ORNL/TM-DRAFT

Computer Science and Mathematics Division

Computational Sciences Section

PRELIMINARY DRAFT
RESULTS OF THE ANALYSIS OF THE BLOOD LYMPHOCYTE PROLIFERATION TEST DATA FROM THE OAK RIDGE Y-12 STUDY

E. L. Frome
D. L. Cragle *
L. G Littlefield *
S. P. Colyer *

* Mathematical Sciences Section
Oak Ridge National Laboratory
Oak Ridge, Tennessee 37831-0117
* Environmental and Health Sciences Division
Oak Ridge Institute for Science
and Education Oak Ridge, Tennessee

Date Published: DRAFT October 1999

Research was supported by the Offices of Occupational Medicine, Environment, Safety and Health, U. S. Department of Energy

Prepared by the
Oak Ridge National Laboratory
Oak Ridge, Tennessee 37831
managed by
Lockheed Martin Energy Research Corp.
for the
U.S. DEPARTMENT OF ENERGY
under Contract No. DE-AC05-96OR22464

Contents

1 Introduction 1
2 Identification of LPTs With Large SIs 1
  2.1 Method 1- Using Distribution of Second Largest SI From Nonexposed Controls and/or Historical Population of Beryllium Exposed Workers 2
  2.2 Method 2- Using The Empirical Distribution of log(SI)s For Each Day and Each Beryllium Concentration 2
3 Graphical Summaries of The Stimulation Index (SI) 3
  3.1 Results for Positive Control LPT Data 3
  3.2 Results for Beryllium Workers(BW) and Nonexposed(NE) Control Data 3
  3.3 Graphical Summaries for SIs For BW and NE LPTs 3
4 Results for Normal and Abnormal LPTs 11
  4.1 Identification of Abnormal LPTs Using Second Largest SI 11
  4.2 Method 2- Using The Empirical Distribution of log(SI) 17
5 Acknowledgments 17
6 References 17

List of Figures

1 Histograms and Probability Plots PHA: Log and Linear Scale 4
2 Histograms and Probability Plots ConA: Log and Linear Scale 5
3 Histograms for SIs for Beryllium Workers and Nonexposed Controls 6
4 Histograms for Log(SI)s for Beryllium Workers and Nonexposed Controls . . . 7
5 Boxplots of Log(SI)s for Beryllium Workers and Nonexposed Controls 8
6 Normal Probability Plots of Log(SI)s For BW and NE 10
7 Histograms an Normal Q-Q Plot Plot For Second Largest Log(SI) 12
8 Histograms of SIs For Abnormal LPTs 13
9 Histograms of SIs For Normal LPTs 14
10 Boxplots of Log(SI)s For Normal and Abnormal LPTs 15
11 Normal Q-Q Plots of Log(SI)s For Normal and Abnormal 16

1. Introduction

THIS IF A PRELIMINARY DRAFT— NOT FOR CITATION OR DISTRIBUTION

This report contains results of the blood beryllium lymphocyte proliferation test (LPT) [1] that were done as part of a study of beryllium workers(BW) at the Y-12 plant in Oak Ridge, Tennessee. The purpose of the Y-12 study was to describe the risk of becoming sensitized to beryllium in this group of workers. Results presented here are limited to LPTs (N=1117) that were done in the same lot of ABi Serum (3040083). If an individual had more that one LPT, only the results for the *FIRST LPT* were include in these analyses . SIs were calculated as described by Frome *et al* [3]. This serum was used in over 95 percent of the LPTs in the Y-12 study. Of the 1117 LPTs 1084 were from the worker population and 33 were nonexposed (control) individuals. The results in this draft report were based on summary statistics computed by the ORISE laboratory. The purposes of this report are i) to present graphical summaries of the SIs in original units and in logarithmic units using histograms, Normal Quantile-Quantile (Q-Q) plots (also refered to as Normal probability plots), and boxplots, and ii) to compare two previously described methods for the identification of an abnormal test. **Throughout this report the expression log(x) represents the** *natural logarithmic function*, i.e., **if y = log x than x = exp(y).**

The results in ths report were presented at The ORISE Advisory Committee Meeting for Beryllium Workers Surveillance Program, November 4, 1999. This document and related material can be viewed at http://www.csm.ornl.gov/~frome/vgLPT/TTLPT/

2. Identification of LPTs With Large SIs

This section describes two approaches that have been proposed identify an “abnormal” LPT. Each of these methods uses the LAV estimates $\tilde{\beta}_j$, j= 1,...,6, of the log(SI)s. Note that $\tilde{\beta}_j$ is a statistical shorthand for the LAV estimate of the jth log(SI).

[1] Abbreviations used: AB,abnormal; Be,beryllium; BW,beryllium worker; CBD,chronic beryllium disease; ConA,concanavalin-A; CV,coefficient of variation; df,degrees of freedom; LAV,least absolute values; LPT,lymphocyte proliferation test; MAD, median absolute deviation; NE,nonexposed; ORISE,Oak Ridge Institute for Science and Education; ORNL,Oak Ridge National Laboratory; PHA,phytohemagglutinin; Q-Q,Quantile-Quantile; SI,stimulation index;

2.1. Method 1- Using Distribution of Second Largest SI From Nonexposed Controls and/or Historical Population of Beryllium Exposed Workers

The second largest log(SI) in a *reference data base* of LPT data sets is used to determine the "cut point". An LPT is considered abnormal if **the second largest** log(SI) exceeds this cut point—i.e.at least two SIs exceed exp(cut). This approach was described by Frome, Newman and Mroz [2] and applied to a test data from the LPT laboratory at National Jewish Center. The steps for this procedure are as follows:

1. Find $\beta_i^\dagger = secondlargest[\tilde{\beta}_{ij}, j = 1, \ldots, 6]$ for $i = 1, \ldots, N$, where N= number of LPTs in *reference data set.*
2. Find $M = median[\beta_i^\dagger, i = 1, \ldots, N]$ and S the median absolute deviation (MAD) estimate of the standard deviation of the $\beta_i^\dagger$.
3. Then calculate *cut* = M + z_pS, where z_p is the pth quantile of the standard normal distribution. If $p = 0.975$ then $z_p = 1.96$.
4. An LPT is defined to be *abnormal if at least two log(SI)s exceed cut.*

The probability of a false positive for this procedure should be about 1-p.

2.2. Method 2- Using The Empirical Distribution of log(SI)s For Each Day and Each Beryllium Concentration

The second approach is the one proposed Frome *et al* [3] (see page 965 or Section 3.6 of ORNL-6818). It is based on the assumption that the log(SIs) are approximately normally distributed (see Figures 5 and 6). **In this report the reference data base consists of all available LPTs in the BE-sample data set and the NE data set.** *In practice this data set would change during the course of a study as new data becomes available.* The first step is to convert each log(SI) into a standardized deviate

$$u_{ij} = \frac{\tilde{\beta}_{ij} - \tilde{\mu}_j}{\tilde{s}_j}$$

using the values of $\tilde{\mu}_j$ and $\tilde{s}_j$. These standardized deviates can be compared with the quantiles of the standard normal distribution, i.e. $\Pr[u < z_p] = p$. If we assume that the log(SIs) are independent then the binomial distribution can be used to calculate an approximate probability of at least k out of six "large" SIs for a given value of z_p. The probability of at least one large SI is $1 - p^6$, and the probability of at least two is $1 - [p^6 + 6(1-p)p^5]$.

In fact, the log(SIs) are positively correlated, so this probability should be a lower bound on the chance of finding a false positive LPT.

3. Graphical Summaries of The Stimulation Index (SI)

3.1. Results for Positive Control LPT Data

Each of the LPTs included two positive controls—PHA and ConA. Graphical summaries of the positive control SIs for 1117 LPTs are shown in Figure 1 and Figure 2 These results strongly support the use of the lognormal distribution to describe the variability of SIs.

