ORNL-6818

Engineering Physics and Mathematics Division
Mathematical Sciences Section

STATISTICAL METHODS FOR THE ANALYSIS OF A SCREENING TEST FOR CHRONIC BERYLLIUM DISEASE

Edward L. Frome †
Matthew H. Smith ‡
L. Gayle Littlefield ‡
Richard L. Neubert †
Shirley P. Colyer ‡

† Mathematical Sciences Section
Engineering Physics and Mathematics Division
Oak Ridge National Laboratory
P. O. Box 2008
Oak Ridge, Tennessee 37831-6367
‡ Medical Sciences Division
Oak Ridge Institute for Science and Education
Oak Ridge, Tennessee

Date Published: October 1994

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

Prepared by the
Oak Ridge National Laboratory
Oak Ridge, Tennessee 37831
operated by
Lockheed Martin Energy Research Corporation
for the
U.S. DEPARTMENT OF ENERGY
under Contract No. DE-AC-05-840R21400

Contents

List of Tables

List of Figures

STATISTICAL METHODS FOR THE ANALYSIS OF A SCREENING TEST FOR CHRONIC BERYLLIUM DISEASE

Edward L. Frome

Matthew H. Smith

L. Gayle Littlefield

Richard L. Neubert

Shirley P. Colyer

Abstract

The lymphocyte proliferation test (LPT) is a noninvasive screening procedure used to identify persons who may have chronic beryllium disease. A practical problem in the analysis of LPT well counts is the occurrence of outlying data values (approximately 7% of the time). A log-linear regression model is used to describe the expected well counts for each set of test conditions. The variance of the well counts is proportional to the square of the expected counts, and two resistant regression methods are used to estimate the parameters of interest. The first approach uses least absolute values (LAV) on the log of the well counts to estimate beryllium stimulation indices (SIs) and the coefficient of variation. The second approach uses a resistant regression version of maximum quasi-likelihood estimation. **A major advantage of the resistant regression methods is that it is not necessary to identify and delete outliers.** These two new methods for the statistical analysis of the LPT data and the outlier rejection method that is currently being used are applied to 173 LPT assays. **We strongly recommend the LAV method for routine analysis of the LPT.**

Outliers are also important when trying to identify individuals with beryllium hypersensitivity, since these individuals have large positive values for their SIs. A new method for identifying large SIs using combined data from the not exposed group and the beryllium workers is proposed. The log(SIs) are described with a Gaussian distribution with location and scale parameters estimated using resistant methods. This approach is applied to the test data and results are compared with those obtained from the current method.

1. Introduction

Chronic beryllium disease (CBD), a disorder mainly affecting the lung, occurs in a small percentage of persons exposed to beryllium dusts. In their discussion of CBD, Stokes and Rossman [24] note that most investigators require evidence of beryllium hypersensitivity as one of several criteria for diagnosis of the disease. They also point out that in vitro proliferation of bronchoalveolar lavage cells to beryllium is extremely sensitive to and specific for the diagnosis of CBD but is not suitable for screening since it is an invasive procedure. A non-invasive procedure based on the proliferative response of blood cells to beryllium has been developed and is referred to as the beryllium specific lymphocyte proliferation test (LPT)[10]. Kreiss et al [12] state that the increasing use of beryllium in several new economic sectors emphasizes the need for medical surveillance in the workplace for CBD. In particular, beryllium has been used in the nuclear industry for a number of years. Kreiss et al [13] have examined the epidemiology of CBD in a stratified sample of workers at a nuclear weapons plant, and discuss the role of the LPT in beryllium disease surveillance in the nuclear industry. The U.S. Department of Energy (DOE) is currently conducting a study of all beryllium exposed workers (approximately 15,000 workers at 20 DOE sites). Each participating beryllium worker will have an LPT at an approved laboratory using a standard protocol developed by the Committee to Accredit Beryllium Sensitization Testing (CABST). The results of each assay are then evaluated and classified as normal, abnormal, or unsatisfactory (see Appendix B).

In this report, statistical methods that can be used to help in the evaluation of each LPT assay are described. The problems considered are the occurrence of "outliers" in the raw data (well counts), procedures for evaluating the quality of the assay results, and identification of "abnormal" LPT results (i.e. beryllium sensitive workers). When the LPT results are analyzed, the concept of outliers is important in two different ways. First, the results of a single LPT assay (usually 56 well counts) can contain multiple outlying counts. Given their undue influence on the parameters of interest(the logSIs), a method for handling outliers is needed.

The "current approach" (as described in the July 1993 version of the CABST protocol, see Appendix B) is to identify and remove the outliers before calculating parameter estimates. **As an alternative to deleting outlying points, we propose using estimation methods that are not sensitive to outliers.** To explain these approaches, we describe the LPT assay using a regression model that relates the expected well counts at each of the three beryllium concentrations to the control well counts. We then show how *resistant* fitting methods are used to estimate the stimulation index (SI) for each concentration of beryllium. A resistant fit is one that is not sensitive to large changes in a few observations. **The main advantage of this approach is that parameter estimates are calculated without having to explicitly identify and delete the outlying well counts.**

Outliers are also important when attempting to identify beryllium exposed workers who exhibit beryllium hypersensitivity. It is anticipated that most (over 90%) of the beryllium workers will have test results similar to those in the group having no known exposure to beryllium. However, even after the use of resistant methods to minimize the effect of outlying well counts, the LPT for some beryllium workers will yield large SIs. **In this case, we want to identify the "outliers" (large SIs) as they represent beryllium workers who exhibit beryllium hypersensitivity.**

2. Statistical Methods

A detailed description of lymphocyte culture methods, quality control measures, and examples of plate maps and print-outs of raw data are included in Appendix F. Following is a brief description of the protocol for the LPT culture assay as established by CABST and implemented by the cytogenetics laboratory at Oak Ridge Institute for Science and Education (ORISE).

1. A 30 ml blood sample is obtained from each patient and mononucleated cells are separated using density gradient centrifugation.

2. Lymphocytes are cultured using standard methods at a final concentration of 2.5×10^5 cells per well in 96 well flat bottom microtiter plates. For each LPT assay 12 replicate control wells, and four replicates for each experimental condition (i.e., 1, 10, and 100 μM of $BeSO_4$, and mitogen stimulated positive controls) are set up.

3. Cells are incubated at $37°C$ for five and seven days and a pulse of tritiated thymidine is delivered prior to harvest. Cells are harvested on filter paper and counts are measured in a Packard Matrix 96 gas ionization counter. Each filter is counted for thirty minutes and the results organized as shown in Table 1 for statistical analysis.

Table 1: Well Counts for LPT Assay (AC153 data shown)

Culture Conditions		j	Replicate Counts			
Day 5	Control	1	965	1173	828	862
	Control	1	1474	7237	1021	976
	Control	1	1500	1729	1672	1992
	Be 1	2	1050	706	1434	687
	Be 10	3	1551	1466	1661	2301
	Be 100	4	3571	5780	4011	5229
Day 7	Control	5	9202	5253	3786	5212
	Control	5	2310	2844	1915	3102
	Control	5	2458	3936	3087	6588
	Be 1	6	714	1135	6084	1097
	Be 10	7	786	846	2757	652
	Be 100	8	6037	8349	6852	10449
Day 5	Pha	9	82425	52954	52669	50487
	Candida	10	35501	21623	21551	22087

2.1. Regression Model for the LPT Data

Let y_{jk} denote the well count for the k^{th} replicate of the j^{th} set of culture conditions. The expected count in each well can be represented by a log-linear regression function:

$$E(y_{jk}) = \lambda_j = exp(\boldsymbol{X_j \beta}), \tag{1}$$

where $j = 1, \ldots, 10$ and $k = 1, \ldots, 12$ for the controls and $k = 1, 2, 3, 4$ for the beryllium stimulated cells and the positive controls. In (1), $\boldsymbol{X_j}$ is a row vector of indicator variables and $\boldsymbol{\beta}$ is the vector of regression parameters (see below). We further assume that the variance of the well counts is proportional to the square of the expected count:

$$Var(y_{jk}) = (\phi\lambda_j)^2. \tag{2}$$

Equations 1 and 2 together are referred to as a generalized linear model with constant coefficient of variation (see [16, chapter 8] for details). The distinct values of the row vectors of covariates $\mathbf{X}_j$, $j = 1, \ldots, 10$ that we use are shown in Table 2.

Table 2: Distinct Rows in the Model Matrix

j	x_{j1}	x_{j2}	x_{j3}	x_{j4}	x_{j5}	x_{j6}	x_{j7}	x_{j8}	x_{j9}	x_{j10}
1	0	0	0	0	0	0	1	0	0	0
2	1	0	0	0	0	0	1	0	0	0
3	0	1	0	0	0	0	1	0	0	0
4	0	0	1	0	0	0	1	0	0	0
5	0	0	0	0	0	0	0	1	0	0
6	0	0	0	1	0	0	0	1	0	0
7	0	0	0	0	1	0	0	1	0	0
8	0	0	0	0	0	1	0	1	0	0
9	0	0	0	0	0	0	1	0	1	0
10	0	0	0	0	0	0	1	0	0	1

With this parameterization, the first three βs represent the log of the SIs for the three concentrations of $BeSO_4$ on harvest day 5 and the next three βs are the corresponding estimates on day 7. The last two βs are the log(SIs) for the positive control wells, and β_7 and β_8 represent the log of the control well counts on day 5 and 7 respectively. We have developed two outlier resistant approaches for estimating the SIs and the coefficient of variation, ϕ.

