify a site for additional attention by NRC inspectors. Unfortunately, the datasets do not support causal inference or meaningful trend analysis using these tests.

The data has small population/sample sizes with very small or zero positive result counts for some test results. These problems are more pronounced when examining data by testing condition (specifically for cause, post-event, and follow-up tests) where nearly two-thirds of sites conducted ten or fewer tests in a given year. Further, 99 percent of data recorded by site and by year has three or fewer positive results for the testing conditions of for cause, post-event, follow-up tests. These data limitations are more pronounced when the SPTF dataset is

[1] There is overlap between measures and metrics. Both can be qualitative or quantitative, but what distinguishes them is important. Measures are concrete, usually measure one thing, and are quantitative in nature (e.g., I have five apples). Metrics describe a quality and require a measurement baseline (e.g., I have five more apples than I did yesterday). Measures are useful for demonstrating workloads and activity, and metrics are useful for evaluating compliance, processes effectiveness, and measuring success against established objectives. https://www.cio.gov/2011/08/18/performance-metrics-and-measures/

[2] DOD, 2011, Status of Drug Use in the Department of Defense Personnel, available at: http://prhome.defense.gov/Portals/52/Documents/RFM/Readiness/DDRP/docs/6b%20FY%202011%20Annual%20Drug%20Use%20Status%20Report.pdf.

subdivided to test PRRs for individual substances or subversion attempt types. For most years at a site, the positive result counts are very small and often zero for a single substance or subversion attempt type. Given such low positive result counts, an incremental difference in the count can result in a dramatic change in the PRR that is not informative regarding FFD program performance. Further, as discussed in Section 3.3, the datasets provided for analysis had gaps. Thus, calculating PRRs for individual substances and the rates for subversion attempts was not feasible and, consequently, was not done.

Under §26.717 and §26.719, licensees must report positive drug testing results and collect, compile, analyze, and report aggregated annual statistical drug and alcohol program performance data (reported on the Single Positive Test Form (Form 890) and Annual Reporting Form for Drug and Alcohol Tests (Form 891)). Data analyzed in this report is a subset of the data reported by licensees on Form 891. Due to limitations in the datasets available for this report (see Section 3.3), an integrated analysis of data from Forms 890 and 891 was not performed nor was data between the two information sources correlated. Collecting and correlating the related data should provide additional insights into drug testing program performance across the industry, over time, and at individual locations.

6.2 Positive Result Rates Do Not Inform Performance

The test results indicate whether a site's PRR is significantly different from another baseline rate (e.g., industrywide average or year-on-year); however, the results do not support conclusions regarding the quality or performance of the licensee's testing programs. This is because it is not possible to determine from PRR alone whether lower rates indicate excellence in program implementation (via screening out potential drug users, effectively discouraging drug use) or whether they indicate lax performance (i.e., failure to prevent and detect subversion or inadequately preventing individuals from anticipating when testing will occur).[3] Further, differences in PRR across sites or over time may reflect differences in the drug- and alcohol-use characteristics of the worker population and/or different sampling techniques used in various types of drug testing, which cannot be attributed to differences in the quality of the FFD programs. As such, PRR alone is an inadequate performance metric against which the merit of FFD programs can be reliably measured.

6.3 Future Performance Metric Identification

Further study is needed to establish performance metrics based on measures different than simple evaluation of PRRs. These metrics may rely upon measures using data reported on Forms 890 or 891 by such means as evaluating data consistency, compliance with the requirements of Part 26, or written comments between sites or operators. Such measures could be those related to detection (e.g., number of tests exceeds fifty percent of covered

[3] There is some indication that pre-employment drug testing using matrices other than urine (i.e., hair) may yield information that supports effectively screening job applicants for drug use. For example, the trucking company JB Hunts experienced a significantly increased PRR in job applicants from hair specimens in comparison to urine (Nunlist 2011).

population),[4] based on goals (e.g., PRR for covered population is less than 0.5%), or possible new measures discovered through additional data analysis.

However, it is conceivable that performance measures might rely on additional information that could be more or less readily available. Examples include plant- and location-specific information already known to the NRC; regional information reported by other governmental agencies; and plant- or corporate-specific information.

6.3.1 Metrics for Compliance

For evaluating compliance, performance metrics could be based on regulation. For example, a performance metric on random testing could rely on several performance measures including such as whether the NPP met:

- the §26.31(d)(2)(vii) requirement that the number of random tests be equal to at least fifty percent of the population subject to the FFD program and
- the §26.31(d)(2)(iv) requirement that all individuals in the population subject to the FFD program have an equal chance of being selected for testing (this measure would be more difficult because the information is not available from the Form 890 and 891 data).

FFD program data other than that reported on Forms 890 and 891 could also be used in performance metric analysis. This data—required for drug and alcohol testing (§26.31, §26.65, §26.67), specimen collection for testing (Subpart E), and laboratories certified by the U.S. Department of Health and Human Services (HHS) (Subpart G)—might serve as the basis for a number of measures for assessing FFD program performance. Investigation into the availability of information about program performance related to implementing procedures in conjunction with PRRs could garner data that provides a stronger basis for identifying high-performing and low-performing FFD programs.

In particular, Subpart G requires use of laboratories certified under the HHS Mandatory Guidelines for Federal Workplace Drug Testing Programs. Testing of urine specimens must meet the requirements of the NRC's FFD and the certified laboratory program.[5,6] Examining the HHS-certified laboratory testing process of urine specimens may represent additional opportunities for identifying sources of error affecting FFD program performance, and therefore offer potential input to performance metrics analyses. See Appendix C for more detail about the types of errors that are likely to occur during the testing process based on clinical laboratory science literature. Note that much of HHS laboratory certification and oversight is under the purview of HHS although the FFD policy has explicit requirements under Subpart G that could appropriately be used in establishing NRC performance measures for drug testing programs.

---

[4] DOD, 2011, Status of Drug Use in the Department of Defense Personnel, available at: http://prhome.defense.gov/Portals/52/Documents/RFM/Readiness/DDRP/docs/6b%20FY%202011%20Annual%20Drug%20Use%20Status%20Report.pdf.

[5] SAMHSA provides current information on drug testing at: https://www.samhsa.gov/workplace/drug-testing#Laboratory%20Resources.

[6] See, for example: https://www.nrc.gov/docs/ML1532/ML15324A366.pdf.

In summary, taking greater advantage of the extant data reported by licensees and other entities and focusing on quantifiable measures based on the FFD regulations might be a useful starting point to identity additional insight into assessing FFD program performance, particularly in terms of program compliance. This approach might be especially helpful for NRC inspectors preparing for onsite inspection.

6.3.2 Metrics for Performance Objectives

Another important consideration for FFD program performance evaluation is to examine the extent to which these programs meet the performance objectives laid out in § 26.23, which states the following:

Fitness-for-duty programs must—
(a) Provide reasonable assurance that individuals are trustworthy and reliable as demonstrated by the avoidance of substance abuse;
(b) Provide reasonable assurance that individuals are not under the influence of any substance, legal or illegal, or mentally or physically impaired from any cause, which in any way adversely affects their ability to safely and competently perform their duties;
(c) Provide reasonable measures for the early detection of individuals who are not fit to perform the duties that require them to be subject to the FFD program;
(d) Provide reasonable assurance that the workplaces subject to this part are free from the presence and effects of illegal drugs and alcohol; and
(e) Provide reasonable assurance that the effects of fatigue and degraded alertness on individuals' abilities to safely and competently perform their duties are managed commensurate with maintaining public health and safety.

To address this issue, the meaning of reasonable assurance would need to be established and approaches to measuring and quantifying a baseline for reasonable assurance need to be developed.

Common facets of performance measures identified in the organizational performance literature (e.g., customer satisfaction, internal business processes (efficiency of operations), quality of service or products, continuous improvement efforts, public responsibility and social commitment, and financial performance) might shed some light on which aspects of FFD program performance could be appraised (Curtright et al. 2000; Kaplan and Norton 1996, 1992). Some key performance parameters, such as technical efficiency (input vs. output), cost effectiveness (cost/unit of output), cost-benefit analysis (cost/unit of outcome), program adequacy (i.e., does the program resolve the problem for which it was created?), would require conceptualization and operationalization with FFD-related data (McDavid et al. 2012). Within the context of FFD programs, it might be informative to examine resource adequacy (e.g. availability of Schedule 1 and 2 drug testing procedures at HHS-certified laboratories used by the licensee or other entity), cost effectiveness (e.g., cost and staff time per test), program capability (e.g., staff credentials and skills, frequency of training, and continued education, FFD program staff turnover rate), operational quality (e.g., condition of onsite collection facility, proper handling of specimen collection, storage, and required documentation, quality and integrity of HHS-certified laboratories contracted for testing, any litigation against FFD

programs). Understandably, such data might not be publically accessible or made available to the NRC by licensees and other entities. However, if available, data on these individual and organizational characteristics could provide additional insight into licensees' and other entities' FFD drug and alcohol deterrence and testing performance.

Further analysis on comprehensive, integrated FFD-related data could yield additional noteworthy trends and patterns in FFD program performance, and identify licensees or other entities whose drug and alcohol testing programs are performing well and those that warrant further attention. Possible tools for discovering new measures include machine learning techniques whose strength lies in their ability to find patterns in large amounts of data even when there are few examples of noteworthy cases.

6.4 Conclusion

Although the analysis results indicate instances in which a number of sites' PRRs were significantly different from a baseline rate (e.g., industrywide average, year-on-year), interpreting these should be approached with caution because the results do not explain *why* a site's PRR is significantly different from the baseline rate. We conclude that PRRs alone are of limited value for evaluating the performance of a licensee's FFD program.

To evaluate the effectiveness of FFD programs holistically, an important consideration is to understand the drug testing program's gatekeeping role as well as its ongoing monitoring responsibilities to ensure that licensee employees and C/Vs are fit for duty and are free from the influence of drugs and alcohol-related impairment at the work place. A different approach to data, analysis, and modeling would be required to understand what good performance means in the context of FFD drug and alcohol testing programs through capturing additional performance measures that could inform FFD program performance metrics both in terms of program compliance and program impact in meeting program performance objectives.

7.0 REFERENCES

10 CFR Part 26. *Code of Federal Regulations*. Title 10, Energy, Part 26, "Fitness for Duty Programs." 73 FR 17117. March 31, 2008. "Fitness for Duty Programs" Federal Register, Nuclear Regulatory Commission, Washington D.C.

29 CFR Part 1904. *Code of Federal Regulations*. Title 10, Energy, Part 26, "Occupational Injury and Illness Recording and Reporting Requirements – NAICS Update and Reporting Revisions." 66 FR 5916. January 19, 2001. Federal Register, Occupational Safety and Health Administration, Washington D.C.

Cattell, R. B. 1966. The Scree Test for the Number of Factors. *Multivariate Behavioral Research* 1, 245-276.

Curtright, J. W., Stolp-Smith, S., and Edell, E. S. 2000. Strategic performance management: Development of a performance measurement system at the mayo clinic. *Journal of Healthcare Management, 45*(1), 58-68. Retrieved on Nov 15, 2017 from https://search.proquest.com/docview/206731011?accountid=28112

D'agostino, R. B. and Russell, H. K. 2005. Scree Test. *Encyclopedia of Biostatistics*. 7.

DiCiccio, T. J. and Efron, B. 1996. Bootstrap Confidence Intervals. *Statistical Science* 11(3):189-228. http://projecteuclid.org/euclid.ss/1032280214.

DOD.2011. Status of Drug Use in the Department of Defense Personnel. Retrieved December 14, 2017 from http://prhome.defense.gov/Portals/52/Documents/RFM/Readiness/DDRP/docs/6b%20FY%202011%20Annual%20Drug%20Use%20Status%20Report.pdf.

Fleiss, J. L., Levin, B., Paik, M. C. and Fleiss, J. 2003. *Statistical Methods for Rates & Proportions*. Hoboken, NJ: Wiley.

Kaplan, R. S., and Norton, D. P. 1996. *The Balanced Scorecard: Translating Strategy into Action.* Boston: Harvard Business School Press.

Kaplan, R. S., and Norton, D. P. 1992. The Balanced Scorecard: Measures that Drive Performance. *Harvard Business Review* 70, no. 1: 71–79

McDavid, J. C., Huse, I., and Hawthorn L. R. L. 2012. Program Evaluation and Performance Measurement: An Introduction to Practice. Thousand Oaks, CA: Sage Publications.

NRC—U.S. Nuclear Regulatory Commission. 2014. Summary of Fitness for Duty Program Performance Reports for Calendar Year 2013. http://pbadupws.nrc.gov/docs/ML1424/ML14246A440.pdf

Nunlist, T, OB Patton. 2011. A Better Way to Drug Test? Hair Testing Reveals Disturbing Facts About Driver Drug Use and DOT Requirements

http://www.truckinginfo.com/channel/safety-compliance/article/story/2011/11/a-better-way-to-drug-test-hair-testing-reveals-disturbing-facts-about-driver-drug-use-and-dot-requi.aspx

Stephenson, W. R., Froelich, A.G., and Duckworth W. M. 2010. Using Resampling to Compare Two Proportions. *Teaching Statistics* 32(3):66-71. http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9639.2009.00379.x/full.

APPENDIX A

TECHNICAL NOTES ON STATISTICAL APPROACHES

APPENDIX A

TECHNICAL NOTES ON STATISTICAL APPROACHES

The exploratory analysis described in Section 3 includes:

- Descriptive statistics and data visualization
- Chi-square test of independence
- Factorial analysis

This appendix includes additional information on the techniques for the chi-square test of independence and factorial analysis. The techniques for the descriptive statistics and data visualization are described in Sections 3.0 through 5.0.

A.1 Chi-Square Test of Independence

Chi-square test of independence was applied to categorical data to estimate the probability that the differences between the observed values for the categories arose by random chance. The chi-square analysis in this report compares drug and alcohol testing results by employee type to determine whether there is a relationship between employee type and the PRR.

Observed counts of positive and negative results (O) are recorded for each employee type in a 2×2 contingency table. Expected counts (E) are then calculated for each cell by multiplying the marginal sums for each employee type-test result combination and dividing by the total sum of counts across the entire table. The expected values are subtracted from the observed values and that value is squared then divided by the expected value ((O-E)^2/E). The sum of those values, Σ(O-E)^2/E, is a test statistic with a chi-square (χ^2) distribution with 1 degree of freedom.

A.2 Factor Analysis

Factor analysis is related to another dimensionality reduction technique, principal components analysis; however, there are important assumptions about the underlying data model that distinguish the two. Principal components analysis is a purely descriptive technique that extracts a number of factors equal to the number of measured variables and completely explains the variance in each measured variable. Factor analysis assumes that there is an underlying causal *common factor model* consisting of common and unique factors. Factor analysis does not completely explain the variance in the measured variables. Rather, it explains only the variance that is shared in common among the measured variables; the remaining variance is attributed to uniqueness in the measured variables. The residual variance accounted for by unique factors is a consequence of the assumption that the common variance can be explained by a number of factors that is strictly smaller than the number of measured variables.

The common factor model can be expressed mathematically as

$$x_{ij} = \mu_j + \lambda_{j1} z_{i1} + \lambda_{j2} z_{i2} \ldots + \lambda_{jm} z_{im} + u_{ij}$$

where x_{ij} is the value of the observation for unit of analysis i on measured variable j, μ_j is the mean of the measured variable j, z_{ik} is the common factor score for unit of analysis i on common factor k, λ_{jk} is the common factor loading for measured variable j on common factor k, and u_{ij} is the unique factor score for unit of analysis i on measured variable j.

In exploratory factor analysis, the goal is to investigate the data to determine whether any interpretable latent structure underlies the data. Unlike in confirmatory factor analysis, exploratory factor analysis does not impose any constraints based on a priori assumptions about the relationship between measured variables. Confirmatory factor analysis is conducted on the basis of well-developed theory. This particular case lacks such a theory; however, some initial observations justify conducting an exploratory factor analysis. Results of the exploratory analysis and chi-square test suggest a statistically significant relationship between the PRR and employee type.

A.3 Performance Metric Analysis and Nonparametric Bootstrap Resampling Confidence Intervals

Typically, the normal approximation test is used to determine whether the difference between two binomial proportions (e.g., PRRs, which are based on only two values, positive and negative) is statistically significant. This test is based on the fact that when the number of observations in the sample is large, the distribution of a binomial variable is reasonably approximated by the normal distribution, and statistical tests based on the properties of the normal distribution may be used to make inferences about the underlying distribution of the population. A determination of statistical significance is made by constructing a confidence interval around the point estimate of the difference between the two proportions, $\Delta\hat{p}$, and observing whether zero falls within the interval.

For the result of the normal approximation test to be statistically valid, the following assumptions must be met:

- The sampling method for each population is simple random sampling.
- Each sample is independent.
- There is a sufficiently large number of observations so that the normal approximation holds.
- The normal approximation may not be valid for values of $\Delta\hat{p}$ near 0 or 1.

For the dataset being analyzed here, these assumptions are frequently violated. In many cases (e.g., follow-up tests), there are very few total tests conducted (n <10) for a site in a given year, and far fewer positive results are reported. In some cases, no positive results are reported for a site in a given year. Therefore, the assumption of normality of the sample proportion $\hat{p}$ frequently fails to hold, and results of the normal approximation test for the difference between two binomial proportions are not statistically valid.

In such cases, an alternative approach for constructing a confidence interval is to use a bootstrap resampling procedure to generate a resampling distribution from which an empirical confidence interval can be derived. The point estimate of the difference in proportions can then be evaluated against this confidence interval, and if the estimate falls outside of the interval, there is evidence to suggest that the true difference between the two PRRs is significantly different from zero.

Graphically, the point estimate of the difference is represented with a dot, while the confidence interval is represented as a pair of error bars. In this report, dots and error bars are color-coded depending on whether a difference is significantly less than zero (blue), significantly greater than zero (red), or not significantly different from zero (black). Figure A.1 illustrates this convention.[1]

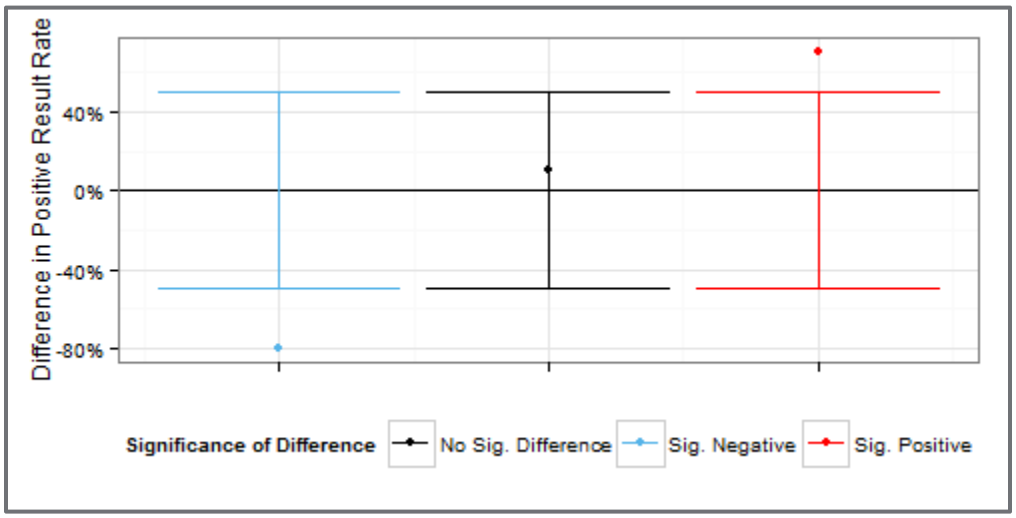


Figure A.1. Illustration of Point Estimates and Confidence Intervals.

To perform the bootstrapping procedure, a 2 × 2 contingency table is constructed to obtain the distribution of positive and negative counts for each comparison group. The confidence interval is constructed empirically on this distribution by selecting the bounds such that they are equal to the α/2 and 1- α /2 percentiles of the distribution, where α is the significance level, or the probability of rejecting the null hypothesis given that it is true. (Consistent with common practice, a significance level of 0.05 is chosen throughout this analysis to evaluate point estimates.) The point estimate of the difference between the two proportions is then evaluated against this interval. If the point estimate of the difference in the PRR falls outside of this interval, strong evidence exists that the result was not obtained by chance alone, and that significant difference exists between the two PRRs.

Table A-1 presents a generalized representation of a 2 × 2 contingency table. A resample of n_1 counts for group 1 is taken, with each draw assigned a probability of a positive result equal to the marginal probability of a positive result over both groups. Then a resample of n_2 counts for group 2 is taken, assigning the same marginal probability of a positive result over both groups to each draw in that group. This resample generates one hypothetical realization of an experiment in which both groups have the same probability of obtaining a positive result. The difference between the PRRs for the two groups is calculated, then the procedure is repeated many more times (e.g., in this analysis, 1,000 resamples were performed to generate each distribution). This generates a distribution of differences under the null hypothesis that there is no underlying difference in the probability of obtaining a positive result between the two groups.

[1] Note that in the final version of the report we will ensure that the graphs reproduce clearly in black and white copying.

The confidence interval is constructed empirically on this distribution by selecting the bounds such that they are equal to the α/2 and 1- α /2 percentiles of the distribution, where α is the significance level, or the probability of rejecting the null hypothesis given that it is true. (Consistent with common practice, a significance level of 0.05 is chosen throughout this analysis to evaluate point estimates.) The point estimate of the difference between the two proportions is then evaluated against this interval. If the point estimate of the difference in the PRR falls outside of this interval, strong evidence exists that the result was not obtained by chance alone, and that significant difference exists between the two PRRs.

Table A-1. Generalized Representation of a 2 x 2 Contingency Table

	Positive Result	Negative Result	Totals
Group 1	pos_1	neg_1	$n_1 = pos_1 + neg_1$
Group 2	pos_2	neg_2	$n_2 = pos_2 + neg_2$
Totals	$pos_{Tot} = pos_1 + pos_2$	$neg_{Tot} = neg_1 + neg_2$	$n_1 + n_2 = pos_{Tot} + neg_{Tot}$
Marginal Probability	$pos_{Tot} / (pos_{Tot} + neg_{Tot})$	$neg_{Tot} / (pos_{Tot} + neg_{Tot})$	

In Figure A.2, a hypothetical bootstrap resampling distribution is shown. The two black vertical lines represent the 2.5 and 97.5 percentiles of the distribution, while the red line represents the point estimate of the difference between the two proportions. In this example, the point estimate lies in the region above the upper cutoff, suggesting that the difference between the two proportions is statistically significant.