3.2. Results for Beryllium Workers(BW) and Nonexposed(NE) Control Data

3.3. Graphical Summaries for SIs For BW and NE LPTs

Histograms for the SIs for each day and Be concentration (BW and NE combined) for the LPT data from the Y-12 study are shown in Figure 3. For plotting purposes SIs greater than four have been set equal to four. These large SIs indicate a response to beryllium. Figure 4 shows the histograms for the Log(SI)s—all logs are base e. The series of twelve boxplots in Figure 5 compares the distributions of the log(SI)s for the beryllium workers(BW) and nonexposed (NE) controls (see Fig. 4 for the ORISE-AC Data in Frome *et al.* [3]). The vertical scale on the right hand side of the plots is in SI units. The SIs for both the NE controls and the workers exposed to beryllium(BW) decrease as the beryllium concentration in the test wells increases. This may be due to a toxic effect of high beryllium concentration that results in ”cell killing”. There is considerably more variability in the log(SI)s for the BW LPTs than for the NE LPTs at each concentration on both day 5 and day 7.

Figure 6 shows lognormal probability plots for the BW and NE SIs for each of the three beryllium concentrations on day 5 and day 7. In each of the six plots the data —ordered values of the log(SI)s—are shown on the vertical scale on the left, and the quantiles of the standard normal distribution are shown on the horizontal scale. A detailed account of the construction and interpretation of normal probability plots is provided by Chambers *et .al* [1].

If the relation between the empirical quantiles (on the vertical axis) and theoretical quantiles (on the horizontal axis) is linear, this indicates that the distribution is Gaussian. Each plot includes the median (labeled M) and a resistant estimate of the standard deviation (labeled S) for the log(SI)s for the BW (shown as circles) and the NE log(SI)s (shown as triangles.) The lines in each plot (solid for NE and dotted for BW) show the relation that is expected

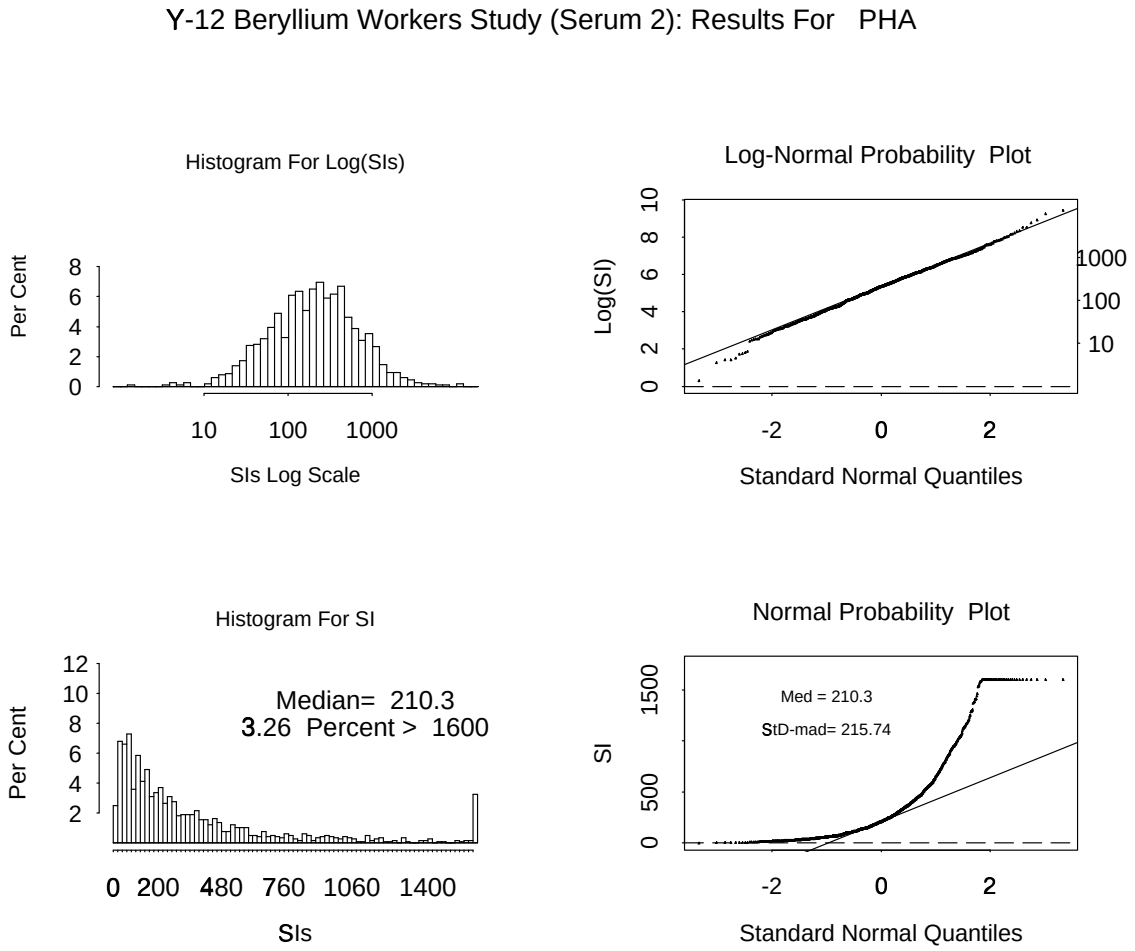

Figure 1: The panels on the left show the histograms of the SIs.The top left is for log(SI)s and the bottom left is for the SIs. The panels on the right are Gaussian probability plots. If the data in the histogram (on the left) is ”normally distributed” the the probability plot (on the right) should look line a straight line. These plots clearly show that log(SI)s follow the normal distribution,i.e. the SIs follow the lognormal distribution.