2.2. Least Absolute Value Regression on Log(y)

The first approach is to take the log of the counts since this is the variance-stabilizing transformation and leads to a linear model in say $z_{jk} = \log(y_{jk})$, i.e.

$$E(z_{jk}) = \boldsymbol{X_j\beta} - \phi^2/2 \text{ and } Var(z_{jk}) \simeq \phi^2.$$

If outliers were not present, applying ordinary least squares to the transformed data would yield consistent estimates for the log(SI) parameters [16]. The effect of outliers is minimized by using least absolute values (or some other robust method) on the z_{jk}. Least absolute value (LAV) regression – also known as L_1 norm, least absolute deviations (LAD) and minimum sum of absolute errors (MSAE) – is well known to be resistant to outliers and is an important particular case of a general class of robust methods known as M-estimators [21, 18]. In general, LAV regression requires special computational resources to calculate parameter estimates [1]. In this situation, however, it is only necessary to find the median of the log of the well counts for each set of design conditions (say $\tilde{z}_j$) and then subtract the control median for each harvest day from the beryllium stimulated medians (see Appendices A and C for details). A resistant estimate of the coefficient of variation can then be obtained as

$$\tilde{\phi}_L = C \times median\{|z_{jk} - \tilde{z}_j|\},$$

where $C = 1.48 \times \sqrt{n/(n-p)}$, $n = 56$, and $p = 10$ (when the assay is complete). The value of C is chosen to make the estimate consistent for the standard deviation for a Gaussian error model and for consistency with the usual least squares results in which the estimated variance is multiplied by the correction factor $n/(n-p)$ – see [14] and S-PLUS function `mad` in [23]. Alternative approaches to estimating ϕ have been discussed in the context of LAV regression (see e.g., [22, 18]) and there is no consensus as to the best approach. In addition to the fact that this parameter is of direct interest in this situation, it is also needed to obtain an estimate of the parameter covariance matrix

$$w^2(\boldsymbol{X}'\boldsymbol{X})^{-1}$$

where $w^2 = [2f(0)]^{-2}$ is the asymptotic variance of the sample median [2]. Following the approach of [17] we assume that the underlying error distribution is Gaussian in the center and use $\tilde{w} = \sqrt{\pi/2}\tilde{\phi}_L$ to obtain an estimate of the standard deviation of the log of the stimulation indices. The appropriate diagonal term from $(\boldsymbol{X}'\boldsymbol{X})^{-1}$ is 4/12, and consequently the estimated standard deviation of log(SI) is $1.25\tilde{\phi}_L(0.58) = 0.72\tilde{\phi}_L$. The results of applying this approach to the data in Table 1 are shown in Table 3.

Table 3: Results of LAV Estimation for log(y) of data in Table 1, $\tilde{\phi}_L = 0.367$

Experimental Conditions		z_{jk}				$\tilde{z}_j$	$(z_{jk} - \tilde{z}_j)/\tilde{\phi}_L$				$\tilde{\beta}$	$\exp(\tilde{\beta})$
Day 5	Controls	6.872	7.067	6.719	6.759	7.182	-0.8	-0.3	-1.3	-1.2		
	Controls	7.296	8.887	6.929	6.883	7.182	0.3	4.6	-0.7	-0.8		
	Controls	7.313	7.455	7.422	7.597	7.182	0.4	0.7	0.7	1.1		
	Be1	6.957	6.560	7.268	6.532	6.758	0.5	-0.5	1.4	-0.6	-0.423	0.655
	Be10	7.347	7.290	7.415	7.741	7.381	-0.1	-0.2	0.1	1.0	0.199	1.221
	Be100	8.181	8.662	8.297	8.562	8.429	-0.7	0.6	-0.4	0.4	1.248	3.483
Day 7	Controls	9.127	8.567	8.239	8.559	8.139	2.7	1.2	0.3	1.1		
	Controls	7.745	7.953	7.557	8.040	8.139	-1.1	-0.5	-1.6	-0.3		
	Controls	7.807	8.278	8.035	8.793	8.139	-0.9	0.4	-0.3	1.8		
	Be1	6.571	7.034	8.713	7.000	7.017	-1.2	0.0	4.6	0.0	-1.122	0.326
	Be10	6.667	6.741	7.922	6.480	6.704	-0.1	0.1	3.3	-0.6	-1.436	0.238
	Be100	8.706	9.030	8.832	9.254	8.931	-0.6	0.3	-0.3	0.9	0.792	2.207
Day 5	Pha	11.320	10.877	10.872	10.829	10.874	1.2	0.0	0.0	-0.1	4.792	120.50
	Candida	10.477	9.982	9.978	10.003	9.992	1.3	0.0	0.0	0.0	3.910	49.880

2.3. Quasi-Likelihood Estimation

In the second approach that we consider, the analysis is done on the original scale and estimation is based on the iterative weighted least squares (IWLS) algorithm. The use of IWLS for generalized linear [19] and nonlinear regression functions [7] leads to maximum likelihood estimates when the dependent variable is in the regular exponential family. McCullagh [15] extended this result to quasi-likelihood estimation which requires specification of the mean and variance function. Extension of the IWLS method to resistant/robust regression has been described by

[9] and [20], and the computational approach described in [6] (see chapter 6) is used here. Similar resistant regression methods have been applied to the analysis of drug concentration-time data encountered in human bioavailability studies [8].

To describe this approach, consider the following weighted sum of squares,

$$\sum_j \sum_k w_{jk}[y_{jk} - \lambda_j]^2, \tag{3}$$

where $\lambda_j = \exp(\boldsymbol{X_j \beta})$ and $w_{jk} \propto 1/var(y_{jk}) = 1/\lambda_j^2$. In the IWLS procedure, we start with an initial estimate, say $\boldsymbol{\beta}^\circ$, of the unknown parameters (see below), replace λ_j in (3) with the first order Taylor series

$$\exp(\boldsymbol{X_j \beta}^\circ) + \boldsymbol{P_j \delta}^\circ,$$

where $\boldsymbol{P_j} = \boldsymbol{X_j}\lambda_j^\circ$, and evaluate the weights at $\boldsymbol{\beta}^\circ$ to obtain

$$\sum w_{jk}^\circ[y_{jk} - (\lambda_j^\circ + \boldsymbol{P_j \delta}^\circ)]^2.$$

The unknown "correction vector", $\boldsymbol{\delta}^\circ$, is then calculated using weighted least squares, i.e. by solving

$$(\boldsymbol{P}^{\circ\prime}\boldsymbol{W}^\circ\boldsymbol{P}^\circ)\boldsymbol{\delta}^\circ = \boldsymbol{P}^{\circ\prime}\boldsymbol{W}^\circ[\mathbf{Y} - \boldsymbol{\lambda}^\circ] \tag{4}$$

for $\boldsymbol{\delta}^\circ$. The estimate of $\boldsymbol{\beta}$ is then updated $\boldsymbol{\beta}^{\mathbf{1}} = \boldsymbol{\beta}^\circ + \boldsymbol{\delta}^\circ$, and the procedure is repeated until convergence (see Appendix D for more details). Following the final iteration, compute the moment estimate of ϕ^2 namely

$$\hat{\phi}^2 = \frac{1}{n-p}\sum\sum\left(\frac{y_{jk} - \hat{y}_{jk}}{\hat{y}_{jk}}\right)^2.$$

The IWLS algorithm described above needs to be modified to adjust for the effect of outliers. This is done by introducing a second weight for each observation

$$w = \begin{cases} 1 & |u| \le k \\ k/|u| & |u| > k \end{cases} \tag{5}$$

where $u = (y_{jk} - \hat{y}_{jk})/\tilde{\phi}\hat{y}_{jk}$ is the standardized residual using the current estimates of $\boldsymbol{\beta}$ and ϕ. This is known as an M-estimator with Huber's loss function. The "tuning constant", k, must be specified and we use $k = 1.345$ which leads to estimates with approximately 95% efficiency [20]. Therefore, we obtain resistant quasi-likelihood estimates by adjusting the weights in the diagonal matrix $\mathbf{W}$ in Equation (4) by multiplying in the Huber weight in (5) (see Appendix D for details). Following the last iteration, the coefficient of variation is estimated using a scaled MAD estimate of the standardized residuals $u_{jk} = (y_{jk} - \hat{y}_{jk})/\hat{y}_{jk}$

$$\tilde{\phi} = 1.48 \times median\{|u_{jk}|\} \times \sqrt{n/(n-p)}.$$

3. Results

The regression model and the estimation methods discussed in Section 2.1 were obtained through analytic reasoning and limited experience with a few data sets. To evaluate the utility of our proposed methods, we have applied them to all available LPT assay results obtained at the ORISE cytogenetics laboratory as of July 1993. The method in use at ORISE (see Appendix B) at that time was also applied to each LPT assay, and is referred to in this report as the current method.

3.1. Description of the Data

Of the 173 test results used in the analysis, 133 are from a group of 120 workers exposed to beryllium and the remaining 40 are from persons having no known exposure to beryllium. The discrepancy between the number of test results and the number of beryllium exposed workers is accounted for by the fact that a second LPT was carried out on 13 workers.

Ideally, there should be 56 observations (well counts) for each assay, but in some cases, well counts are missing due to lack of sufficient cells to set up a complete test series or technical errors (see Appendix F.1). When an assay is incomplete, parameters are estimated (if possible) based on the reduced data set.

3.2. Comparison of Moment and Resistant Estimates of the Coefficient of Variation for Control Wells

An important assumption that we are making is that the standard deviation of the well counts is proportional to the mean as implied by Equation (2). Since each of the 173 assays contains 12 replicate control wells on both day 5 and day 7 we can evaluate this assumption by computing location and scale estimates for each assay on day 5 and day 7.