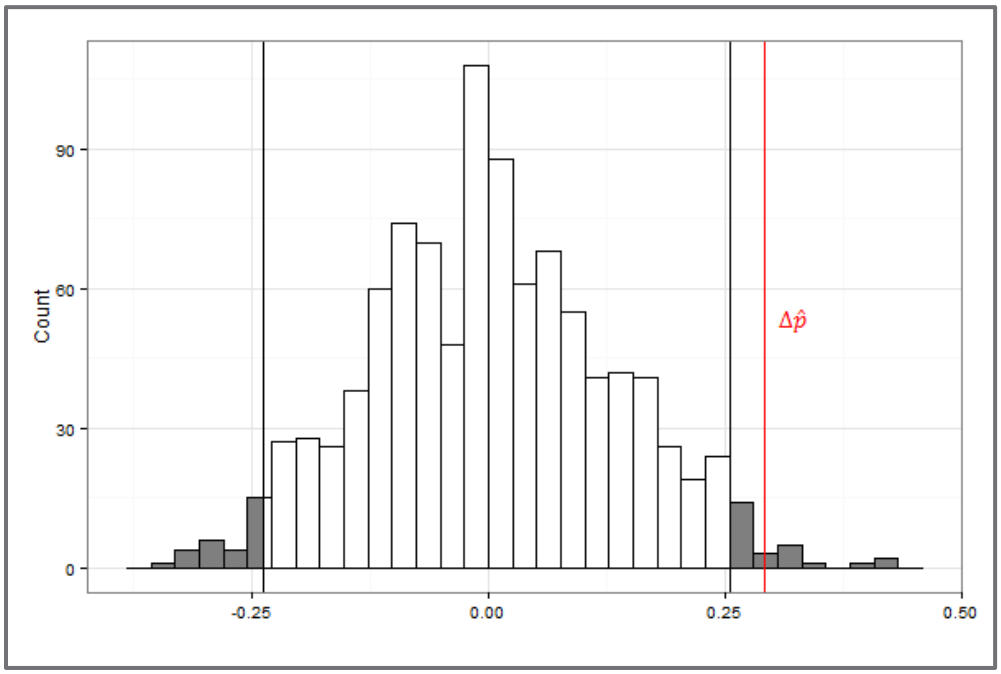


Figure A.2. Hypothetical Bootstrap Resampling Distribution with Associated Confidence Interval

APPENDIX B

PERFORMANCE METRIC ANALYSIS GRAPHS

APPENDIX B

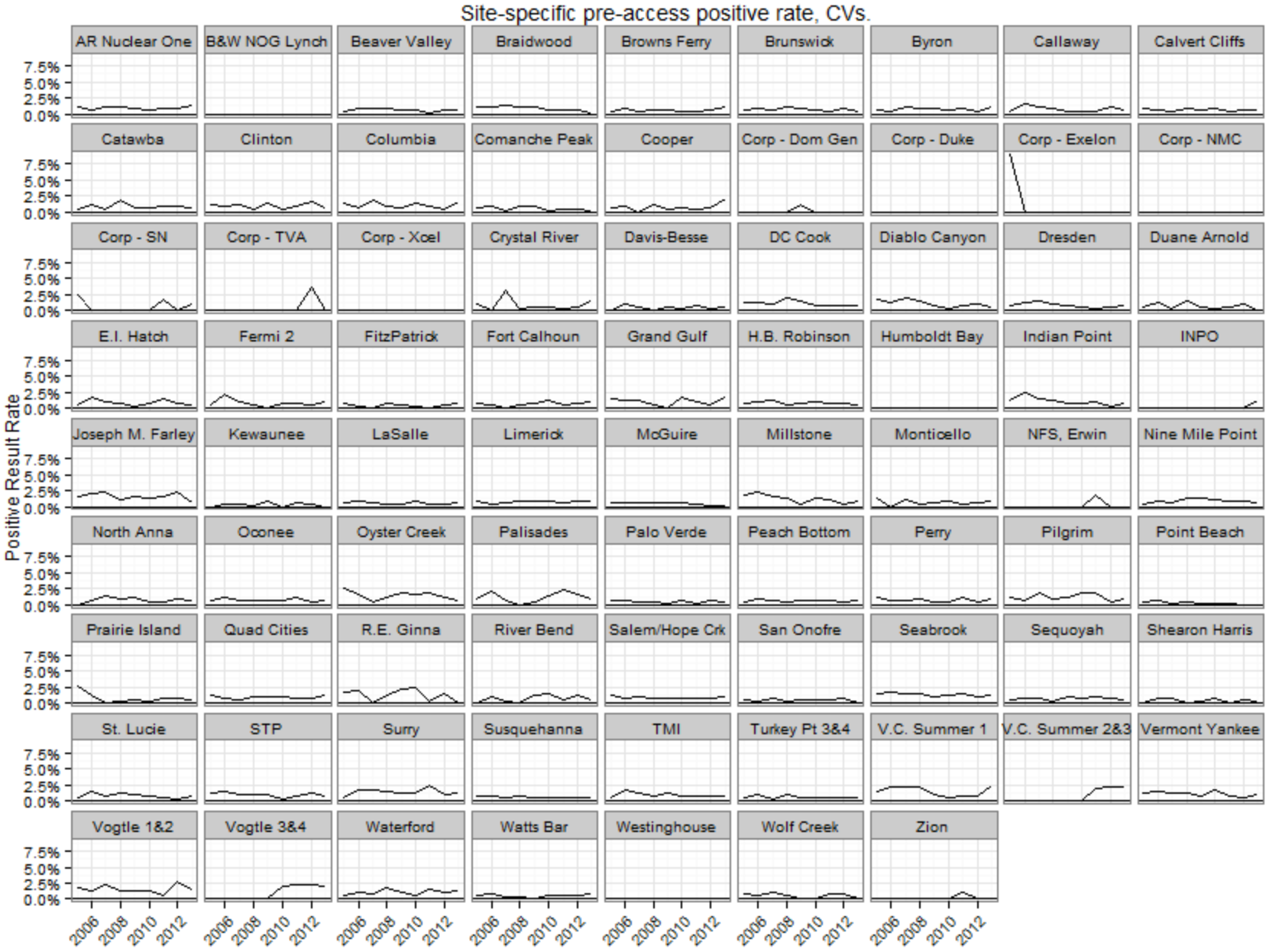
PERFORMANCE METRIC ANALYSIS GRAPHS

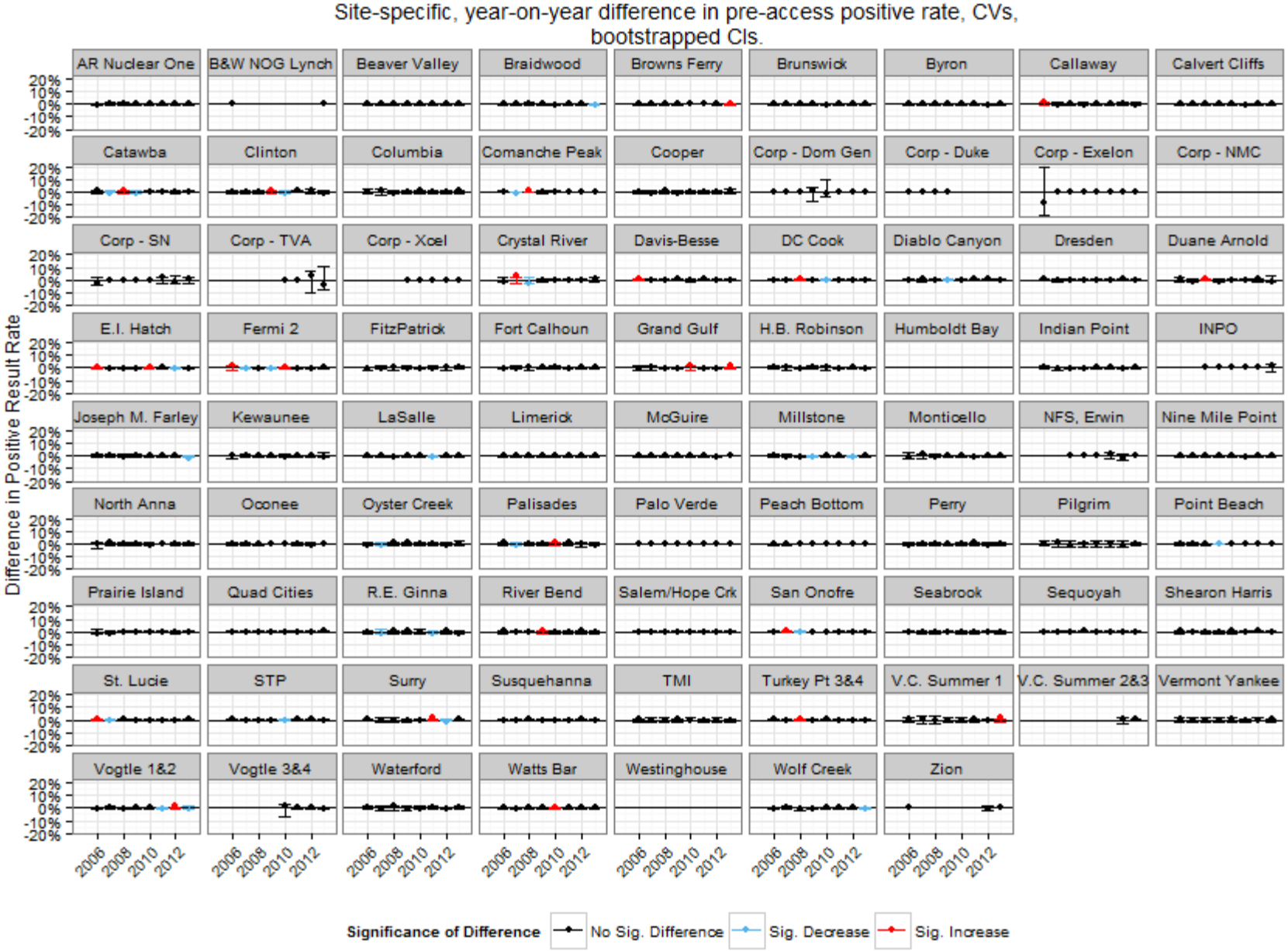
The nonparametric bootstrap resampling confidence interval results discussed in Section 5 are included in this appendix. The nonparametric bootstrap resampling process is described in Appendix A.3

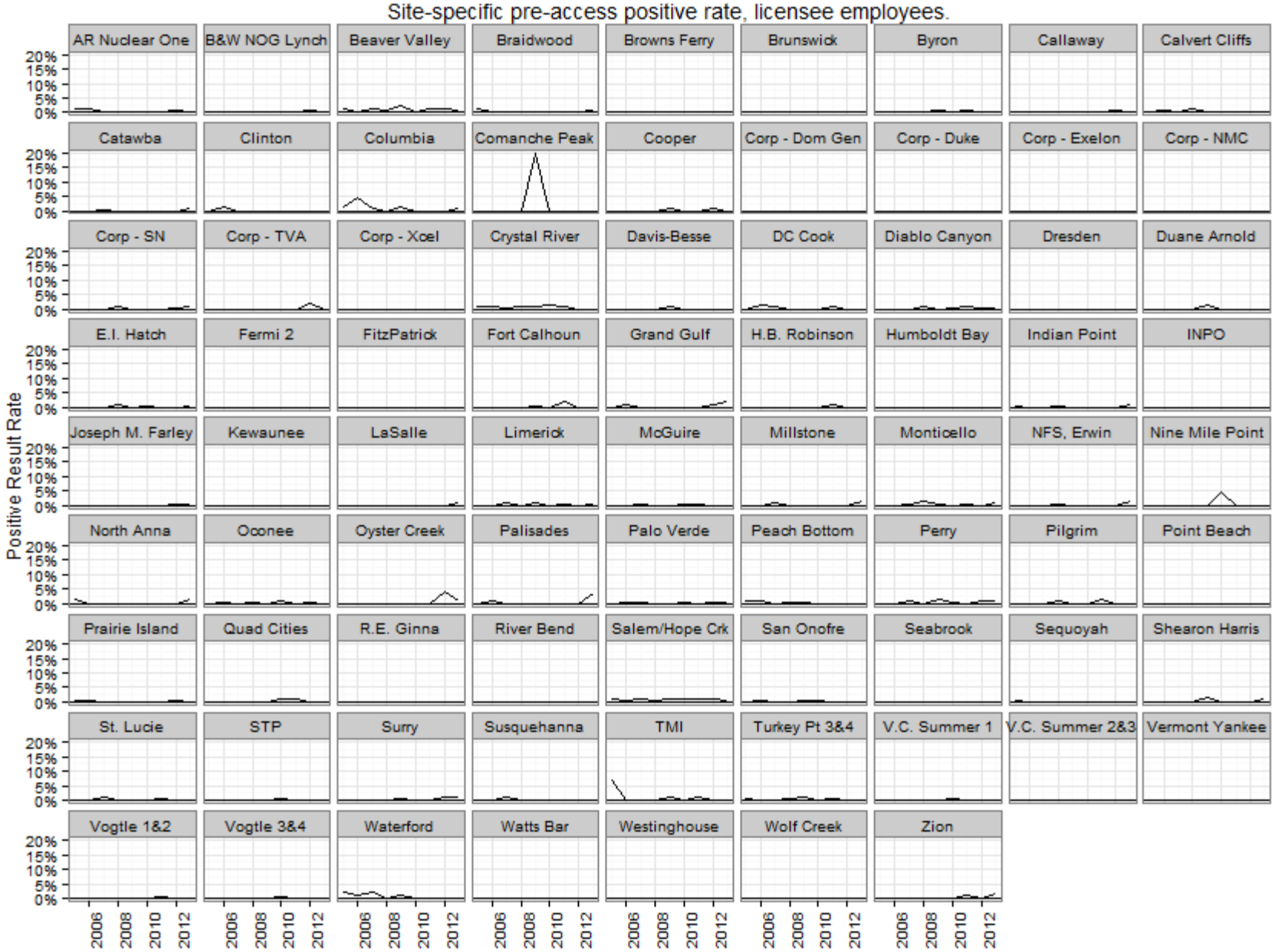
A note on symbols in the graphs of differences in positive result rates (see Figure A.1):

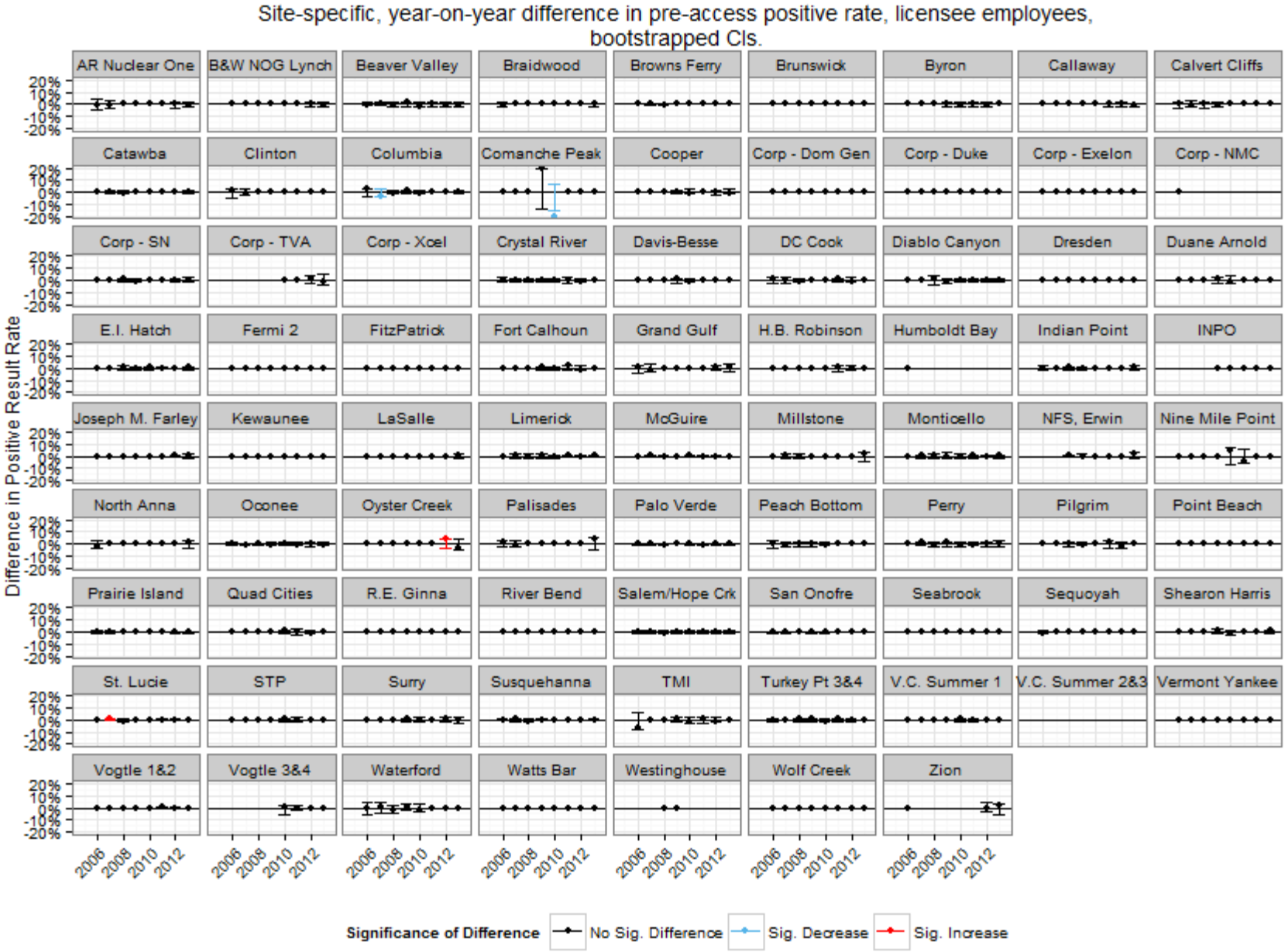
- A dot with error bars indicates that at least one of the two proportions being compared is greater than zero.
- A dot without error bars indicates that both proportions being compared are equal to zero; these dots always lie exactly on the line at the origin.
- A missing dot indicates that at least one of the proportions being compared is undefined because no tests were conducted for that year; since the denominator of the fraction is zero, its value is undefined, and consequently all statistics computed from it are undefined and cannot be plotted.

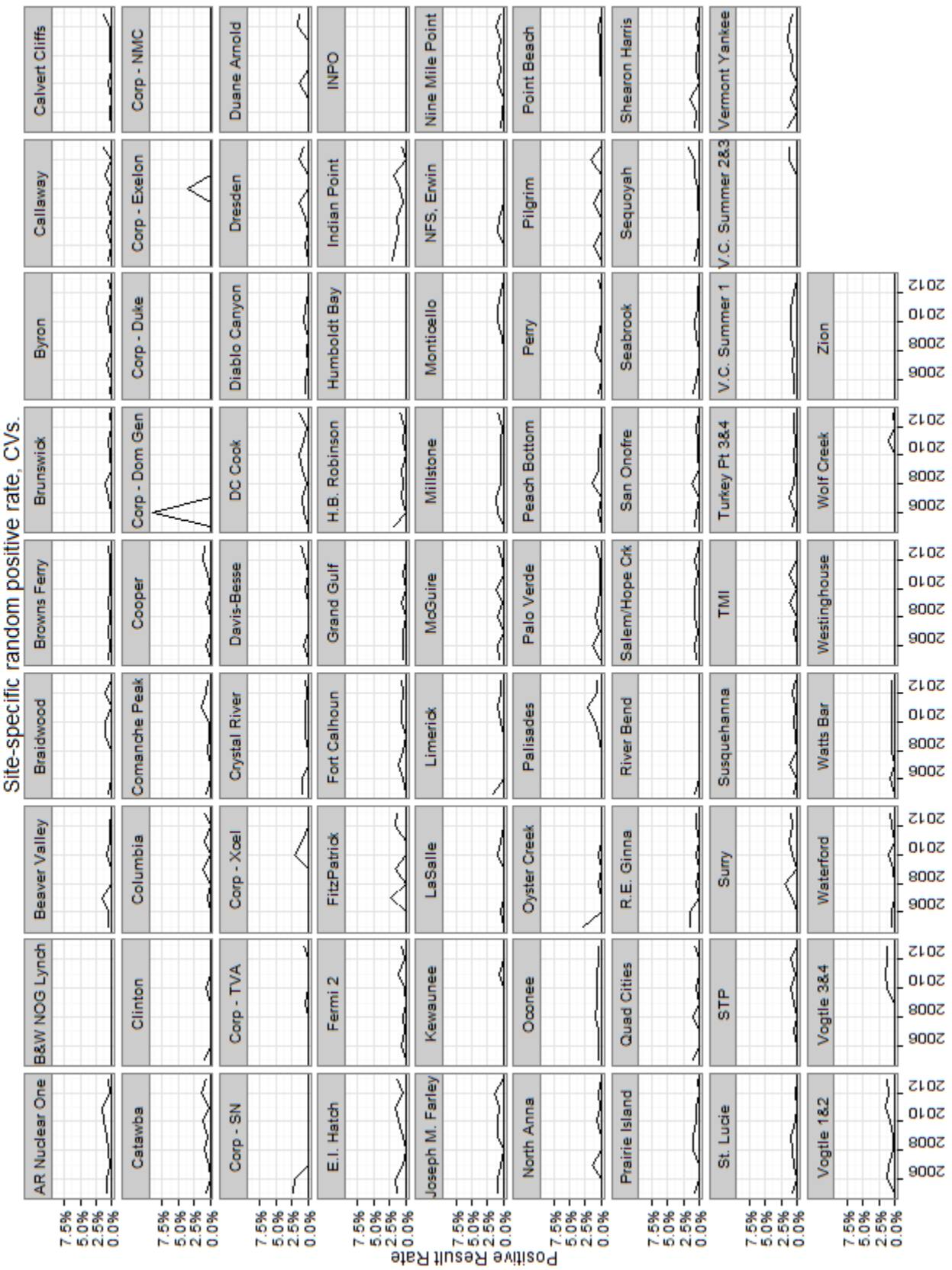
The bootstrap resampling R code used for preparing these graphs was adapted from: http://www.public.iastate.edu/~wrstephe/Resampling_Final.pdf