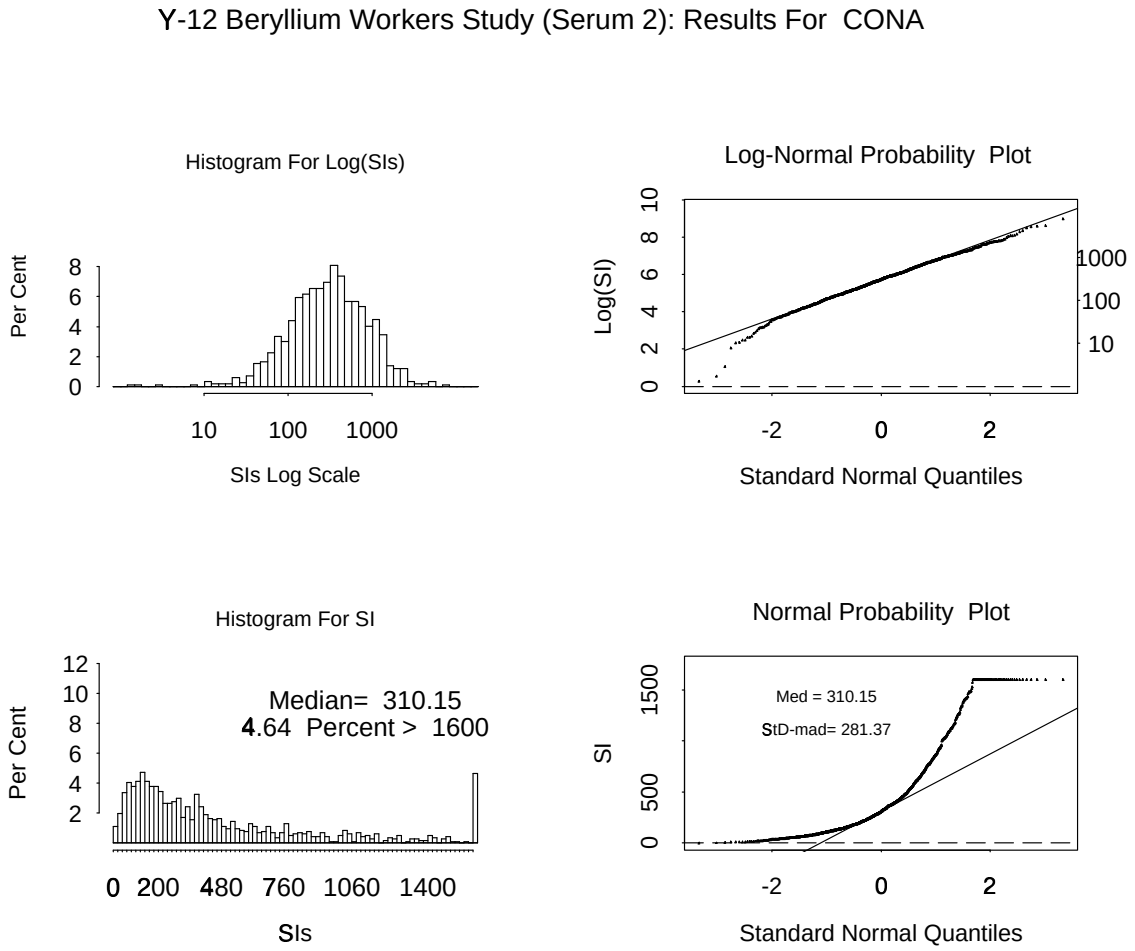

Figure 2: The panels on the left show the histograms of the SIs.The top left is for log(SI)s and the bottom left is for the SIs. The panels on the right are Gaussian probability plots. If the data in the histogram (on the left) is ”normally distributed” the the probability plot (on the right) should look line a straight line. These plots clearly show that log(SI)s follow the normal distribution,i.e. the SIs follow the lognormal distribution.

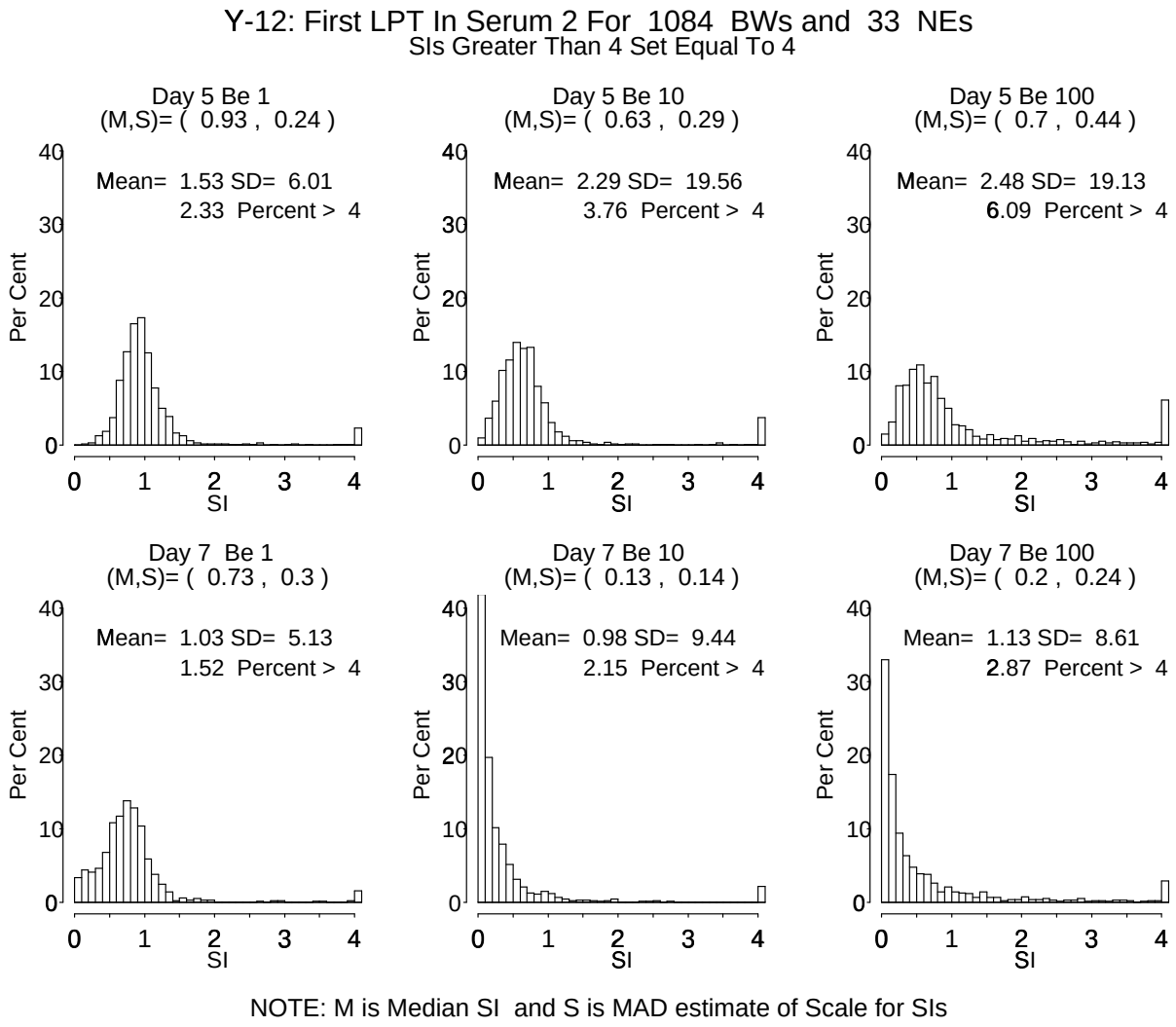

Figure 3: Histograms of the SIs for the BW and NE LPTs. Numbers in parenthesis are the median and MAD estimate of location . The mean and standard deviation (SD) for each distribution are also given.