Figure 1 (top) shows the relationship between the moment estimator of location ($\bar{y}$, the sample mean) and the moment estimator of scale (s, the sample standard deviation) for the day 5 control wells. This plot also shows the resulting line when the standard deviation is regressed on the mean. The least squares equation for this line is $\hat{\sigma} = 0.448\bar{y}$, and the slope (0.448) is an estimate of the coefficient of variation for day 5. Figure 1 (bottom) is a similar plot but resistant estimates are used in place of moment estimates. Specifically, the sample median ($\tilde{y}$) replaces the sample mean, the MAD estimate ($\tilde{\sigma}$) replaces s, and LAV is used to regress $\tilde{\sigma}$ on $\tilde{y}$. The solution to this resistant fit is $\tilde{\sigma} = 0.34\tilde{y}$ and the slope (0.34) is a resistant estimate of the coefficient of variation. Given the presence of outliers in the LPT data, we prefer the resistant coefficient of variation.

Figure 2 shows the relationship between the resistant estimates of location and scale for the day 5 control wells (top) and the day 7 control wells (bottom) on a log-log scale. Note that if the standard deviation is proportional to the mean (i.e. constant coefficient of variation), the log-log plot should be linear with a slope of 1. The LAV fit is $\tilde{\sigma} = 0.361\tilde{y}$ for the day 7 control wells and the solid line in Figure 2 (bottom) is $\log(\tilde{\sigma}) = \log(0.361) + \log(\tilde{y})$. Comparing this fit to the resistant fit for the day 5 control wells ($\tilde{\sigma} = 0.34\tilde{y}$) reveals that the results on both days are quite similar. The main difference in the day 5 and day 7 results is that the day 7 results are shifted to the right since the control well counts are generally higher on day 7 than those on day 5. The median of the $\tilde{y}$s on day 5 is 1247 compared with 1840 for day 7. **These results are consistent with the laboratory observation that day 7 results are generally higher and**

show greater variability than well counts on day 5. They also support the regression model assumption discussed in Section 2.1 that the variance of the well counts is proportional to the square of the expected counts.

3.3. Summary of Results for Three Methods

The two methods of analysis described in Section 2 and the method currently being used (see Appendix B) were applied to the data described in Section 3.1. The results are summarized in six graphical displays (two for each method) in Figures 3 - 8. We will describe the two plots for the LAV method in detail.

The first graphical display (see Figure 3) is a series of 12 boxplots (see [26, 17, 23]) placed side by side for the log(SIs) – the vertical axis on the right shows the untransformed SIs. The ends of the box correspond to the 25^{th} and 75^{th} percentile so that 50 percent of the log(SIs) are contained in the box for each group. The vertical dotted lines are drawn to the nearest value not beyond a standard span – 1.5×(Inter-Quartile Range) – from the quartiles. The outlying values are shown individually for each group of data. There are two boxplots for each beryllium concentration on day 5 and day 7. The first one in each pair is labeled "BW" for beryllium workers, and the second one is labeled "NE" for not exposed. **Consequently, each pair of boxplots provides a comparison of the distribution of the SIs for the beryllium group and the not exposed group for each of the six culture conditions.** Consider, for example, the first two boxplots in Figure 3 which are for beryllium concentration 1 on day 5 (BW-d5be1 and NE-d5be1) for the LAV estimates. Both distributions are centered near zero (for log(SI)), and the not exposed group is a little more spread out in the center. The beryllium workers group shows nine outlying values in the positive direction and one in the negative direction. The notches (which represent confidence limits for the sample median) in the boxplots overlap, indicating that the difference in the location of the two distributions is not significant at a rough 5% level. The dashed horizontal line corresponds to log(SI) equal to zero, and passes through both notches indicating that both distributions are centered near

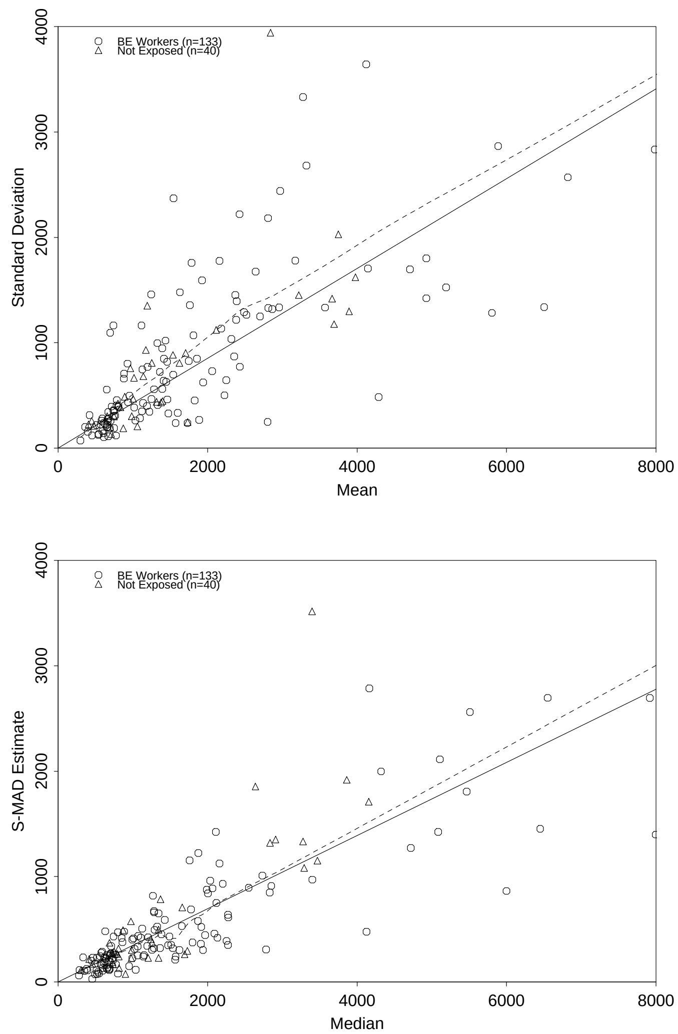

Figure 1: (top) Relationship between the mean, $\bar{y}$, and the standard deviation , $\hat{\sigma}$, for day 5 control wells. The solid line is the ordinary least squares fit, $\hat{\sigma} = 0.448\bar{y}$. (bottom) Relationship between the Median, $\tilde{y}$, and the MAD, $\tilde{\sigma}$. The L_1 solution, the solid line, is $\tilde{\sigma} = 0.34\tilde{y}$. The dashed lines are the result of applying scatterplot smoothers to the data.

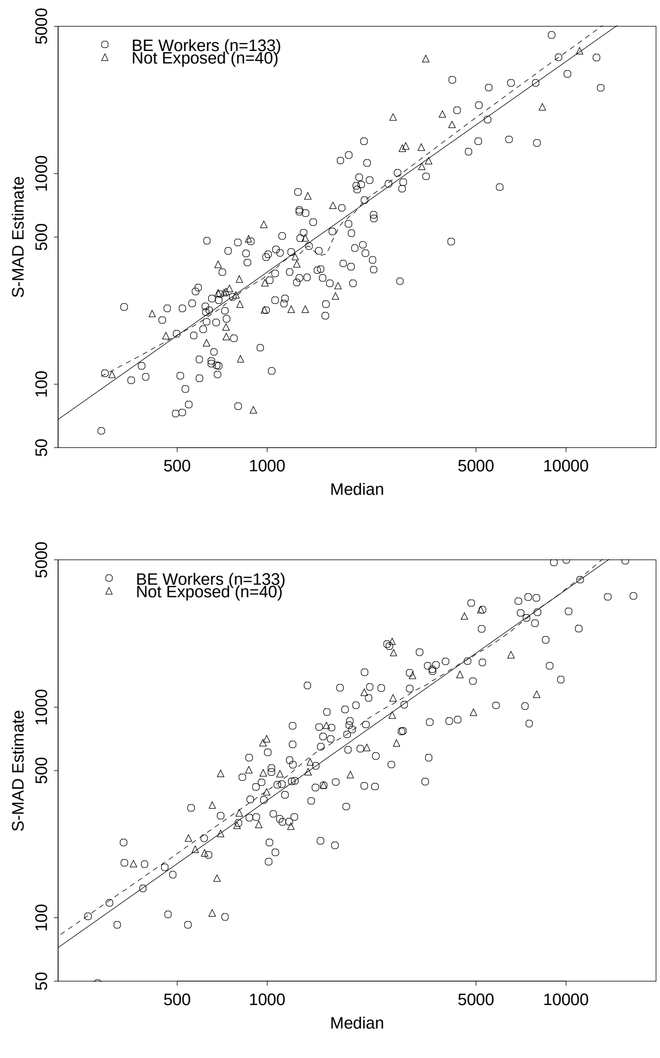

Figure 2: Relationship between the Median, $\tilde{y}$, and the MAD, $\tilde{\sigma}$, ($\tilde{\sigma} = 0.34\tilde{y}$) for the day 5 (top) and day 7 (bottom) control wells ($\tilde{\sigma} = 0.361\tilde{y}$) shown on a log scale. Note that since the results are shown on a log scale, the slope of the line is 1.

zero on the log scale.

Each of the boxplots in Figure 3 is centered near zero and is spread out evenly in both directions. As the beryllium concentration increases on each day the variability (as indicated by the length of the boxes) increases. Results on day 7 are more variable than those on day 5 for each of the three beryllium concentrations. Also note that on day 7 the SIs for the beryllium workers are generally smaller that those for the not exposed group, and that the median log(SIs) are less than zero except for the ne-d7be100 group.

The second graphical display (Figure 4) shows a normal (Gaussian) probability plot for the combined 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(SIs)) 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 [5]. In our situation statistical theory indicates that the log(SIs) should be approximately normally distributed, and the large sample standard deviation should be about 0.28 if the coefficient of variation is 0.4. If the relation between the empirical and theoretical quantiles is linear, this indicates that the distribution is Gaussian. In each plot we have included the median (labeled M) and a resistant estimate of the standard deviation (labeled S) for the log SIs. The solid line in each plot shows the relation that is expected if the log SI values are from a normal distribution with mean M (which determines the intercept) and standard deviation S (which determines the slope). (The values of M and S are also shown in Table 4). Note that we have used resistant methods to estimate the mean and standard deviation for the combined data from the BE and NE groups. This reflects our assumption that most of beryllium workers do not show an abnormal response, i.e. they look like the not exposed group. For example, consider the plot for day 5 Be-1 in Figure 4. The log(SIs) appear to be approximately normal in the center, but there are several values that are larger than expected (these are the points above the line). These "outliers" are SIs that indicate hypersensitivity to beryllium.