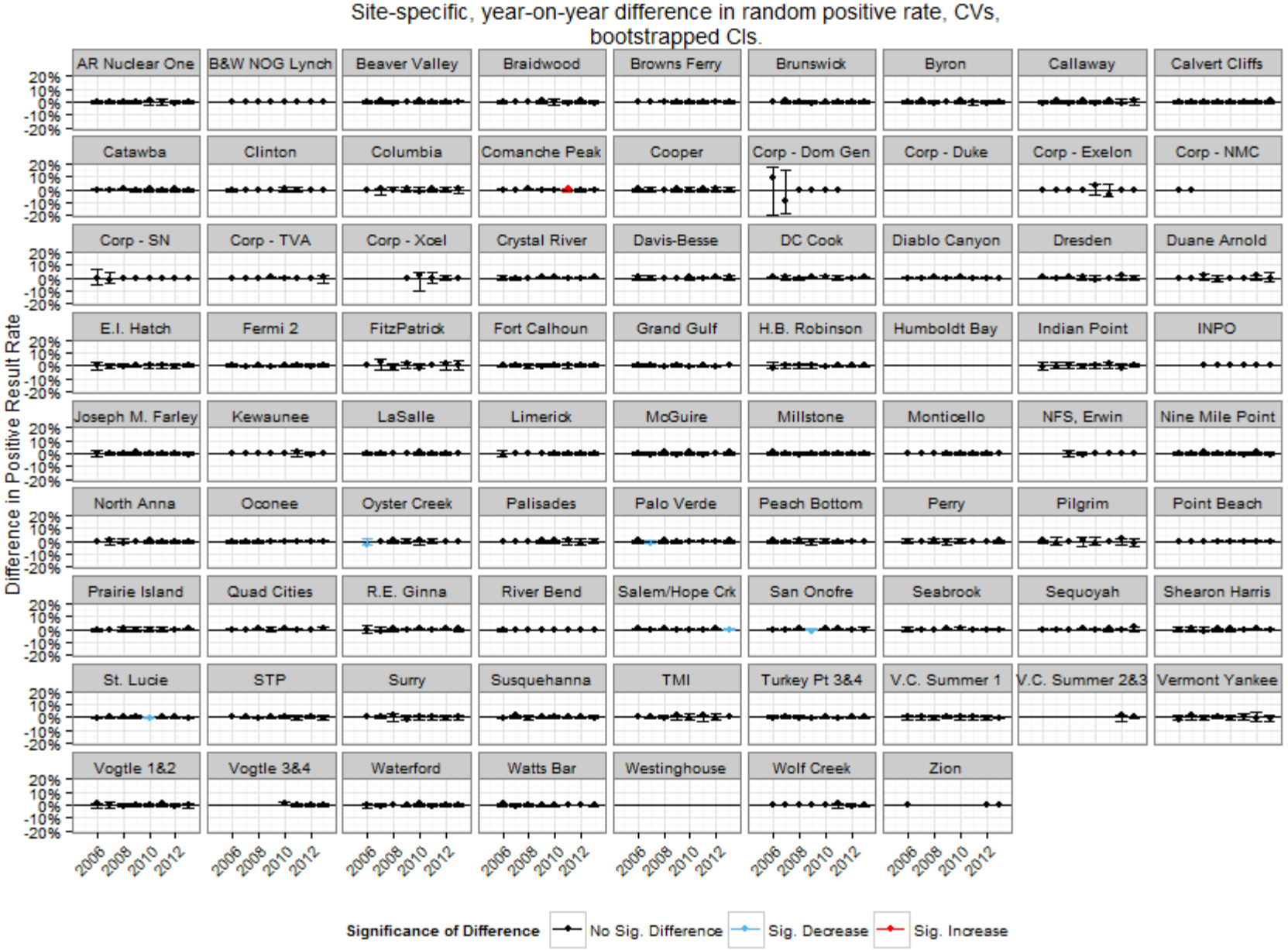

Site-specific pre-access positive rate, CVs.
Positive Result Rate
AR Nuclear One
B&W NOG Lynch
Beaver Valley
Braidwood
Browns Ferry
Brunswick
Byron
Callaway
Calvert Cliffs
Catawba
Clinton
Columbia
Comanche Peak
Cooper
Corp - Dom Gen
Corp - Duke
Corp - Exelon
Corp - NMC
Corp - SN
Corp - TVA
Corp - Xcel
Crystal River
Davis-Besse
DC Cook
Diablo Canyon
Dresden
Duane Arnold
E.I. Hatch
Fermi 2
FitzPatrick
Fort Calhoun
Grand Gulf
H.B. Robinson
Humboldt Bay
Indian Point
INPO
Joseph M. Farley
Kewaunee
LaSalle
Limerick
McGuire
Millstone
Monticello
NFS, Erwin
Nine Mile Point
North Anna
Oconee
Oyster Creek
Palisades
Palo Verde
Peach Bottom
Perry
Pilgrim
Point Beach
Prairie Island
Quad Cities
R.E. Ginna
River Bend
Salem/Hope Crk
San Onofre
Seabrook
Sequoyah
Shearon Harris
St. Lucie
STP
Surry
Susquehanna
TMI
Turkey Pt 3&4
V.C. Summer 1
V.C. Summer 2&3
Vermont Yankee
Vogtle 1&2
Vogtle 3&4
Waterford
Watts Bar
Westinghouse
Wolf Creek
Zion
7.5%
5.0%
2.5%
0.0%
2006
2008
2010
2012

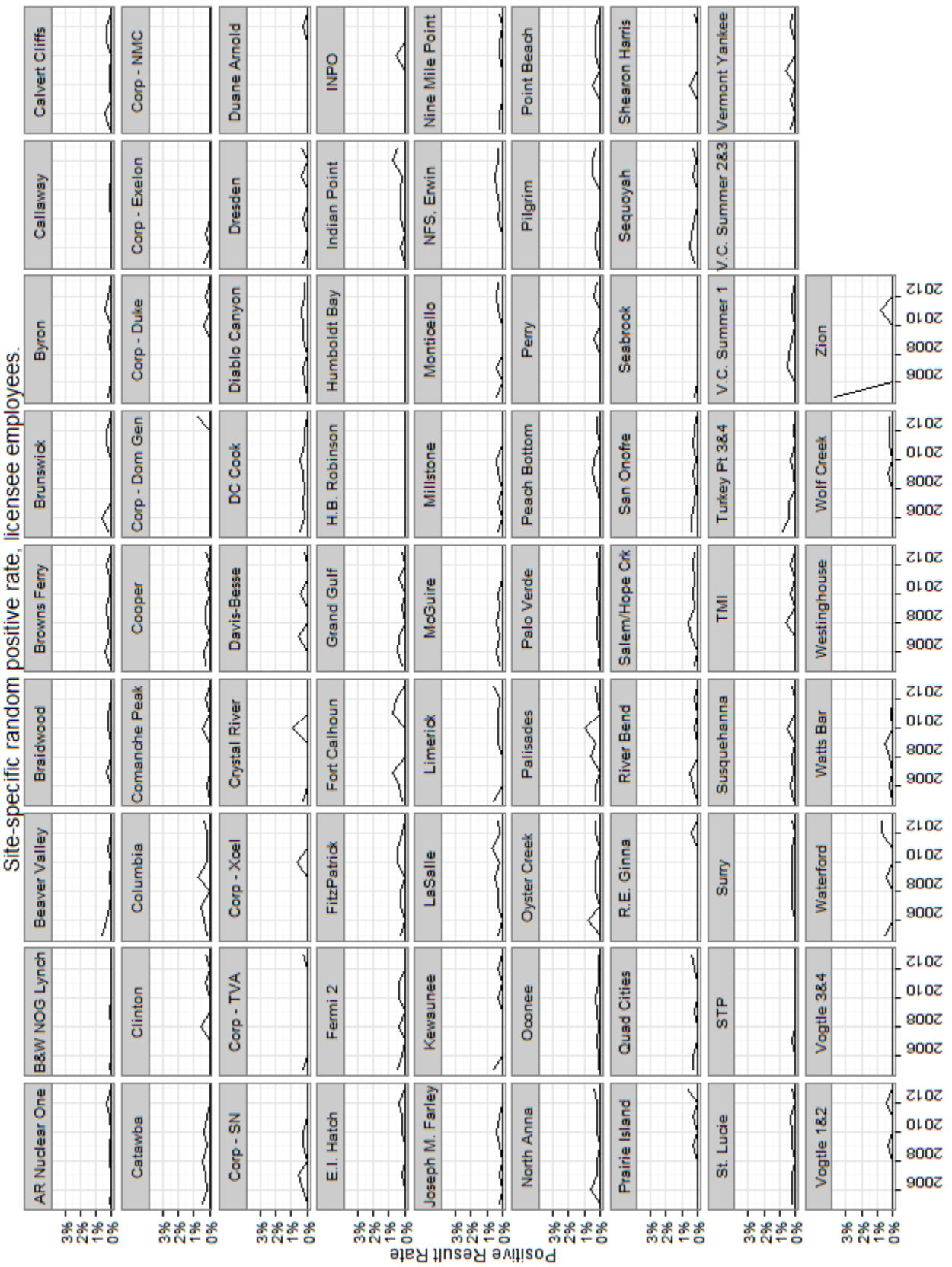

Site-specific, year-on-year difference in pre-access positive rate, CVs, bootstrapped CIs.
Difference in Positive Result Rate
AR Nuclear One
B&W NOG Lynch
Beaver Valley
Braidwood
Browns Ferry
Brunswick
Byron
Callaway
Calvert Cliffs
Catawba
Clinton
Columbia
Comanche Peak
Cooper
Corp - Dom Gen
Corp - Duke
Corp - Exelon
Corp - NMC
Corp - SN
Corp - TVA
Corp - Xcel
Crystal River
Davis-Besse
DC Cook
Diablo Canyon
Dresden
Duane Arnold
E.I. Hatch
Fermi 2
FitzPatrick
Fort Calhoun
Grand Gulf
H.B. Robinson
Humboldt Bay
Indian Point
INPO
Joseph M. Farley
Kewaunee
LaSalle
Limerick
McGuire
Millstone
Monticello
NFS, Erwin
Nine Mile Point
North Anna
Oconee
Oyster Creek
Palisades
Palo Verde
Peach Bottom
Perry
Pilgrim
Point Beach
Prairie Island
Quad Cities
R.E. Ginna
River Bend
Salem/Hope Crk
San Onofre
Seabrook
Sequoyah
Shearon Harris
St. Lucie
STP
Surry
Susquehanna
TMI
Turkey Pt 3&4
V.C. Summer 1
V.C. Summer 2&3
Vermont Yankee
Vogtle 1&2
Vogtle 3&4
Waterford
Watts Bar
Westinghouse
Wolf Creek
Zion
20%
10%
0%
-10%
-20%
2006
2008
2010
2012
Significance of Difference
No Sig. Difference
Sig. Decrease
Sig. Increase

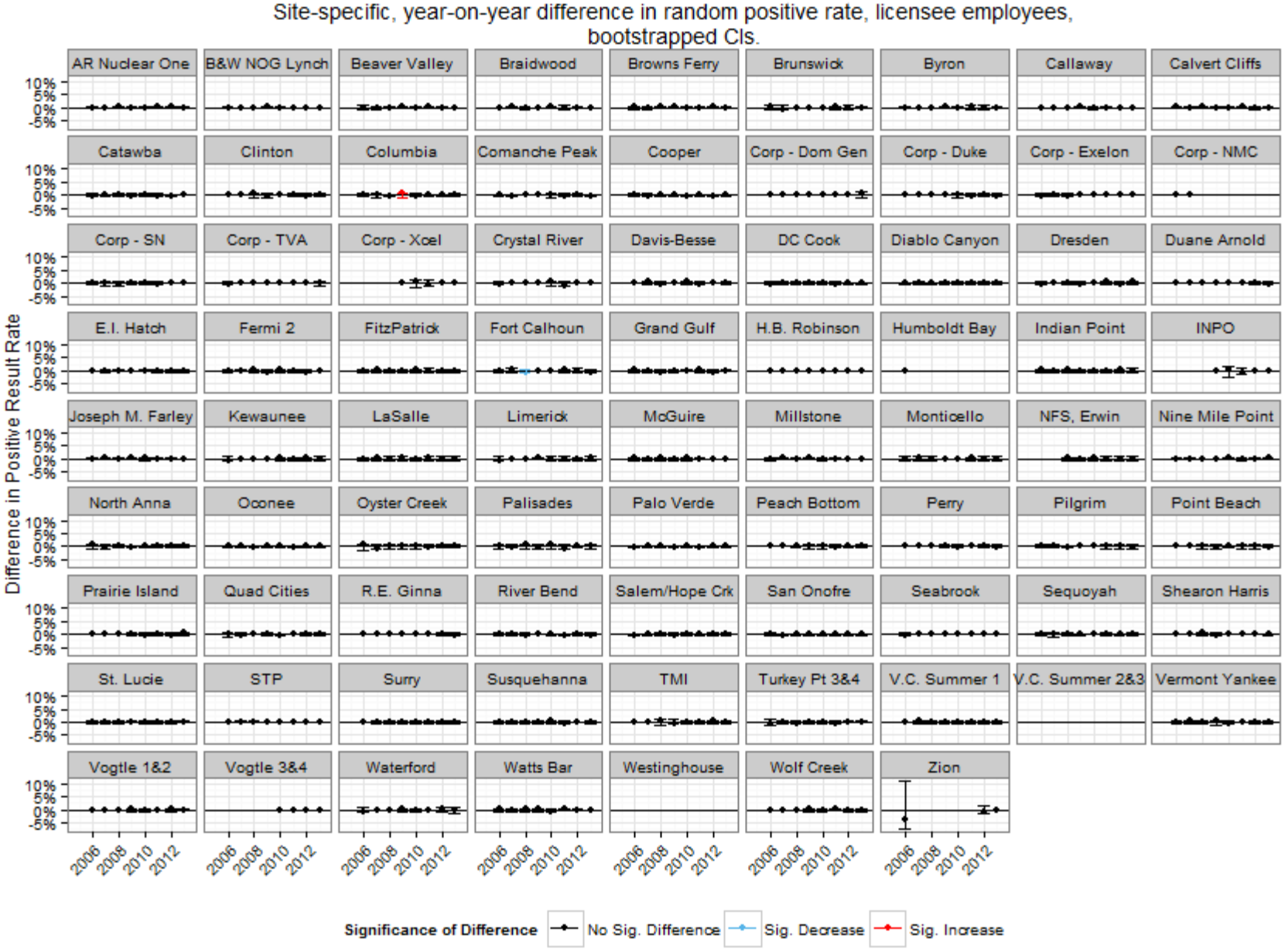

Site-specific pre-access positive rate, licensee employees.
Positive Result Rate
AR Nuclear One
B&W NOG Lynch
Beaver Valley
Braidwood
Browns Ferry
Brunswick
Byron
Callaway
Calvert Cliffs
Catawba
Clinton
Columbia
Comanche Peak
Cooper
Corp - Dom Gen
Corp - Duke
Corp - Exelon
Corp - NMC
Corp - SN
Corp - TVA
Corp - Xcel
Crystal River
Davis-Besse
DC Cook
Diablo Canyon
Dresden
Duane Arnold
E.I. Hatch
Fermi 2
FitzPatrick
Fort Calhoun
Grand Gulf
H.B. Robinson
Humboldt Bay
Indian Point
INPO
Joseph M. Farley
Kewaunee
LaSalle
Limerick
McGuire
Millstone
Monticello
NFS, Erwin
Nine Mile Point
North Anna
Oconee
Oyster Creek
Palisades
Palo Verde
Peach Bottom
Perry
Pilgrim
Point Beach
Prairie Island
Quad Cities
R.E. Ginna
River Bend
Salem/Hope Crk
San Onofre
Seabrook
Sequoyah
Shearon Harris
St. Lucie
STP
Surry
Susquehanna
TMI
Turkey Pt 3&4
V.C. Summer 1
V.C. Summer 2&3
Vermont Yankee
Vogtle 1&2
Vogtle 3&4
Waterford
Watts Bar
Westinghouse
Wolf Creek
Zion
20%
15%
10%
5%
0%
2006
2008
2010
2012

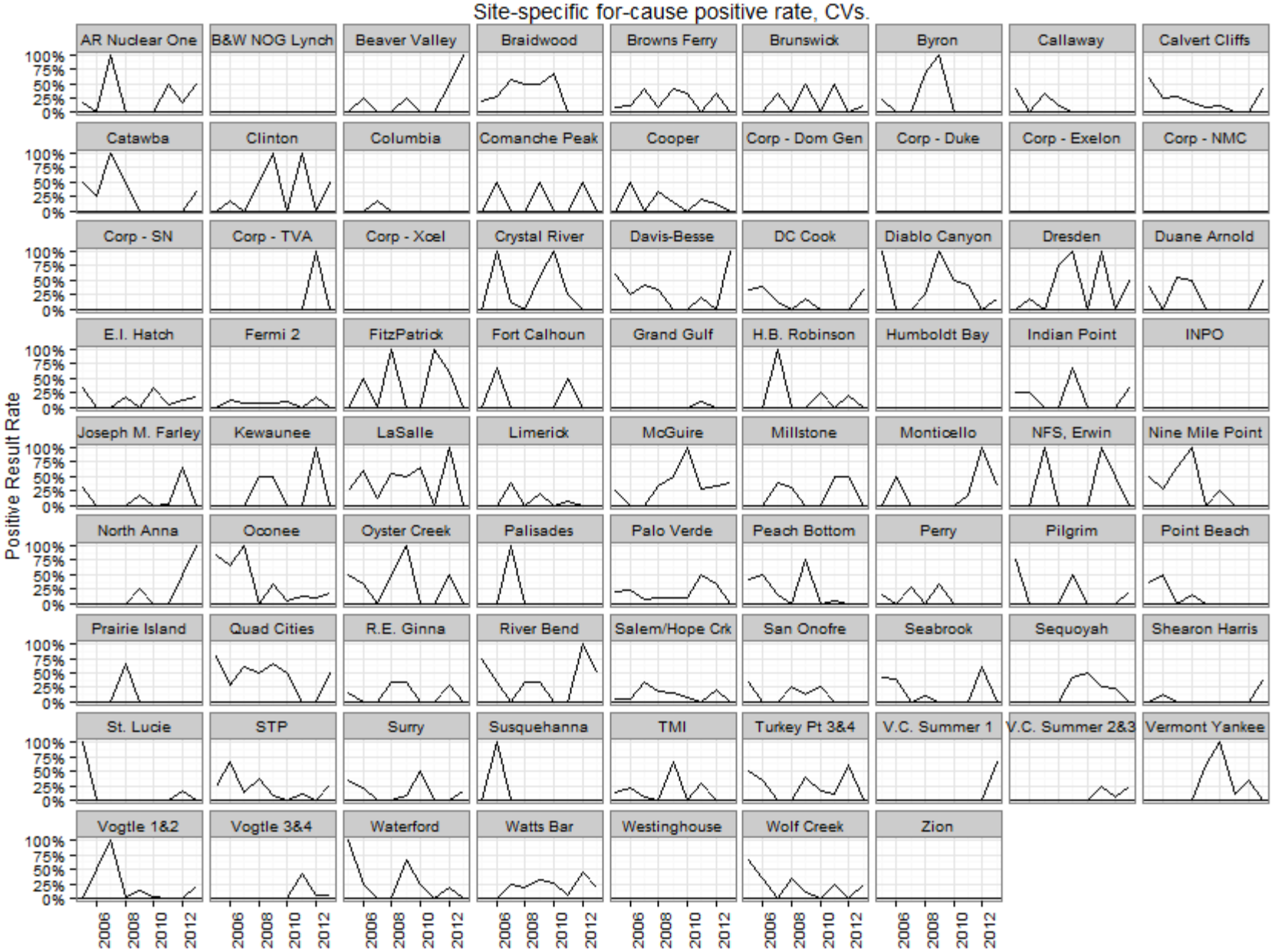

Site-specific, year-on-year difference in pre-access positive rate, licensee employees, bootstrapped CIs.
Difference in Positive Result Rate
AR Nuclear One
B&W NOG Lynch
Beaver Valley
Braidwood
Browns Ferry
Brunswick
Byron
Callaway
Calvert Cliffs
Catawba
Clinton
Columbia
Comanche Peak
Cooper
Corp - Dom Gen
Corp - Duke
Corp - Exelon
Corp - NMC
Corp - SN
Corp - TVA
Corp - Xcel
Crystal River
Davis-Besse
DC Cook
Diablo Canyon
Dresden
Duane Arnold
E.I. Hatch
Fermi 2
FitzPatrick
Fort Calhoun
Grand Gulf
H.B. Robinson
Humboldt Bay
Indian Point
INPO
Joseph M. Farley
Kewaunee
LaSalle
Limerick
McGuire
Millstone
Monticello
NFS, Erwin
Nine Mile Point
North Anna
Oconee
Oyster Creek
Palisades
Palo Verde
Peach Bottom
Perry
Pilgrim
Point Beach
Prairie Island
Quad Cities
R.E. Ginna
River Bend
Salem/Hope Crk
San Onofre
Seabrook
Sequoyah
Shearon Harris
St. Lucie
STP
Surry
Susquehanna
TMI
Turkey Pt 3&4
V.C. Summer 1
V.C. Summer 2&3
Vermont Yankee
Vogtle 1&2
Vogtle 3&4
Waterford
Watts Bar
Westinghouse
Wolf Creek
Zion
20%
10%
0%
-10%
-20%
2006
2008
2010
2012
Significance of Difference
No Sig. Difference
Sig. Decrease
Sig. Increase