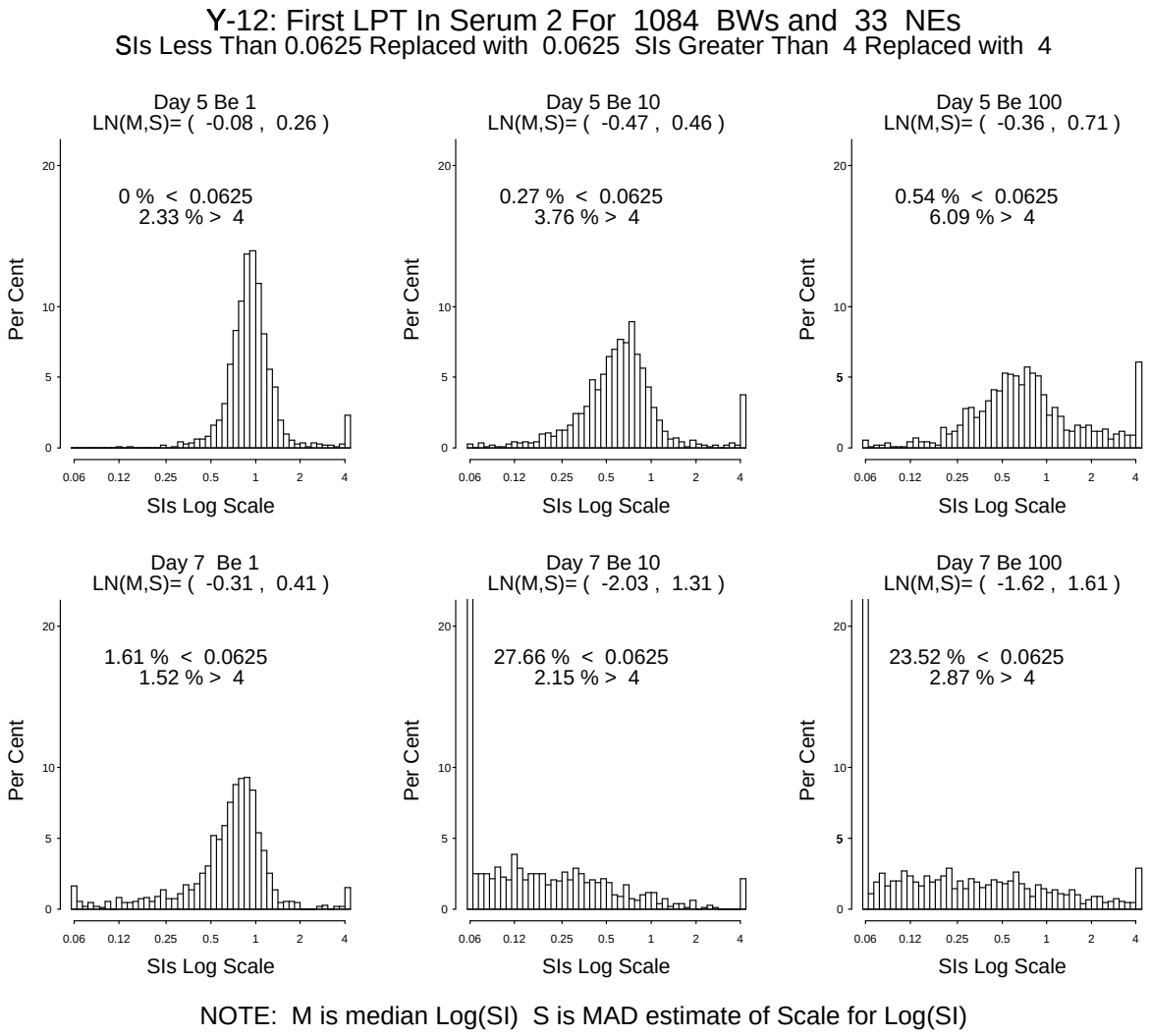

Figure 4: Histograms of the Log(SI)s for the BWs and NEs LPTs. The the location (M median) and scale (S-MAD) parameters of the Log(SI)s for each distribution are given in parenthesis.

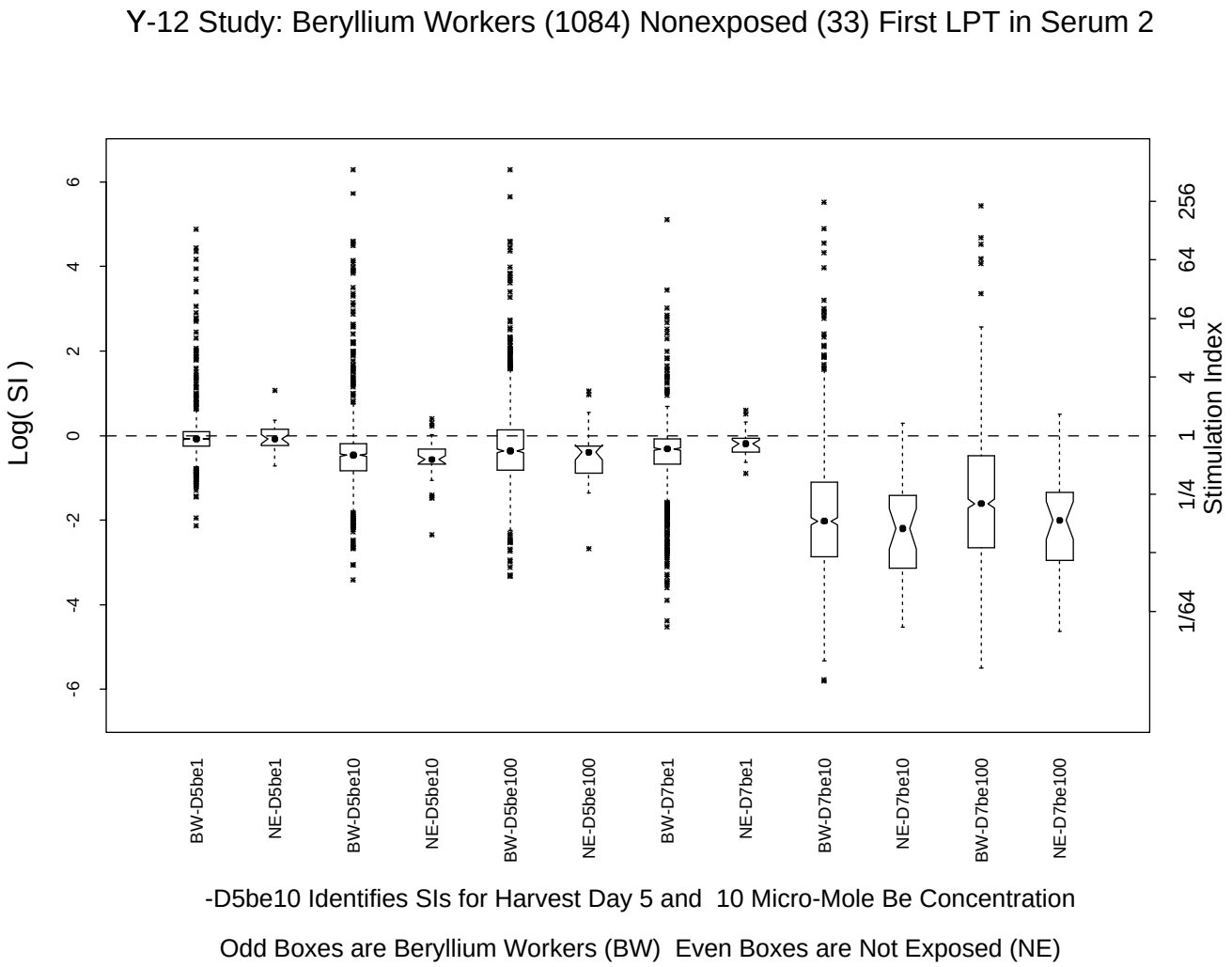

Figure 5: Boxplots of Log SIs for Sample of Beryllium Exposed Workers and Nonexposed Controls. There are Three Beryllium Concentrations on Harvest Days 5 and 7.

if the log(SI) values are from a normal distribution with location parameter M (which determines the intercept) and standard deviation S (which determines the slope) . Resistant methods were used to estimate the location and scale parameters . This reflects the assumption that most of beryllium exposed workers do not show an abnormal response, i.e. they look like the nonexposed group. For example, consider the plot for day 5 Be-1 in Figure 6. The log(SI)s appear to be approximately normal in the center, for both the NE controls and the beryllium workers. There are several values that are larger than expected (these are the points above the lines). These ”outliers” are SIs that indicate hypersensitivity to beryllium. There are also several points below the line which indicate cell killing. Compare these results to similar plots for earlier ORISE data—see [3] Fig. 5, and similar LPT test results from The National Jewish Center [2] Figures 5 and 6.