The results in Figures 3 and 4 indicate that the log(SIs) are approximately normally distributed. The center of each log(SI) distribution is greater that zero for each beryllium concentration on day 5 and is less than zero for each beryllium concentration on day 7. Note that the untransformed SI units are shown on the vertical scale on the right side of each plot. The estimated standard deviations increase with beryllium concentration on each day, and are larger on day 7 than on day 5. The median and $\tilde{s}$ values for each method are summarized in Table 4

Table 4: Median Estimates ($\tilde{\mu}$) and Resistant Estimates ($\tilde{s}$) of the standard deviation (shown in parenthesis) of log(SI) for LPT Data.

Method	**Day 5**			**Day 7**		
	Be 1	Be 10	Be 100	Be 1	Be 10	Be 100
Current - $\tilde{\mu}_j$	0.069	0.104	0.280	-0.191	-0.330	-0.163
$\tilde{s}_j$	(0.300)	(0.568)	(0.802)	(0.514)	(0.857)	(1.066)
LAV - $\tilde{\mu}_j$	0.066	0.152	0.284	-0.211	-0.388	-0.139
$\tilde{s}_j$	(0.317)	(0.531)	(0.770)	(0.599)	(0.883)	(1.113)
QL - $\tilde{\mu}_j$	0.049	0.102	0.258	-0.211	-0.375	-0.226
$\tilde{s}_j$	(0.337)	(0.561)	(0.795)	(0.665)	(0.918)	(1.155)

3.4. Comparison of Current Method and LAV Method

To compare the current method and the LAV approach, we subtract the six LAV log(SIs) for each of the 173 LPTs from the corresponding log(SIs) calculated using the current method, multiply by 100, and take the average of the differences, i.e.

$$\text{avedif12} = \text{mean}[100 * (\tilde{\beta}_{ij}^{CM} - \tilde{\beta}_{ij}^{L}), j = 1, \ldots, 6].$$

For example, for AC147 (see Table 5 and Appendix E, the current method day 5 be100 SI is $\exp(1.41) = 4.10$ and $\exp(1.45) = 4.26$ for the LAV procedure, and $100 \times \log(4.10/4.26) = 100 \times (1.41 - 1.45) = -4L\%$ where L% stands for the logarithmic percent (see [25]). The SI for the current method is 96 percent of the LAV SI, i.e. about 4% smaller. The average difference between the current

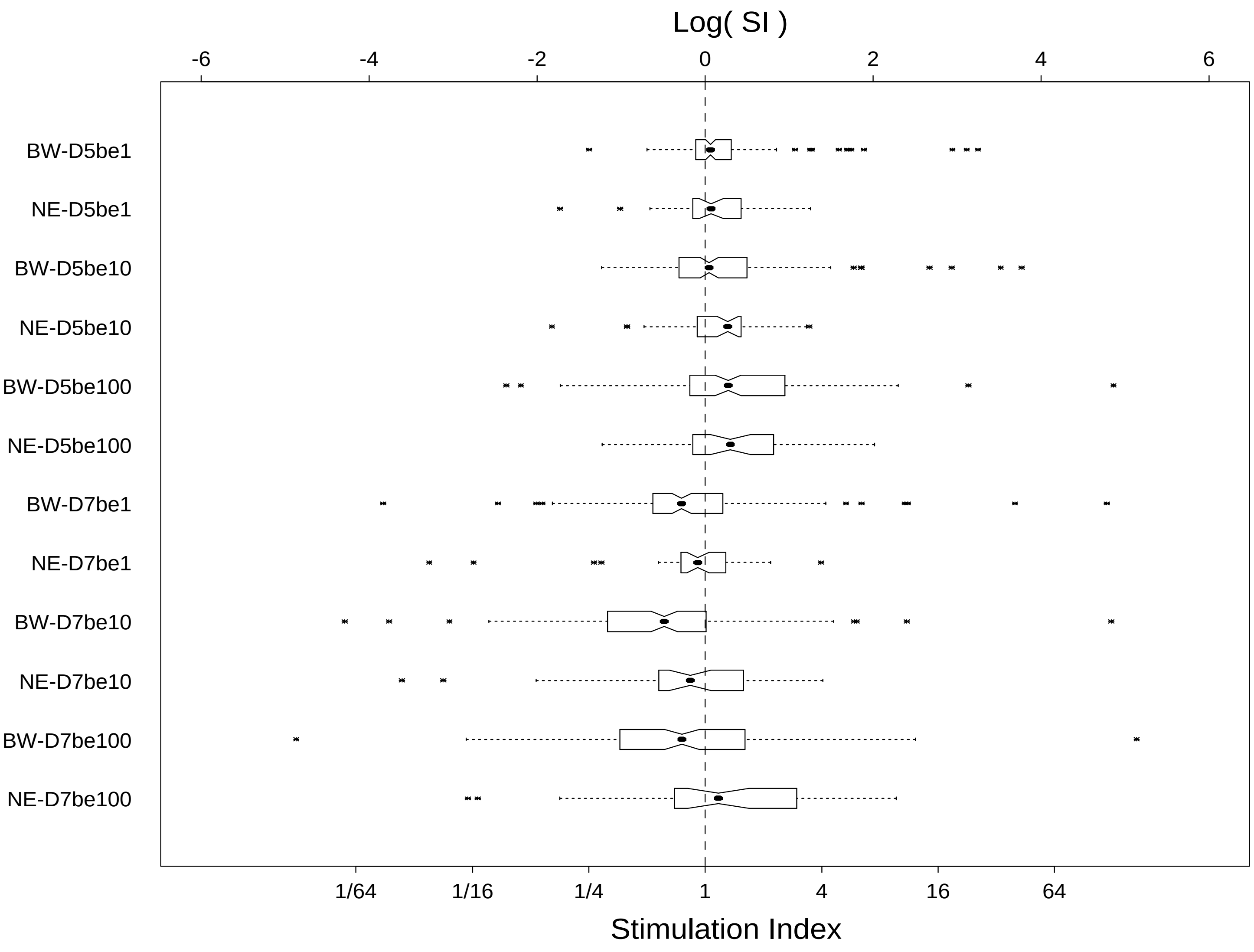

Figure 3: Boxplots for beryllium workers (BW) and not exposed (NE) control group. LAV SIs by day and beryllium concentration

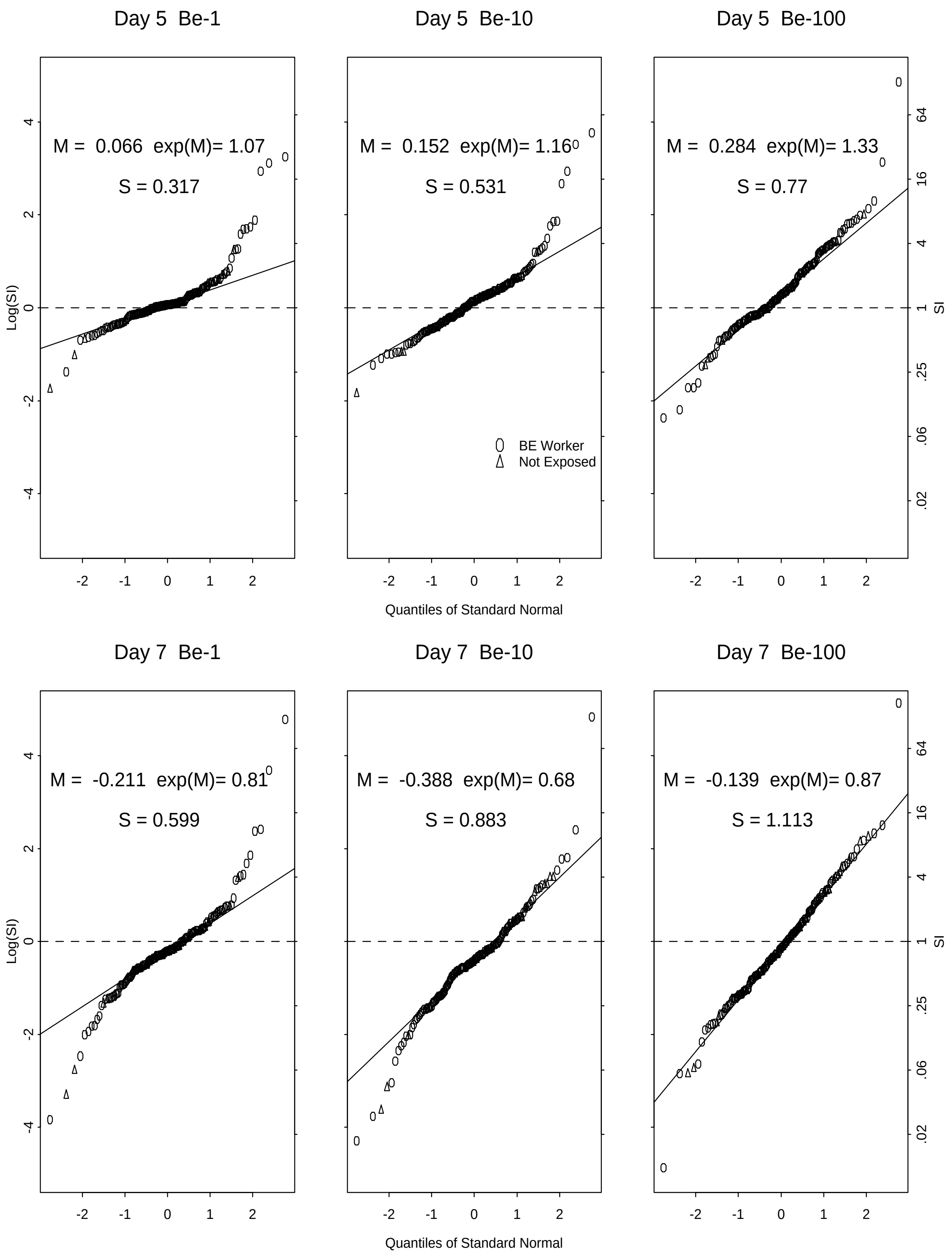

Figure 4: Normal Probability Plots of LAV log(SIs)