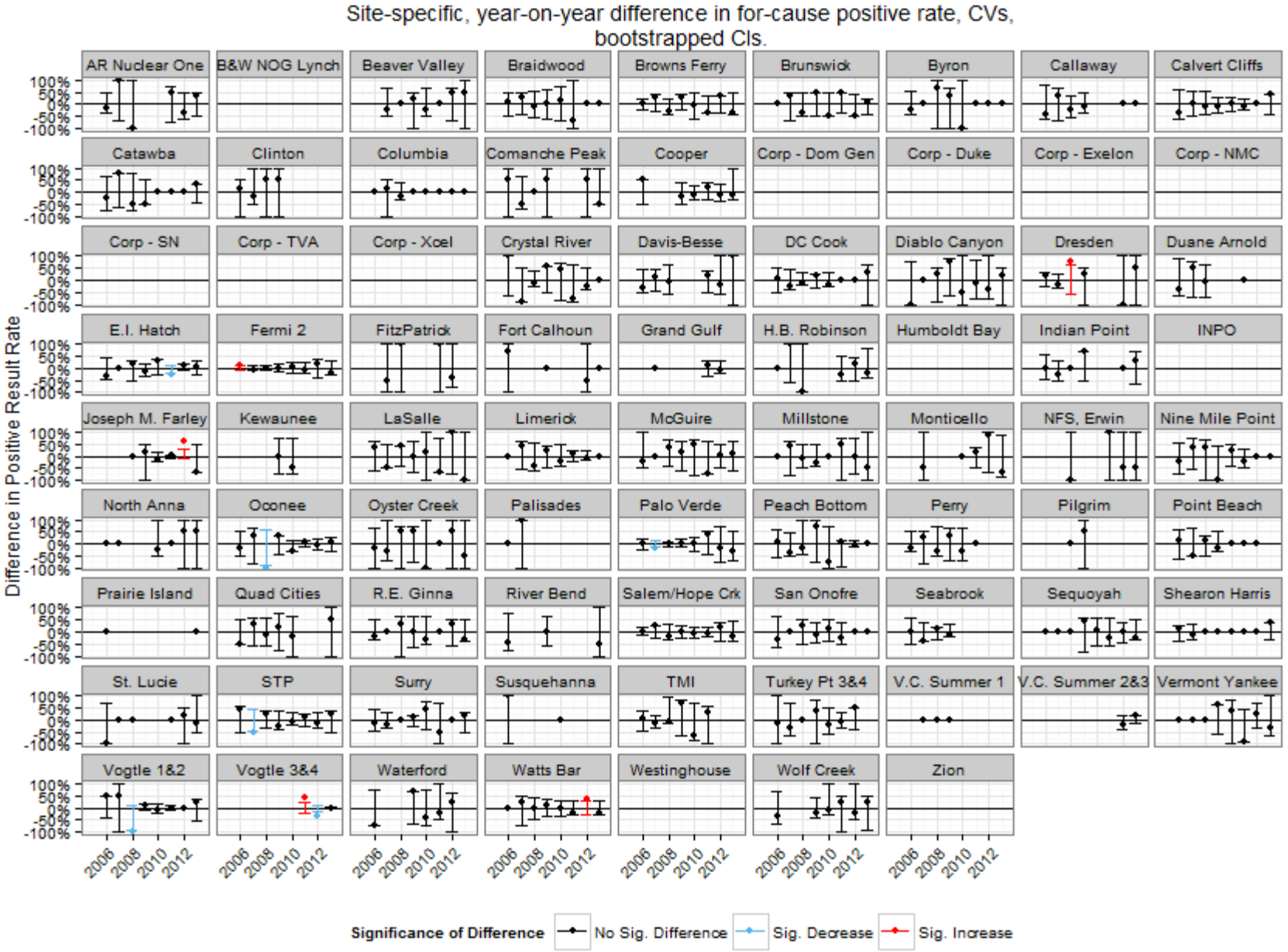

Site-specific random positive rate, CVs.
Positive Result Rate
AR Nuclear One
B&W NOG Lynch
Beaver Valley
Braidwood
Browns Ferry
Brunswick
Byron
Callaway
Calvert Cliffs
Catawba
Clinton
Columbia
Comanche Peak
Cooper
Corp - Dom Gen
Corp - Duke
Corp - Exelon
Corp - NMC
Corp - SN
Corp - TVA
Corp - Xoel
Crystal River
Davis-Besse
DC Cook
Diablo Canyon
Dresden
Duane Arnold
E.I. Hatch
Fermi 2
FitzPatrick
Fort Calhoun
Grand Gulf
H.B. Robinson
Humboldt Bay
Indian Point
INPO
Joseph M. Farley
Kewaunee
LaSalle
Limerick
McGuire
Millstone
Monticello
NFS, Erwin
Nine Mile Point
North Anna
Oconee
Oyster Creek
Palisades
Palo Verde
Peach Bottom
Perry
Pilgrim
Point Beach
Prairie Island
Quad Cities
R.E. Ginna
River Bend
Salem/Hope Crk
San Onofre
Seabrook
Sequoyah
Shearon Harris
St. Lucie
STP
Surry
Susquehanna
TMI
Turkey Pt 3&4
V.C. Summer 1
V.C. Summer 2&3
Vermont Yankee
Vogtle 1&2
Vogtle 3&4
Waterford
Watts Bar
Westinghouse
Wolf Creek
Zion
7.5%
5.0%
2.5%
0.0%
2006
2008
2010
2012

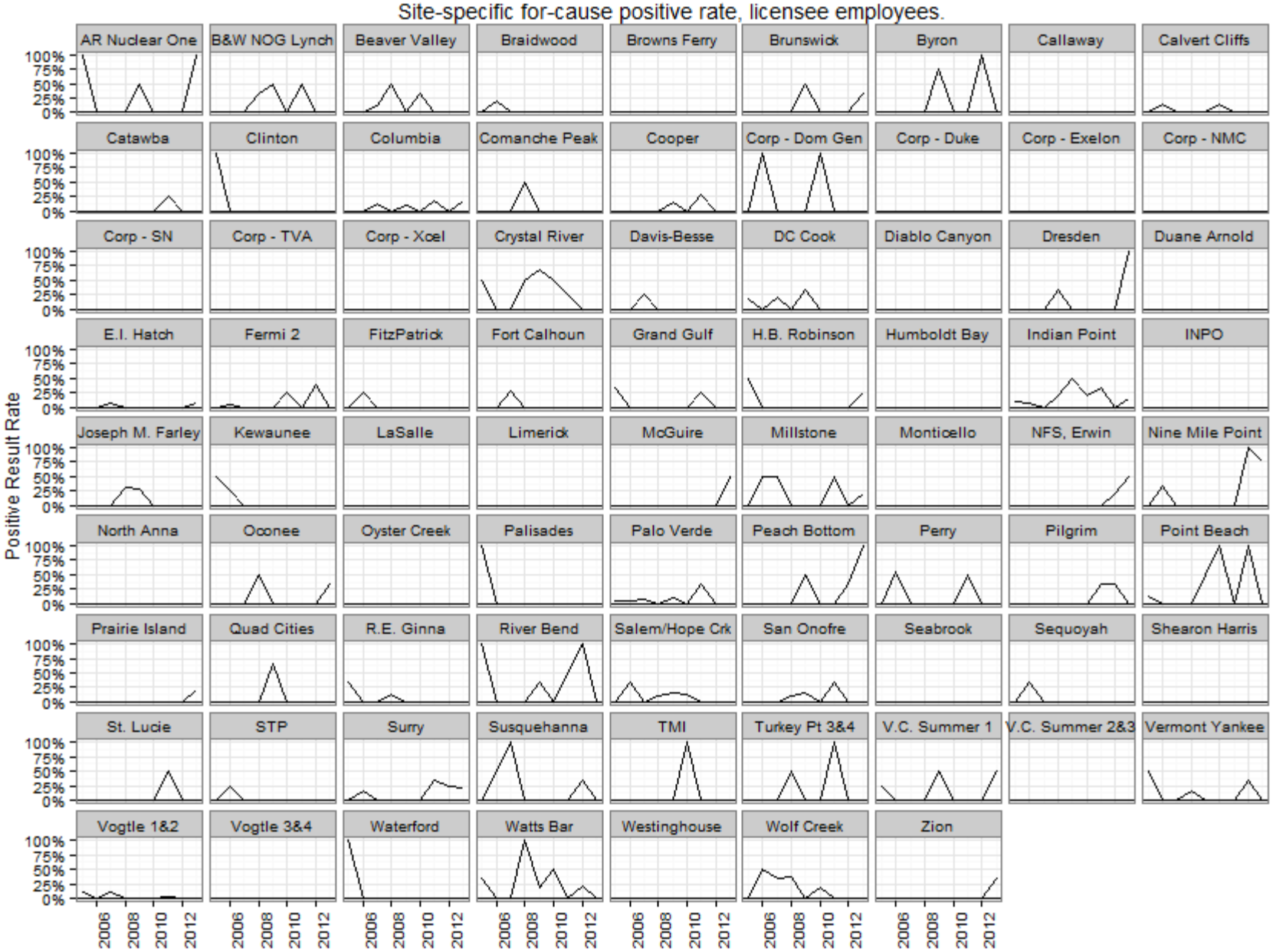

Site-specific, year-on-year difference in random positive rate, CVs, bootstrapped CIs.
Difference in Positive Result Rate
AR Nuclear One
B&W NOG Lynch
Beaver Valley
Braidwood
Browns Ferry
Brunswick
Byron
Callaway
Calvert Cliffs
Catawba
Clinton
Columbia
Comanche Peak
Cooper
Corp - Dom Gen
Corp - Duke
Corp - Exelon
Corp - NMC
Corp - SN
Corp - TVA
Corp - Xcel
Crystal River
Davis-Besse
DC Cook
Diablo Canyon
Dresden
Duane Arnold
E.I. Hatch
Fermi 2
FitzPatrick
Fort Calhoun
Grand Gulf
H.B. Robinson
Humboldt Bay
Indian Point
INPO
Joseph M. Farley
Kewaunee
LaSalle
Limerick
McGuire
Millstone
Monticello
NFS, Erwin
Nine Mile Point
North Anna
Oconee
Oyster Creek
Palisades
Palo Verde
Peach Bottom
Perry
Pilgrim
Point Beach
Prairie Island
Quad Cities
R.E. Ginna
River Bend
Salem/Hope Crk
San Onofre
Seabrook
Sequoyah
Shearon Harris
St. Lucie
STP
Surry
Susquehanna
TMI
Turkey Pt 3&4
V.C. Summer 1
V.C. Summer 2&3
Vermont Yankee
Vogtle 1&2
Vogtle 3&4
Waterford
Watts Bar
Westinghouse
Wolf Creek
Zion
20%
10%
0%
-10%
-20%
2006
2008
2010
2012
Significance of Difference
No Sig. Difference
Sig. Decrease
Sig. Increase

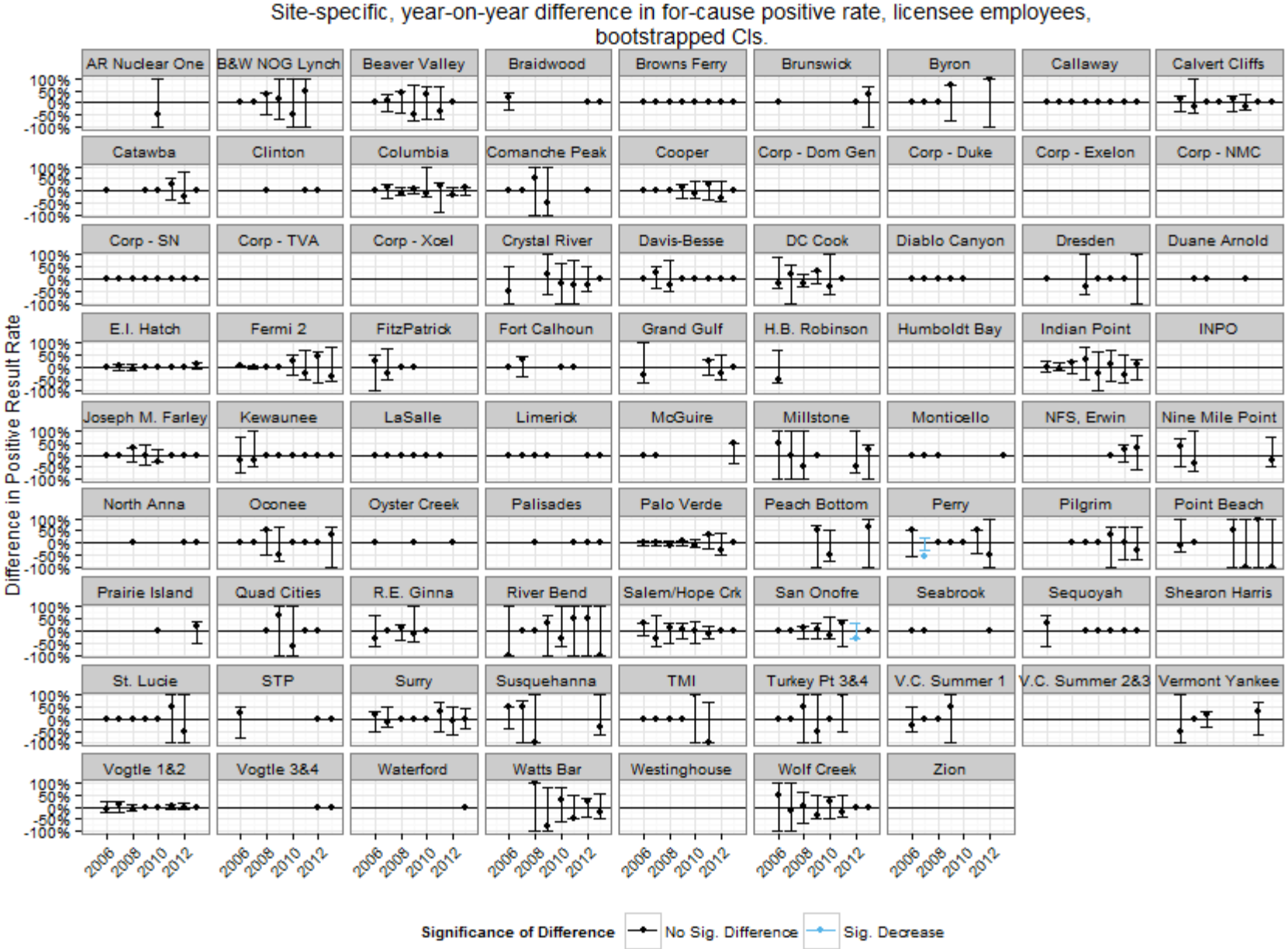

Site-specific random positive rate, licensee employees.
Positive Result Rate
AR Nuclear One
B&W NOG Lynch
Beaver Valley
Braidwood
Browns Ferry
Brunswick
Byron
Callaway
Calvert Cliffs
Catawba
Clinton
Columbia
Comanche Peak
Cooper
Corp - Dom Gen
Corp - Duke
Corp - Exelon
Corp - NMC
Corp - SN
Corp - TVA
Corp - Xcel
Crystal River
Davis-Besse
DC Cook
Diablo Canyon
Dresden
Duane Arnold
E.I. Hatch
Fermi 2
FitzPatrick
Fort Calhoun
Grand Gulf
H.B. Robinson
Humboldt Bay
Indian Point
INPO
Joseph M. Farley
Kewaunee
LaSalle
Limerick
McGuire
Millstone
Monticello
NFS, Erwin
Nine Mile Point
North Anna
Oconee
Oyster Creek
Palisades
Palo Verde
Peach Bottom
Perry
Pilgrim
Point Beach
Prairie Island
Quad Cities
R.E. Ginna
River Bend
Salem/Hope Crk
San Onofre
Seabrook
Sequoyah
Shearon Harris
St. Lucie
STP
Surry
Susquehanna
TMI
Turkey Pt 3&4
V.C. Summer 1
V.C. Summer 2&3
Vermont Yankee
Vogtle 1&2
Vogtle 3&4
Waterford
Watts Bar
Westinghouse
Wolf Creek
Zion
0%
1%
2%
3%
2006
2008
2010
2012

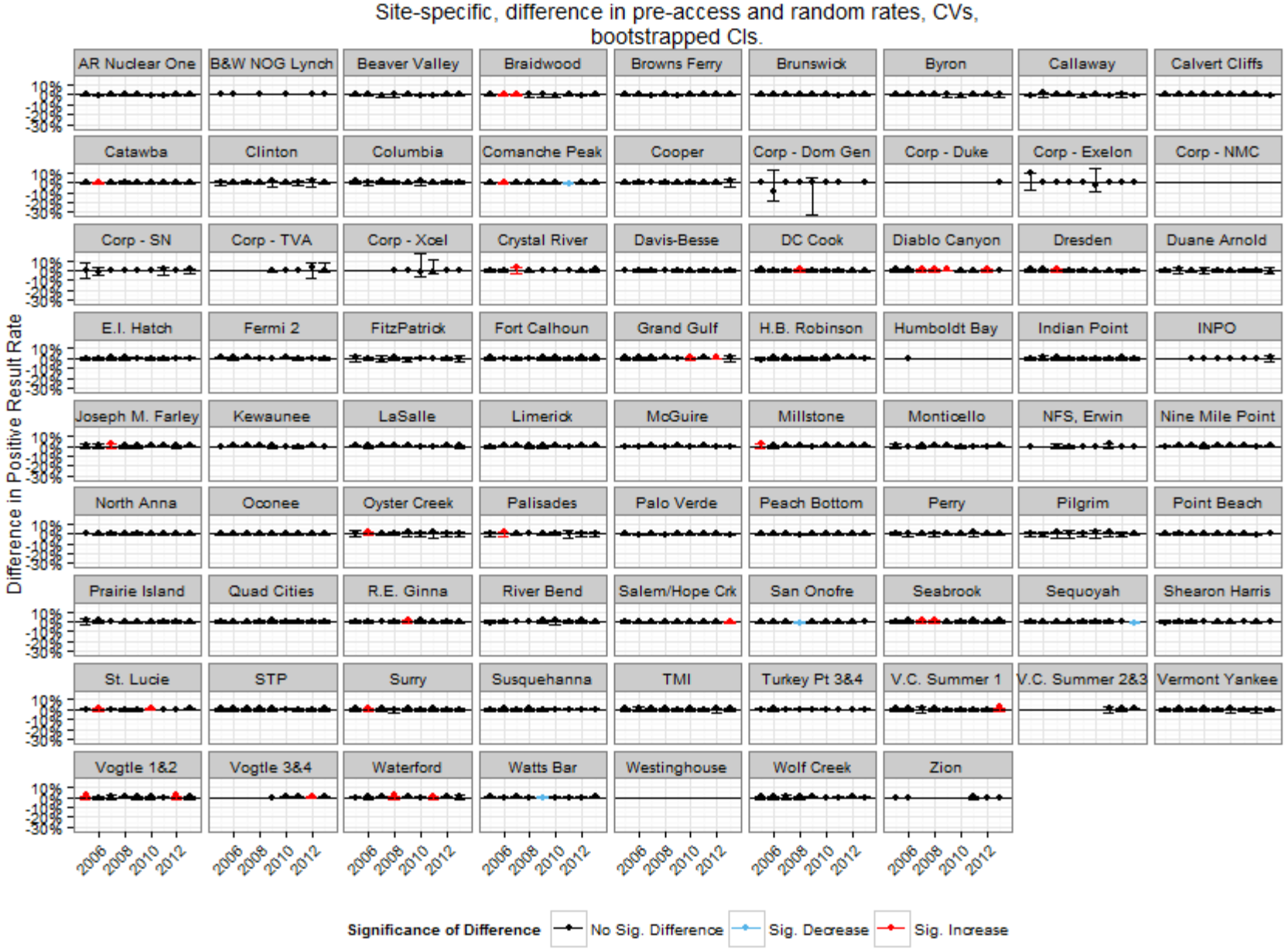

Site-specific, year-on-year difference in random positive rate, licensee employees, bootstrapped CIs.
Difference in Positive Result Rate
AR Nuclear One
B&W NOG Lynch
Beaver Valley
Braidwood
Browns Ferry
Brunswick
Byron
Callaway
Calvert Cliffs
Catawba
Clinton
Columbia
Comanche Peak
Cooper
Corp - Dom Gen
Corp - Duke
Corp - Exelon
Corp - NMC
Corp - SN
Corp - TVA
Corp - Xcel
Crystal River
Davis-Besse
DC Cook
Diablo Canyon
Dresden
Duane Arnold
E.I. Hatch
Fermi 2
FitzPatrick
Fort Calhoun
Grand Gulf
H.B. Robinson
Humboldt Bay
Indian Point
INPO
Joseph M. Farley
Kewaunee
LaSalle
Limerick
McGuire
Millstone
Monticello
NFS, Erwin
Nine Mile Point
North Anna
Oconee
Oyster Creek
Palisades
Palo Verde
Peach Bottom
Perry
Pilgrim
Point Beach
Prairie Island
Quad Cities
R.E. Ginna
River Bend
Salem/Hope Crk
San Onofre
Seabrook
Sequoyah
Shearon Harris
St. Lucie
STP
Surry
Susquehanna
TMI
Turkey Pt 3&4
V.C. Summer 1
V.C. Summer 2&3
Vermont Yankee
Vogtle 1&2
Vogtle 3&4
Waterford
Watts Bar
Westinghouse
Wolf Creek
Zion
10%
5%
0%
-5%
2006
2008
2010
2012
Significance of Difference
No Sig. Difference
Sig. Decrease
Sig. Increase

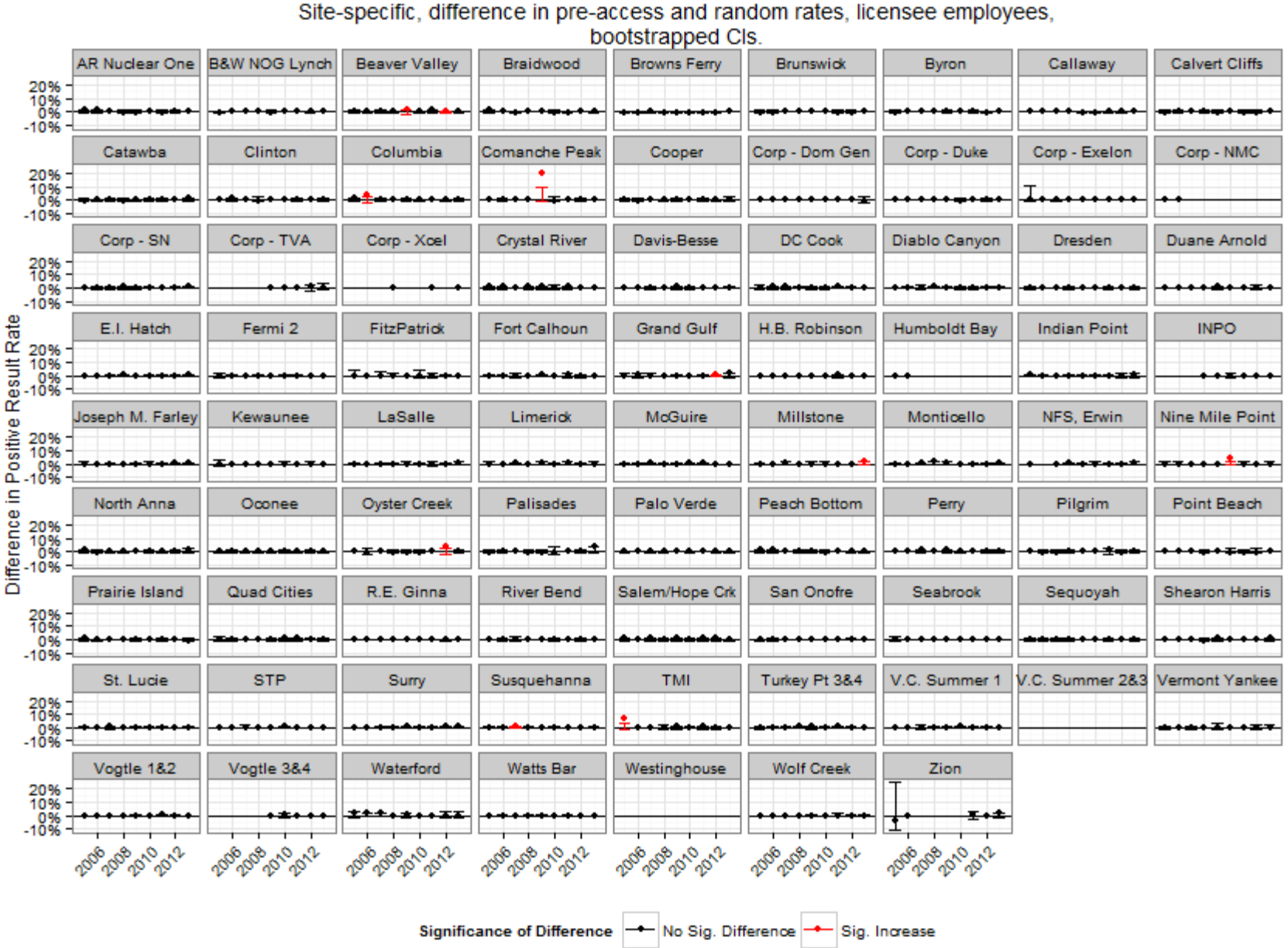

Site-specific for-cause positive rate, CVs.
Positive Result Rate
AR Nuclear One
B&W NOG Lynch
Beaver Valley
Braidwood
Browns Ferry
Brunswick
Byron
Callaway
Calvert Cliffs
Catawba
Clinton
Columbia
Comanche Peak
Cooper
Corp - Dom Gen
Corp - Duke
Corp - Exelon
Corp - NMC
Corp - SN
Corp - TVA
Corp - Xcel
Crystal River
Davis-Besse
DC Cook
Diablo Canyon
Dresden
Duane Arnold
E.I. Hatch
Fermi 2
FitzPatrick
Fort Calhoun
Grand Gulf
H.B. Robinson
Humboldt Bay
Indian Point
INPO
Joseph M. Farley
Kewaunee
LaSalle
Limerick
McGuire
Millstone
Monticello
NFS, Erwin
Nine Mile Point
North Anna
Oconee
Oyster Creek
Palisades
Palo Verde
Peach Bottom
Perry
Pilgrim
Point Beach
Prairie Island
Quad Cities
R.E. Ginna
River Bend
Salem/Hope Crk
San Onofre
Seabrook
Sequoyah
Shearon Harris
St. Lucie
STP
Surry
Susquehanna
TMI
Turkey Pt 3&4
V.C. Summer 1
V.C. Summer 2&3
Vermont Yankee
Vogtle 1&2
Vogtle 3&4
Waterford
Watts Bar
Westinghouse
Wolf Creek
Zion
100%
75%
50%
25%
0%
2006
2008
2010
2012