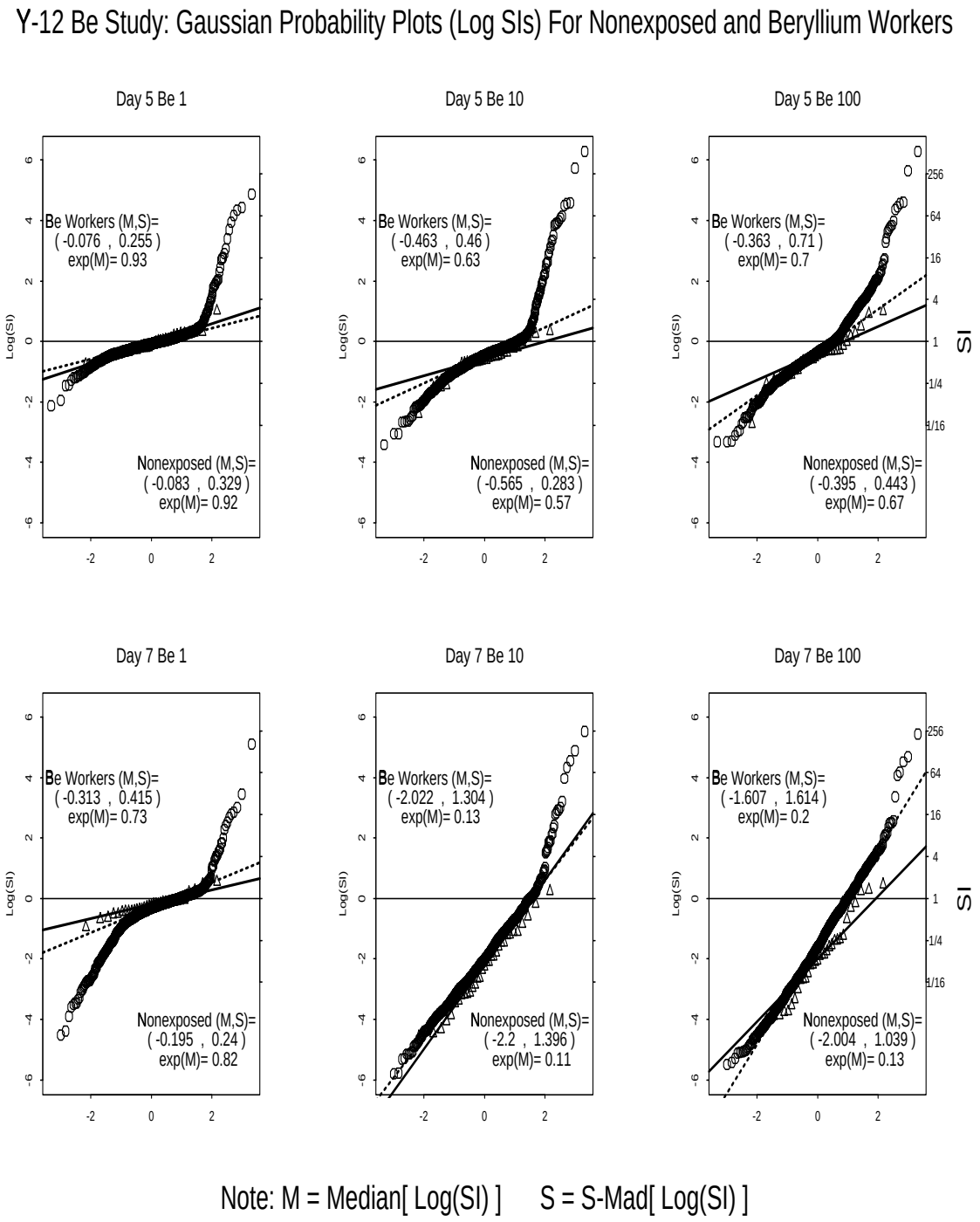

Figure 6: Normal Q-Q Plots of Log(SI)s for Each Beryllium Concentration on Day 5 and Day 7 for BW and NE controls. The data values are shown on the vertical axis. The Median (M), MAD scale estimate(S) of the log(SI)s and exp(M) are Listed on Each Plot. Values of M and S for BWs are in upper left and NE controls are in lower right of each panel.

4. Results for Normal and Abnormal LPTs

4.1. Identification of Abnormal LPTs Using Second Largest SI

The method described in Section 2.1 (with p= 0.001) was used to identify ”Abnormal” LPT. **All LPTs from the BW and NE controls were used as the reference data base**. There were 33 LPTs for nonexposed individuals, and 1084 beryllium exposed workers. The histogram of the 2nd largest log(SI)s is shown in the left hand panel of Figure 7. The lognormal probability plot is shown in the right hand panel.

The cut point for an abnormal LPT is cut = M + $z_{.999}$*S=-0.155+3.09*0.325= 0.849. There were 84 abnormal LPTs among the 1117 LPTs in the **reference data set**. Figure 8 shows the SI histograms for the abnormal LPTs and Figure 9 shows the SI histograms for the normal LPTs.

The series of 12 boxplots in Figure 10 compares the distributions of the log(SI)s for the normal and abnormal LPTs.

Figure 11 shows lognormal probability plots for the three beryllium concentrations on day 5 and day 7

Y-12 Beryllium Workers Study (Serum 2): Results For 2nd Largest SI

NOTE:SIs Less Than 0.25 Replaced with 0.25 SIs Greater Than 4 Replaced with 4

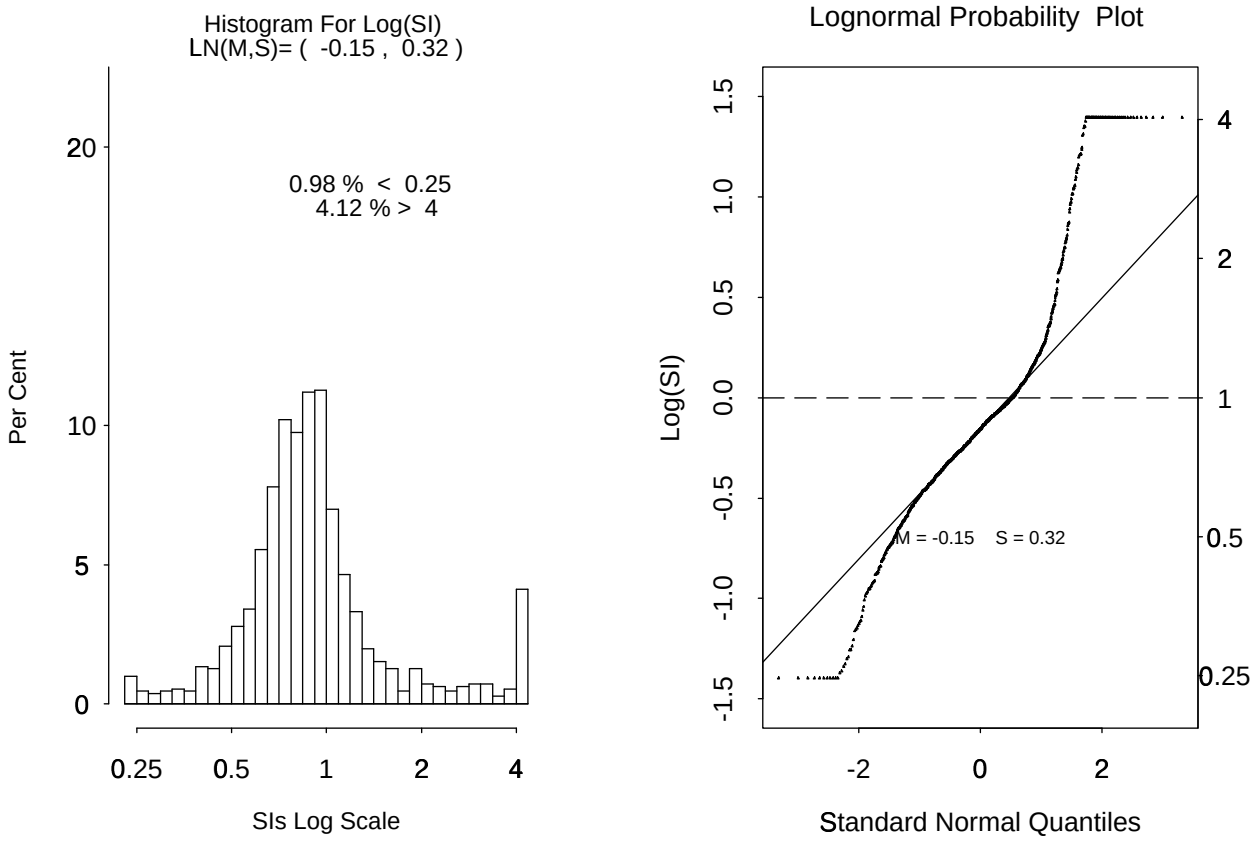

Figure 7: Normal Q-Q Plots for Second Largest Log(SI) for BW and NE combined. The Median (M), MAD scale estimate(S) of the Log(SI)s are shown. The cut point for an abnormal LPT (with p= 0.001) using the 2nd largest log(SI) is cut = M + 3.09*S = 0.8487—note exp(0.8487)= 2.34. There were 84 abnormals in the BW LPTs.