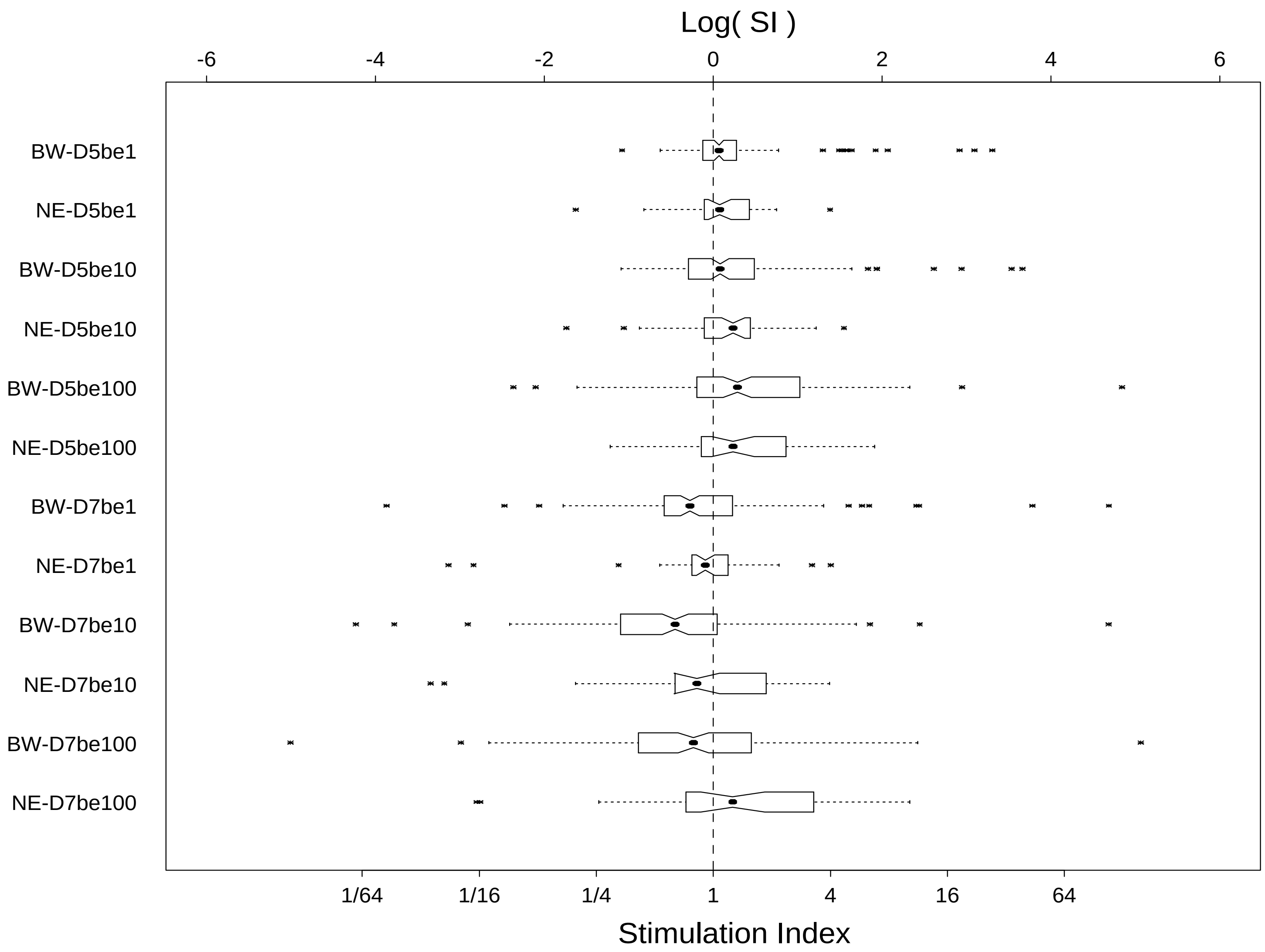

Figure 5: Boxplots for beryllium workers (BW) and not exposed (NE) control group. Current Method SIs by day and beryllium concentration

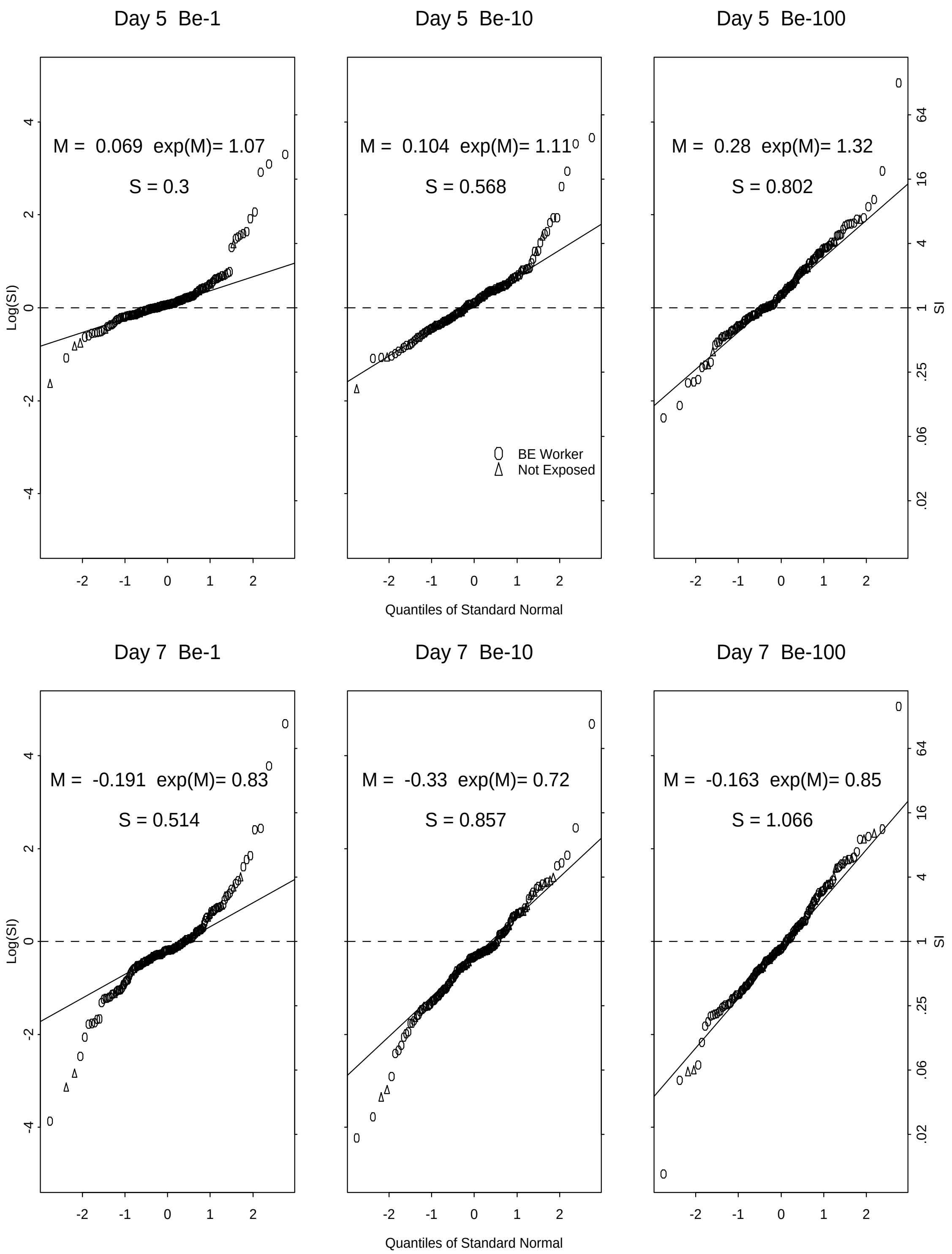

Figure 6: Normal Probability Plots of Current Method log(SIs)