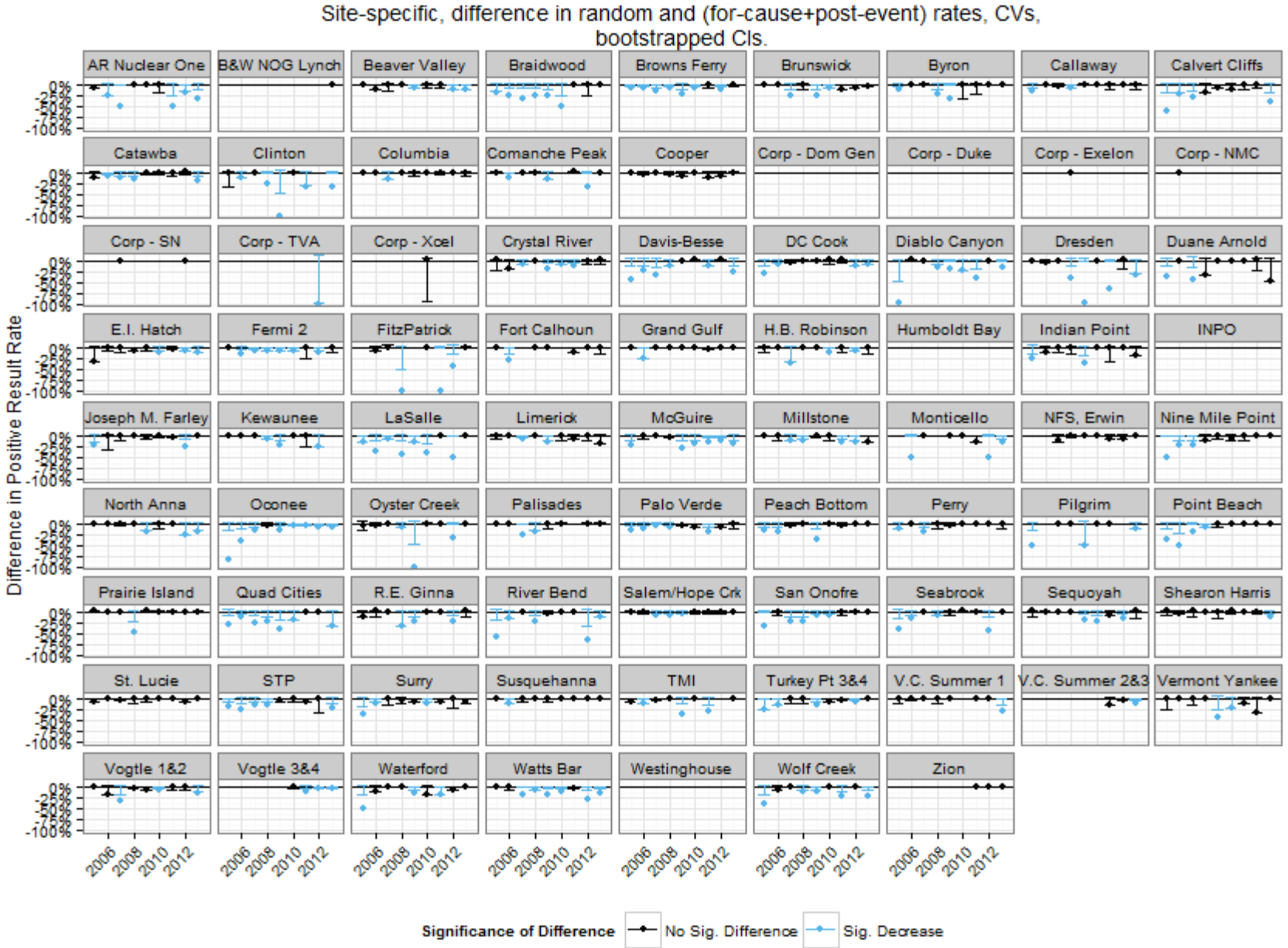

Site-specific, year-on-year difference in for-cause positive rate, CVs, bootstrapped CIs.
Difference in Positive Result Rate
AR Nuclear One
B&W NOG Lynch
Beaver Valley
Braidwood
Browns Ferry
Brunswick
Byron
Callaway
Calvert Cliffs
Catawba
Clinton
Columbia
Comanche Peak
Cooper
Corp - Dom Gen
Corp - Duke
Corp - Exelon
Corp - NMC
Corp - SN
Corp - TVA
Corp - Xcel
Crystal River
Davis-Besse
DC Cook
Diablo Canyon
Dresden
Duane Arnold
E.I. Hatch
Fermi 2
FitzPatrick
Fort Calhoun
Grand Gulf
H.B. Robinson
Humboldt Bay
Indian Point
INPO
Joseph M. Farley
Kewaunee
LaSalle
Limerick
McGuire
Millstone
Monticello
NFS, Erwin
Nine Mile Point
North Anna
Oconee
Oyster Creek
Palisades
Palo Verde
Peach Bottom
Perry
Pilgrim
Point Beach
Prairie Island
Quad Cities
R.E. Ginna
River Bend
Salem/Hope Crk
San Onofre
Seabrook
Sequoyah
Shearon Harris
St. Lucie
STP
Surry
Susquehanna
TMI
Turkey Pt 3&4
V.C. Summer 1
V.C. Summer 2&3
Vermont Yankee
Vogtle 1&2
Vogtle 3&4
Waterford
Watts Bar
Westinghouse
Wolf Creek
Zion
100%
50%
0%
-50%
-100%
2006
2008
2010
2012
Significance of Difference
No Sig. Difference
Sig. Decrease
Sig. Increase

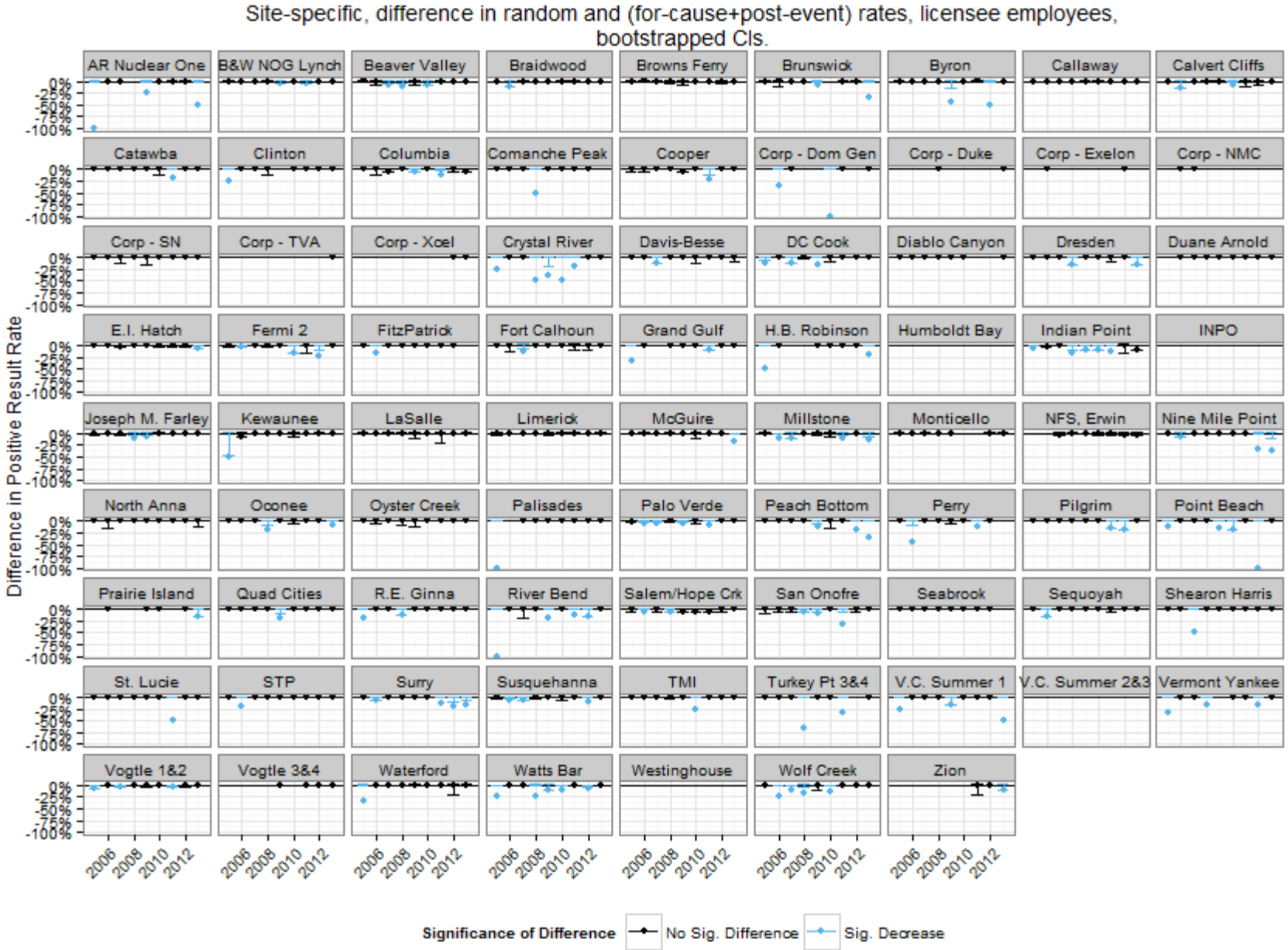

Site-specific for-cause positive rate, licensee employees.
Positive Result Rate
AR Nuclear One
B&W NOG Lynch
Beaver Valley
Braidwood
Browns Ferry
Brunswick
Byron
Callaway
Calvert Cliffs
Catawba
Clinton
Columbia
Comanche Peak
Cooper
Corp - Dom Gen
Corp - Duke
Corp - Exelon
Corp - NMC
Corp - SN
Corp - TVA
Corp - Xcel
Crystal River
Davis-Besse
DC Cook
Diablo Canyon
Dresden
Duane Arnold
E.I. Hatch
Fermi 2
FitzPatrick
Fort Calhoun
Grand Gulf
H.B. Robinson
Humboldt Bay
Indian Point
INPO
Joseph M. Farley
Kewaunee
LaSalle
Limerick
McGuire
Millstone
Monticello
NFS, Erwin
Nine Mile Point
North Anna
Oconee
Oyster Creek
Palisades
Palo Verde
Peach Bottom
Perry
Pilgrim
Point Beach
Prairie Island
Quad Cities
R.E. Ginna
River Bend
Salem/Hope Crk
San Onofre
Seabrook
Sequoyah
Shearon Harris
St. Lucie
STP
Surry
Susquehanna
TMI
Turkey Pt 3&4
V.C. Summer 1
V.C. Summer 2&3
Vermont Yankee
Vogtle 1&2
Vogtle 3&4
Waterford
Watts Bar
Westinghouse
Wolf Creek
Zion
100%
75%
50%
25%
0%
2006
2008
2010
2012

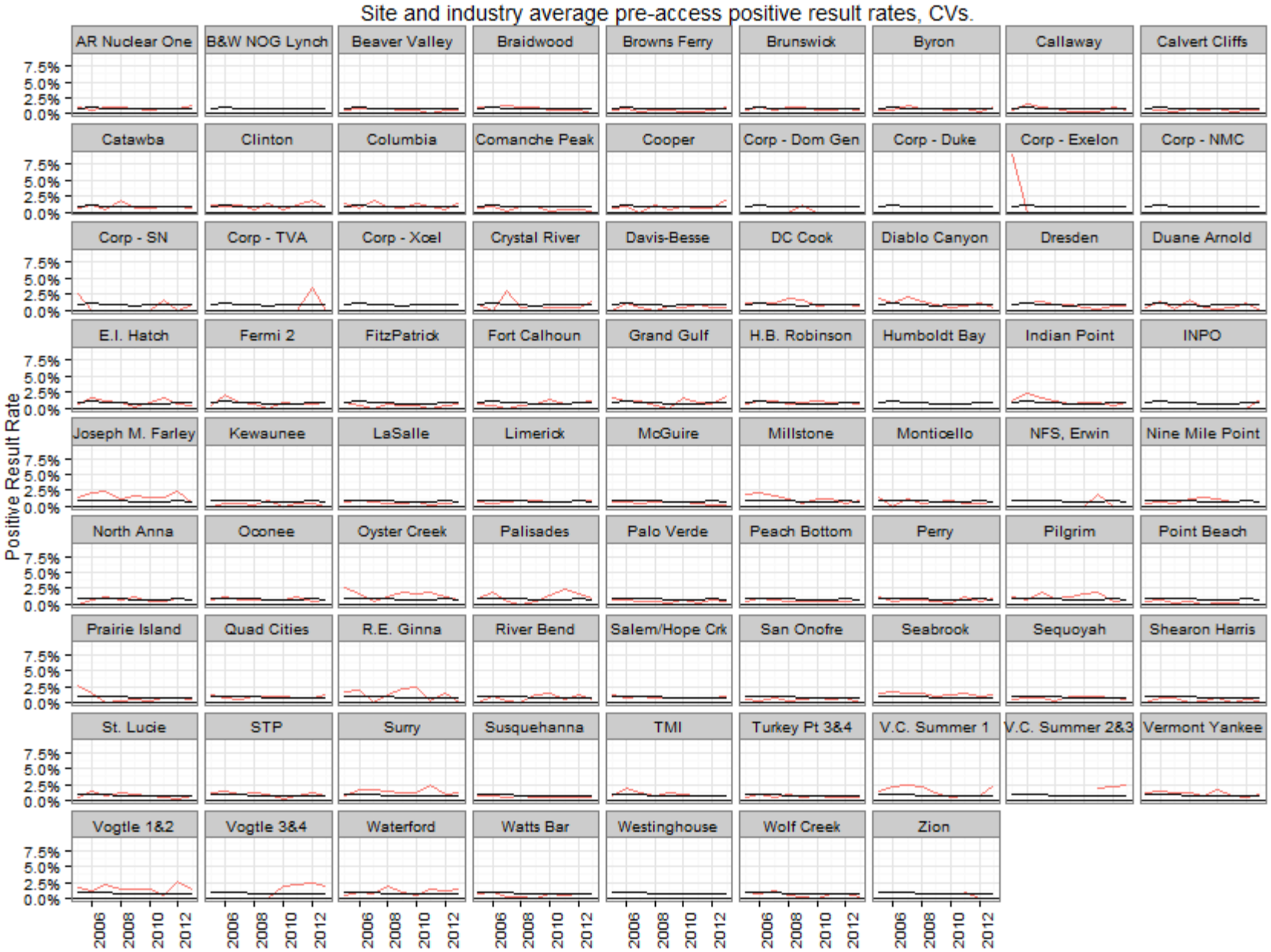

Site-specific, year-on-year difference in for-cause positive rate, licensee employees, bootstrapped CIs.
Difference in Positive Result Rate
AR Nuclear One
B&W NOG Lynch
Beaver Valley
Braidwood
Browns Ferry
Brunswick
Byron
Callaway
Calvert Cliffs
Catawba
Clinton
Columbia
Comanche Peak
Cooper
Corp - Dom Gen
Corp - Duke
Corp - Exelon
Corp - NMC
Corp - SN
Corp - TVA
Corp - Xcel
Crystal River
Davis-Besse
DC Cook
Diablo Canyon
Dresden
Duane Arnold
E.I. Hatch
Fermi 2
FitzPatrick
Fort Calhoun
Grand Gulf
H.B. Robinson
Humboldt Bay
Indian Point
INPO
Joseph M. Farley
Kewaunee
LaSalle
Limerick
McGuire
Millstone
Monticello
NFS, Erwin
Nine Mile Point
North Anna
Oconee
Oyster Creek
Palisades
Palo Verde
Peach Bottom
Perry
Pilgrim
Point Beach
Prairie Island
Quad Cities
R.E. Ginna
River Bend
Salem/Hope Crk
San Onofre
Seabrook
Sequoyah
Shearon Harris
St. Lucie
STP
Surry
Susquehanna
TMI
Turkey Pt 3&4
V.C. Summer 1
V.C. Summer 2&3
Vermont Yankee
Vogtle 1&2
Vogtle 3&4
Waterford
Watts Bar
Westinghouse
Wolf Creek
Zion
100%
50%
0%
-50%
-100%
2006
2008
2010
2012
Significance of Difference
No Sig. Difference
Sig. Decrease

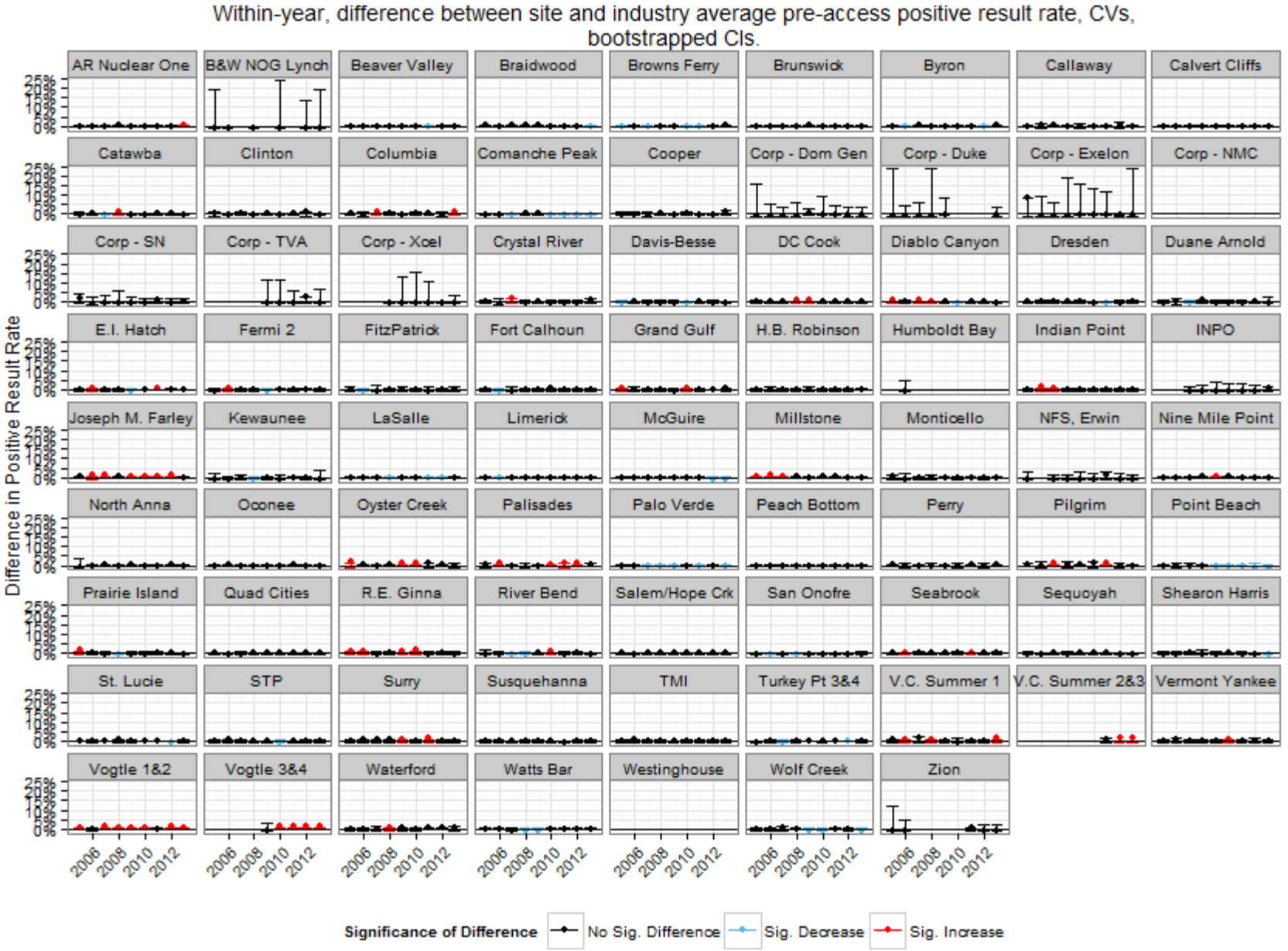

Site-specific, difference in pre-access and random rates, CVs, bootstrapped CIs.
Difference in Positive Result Rate
AR Nuclear One
B&W NOG Lynch
Beaver Valley
Braidwood
Browns Ferry
Brunswick
Byron
Callaway
Calvert Cliffs
Catawba
Clinton
Columbia
Comanche Peak
Cooper
Corp - Dom Gen
Corp - Duke
Corp - Exelon
Corp - NMC
Corp - SN
Corp - TVA
Corp - Xcel
Crystal River
Davis-Besse
DC Cook
Diablo Canyon
Dresden
Duane Arnold
E.I. Hatch
Fermi 2
FitzPatrick
Fort Calhoun
Grand Gulf
H.B. Robinson
Humboldt Bay
Indian Point
INPO
Joseph M. Farley
Kewaunee
LaSalle
Limerick
McGuire
Millstone
Monticello
NFS, Erwin
Nine Mile Point
North Anna
Oconee
Oyster Creek
Palisades
Palo Verde
Peach Bottom
Perry
Pilgrim
Point Beach
Prairie Island
Quad Cities
R.E. Ginna
River Bend
Salem/Hope Crk
San Onofre
Seabrook
Sequoyah
Shearon Harris
St. Lucie
STP
Surry
Susquehanna
TMI
Turkey Pt 3&4
V.C. Summer 1
V.C. Summer 2&3
Vermont Yankee
Vogtle 1&2
Vogtle 3&4
Waterford
Watts Bar
Westinghouse
Wolf Creek
Zion
10%
0%
-10%
-20%
-30%
2006
2008
2010
2012
Significance of Difference
No Sig. Difference
Sig. Decrease
Sig. Increase