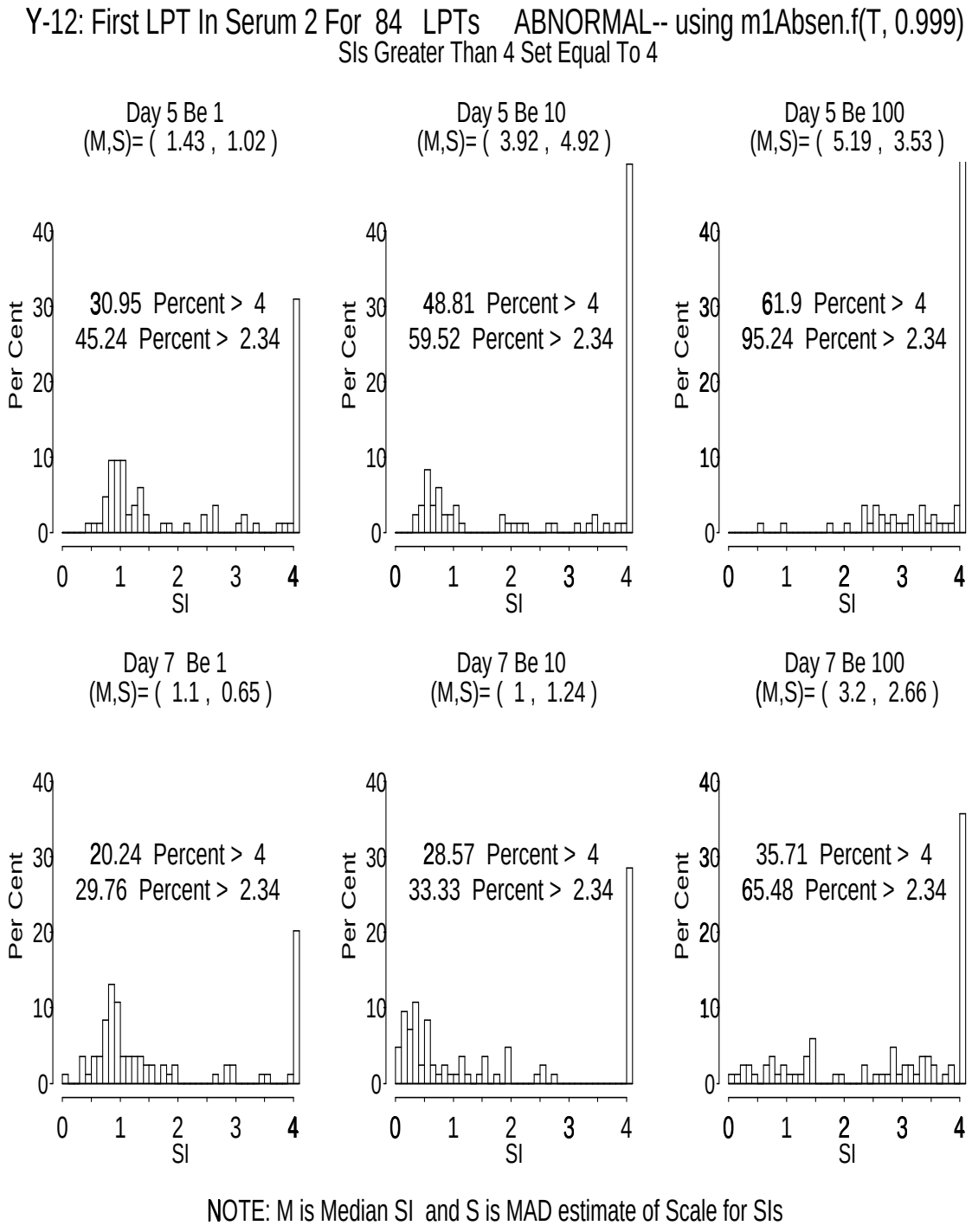

Figure 8: Histograms of SIs for Each Beryllium Concentration on Day 5 and Day 7 for Abnormal LPTs. The Median (M), MAD scale estimate(S) of the SI are Listed on Each Plot.