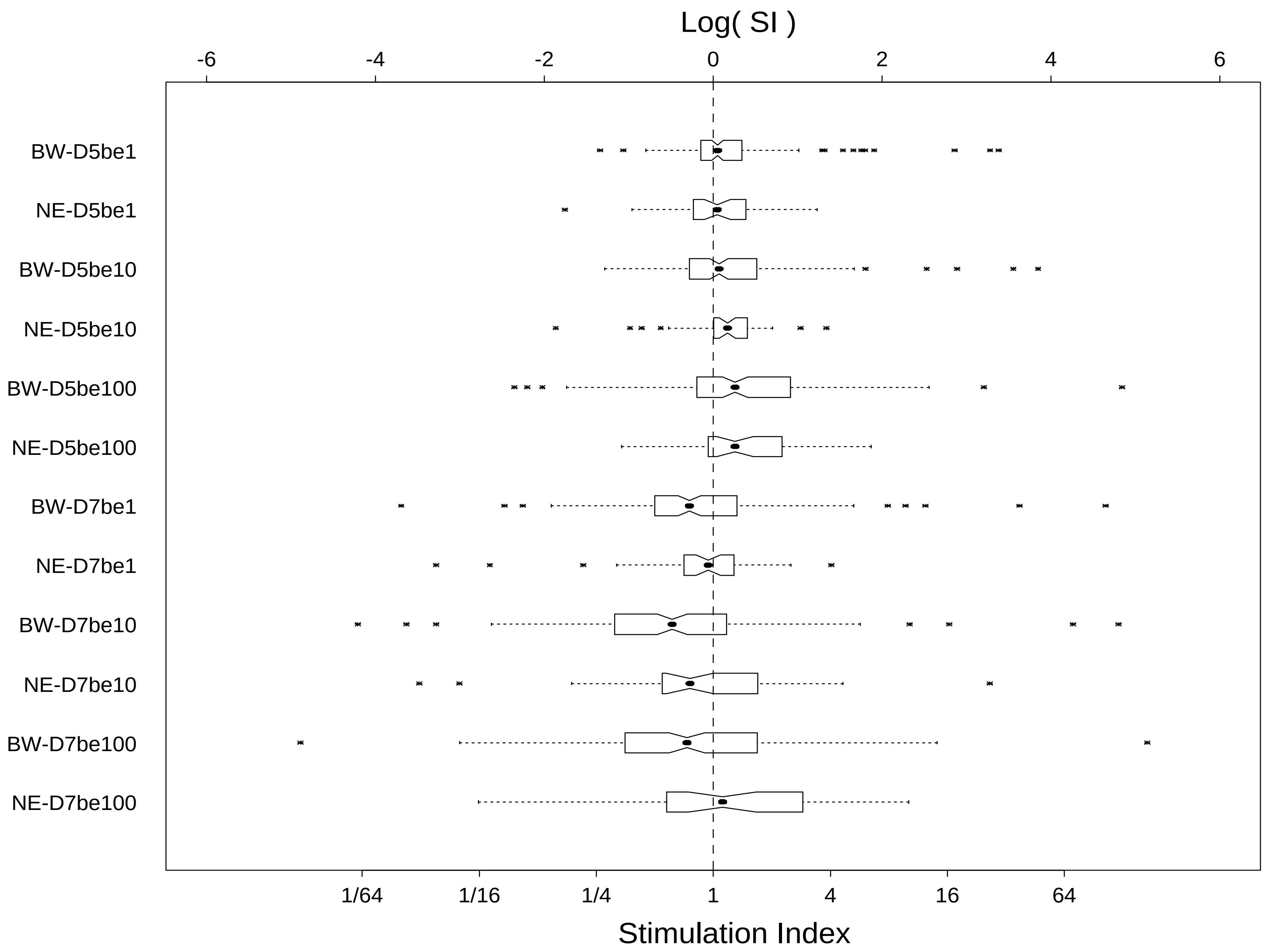

Figure 7: Boxplots for beryllium workers (BW) and not exposed (NE) control group. QL SIs by day and beryllium concentration

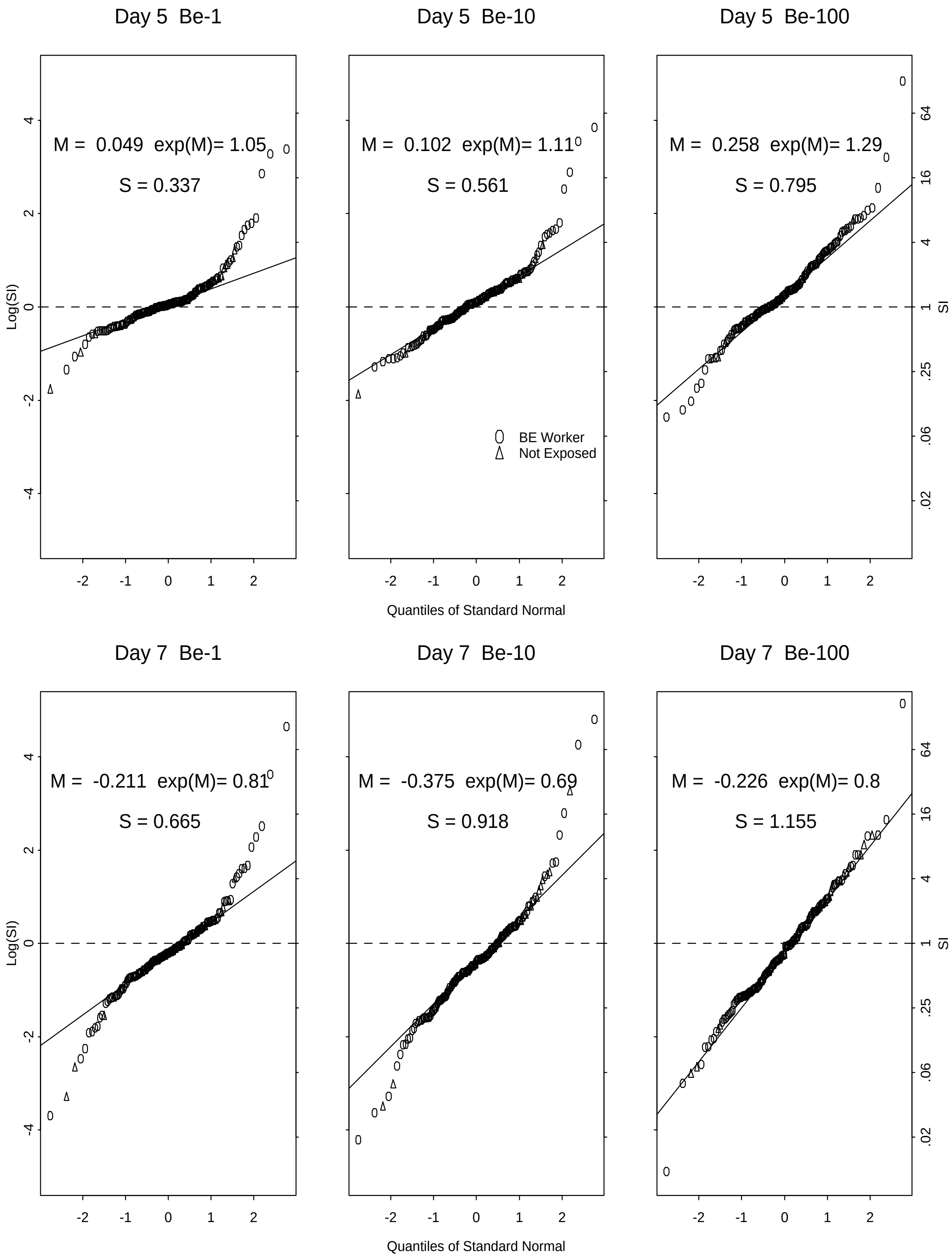

Figure 8: Normal Probability Plots of QL log(SIs)

method and LAV method for AC147 is -1.7L% . Table 5 compares AC147's log(SIs) for all three methods.

Table 5: Comparison of Log(SIs) for AC147

Method	**Day 5**			**Day 7**		
	Be 1	Be 10	Be 100	Be 1	Be 10	Be 100
Current	0.33	1.83	1.41	-0.17	1.86	1.81
LAV	0.26	1.85	1.45	-0.01	1.81	1.81
QL	0.16	1.67	1.39	-0.03	1.72	1.67
$100 * (CM - LAV)$	7	-2	-4	-16	5	0
$100 * (QL - LAV)$	-10	-18	-6	-2	-9	-14

Figure 9 shows a boxplot of avedif12 (as defined above) for the 173 LPTs – note that the average difference is between -3L% and 9L% fifty percent of the time and the mean of the average difference is 3L%. A large, positive value of avedif12 indicates that the current method log(SIs) for an LPT are greater than the LAV log(SIs).

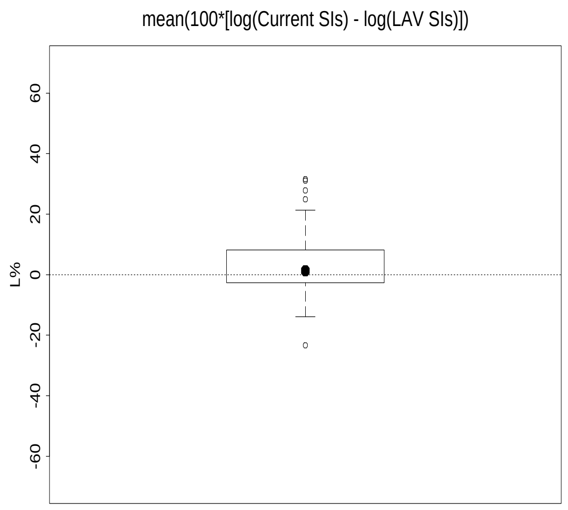

Figure 9: Average Difference of the Current Method and LAV log(SIs)

3.5. Comparison of Resistant Quasi-Likelihood Method and LAV Method

The average difference is also used to compare the quasi-likelihood and LAV methods. This quantity is defined as

$$\text{avedif32} = \text{mean}[100 * (\tilde{\beta}_{ij}^{QL} - \tilde{\beta}_{ij}^{L}), j = 1, \ldots, 6],$$

and a large positive value here indicates that the QL log(SIs) for an LPT are greater than the LAV log(SIs). Figure 10 shows a boxplot of these values. The average difference is between -6L% and 3L% fifty percent of the time, and the mean of the average difference is -0.6L% .