Site-specific, difference in pre-access and random rates, licensee employees, bootstrapped CIs.
Difference in Positive Result Rate
AR Nuclear One
B&W NOG Lynch
Beaver Valley
Braidwood
Browns Ferry
Brunswick
Byron
Callaway
Calvert Cliffs
Catawba
Clinton
Columbia
Comanche Peak
Cooper
Corp - Dom Gen
Corp - Duke
Corp - Exelon
Corp - NMC
Corp - SN
Corp - TVA
Corp - Xcel
Crystal River
Davis-Besse
DC Cook
Diablo Canyon
Dresden
Duane Arnold
E.I. Hatch
Fermi 2
FitzPatrick
Fort Calhoun
Grand Gulf
H.B. Robinson
Humboldt Bay
Indian Point
INPO
Joseph M. Farley
Kewaunee
LaSalle
Limerick
McGuire
Millstone
Monticello
NFS, Erwin
Nine Mile Point
North Anna
Oconee
Oyster Creek
Palisades
Palo Verde
Peach Bottom
Perry
Pilgrim
Point Beach
Prairie Island
Quad Cities
R.E. Ginna
River Bend
Salem/Hope Crk
San Onofre
Seabrook
Sequoyah
Shearon Harris
St. Lucie
STP
Surry
Susquehanna
TMI
Turkey Pt 3&4
V.C. Summer 1
V.C. Summer 2&3
Vermont Yankee
Vogtle 1&2
Vogtle 3&4
Waterford
Watts Bar
Westinghouse
Wolf Creek
Zion
20%
10%
0%
-10%
2006
2008
2010
2012
Significance of Difference
No Sig. Difference
Sig. Increase

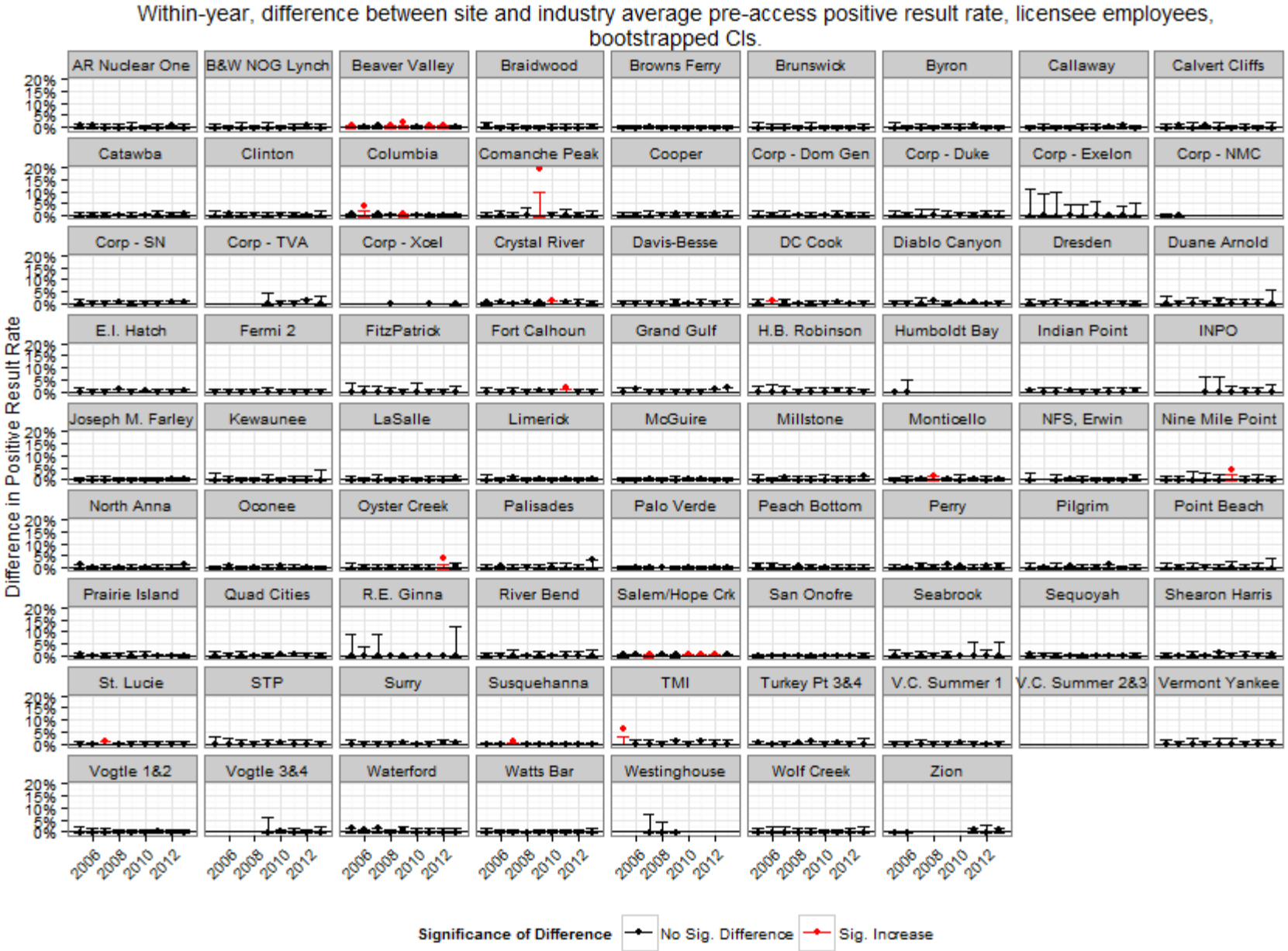

Site-specific, difference in random and (for-cause+post-event) rates, CVs, bootstrapped CIs.
Difference in Positive Result Rate
AR Nuclear One
B&W NOG Lynch
Beaver Valley
Braidwood
Browns Ferry
Brunswick
Byron
Callaway
Calvert Cliffs
Catawba
Clinton
Columbia
Comanche Peak
Cooper
Corp - Dom Gen
Corp - Duke
Corp - Exelon
Corp - NMC
Corp - SN
Corp - TVA
Corp - Xcel
Crystal River
Davis-Besse
DC Cook
Diablo Canyon
Dresden
Duane Arnold
E.I. Hatch
Fermi 2
FitzPatrick
Fort Calhoun
Grand Gulf
H.B. Robinson
Humboldt Bay
Indian Point
INPO
Joseph M. Farley
Kewaunee
LaSalle
Limerick
McGuire
Millstone
Monticello
NFS, Erwin
Nine Mile Point
North Anna
Oconee
Oyster Creek
Palisades
Palo Verde
Peach Bottom
Perry
Pilgrim
Point Beach
Prairie Island
Quad Cities
R.E. Ginna
River Bend
Salem/Hope Crk
San Onofre
Seabrook
Sequoyah
Shearon Harris
St. Lucie
STP
Surry
Susquehanna
TMI
Turkey Pt 3&4
V.C. Summer 1
V.C. Summer 2&3
Vermont Yankee
Vogtle 1&2
Vogtle 3&4
Waterford
Watts Bar
Westinghouse
Wolf Creek
Zion
0%
-25%
-50%
-75%
-100%
2006
2008
2010
2012
Significance of Difference
No Sig. Difference
Sig. Decrease

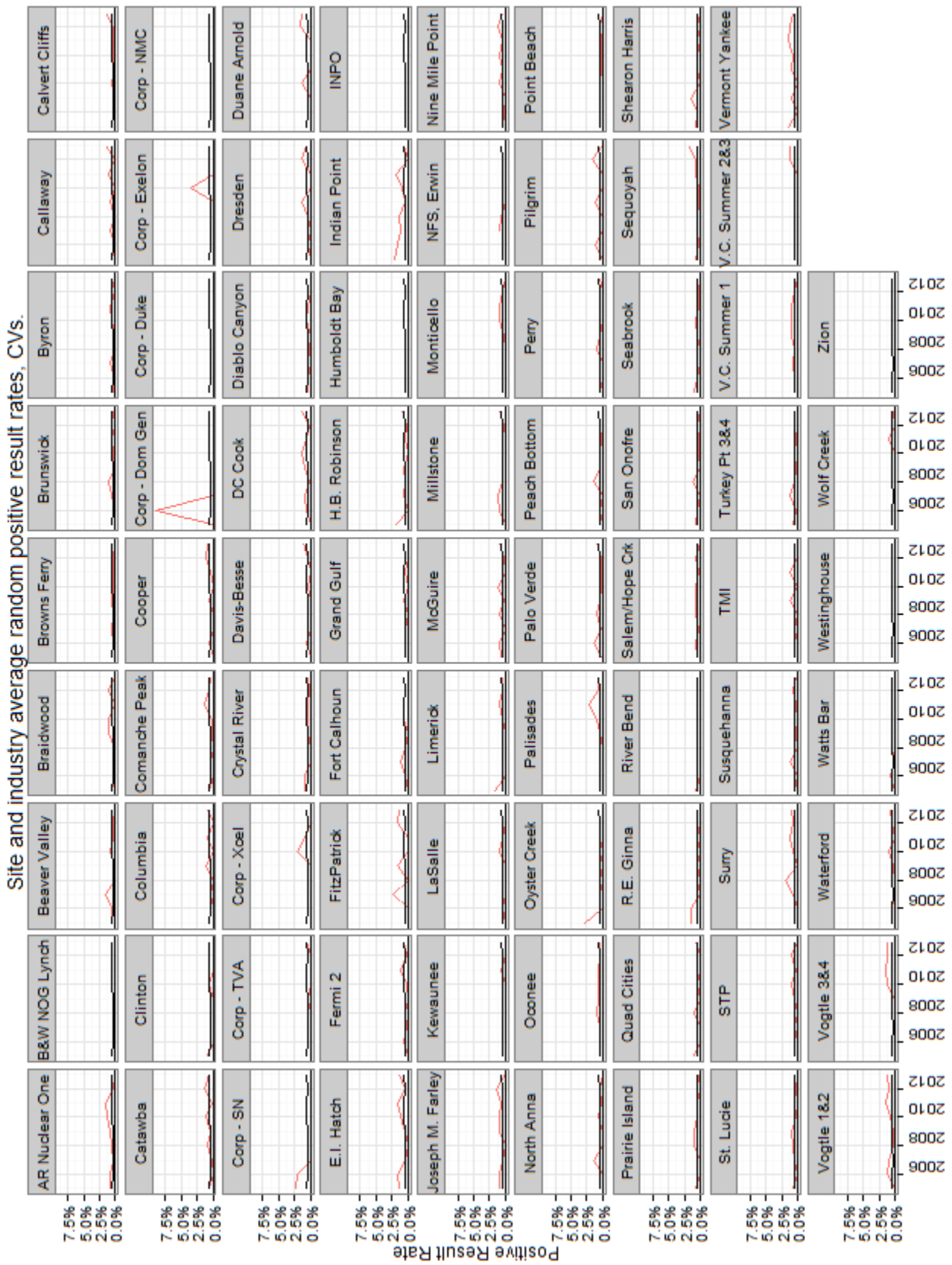

Site-specific, difference in random and (for-cause+post-event) rates, licensee employees, bootstrapped CIs.
Difference in Positive Result Rate
AR Nuclear One
B&W NOG Lynch
Beaver Valley
Braidwood
Browns Ferry
Brunswick
Byron
Callaway
Calvert Cliffs
Catawba
Clinton
Columbia
Comanche Peak
Cooper
Corp - Dom Gen
Corp - Duke
Corp - Exelon
Corp - NMC
Corp - SN
Corp - TVA
Corp - Xcel
Crystal River
Davis-Besse
DC Cook
Diablo Canyon
Dresden
Duane Arnold
E.I. Hatch
Fermi 2
FitzPatrick
Fort Calhoun
Grand Gulf
H.B. Robinson
Humboldt Bay
Indian Point
INPO
Joseph M. Farley
Kewaunee
LaSalle
Limerick
McGuire
Millstone
Monticello
NFS, Erwin
Nine Mile Point
North Anna
Oconee
Oyster Creek
Palisades
Palo Verde
Peach Bottom
Perry
Pilgrim
Point Beach
Prairie Island
Quad Cities
R.E. Ginna
River Bend
Salem/Hope Crk
San Onofre
Seabrook
Sequoyah
Shearon Harris
St. Lucie
STP
Surry
Susquehanna
TMI
Turkey Pt 3&4
V.C. Summer 1
V.C. Summer 2&3
Vermont Yankee
Vogtle 1&2
Vogtle 3&4
Waterford
Watts Bar
Westinghouse
Wolf Creek
Zion
0%
-25%
-50%
-75%
-100%
2006
2008
2010
2012
Significance of Difference
No Sig. Difference
Sig. Decrease

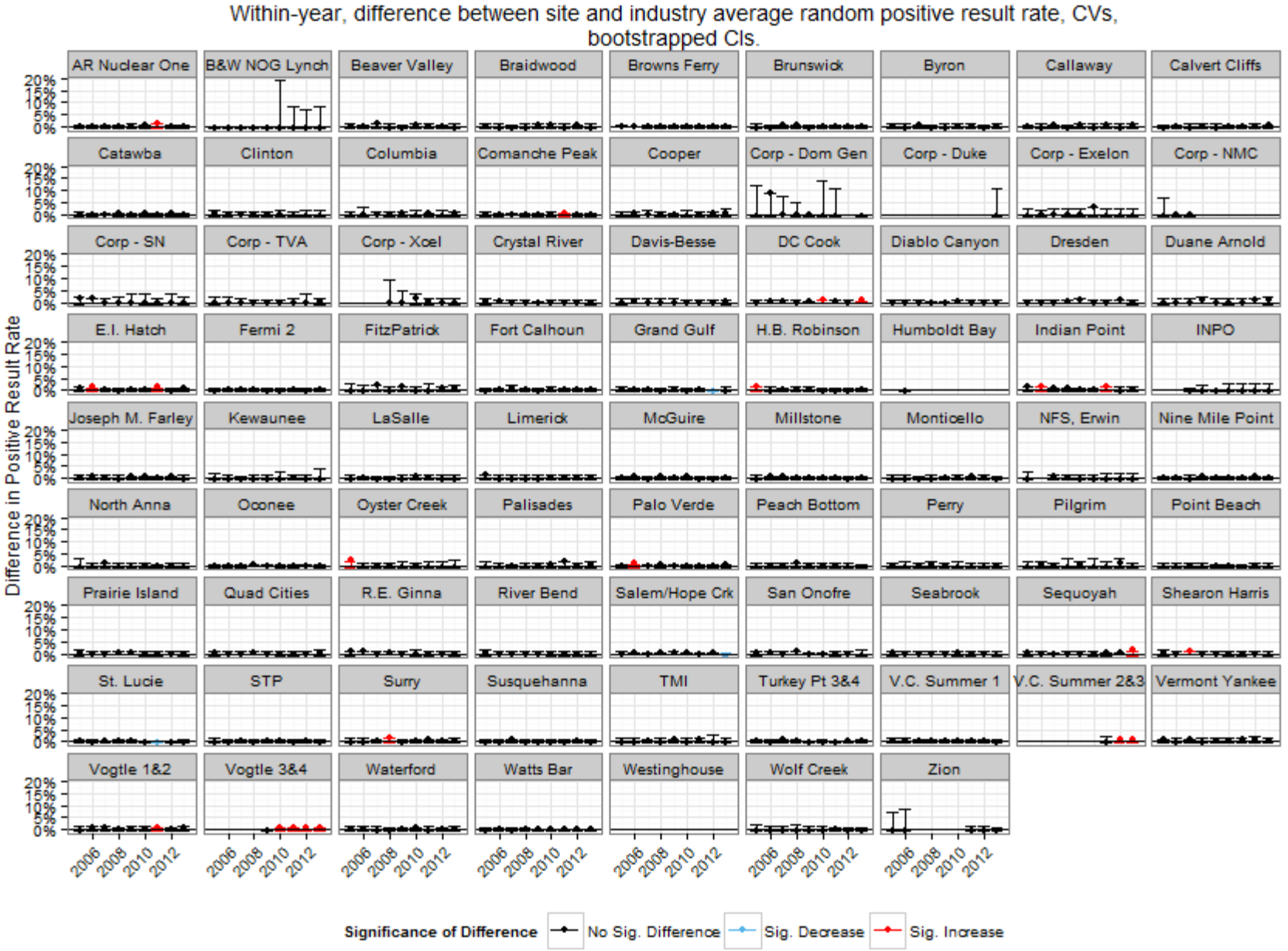

Site and industry average pre-access positive result rates, CVs.
Positive Result Rate
AR Nuclear One
B&W NOG Lynch
Beaver Valley
Braidwood
Browns Ferry
Brunswick
Byron
Callaway
Calvert Cliffs
Catawba
Clinton
Columbia
Comanche Peak
Cooper
Corp - Dom Gen
Corp - Duke
Corp - Exelon
Corp - NMC
Corp - SN
Corp - TVA
Corp - Xcel
Crystal River
Davis-Besse
DC Cook
Diablo Canyon
Dresden
Duane Arnold
E.I. Hatch
Fermi 2
FitzPatrick
Fort Calhoun
Grand Gulf
H.B. Robinson
Humboldt Bay
Indian Point
INPO
Joseph M. Farley
Kewaunee
LaSalle
Limerick
McGuire
Millstone
Monticello
NFS, Erwin
Nine Mile Point
North Anna
Oconee
Oyster Creek
Palisades
Palo Verde
Peach Bottom
Perry
Pilgrim
Point Beach
Prairie Island
Quad Cities
R.E. Ginna
River Bend
Salem/Hope Crk
San Onofre
Seabrook
Sequoyah
Shearon Harris
St. Lucie
STP
Surry
Susquehanna
TMI
Turkey Pt 3&4
V.C. Summer 1
V.C. Summer 2&3
Vermont Yankee
Vogtle 1&2
Vogtle 3&4
Waterford
Watts Bar
Westinghouse
Wolf Creek
Zion
7.5%
5.0%
2.5%
0.0%
2006
2008
2010
2012

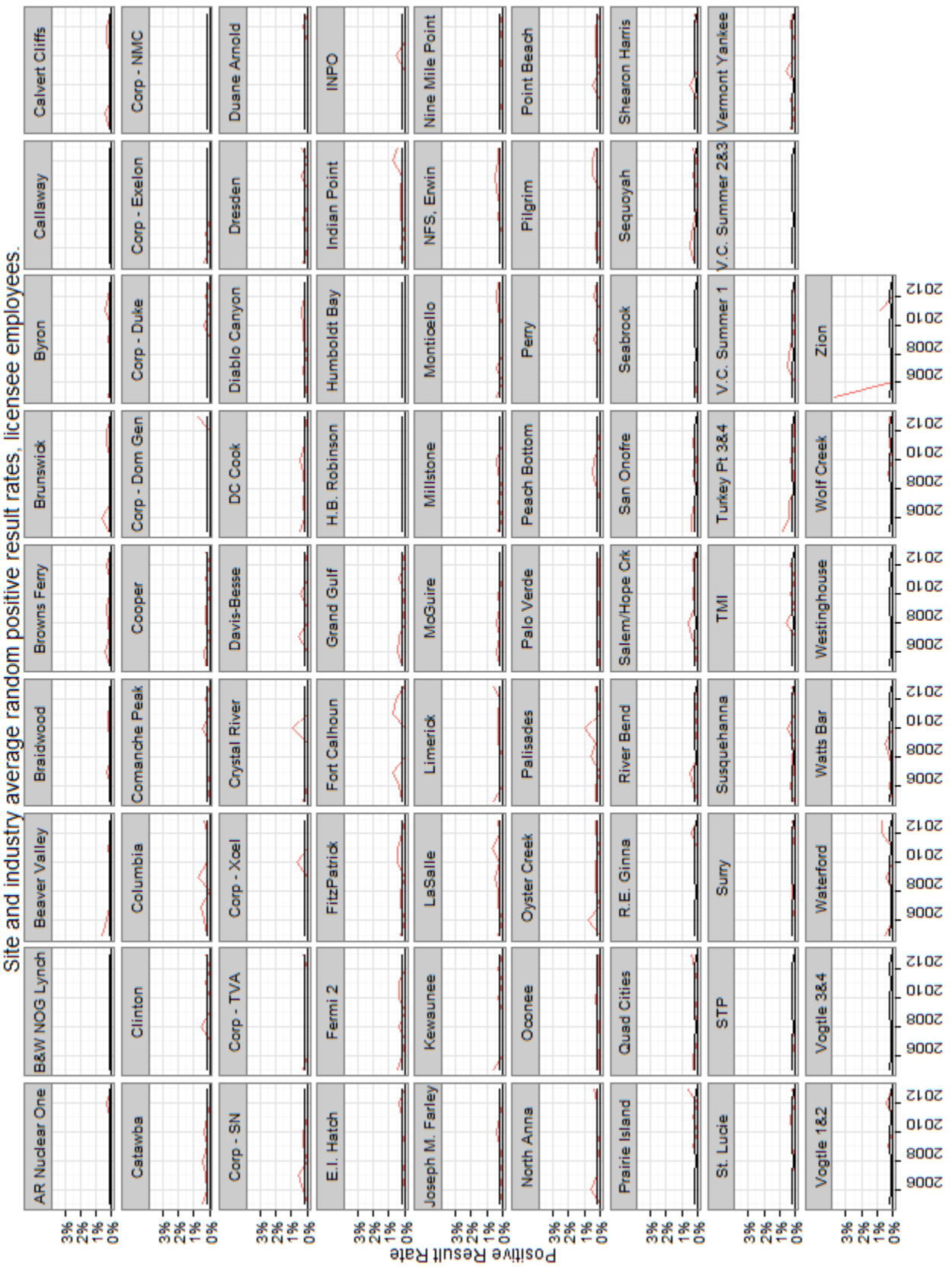

Within-year, difference between site and industry average pre-access positive result rate, CVs, bootstrapped CIs.
Difference in Positive Result Rate
AR Nuclear One
B&W NOG Lynch
Beaver Valley
Braidwood
Browns Ferry
Brunswick
Byron
Callaway
Calvert Cliffs
Catawba
Clinton
Columbia
Comanche Peak
Cooper
Corp - Dom Gen
Corp - Duke
Corp - Exelon
Corp - NMC
Corp - SN
Corp - TVA
Corp - Xcel
Crystal River
Davis-Besse
DC Cook
Diablo Canyon
Dresden
Duane Arnold
E.I. Hatch
Fermi 2
FitzPatrick
Fort Calhoun
Grand Gulf
H.B. Robinson
Humboldt Bay
Indian Point
INPO
Joseph M. Farley
Kewaunee
LaSalle
Limerick
McGuire
Millstone
Monticello
NFS, Erwin
Nine Mile Point
North Anna
Oconee
Oyster Creek
Palisades
Palo Verde
Peach Bottom
Perry
Pilgrim
Point Beach
Prairie Island
Quad Cities
R.E. Ginna
River Bend
Salem/Hope Crk
San Onofre
Seabrook
Sequoyah
Shearon Harris
St. Lucie
STP
Surry
Susquehanna
TMI
Turkey Pt 3&4
V.C. Summer 1
V.C. Summer 2&3
Vermont Yankee
Vogtle 1&2
Vogtle 3&4
Waterford
Watts Bar
Westinghouse
Wolf Creek
Zion
25%
20%
15%
10%
5%
0%
2006
2008
2010
2012
Significance of Difference
No Sig. Difference
Sig. Decrease
Sig. Increase