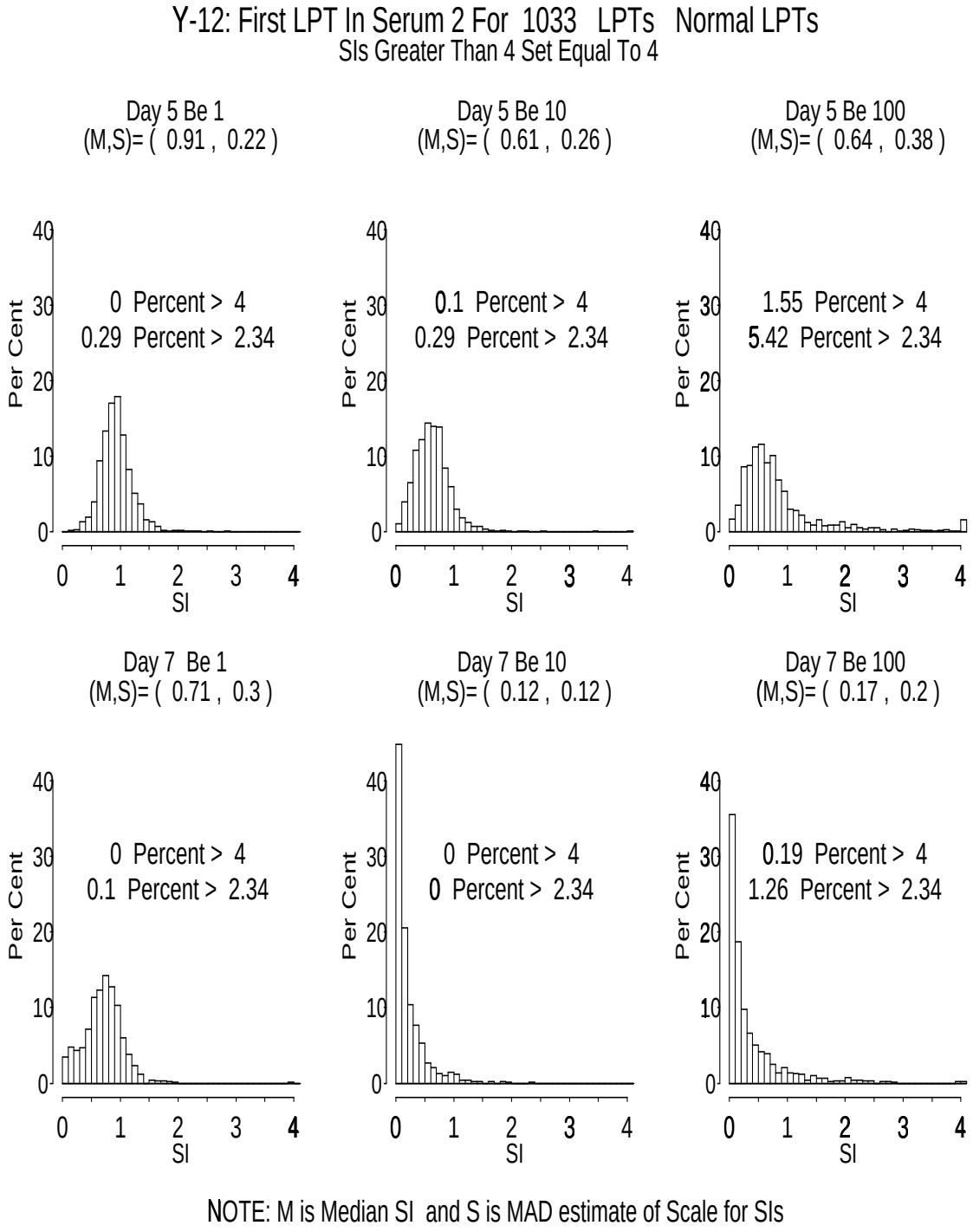

Figure 9: Gaussian Probability Plots of SIs for Each Beryllium Concentration on Day 5 and Day 7 for Normal LPTs. The Median (M), MAD scale estimate(S) of the SIs are listed on each plot.

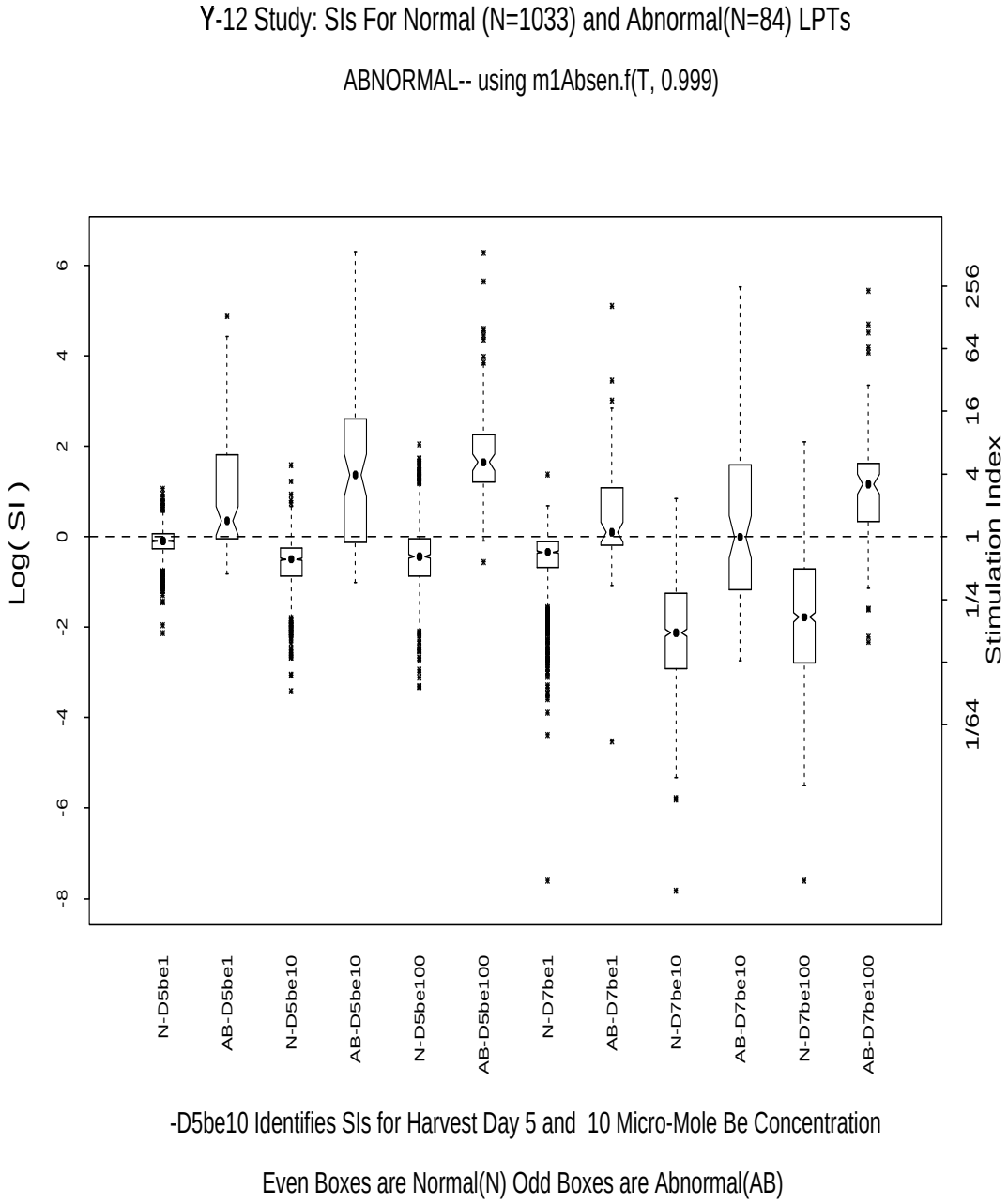

Figure 10: Boxplots of Log(SI)s for Each Beryllium Concentration on Day 5 and Day 7 for Normal and Abnormal LPTs. Abnormal LPTs were identified using method 1A described in Section 2.1 (with p= 0.001).