Figure 11 compares the distribution of $\tilde{\phi}_L$ (the LAV coefficient of variation, see Section 2.2) and $\tilde{\phi}_Q$ (the QL coefficient of variation, see Section 2.3). For both methods, the estimated coefficient of variation is between 0.25 and 0.40 most of the time. The median value of $\tilde{\phi}_L$ is 0.321 and the median value of $\tilde{\phi}_Q$ is 0.329

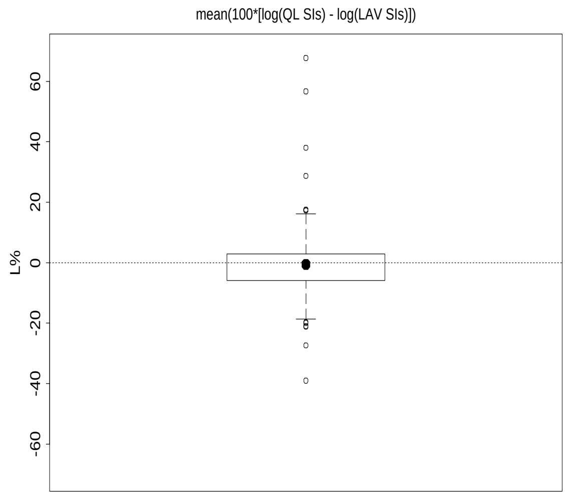

Figure 10: Average Difference between QL and LAV log(SIs)

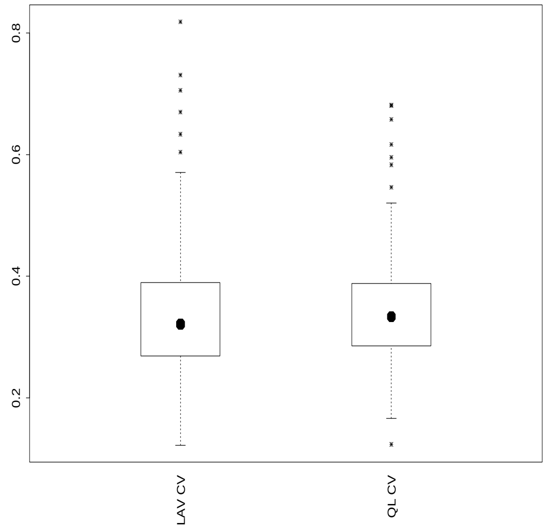

Figure 11: Comparison of QL and LAV Coefficients of Variation ($\tilde{\phi}s$)

3.6. Identification of LPT Results With Large SIs

The method that we suggest for identifying an abnormal LPT result is based on the assumption that the log(SIs) are approximately normally distributed (see Figure 4). The first step is to convert each log(SI) into a standardized deviate

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

using the values of $\tilde{\mu}_j$ and $\tilde{s}_j$ given in Table 4. Results of this computation for the LAV log(SIs) for AC147 (see Table 5) are given in Table 6. These standardized deviates can be compared with the quantiles of the standard normal distribution, i.e. $\Pr[u < z_p] = p$. For example, with $z_{.975} = 1.96$, two of the SIs for AC147(day 5 Be10 and day 7 Be10) would be considered large (at the .025 significance level). 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 $1 - p^6 = .141$ (for $p = .975$). The probability of at least two is $1 - [p^6 + 6(1-p)p^5] = .009$

(for $p = .975$). 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.

Table 6: Values of Standardized Deviates for AC147 using LAV Method

Day 5			Day 7		
Be 1	Be 10	Be 100	Be 1	Be 10	Be 100
0.61	3.21	1.52	0.33	2.48	1.75

The results of applying this procedure to the 173 LPT assays using $z_{.975} = 1.96$ are shown in Table 7.The last row in Table 7 gives the values of

$$\mathrm{SI}_j^* = \exp(\tilde{\mu}_j + z_p \tilde{s}_j), j = 1, \ldots, 6$$

that must be exceeded for that SI to be considered large. These were obtained using the values of $\tilde{\mu}_j$ and $\tilde{s}_j$ for the LAV method in Table 4.

The method that is currently being used at ORISE to identify workers with large SIs is based on the distribution of the maximum SI for each individual in the not exposed group, i.e.

$$\mathrm{MSI}(i) = \max[\exp(\hat{\beta}_{ij}), j = 1, \ldots, 6], i = 1, \ldots, Ne,$$

where Ne = number in the not exposed group. The boxplots in Figure 12 shows this distribution for the not exposed group and the beryllium workers. The mean and standard deviation of the MSI for the not exposed group are used to calculate the value of SI* = mean+2(standard deviations). An LPT for a beryllium worker is define to be "abnormal" if at least two SIs exceed SI* (currently equal to 5.65). Using this approach leads to the identification of AC147, AC149, AC235, and AC236 as abnormal. (Six LPTs that had 2 or more SIs greater than 5.65 were found to be unacceptable based on the values of the within group CVs as described in Appendix B. They are AC161, AC171, AC174, AC196, AC208, and AC225.)

In some situations the quality of the LPT results may be questionable. This could be indicated in the LAV approach by the value of the coefficient of variation,

Table 7: Values of the standardized deviates(u_{ij}s) for LPTs with at Least Two $> z_{.975}$

	Day 5			**Day 7**			
	Be 1	Be 10	Be 100	Be 1	Be 10	Be 100	
ID	u_1	u_2	u_3	u_4	u_5	u_6	$\hat{\phi}_L$
AC128	2.10	0.65	2.11	-1.23	-0.63	0.49	0.30
AC147	0.61	3.21	1.52	0.33	2.48	1.75	0.26
AC161	9.61	6.81	5.95	8.33	5.91	4.74	0.42
AC171	2.47	2.53	2.22	1.90	2.45	2.21	0.18
AC174	-0.13	3.22	2.40	-0.50	1.42	1.07	0.27
AC182	0.87	1.20	1.73	3.45	2.17	0.70	0.33
AC187	3.74	1.96	-0.61	1.62	0.07	-0.98	0.71
AC196	9.08	5.25	2.07	6.50	3.16	2.08	0.38
AC208	5.76	2.15	0.91	4.38	1.04	0.72	0.42
AC209	3.79	0.96	1.99	2.55	1.82	1.63	0.43
AC218	5.28	-0.07	-0.72	2.75	-1.16	-1.07	0.40
AC225	4.81	4.75	1.47	4.32	1.73	1.03	0.82
AC235	5.17	2.23	2.62	2.71	-0.27	1.59	0.39
AC236	10.04	6.34	3.70	3.15	1.46	0.73	0.21
SI*	2.00	3.30	6.00	2.62	3.83	7.72	

Note: $\mathrm{SI}_j^* = \exp(\tilde{\mu}_j + z_p \tilde{s}_j)$ is the SI value that must be exceeded (see text)

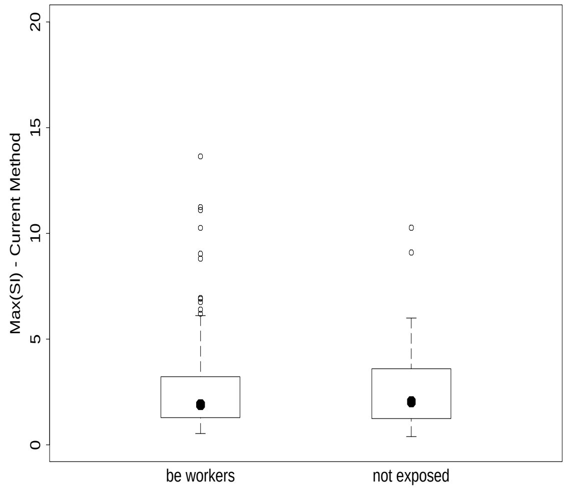

Figure 12: Distribution of Maximum SI from Current Method

$\tilde{\phi}_L$, that is computed for each LPT. Note that $\tilde{\phi}_L$ is a resistant estimate of the "within group" standard deviation of the log well counts. This means that it is not inflated by "outliers" in the well counts (that could be caused by measurement error) and suggests that $\tilde{\phi}_L$ may be reflecting some intrinsic biological variables associated with the lymphocyte proliferation response in certain cell donors.

Figure 13 shows a normal probability plot for $\log(\tilde{\phi}_L)$. The resistant estimates of the mean and standard deviation of $\log(\tilde{\phi}_L)$ are -1.136 and 0.285. From this we compute the 99^{th} percentile $\phi^* = 0.623$. Five of the LPTs in our database have values of $\tilde{\phi}_L > 0.623$. They are AC242, AC223, AC187, AC211, and AC225.

4. Conclusions

The three methods that were considered are the "current" method (in use at ORISE in July, 1993), and two new methods (LAV and QL) that are based on resistant regression techniques. Both of the new methods are highly resistant to outliers (in the well counts), have well known statistical properties, and provide

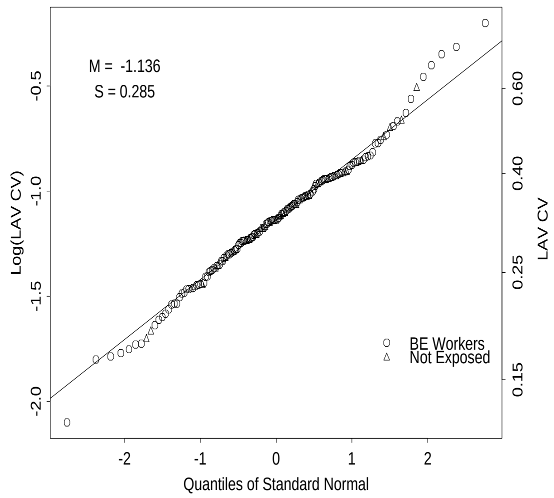

Figure 13: Normal Probability Plot of $\log(\tilde{\phi}_L)$

a "pooled" estimate of the coefficient of variation (ϕ) for each LPT. **The LAV method is also very easy to compute and is recommended for routine analysis of the LPT.**

Each of the three methods was applied to a data base of 173 LPTs (133 from beryllium workers and 40 from individuals with no beryllium exposure). Graphical and numerical summaries show that the three methods are generally in close agreement. Estimates of the log(SIs) are found to be approximately normally distributed. The log(SI) distributions are centered near zero for each of the three concentrations of $BeSO_4$ on harvest day 5 and 7. The variability is greater on day 7 than on day 5, and increases with concentration on each day. Resistant estimates of the location and scale parameters for each of the six log(SI) distributions are used to define "large" SIs, which are used to identify "abnormal" test results. Results of this preliminary approach to identify abnormal LPTs were compared with results obtained using the current method.