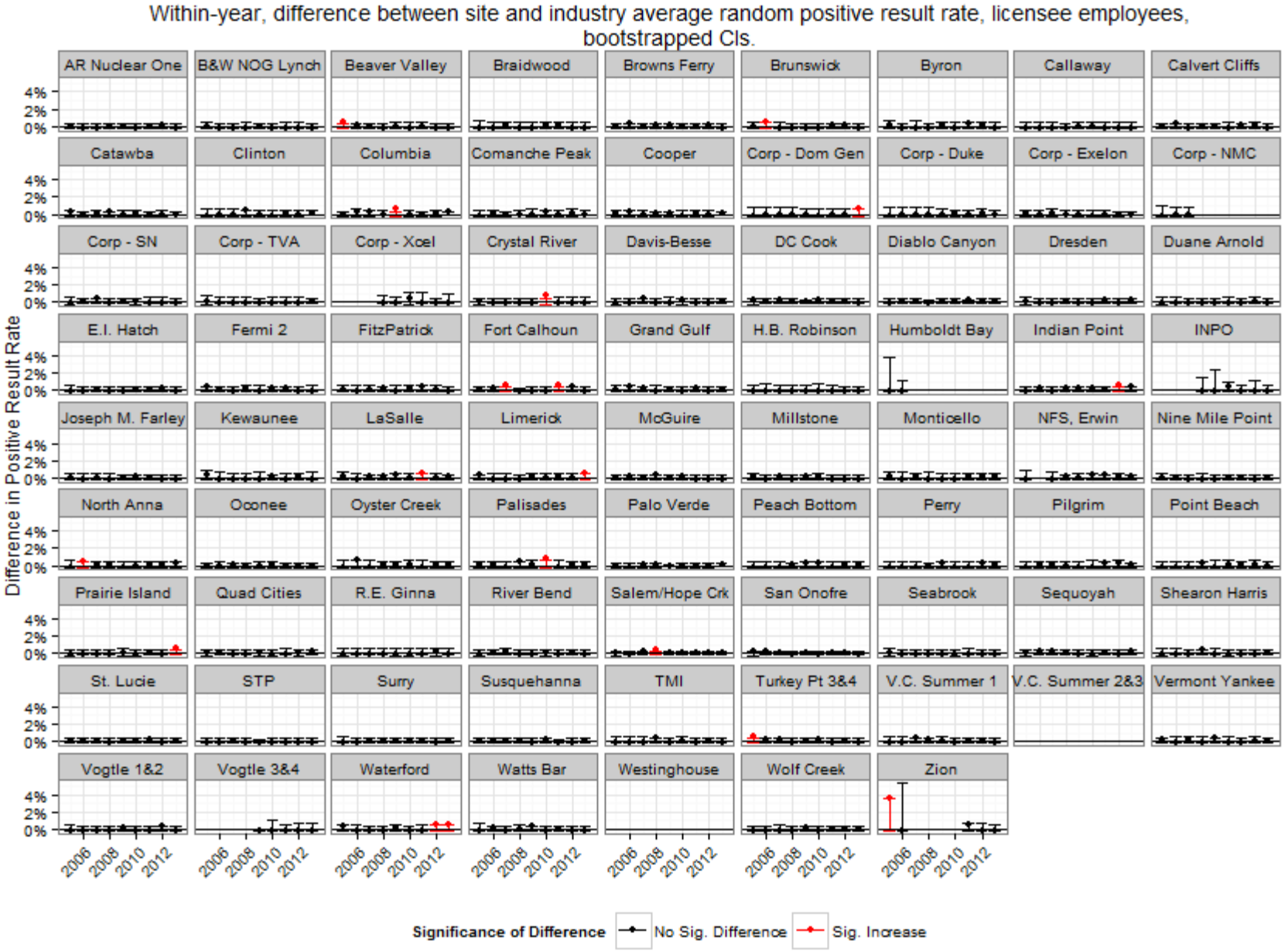

Site and industry average pre-access positive result rates, licensee employees.
Positive Result Rate
AR Nuclear One
B&W NOG Lynch
Beaver Valley
Braidwood
Browns Ferry
Brunswick
Byron
Callaway
Calvert Cliffs
Catawba
Clinton
Columbia
Comanche Peak
Cooper
Corp - Dom Gen
Corp - Duke
Corp - Exelon
Corp - NMC
Corp - SN
Corp - TVA
Corp - Xcel
Crystal River
Davis-Besse
DC Cook
Diablo Canyon
Dresden
Duane Arnold
E.I. Hatch
Fermi 2
FitzPatrick
Fort Calhoun
Grand Gulf
H.B. Robinson
Humboldt Bay
Indian Point
INPO
Joseph M. Farley
Kewaunee
LaSalle
Limerick
McGuire
Millstone
Monticello
NFS, Erwin
Nine Mile Point
North Anna
Oconee
Oyster Creek
Palisades
Palo Verde
Peach Bottom
Perry
Pilgrim
Point Beach
Prairie Island
Quad Cities
R.E. Ginna
River Bend
Salem/Hope Crk
San Onofre
Seabrook
Sequoyah
Shearon Harris
St. Lucie
STP
Surry
Susquehanna
TMI
Turkey Pt 3&4
V.C. Summer 1
V.C. Summer 2&3
Vermont Yankee
Vogtle 1&2
Vogtle 3&4
Waterford
Watts Bar
Westinghouse
Wolf Creek
Zion
20%
15%
10%
5%
0%
2006
2008
2010
2012

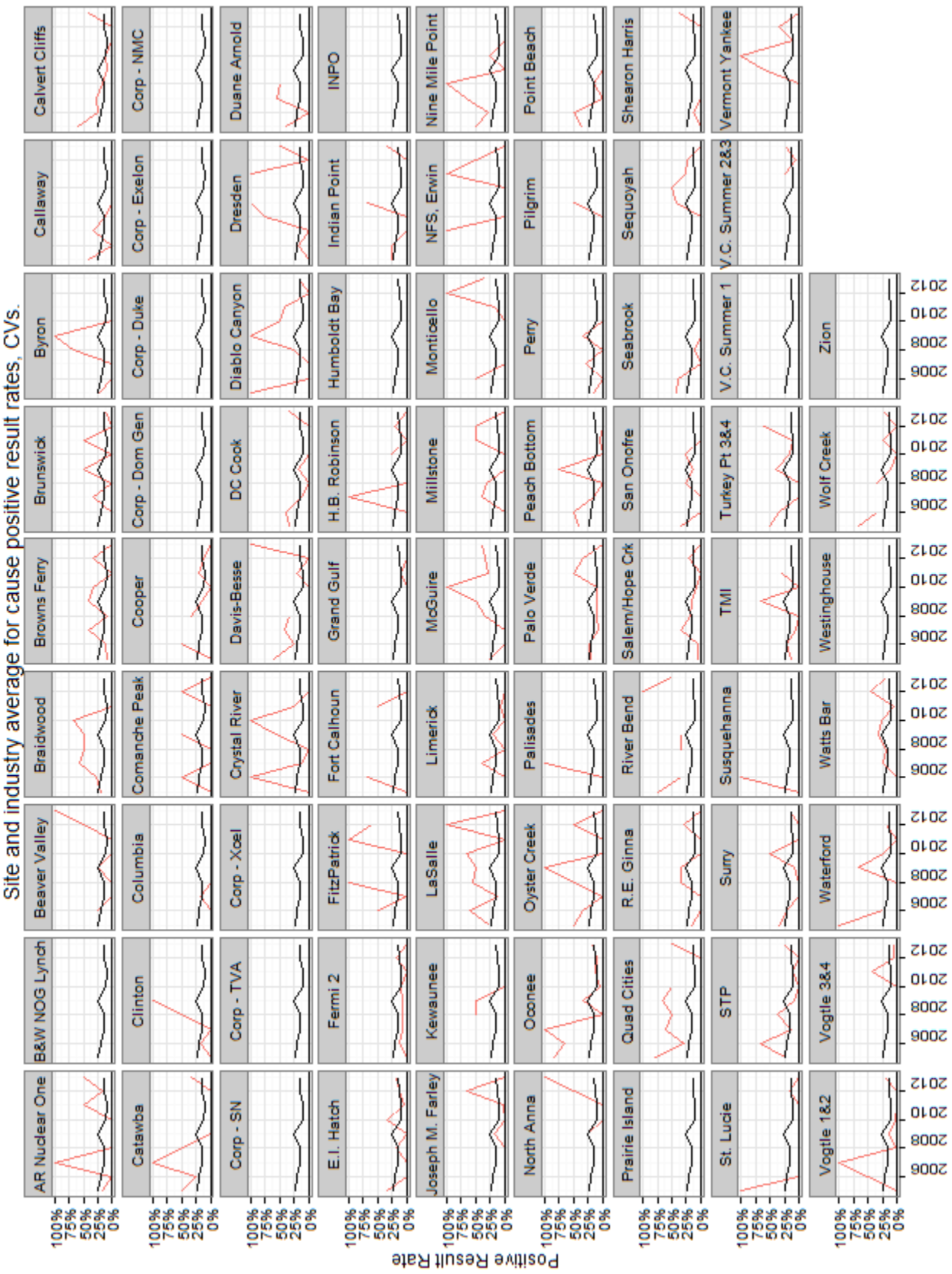

Within-year, difference between site and industry average pre-access positive result rate, licensee employees, bootstrapped CIs.
Difference in Positive Result Rate
AR Nuclear One
B&W NOG Lynch
Beaver Valley
Braidwood
Browns Ferry
Brunswick
Byron
Callaway
Calvert Cliffs
Catawba
Clinton
Columbia
Comanche Peak
Cooper
Corp - Dom Gen
Corp - Duke
Corp - Exelon
Corp - NMC
Corp - SN
Corp - TVA
Corp - Xcel
Crystal River
Davis-Besse
DC Cook
Diablo Canyon
Dresden
Duane Arnold
E.I. Hatch
Fermi 2
FitzPatrick
Fort Calhoun
Grand Gulf
H.B. Robinson
Humboldt Bay
Indian Point
INPO
Joseph M. Farley
Kewaunee
LaSalle
Limerick
McGuire
Millstone
Monticello
NFS, Erwin
Nine Mile Point
North Anna
Oconee
Oyster Creek
Palisades
Palo Verde
Peach Bottom
Perry
Pilgrim
Point Beach
Prairie Island
Quad Cities
R.E. Ginna
River Bend
Salem/Hope Crk
San Onofre
Seabrook
Sequoyah
Shearon Harris
St. Lucie
STP
Surry
Susquehanna
TMI
Turkey Pt 3&4
V.C. Summer 1
V.C. Summer 2&3
Vermont Yankee
Vogtle 1&2
Vogtle 3&4
Waterford
Watts Bar
Westinghouse
Wolf Creek
Zion
20%
15%
10%
5%
0%
2006
2008
2010
2012
Significance of Difference
No Sig. Difference
Sig. Increase

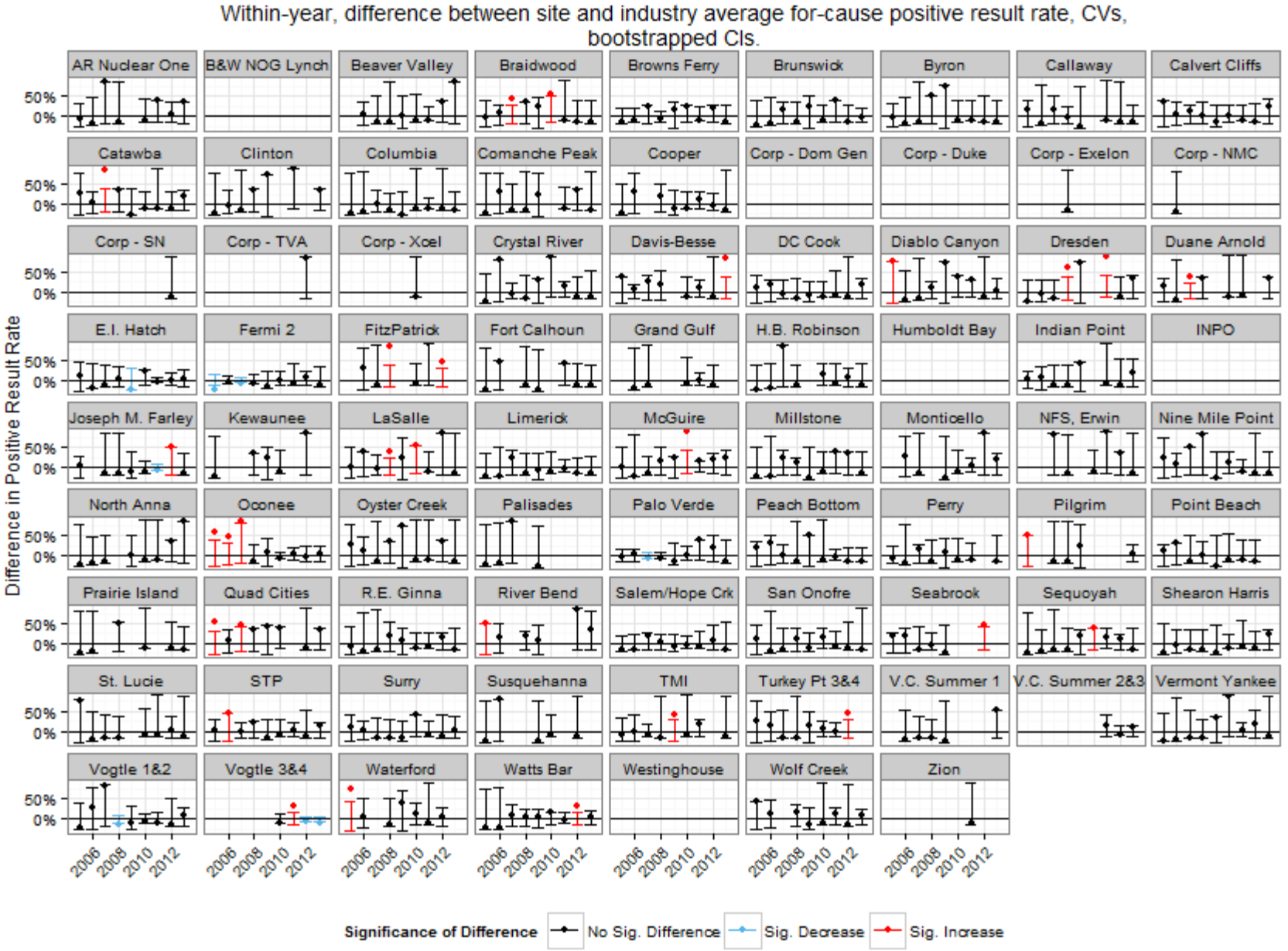

Site and industry average random positive result rates, CVs.
Positive Result Rate
AR Nuclear One
B&W NOG Lynch
Beaver Valley
Braidwood
Browns Ferry
Brunswick
Byron
Callaway
Calvert Cliffs
Catawba
Clinton
Columbia
Comanche Peak
Cooper
Corp - Dom Gen
Corp - Duke
Corp - Exelon
Corp - NMC
Corp - SN
Corp - TVA
Corp - Xcel
Crystal River
Davis-Besse
DC Cook
Diablo Canyon
Dresden
Duane Arnold
E.I. Hatch
Fermi 2
FitzPatrick
Fort Calhoun
Grand Gulf
H.B. Robinson
Humboldt Bay
Indian Point
INPO
Joseph M. Farley
Kewaunee
LaSalle
Limerick
McGuire
Millstone
Monticello
NFS, Erwin
Nine Mile Point
North Anna
Oconee
Oyster Creek
Palisades
Palo Verde
Peach Bottom
Perry
Pilgrim
Point Beach
Prairie Island
Quad Cities
R.E. Ginna
River Bend
Salem/Hope Crk
San Onofre
Seabrook
Sequoyah
Shearon Harris
St. Lucie
STP
Surry
Susquehanna
TMI
Turkey Pt 3&4
V.C. Summer 1
V.C. Summer 2&3
Vermont Yankee
Vogtle 1&2
Vogtle 3&4
Waterford
Watts Bar
Westinghouse
Wolf Creek
Zion
7.5%
5.0%
2.5%
0.0%
2006
2008
2010
2012

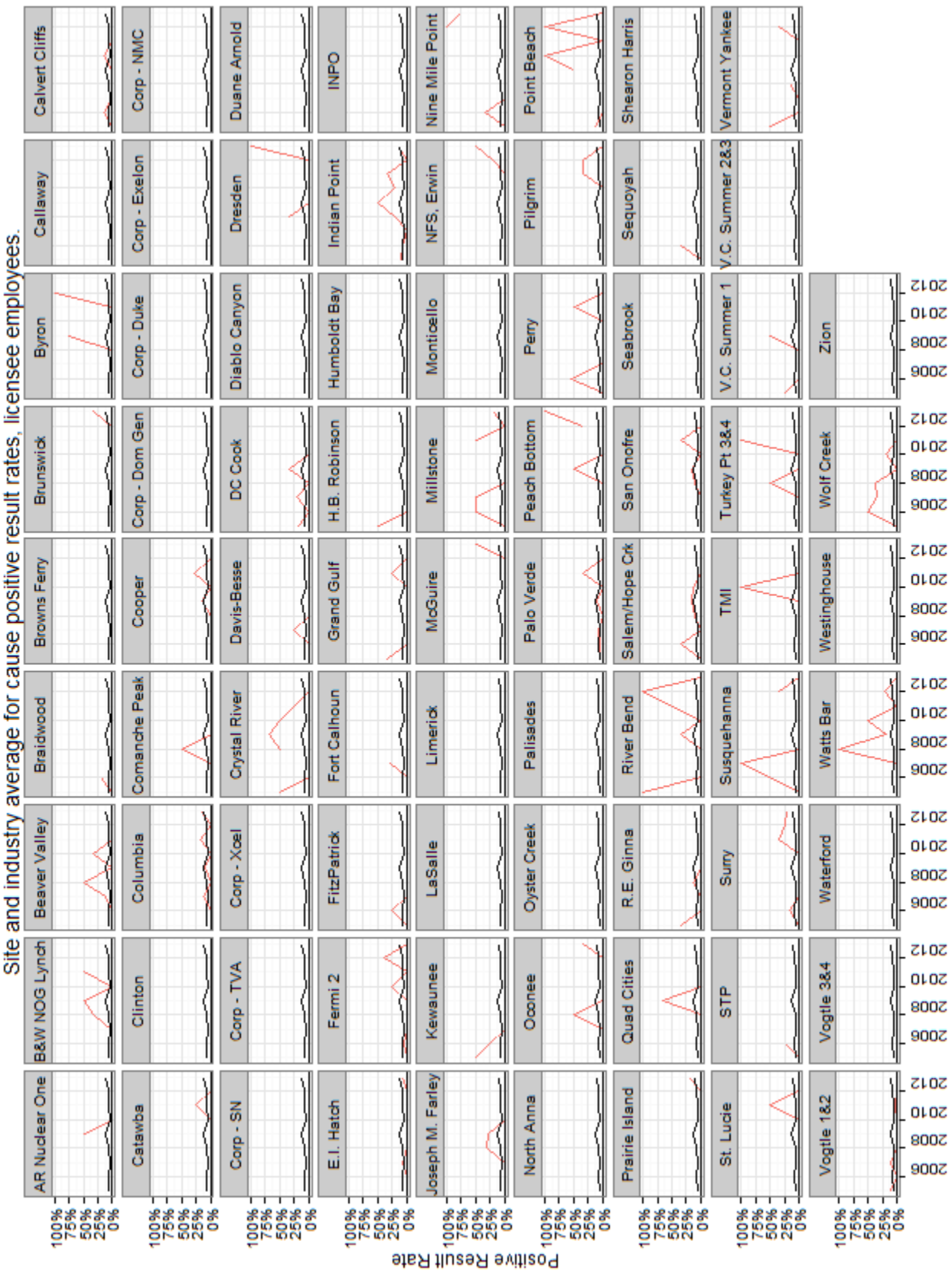

Within-year, difference between site and industry average random positive result rate, CVs, bootstrapped CIs.
Difference in Positive Result Rate
AR Nuclear One
B&W NOG Lynch
Beaver Valley
Braidwood
Browns Ferry
Brunswick
Byron
Callaway
Calvert Cliffs
Catawba
Clinton
Columbia
Comanche Peak
Cooper
Corp - Dom Gen
Corp - Duke
Corp - Exelon
Corp - NMC
Corp - SN
Corp - TVA
Corp - Xcel
Crystal River
Davis-Besse
DC Cook
Diablo Canyon
Dresden
Duane Arnold
E.I. Hatch
Fermi 2
FitzPatrick
Fort Calhoun
Grand Gulf
H.B. Robinson
Humboldt Bay
Indian Point
INPO
Joseph M. Farley
Kewaunee
LaSalle
Limerick
McGuire
Millstone
Monticello
NFS, Erwin
Nine Mile Point
North Anna
Oconee
Oyster Creek
Palisades
Palo Verde
Peach Bottom
Perry
Pilgrim
Point Beach
Prairie Island
Quad Cities
R.E. Ginna
River Bend
Salem/Hope Crk
San Onofre
Seabrook
Sequoyah
Shearon Harris
St. Lucie
STP
Surry
Susquehanna
TMI
Turkey Pt 3&4
V.C. Summer 1
V.C. Summer 2&3
Vermont Yankee
Vogtle 1&2
Vogtle 3&4
Waterford
Watts Bar
Westinghouse
Wolf Creek
Zion
20%
15%
10%
5%
0%
2006
2008
2010
2012
Significance of Difference
No Sig. Difference
Sig. Decrease
Sig. Increase