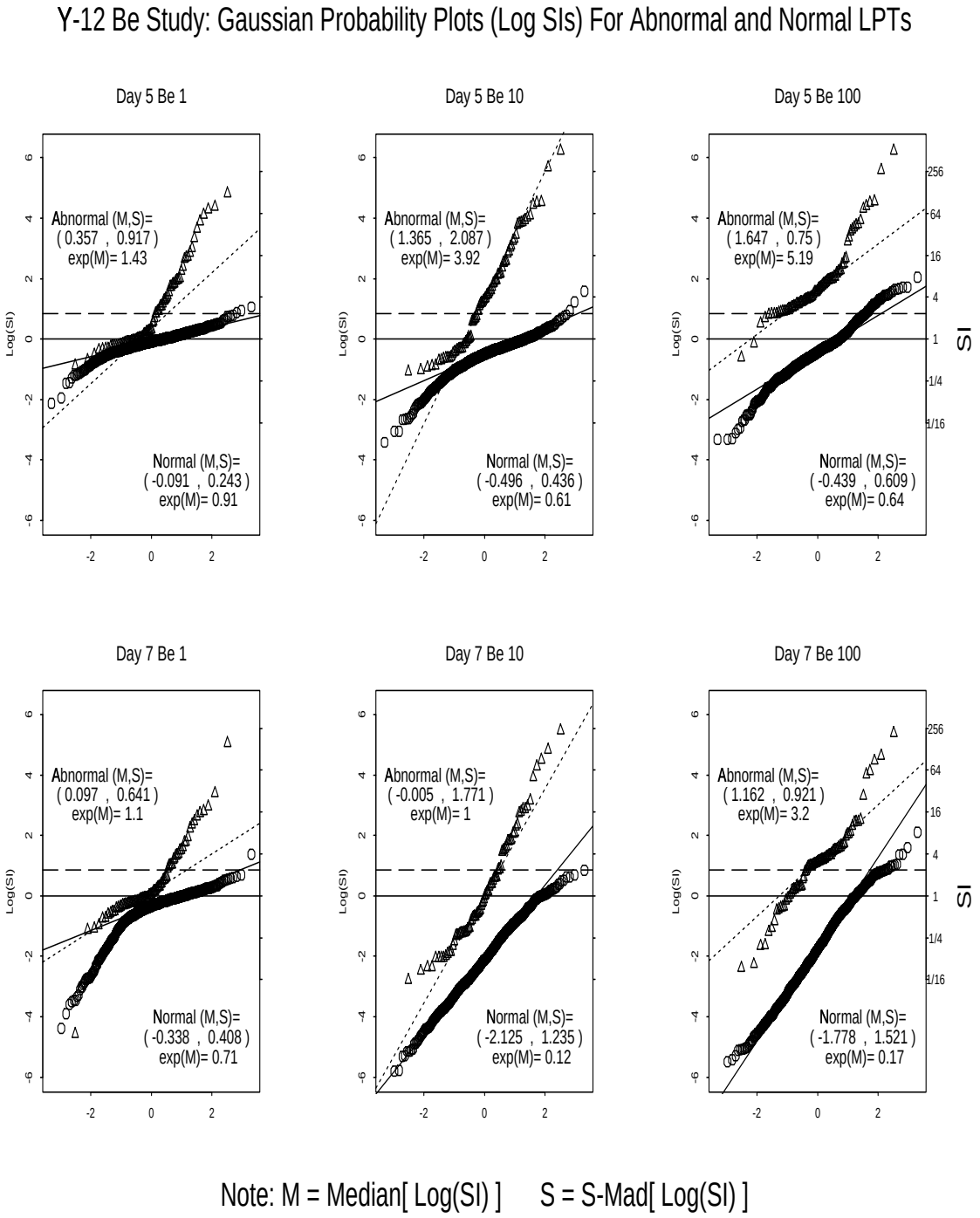

Figure 11: Normal Q-Q Plots of Log(SI)s for Each Beryllium Concentration on Day 5 and Day 7 for Normal and Abnormal LPTs. Values of M and S for abnormal LPTs are in upper left and for Normal LPTs are in lower right of each panel. The cut point for an abnormal LPT using the 2nd largest log(SI) is cut = M + 3.09*S = 0.8487—note exp(0.8487)= 2.34. There were 84 abnormals in the BW LPTs.

4.2. Method 2- Using The Empirical Distribution of log(SI)

This results using Method 2 (see Section 2.2) are shown in the Table 1. The values of $\tilde{\mu}_j$ (Med) and $\tilde{s}_j$ (Smad) are listed at the top of the Table. The cut points for each log(SI) were defined so that the probability of at least two large SIs is about 0.001. The corresponding SI cut points for each day and concentration are shown at the top of the Table. All of the LPTs in this table have at least two large SIs on Day 5.

5. Acknowledgments

This research was supported by the Offices of Occupational Medicine, Environment, Safety and Health, U. S. Department of Energy under contract DE-AC05-96OR22464 with Lockheed Martin Energy Research Corp. The authors thank ...

The work has been authored by a contractor of the U.S. Government. Accordingly, the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this work, or to allow others to do so for U. S. Government purposes.

6. References

[1] J.M. Chambers, W.S. Cleveland, B. Kleiner, and P.A. Tukey. *Graphical Methods for Data Analysis.* Duxbury Press, Boston, 1983.

[2] E.L. Frome, L.S. Newman, and M.M. Morz. Results of the analysis of the blood lymphocyte proliferation test data from the national jewish center,1997. Technical Report ORNL/TM-13338, Oak Ridge National Laboratory, 1997.

[3] E.L. Frome, M.H. Smith, L.G. Littlefield, R.L. Neubert, and S.P. Colyer. Statistical methods for the blood beryllium lymphocyte proliferation test. *Environmental Health Perspectives Supplement,* 104:957–968, 1996.

Table 1 SIs for Abnormal LPTs From Method 2

Med	-0.08	-0.47	-0.36	-0.31	-2.03	-1.62
Smad	0.26	0.46	0.71	0.41	1.31	1.61
Cut	1.72	1.87	3.78	1.95	3.01	9.37

	Day 5 SIs			Day 7 SIs		
	Be1	Be10	Be100	Be1	Be10	Be100
ybw0078	6.46	306.36	282.67	2.98	52.80	91.94
ybw0106	2.69	0.43	4.66	2.61	0.20	1.16
ybw0139	1.24	3.47	0.57	3.45	0.13	0.11
ybw0157	3.14	13.20	30.03	1.10	2.52	4.96
ybw0182	6.33	46.67	41.98	3.95	24.66	28.71
ybw0191	9.99	53.85	78.33	9.84	249.17	230.37
ybw0192	1.39	4.06	9.88	0.87	1.21	8.02
ybw0206	18.18	48.63	86.12	7.35	6.43	5.12
ybw0215	1.43	6.78	8.21	0.94	0.31	0.36
ybw0240	7.17	49.34	53.35	1.73	18.59	12.09
ybw0241	1.80	5.25	3.37	1.60	0.53	0.90
ybw0243	1.90	4.85	2.00	0.55	0.38	0.65
ybw0296	3.15	4.95	15.37	4.80	8.48	11.92
ybw0302	1.09	3.15	5.55	0.89	0.94	0.70
ybw0313	3.40	13.06	4.94	0.72	0.75	0.98
ybw0323	1.35	3.86	4.13	0.84	1.91	0.20
ybw0353	1.06	2.24	7.61	1.39	0.59	13.10
ybw0364	1.06	4.62	4.28	0.96	4.69	4.36
ybw0387	1.09	3.48	9.17	0.78	0.30	4.02
ybw0464	3.99	2.66	1.76	1.57	0.53	0.32
ybw0468	7.45	4.49	6.25	1.32	0.13	1.26
ybw0472	2.67	13.77	8.29	4.06	1.94	3.02
ybw0492	7.83	28.82	2.42	1.99	8.23	1.31
ybw0568	4.15	1.97	6.78	0.82	0.52	0.96
ybw0593	1.20	3.97	5.26	0.78	1.11	0.74
ybw0597	5.98	23.06	12.79	2.93	10.32	4.87
ybw0616	21.42	89.14	99.48	16.18	133.13	108.28
ybw0628	84.40	98.40	95.61	6.20	11.11	5.56
ybw0629	130.81	534.28	534.16	31.44	75.91	65.50
ybw0670	5.97	27.08	3.61	3.59	5.38	1.09
ybw0683	11.50	59.46	37.16	4.73	20.17	7.43
ybw0722	6.57	9.05	2.09	6.26	1.74	1.44
ybw0731	40.25	62.76	45.00	17.30	17.03	8.73
ybw0811	2.11	13.89	39.70	0.99	5.01	6.06
ybw0875	2.43	3.62	5.12	4.26	0.10	3.20
ybw0879	2.45	7.38	7.73	1.95	1.05	1.40
ybw0961	7.49	7.34	2.60	1.38	0.37	0.65
ybw0968	29.75	32.99	12.29	20.32	6.51	3.10
ybw0972	4.94	9.23	3.94	2.88	6.39	1.32
ybw0975	3.72	2.70	6.00	1.84	0.60	0.20
ybw0979	51.78	50.28	46.73	14.54	18.68	5.30
ybw0982	77.37	96.48	26.50	164.27	95.22	58.51
ybw0994	1.86	5.88	3.58	2.82	4.80	3.46
ybw1007	14.82	11.13	7.68	1.79	1.42	0.68
ybw1057	16.01	19.05	5.75	12.37	4.51	1.44
ybw1116	0.96	6.58	4.96	0.78	16.10	12.19
ybw1139	3.07	8.93	7.38	0.35	1.20	1.42
ybw1140	15.78	21.98	9.21	5.22	1.91	2.31
ybw1220	64.42	17.74	7.50	11.17	2.53	0.72
ybw1256	0.63	4.60	6.93	0.98	6.71	10.57
ybw1291	0.86	7.22	7.23	0.34	1.53	1.95