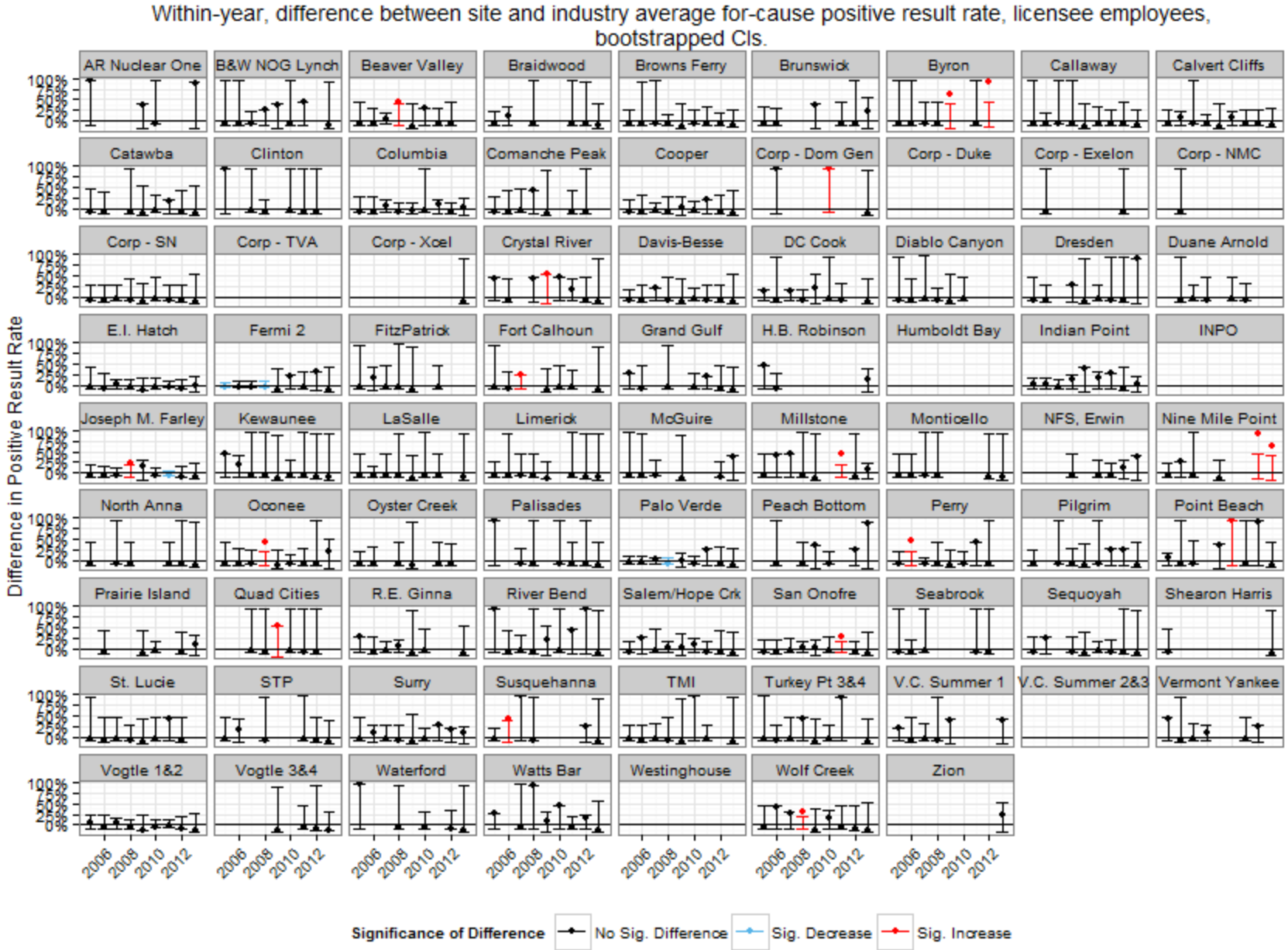

Site and industry average random positive result rates, licensee employees.
Positive Result Rate
AR Nuclear One
B&W NOG Lynch
Beaver Valley
Braidwood
Browns Ferry
Brunswick
Byron
Callaway
Calvert Cliffs
Catawba
Clinton
Columbia
Comanche Peak
Cooper
Corp - Dom Gen
Corp - Duke
Corp - Exelon
Corp - NMC
Corp - SN
Corp - TVA
Corp - Xoel
Crystal River
Davis-Besse
DC Cook
Diablo Canyon
Dresden
Duane Arnold
E.I. Hatch
Fermi 2
FitzPatrick
Fort Calhoun
Grand Gulf
H.B. Robinson
Humboldt Bay
Indian Point
INPO
Joseph M. Farley
Kewaunee
LaSalle
Limerick
McGuire
Millstone
Monticello
NFS, Erwin
Nine Mile Point
North Anna
Oconee
Oyster Creek
Palisades
Palo Verde
Peach Bottom
Perry
Pilgrim
Point Beach
Prairie Island
Quad Cities
R.E. Ginna
River Bend
Salem/Hope Crk
San Onofre
Seabrook
Sequoyah
Shearon Harris
St. Lucie
STP
Surry
Susquehanna
TMI
Turkey Pt 3&4
V.C. Summer 1
V.C. Summer 2&3
Vermont Yankee
Vogtle 1&2
Vogtle 3&4
Waterford
Watts Bar
Westinghouse
Wolf Creek
Zion
3%
2%
1%
0%
2006
2008
2010
2012


Within-year, difference between site and industry average random positive result rate, licensee employees, bootstrapped CIs.
Difference in Positive Result Rate
AR Nuclear One
B&W NOG Lynch
Beaver Valley
Braidwood
Browns Ferry
Brunswick
Byron
Callaway
Calvert Cliffs
Catawba
Clinton
Columbia
Comanche Peak
Cooper
Corp - Dom Gen
Corp - Duke
Corp - Exelon
Corp - NMC
Corp - SN
Corp - TVA
Corp - Xcel
Crystal River
Davis-Besse
DC Cook
Diablo Canyon
Dresden
Duane Arnold
E.I. Hatch
Fermi 2
FitzPatrick
Fort Calhoun
Grand Gulf
H.B. Robinson
Humboldt Bay
Indian Point
INPO
Joseph M. Farley
Kewaunee
LaSalle
Limerick
McGuire
Millstone
Monticello
NFS, Erwin
Nine Mile Point
North Anna
Oconee
Oyster Creek
Palisades
Palo Verde
Peach Bottom
Perry
Pilgrim
Point Beach
Prairie Island
Quad Cities
R.E. Ginna
River Bend
Salem/Hope Crk
San Onofre
Seabrook
Sequoyah
Shearon Harris
St. Lucie
STP
Surry
Susquehanna
TMI
Turkey Pt 3&4
V.C. Summer 1
V.C. Summer 2&3
Vermont Yankee
Vogtle 1&2
Vogtle 3&4
Waterford
Watts Bar
Westinghouse
Wolf Creek
Zion
4%
2%
0%
2006
2008
2010
2012
Significance of Difference
No Sig. Difference
Sig. Increase


Site and industry average for cause positive result rates, CVs.
Positive Result Rate
AR Nuclear One
B&W NOG Lynch
Beaver Valley
Braidwood
Browns Ferry
Brunswick
Byron
Callaway
Calvert Cliffs
Catawba
Clinton
Columbia
Comanche Peak
Cooper
Corp - Dom Gen
Corp - Duke
Corp - Exelon
Corp - NMC
Corp - SN
Corp - TVA
Corp - Xcel
Crystal River
Davis-Besse
DC Cook
Diablo Canyon
Dresden
Duane Arnold
E.I. Hatch
Fermi 2
FitzPatrick
Fort Calhoun
Grand Gulf
H.B. Robinson
Humboldt Bay
Indian Point
INPO
Joseph M. Farley
Kewaunee
LaSalle
Limerick
McGuire
Millstone
Monticello
NFS, Erwin
Nine Mile Point
North Anna
Oconee
Oyster Creek
Palisades
Palo Verde
Peach Bottom
Perry
Pilgrim
Point Beach
Prairie Island
Quad Cities
R.E. Ginna
River Bend
Salem/Hope Crk
San Onofre
Seabrook
Sequoyah
Shearon Harris
St. Lucie
STP
Surry
Susquehanna
TMI
Turkey Pt 3&4
V.C. Summer 1
V.C. Summer 2&3
Vermont Yankee
Vogtle 1&2
Vogtle 3&4
Waterford
Watts Bar
Westinghouse
Wolf Creek
Zion
100%
75%
50%
25%
0%
2006
2008
2010
2012


Within-year, difference between site and industry average for-cause positive result rate, CVs, bootstrapped CIs.
Difference in Positive Result Rate
AR Nuclear One
B&W NOG Lynch
Beaver Valley
Braidwood
Browns Ferry
Brunswick
Byron
Callaway
Calvert Cliffs
Catawba
Clinton
Columbia
Comanche Peak
Cooper
Corp - Dom Gen
Corp - Duke
Corp - Exelon
Corp - NMC
Corp - SN
Corp - TVA
Corp - Xcel
Crystal River
Davis-Besse
DC Cook
Diablo Canyon
Dresden
Duane Arnold
E.I. Hatch
Fermi 2
FitzPatrick
Fort Calhoun
Grand Gulf
H.B. Robinson
Humboldt Bay
Indian Point
INPO
Joseph M. Farley
Kewaunee
LaSalle
Limerick
McGuire
Millstone
Monticello
NFS, Erwin
Nine Mile Point
North Anna
Oconee
Oyster Creek
Palisades
Palo Verde
Peach Bottom
Perry
Pilgrim
Point Beach
Prairie Island
Quad Cities
R.E. Ginna
River Bend
Salem/Hope Crk
San Onofre
Seabrook
Sequoyah
Shearon Harris
St. Lucie
STP
Surry
Susquehanna
TMI
Turkey Pt 3&4
V.C. Summer 1
V.C. Summer 2&3
Vermont Yankee
Vogtle 1&2
Vogtle 3&4
Waterford
Watts Bar
Westinghouse
Wolf Creek
Zion
50%
0%
2006
2008
2010
2012
Significance of Difference
No Sig. Difference
Sig. Decrease
Sig. Increase


Site and industry average for cause positive result rates, licensee employees.
Positive Result Rate
AR Nuclear One
B&W NOG Lynch
Beaver Valley
Braidwood
Browns Ferry
Brunswick
Byron
Callaway
Calvert Cliffs
Catawba
Clinton
Columbia
Comanche Peak
Cooper
Corp - Dom Gen
Corp - Duke
Corp - Exelon
Corp - NMC
Corp - SN
Corp - TVA
Corp - Xcel
Crystal River
Davis-Besse
DC Cook
Diablo Canyon
Dresden
Duane Arnold
E.I. Hatch
Fermi 2
FitzPatrick
Fort Calhoun
Grand Gulf
H.B. Robinson
Humboldt Bay
Indian Point
INPO
Joseph M. Farley
Kewaunee
LaSalle
Limerick
McGuire
Millstone
Monticello
NFS, Erwin
Nine Mile Point
North Anna
Oconee
Oyster Creek
Palisades
Palo Verde
Peach Bottom
Perry
Pilgrim
Point Beach
Prairie Island
Quad Cities
R.E. Ginna
River Bend
Salem/Hope Crk
San Onofre
Seabrook
Sequoyah
Shearon Harris
St. Lucie
STP
Surry
Susquehanna
TMI
Turkey Pt 3&4
V.C. Summer 1
V.C. Summer 2&3
Vermont Yankee
Vogtle 1&2
Vogtle 3&4
Waterford
Watts Bar
Westinghouse
Wolf Creek
Zion
100%
75%
50%
25%
0%
2006
2008
2010
2012


Within-year, difference between site and industry average for-cause positive result rate, licensee employees, bootstrapped CIs.
Difference in Positive Result Rate
AR Nuclear One
B&W NOG Lynch
Beaver Valley
Braidwood
Browns Ferry
Brunswick
Byron
Callaway
Calvert Cliffs
Catawba
Clinton
Columbia
Comanche Peak
Cooper
Corp - Dom Gen
Corp - Duke
Corp - Exelon
Corp - NMC
Corp - SN
Corp - TVA
Corp - Xcel
Crystal River
Davis-Besse
DC Cook
Diablo Canyon
Dresden
Duane Arnold
E.I. Hatch
Fermi 2
FitzPatrick
Fort Calhoun
Grand Gulf
H.B. Robinson
Humboldt Bay
Indian Point
INPO
Joseph M. Farley
Kewaunee
LaSalle
Limerick
McGuire
Millstone
Monticello
NFS, Erwin
Nine Mile Point
North Anna
Oconee
Oyster Creek
Palisades
Palo Verde
Peach Bottom
Perry
Pilgrim
Point Beach
Prairie Island
Quad Cities
R.E. Ginna
River Bend
Salem/Hope Crk
San Onofre
Seabrook
Sequoyah
Shearon Harris
St. Lucie
STP
Surry
Susquehanna
TMI
Turkey Pt 3&4
V.C. Summer 1
V.C. Summer 2&3
Vermont Yankee
Vogtle 1&2
Vogtle 3&4
Waterford
Watts Bar
Westinghouse
Wolf Creek
Zion
100%
75%
50%
25%
0%
2006
2008
2010
2012
Significance of Difference
No Sig. Difference
Sig. Decrease
Sig. Increase

APPENDIX C

ADDITIONAL CONSIDERATIONS ON FFD PROGRAM PERFORMANCE

APPENDIX C

ADDITIONAL CONSIDERATIONS ON FFD PROGRAM PERFORMANCE

The FFD program's drug testing success relies on effective quality assurance. Although 10 CFR Subpart F has provisions for licensee testing facilities, all licensees have elected to conduct urine testing under 10 CFR Subpart G, which requires use of laboratories certified under the U.S. Department of Health and Human Services (HHS) Mandatory Guidelines for Federal Workplace Drug Testing Programs (also known as the National Laboratory Certification Program). Substance Abuse and Mental Health Services Administration's (SAMHSA's) Division of Workplace Programs oversees HHS-certified laboratories that perform forensic drug testing for federal agencies and federally regulated industries. Testing of urine specimens must meet the requirements of the NRC's FFD policy and HHS' certified laboratory program.[1,2]

Under §26.717 and §26.719, licensees must collect, compile, analyze, and report aggregated annual statistical drug and alcohol program performance data. This data is typically reported on the Single Positive Test Form (Form 890) and Annual Reporting Form for Drug and Alcohol Tests (Form 891). Data analyzed in this report was reported by licensees on Form 891. These forms focus on findings from testing program implementation and specimen test results. The regulations in Subpart G include detailed, specific requirements on the following:

26.155 Laboratory personnel
26.157 Procedures
26.159 Assuring specimen security, chain of custody, and preservation
26.161 Cutoff levels for validity testing
26.163 Cutoff levels for drugs and drug metabolites
26.165 Testing split specimens and retesting single specimens
26.167 Quality assurance and quality control
26.168 Blind performance testing
26.169 Reporting results.

These requirements include a number of measurable elements that may represent additional measurable sources of error for FFD program performance, and therefore potential input to performance metrics analyses. If available, data on these could provide NRC with additional insight into licensees and other entities' FFD drug and alcohol deterrence and testing performance and help inform the inspection process. The remainder of this discussion focuses on quality measures in use for clinical laboratories to input to defining measures that could inform FFD drug testing program performance metrics.

In the field of clinical laboratory science, the sequence of steps beginning with a clinical encounter between a physician and a patient and concluding with the impact of a test result on

[1] SAMHSA provides current information on drug testing at: https://www.samhsa.gov/workplace/drug-testing#Laboratory%20Resources.
[2] See, for example: https://www.nrc.gov/docs/ML1532/ML15324A366.pdf.

patient care is known as the total testing process (Schumacher and Barr 1998). This is analogous to the testing process in the FFD program, with the key distinctions being that the decision to order a test is made not at the discretion of a physician but through administrative requirements (as specified in the rule), and that the specimen donor is not a patient but an individual subject to FFD drug and alcohol testing requirements. The total testing process is generally divided into three phases: pre-analytical, analytical, and post-analytical. ISO 15189:2012 *Medical Laboratories — Requirements for Quality and Competence* defines the pre-analytical phase as:

> processes that start, in chronological order, from the clinician's request and include the examination request, preparation and identification of the patient, collection of the primary sample(s), and transportation to and within the laboratory, and end when the analytical examination begins.

and the post-analytical phase as:

> processes following the examination including review of results, retention and storage of clinical material, sample (and waste) disposal, and formatting, releasing, reporting and retention of examination results.

In both definitions, "examination" is understood to mean "assay or test."[1]

There are many parallel elements in the pre- and post-analytical phases in both the clinical and workplace drug testing contexts. Overlapping pre-analytical elements include collection, donor identification and specimen labeling, control and custody, transportation, storage, and administrative processing of specimens for analytical examination; post-analytical elements include analytical result interpretation, reporting of results, storage and disposal of specimens, and retention of records for future review. Each of these elements comprises multiple actions performed by multiple actors, in some cases spanning multiple physical locations. The challenges of consistent coordination and communication across this complex set of steps of the total testing process substantially increase the opportunities for errors that may compromise the performance of an FFD program.

A review of studies on the incidence of errors in the total testing process finds that "the proportion of errors associated with the two extra-analytical phases is 4 to 5 times that seen in the analytical phase, with the pre-analytical phase consistently representing over half of all errors in published studies" (Hawkins 2012, p. 6). Because the NRC does not require licensees and entities to report aggregate statistical data on elements of the pre- and post-analytical phases, this research suggests that it may not receive information on more than half of the errors occurring within the total testing process of licensees and other entities' FFD programs.

A study by Plebani (2010) provides more detail on the types of errors committed in the total testing process and the relative frequency of errors observed (see Table C-1). Note that Plebani (2010) subdivides the pre- and post-analytical phases to include a pre-pre- and a post-post-analytical phase.

[1] International Organization for Standardization, "ISO 15189:2012."

Table C-1. Types and Relative Frequency of Errors in the Different Phases of the Total Testing Process

Phase of Total Testing Process	Type of Error	Relative Frequency (%)
Pre-pre-analytical	Inappropriate test request Order entry Patient/specimen misidentification Sample collection Inappropriate container Handling, storage, and transportation	46–68
Pre-analytical	Sorting and routing Pour-off Aliquoting, pipetting, and labeling Centrifugation (time and/or speed)	3–5
Analytical	Equipment malfunction Sample mix-ups Interference (endogenous or exogenous) Undetected failure in quality control	7–13
Post-analytical	Erroneous validation of analytical data Failure in reporting/addressing the report Excessive turn-around time Improper data entry/manual transcription error Failure/delay in reporting critical values	13–20
Post-post-analytical	Delayed/missed reaction to laboratory reporting Incorrect interpretation	25–46

Note. Adapted from Plebani (2010)

In the field of clinical laboratory science, there is an ongoing effort to define, standardize, and harmonize performance measures spanning the total testing process. In that field, such measures are known as "quality measures" or "quality indicators." The use of the term "quality" is intended to emphasize that the ultimate goal of clinical laboratory testing is to ensure the highest quality of patient care through the correct performance of each step of the total testing process (Plebani 2012). Quality indicators are generally defined as rates, indicating the number of laboratory errors committed per total number of actions taken for a given period of time.[1] Although the objectives of the FFD program differ from those in clinical laboratory science, there is a common interest in ensuring that all steps are performed correctly and are free of errors.

Appendix C References

10 CFR Part 26. Code of Federal Regulations. Title 10, Energy, Part 26, "Fitness for Duty Programs." 73 FR 17117. March 31, 2008. "Fitness for Duty Programs" Federal Register, Nuclear Regulatory Commission, Washington D.C.

[1] ISO/TS 22367:2008 *Medical laboratories — Reduction of error through risk management and continual improvement* defines laboratory error as: a "failure of a planned action to be completed as intended, or use of a wrong plan to achieve an aim, occurring at any part of the laboratory cycle, from ordering examinations to reporting results and appropriately interpreting and reacting to them."

Hawkins, R. 2012. Managing the Pre- and Post-analytical Phases of the Total Testing Process. *Annals of Laboratory Medicine* 32(1): 5–16.

International Organization for Standardization. 2012. ISO 15189:2012 Medical Laboratories — Requirements for Quality and Competence. https://www.iso.org/obp/ui/#iso:std:iso:15189:ed-3:v2:en. Accessed October 29, 2015.

International Organization for Standardization. 2008. SO/TS 22367:2008 Medical Laboratories — Reduction of Error through Risk Management and Continual Improvement. http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=40918. Accessed October 29, 2015.

Plebani, M. 2010. The Detection and Prevention of Errors in Laboratory Medicine. *Annals of Clinical Biochemistry* 47(Pt 2):101-10.

Plebani, M. 2012. Quality Indicators to Detect Pre-Analytical Errors in Laboratory Testing. *Clinical Biochemist Reviews* 33(3):85-8.

Schumacher, G.E., and Barr, J.T. 1998. Total Testing Process Applied to Therapeutic Drug Monitoring: Impact on Patients' Outcomes and Economics. *Clinical Chemistry* 44(2): 370-374.

Pacific Northwest
NATIONAL LABORATORY

Proudly Operated by **Battelle** *Since 1965*

U.S. DEPARTMENT OF
ENERGY

902 Battelle Boulevard
P.O. Box 999
Richland, WA 99352
1-888-375-PNNL (7665)

www.pnnl.gov