In a subsequent report further consideration will be given to the use of the LAV approach to address the following questions:

i) how should "abnormal" LPTs be identified; and

ii) how to determine if an LPT result should be considered "uninterpretable" using the resistant estimate of the coefficient of variation ($\tilde{\phi}_L$).

The approaches(s) that are developed will be applied to a much larger data base of LPTs obtained from ORISE and at least one additional laboratory that is currently using the LPT assay to identify persons who may have CBD.

5. Acknowledgments

The authors thank B. L. Gangaware and L. R. Moore for their assistance in data collection. A draft version of this report was reviewed by several members of the BISAC committee and their colleagues, and by R. L. Schmoyer, J. P. Watkins, and D. L. Cragle. We greatly appreciate their useful comments and we will address points that are beyond the scope of the present study in a subsequent report.

6. References

[1] R.A. Armstrong, E.L. Frome, and D.S. Kung. A revised simplex algorithm for the absolute deviation curve fitting problem. *Commun. Statist.*, B8:175–190, 1979.

[2] G. Basset, Jr. and R. Koenker. Asymptotic theory of least absolute error regression. *Journal of the American Statistical Association*, 73:618–622, 1978.

[3] N.E. Breslow. Extra-poisson variation in log-linear models. *Applied Statistics*, 33:38–44, 1984.

[4] P.J. Burns and R.D Martin. One-sample robust scale estimation in contamination models. submitted, 1992.

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

[6] J.M. Chambers and T.J. Hastie. *Statistical Models in S.* Wadsworth and Brooks/Cole, California, 1992.

[7] A. Charnes, E.L. Frome, and P.L. Yu. The equivalence of generalized least squares and maximum likelihood estimates in the exponential family. *Journal of the American Statistical Association*, 71:214–222, 1976.

[8] E.L. Frome and G.J. Yakatan. Statistical estimation of the pharmocokinetic parameters in the one compartment open model. *Commun. Statist.*, B9:201–222, 1980.

[9] P.J. Green. Iteratively reweighted least squares for maximum likelihood estimation, and some robust and resistant alternatives. *The Journal of the Royal Statistical Society, B*, 46:149–192, 1984.

[10] J.M. Hanifin, W.L. Epstein, and M.J. Cline. In vitro studies of granulomatous hypersensitivity to beryllium. *Journal of Investigative Dermatology*, 55:284–288, 1970.

[11] R.M. Heiberger and R.A. Becker. Design of an *S* function for robust regression using iteratively reweighted least squares. *Journal of Computational and Graphical Statistics*, 1:181–196, 1992.

[12] K. Kreiss, M.M. Mroz, B. Zhen, J.W. Martyny, and L.S. Newman. Epidemiology of beryllium sensitization and disease in nuclear workers. *American Review of Respiratory Disease*, 148:985–991, 1993.

[13] K. Kreiss, L.S. Newman, M.M. Mroz, and P.A. Campbell. Screening blood test identifies subclinical beryllium disease. *Journal of Occupational Medicine*, 31:603–608, 1989.

[14] G. Li. Robust regression. In D.C Hoaglin, F. Mosteller, and J.W. Tukey, editors, *Exploring Data Tables, Trends, and Shapes*, chapter 8. John Wiley & Sons, 1985.

[15] P. McCullagh. Quasi-likelihood functions. *Ann. Statist.*, 11:59–67, 1983.

[16] P. McCullagh and J.A. Nelder. *Generalized Linear Models.* Chapman and Hall, London, 1989.

[17] R. McGill, J.W. Tukey, and W.A. Larsen. Variations of box plots. *The American Statistician*, 32:12–16, 1978.

[18] S.C. Narula. The minimum sum of absolute errors regression. *Journal of Quality Technology*, 19:37–45, 1987.

[19] J.A. Nelder and R.W.M. Wedderburn. Generalized linear models. *The Journal of the Royal Statistical Society, A*, 135:370–384, 1972.

[20] Daryl Pregibon. Resistant fits for some commonly used logistic models with medical applications. *Biometrics*, 38:485–498, 1982.

[21] C.R. Rao. Methodology based on the L_1-Norm, in statistical inference. *Sankhyā*, 50:289–313, 1988.

[22] R.M. Schrader and J.W. McKean. Small sample properties of least absolute errors analysis of variance. In Y. Dodge, editor, *Statistical Data Analysis Based on the L_1-Norm and Related Methods*, pages 307–321. Elsevier Science Publishers, North-Holland, 1987.

[23] Statistical Sciences, Inc., Seattle. *S-PLUS reference manual*, 1991.

[24] R.F. Stokes and M.D. Rossman. Blood cell proliferation response to beryllium: analysis by receiver-operating characteristics. *Journal of Occupational Medicine*, 33:23–28, 1991.

[25] L. Tornquist, P. Vartia, and Y. Vartia. How should relative changes be measured ? *The American Statistician*, 39:43–46, 1985.

[26] J.W. Tukey. *Exploratory Data Analysis.* Addison-Wesley, Reading, Mass, 1977.

A. Calculating Resistant Statistics

Here, the resistant statistics used to calculate the Stimulation Indices and other information are defined.

A.1. The Median

The median of a data set has the property that half the points have a value larger than it and half the points have a value less than it. One way to calculate the median is to sort the data and find the middle term. In general, if there are an odd number of values in a data set, the median is the $(n+1)/2$ term when the data set is sorted. When there are an even number of values in a data set, the median is the average of the $(n+1)/2$ and $((n+1)/2)+1$ terms when the data are sorted. As an example, consider the log of the day 7 be10 data for patient

AC234: `(8.75, 8.06, 8.84, 8.86)`. When the values are sorted, the data set is `(8.06, 8.75, 8.84, 8.86)` and the median is $(8.75 + 8.84)/2 = 8.80$.

A.2. The Median Absolute Deviation (MAD)

The Median Absolute Deviation (MAD) for a data set $z_1, \ldots, z_n$ is defined as

$$MAD = median(|z_i - \tilde{z}|), i = 1, \ldots, n,$$

where $\tilde{z}$ is the median of the z_is. To calculate a resistant scale estimate, follow these steps:

1. Calculate the median ($\tilde{z}$) of the data set
2. Calculate the deviations $z_i - \tilde{z}$ by subtracting the median from each data value
3. Find the median of the absolute values of the deviations calculated in Step 2 to obtain the MAD estimate
4. Finally, multiply the MAD estimate by $C = 1.48 \times \sqrt{n/(n-p)}$ to obtain a resistant estimate of the standard deviation ($\tilde{s}$). See Section 2.2.

Continuing with our sample data set shown in Section A.1, the resistant scale estimate is calculated as shown in Table 8.

Table 8: Calculating a Resistant Scale Parameter for a Sample Data Set

$y_i - \tilde{y}$	$\|y_i - \tilde{y}\|$	MAD	$\tilde{s} = C \times MAD$
$8.06 - 8.80 = -0.74$	0.74		
$8.75 - 8.80 = -0.05$	0.05		
$8.84 - 8.80 = 0.04$	0.04	0.055	$1.71 \times 0.055 = 0.094$
$8.86 - 8.80 = 0.06$	0.06		

B. Current Method of Analysis

Currently, Stimulation Indices are calculated based on means and the "reliability" of the data is gauged by the coefficient of variation. Outliers are dealt with by deleting points (no more than one third for each set of culture conditions) until the coefficient of variation is less than or equal to .30. A patient's data is acceptable if the coefficient of variation is less than or equal to .30 for both day 5 and day 7 control data and 4 of the 6 sets of treatment data. A patient is deemed beryllium sensitive if 2 or more stimulation indices exceed the mean peak SI + $2\hat{\sigma}$ for people known to have never been exposed to beryllium. The stimulation indices for this procedure are ratios of the treatment means and the corresponding control data, i.e.

$$SI = \frac{\text{mean(treated)}}{\text{mean(control)}}.$$

As positive control wells are counted for only 10 minutes, their stimulation indices are defined as

$$SI = 3 \times \frac{\text{mean(positivecontrol)}}{\text{mean(control)}}.$$

The well counts for AC234 are shown in Table 9 and Table 10 shows the results of the current analysis using this data.

C. Least Absolute Values Method

The steps necessary to calculate the LAV SIs are outlined below. The procedure is as follows:

- Take natural logs of the well counts
- Calculate medians for the following data subsets (culture conditions): day 5 Controls, day 5 be1, day 5 be10, day 5 be100, day 7 Controls, day 7 be1, day 7 be10, day 7 be100, and positive control data
- Calculate the log of the Stimulation Index defined as:

$$\log(SI) = \text{Median(treated)} - \text{Median(controls)}$$

Table 9: Well Counts for AC234

Data Subset	Raw Data			
day5 controls	2247	1257	2397	2302
day5 controls	2639	1753	3225	2432
day5 controls	3412	2006	1814	1489
day5 be1	3162	2358	1878	2546
day5 be10	3416	4684	4040	4571
day5 be100	7990	10050	7351	11334
day7 controls	1774	2043	2239	4929
day7 controls	1491	5155	1601	3254
day7 controls	1666	2864	1935	2716
day7 be1	1888	1899	1079	2253
day7 be10	6340	3181	6919	7074
day7 be100	13397	13242	10397	13476
pha	185261	99187	127343	147382
candida	10584	16998	23131	11299

Table 10: Current Analysis for AC234

Data Subset	Reps	Time	Avg	CV	SI	logSI
day5 controls	12.00	30.00	2247.75	0.29		
day5 be1	4.00	30.00	2486.00	0.21	1.11	0.10
day5 be10	4.00	30.00	4177.75	0.14	1.86	0.62
day5 be100	4.00	30.00	9181.25	0.20	4.08	1.41
day7 controls	10.00	30.00	2158.30	0.28		
day7 be1	4.00	30.00	1779.75	0.28	0.82	-0.19
day7 be10	3.00	30.00	6777.67	0.06	3.14	1.14
day7 be100	4.00	30.00	12628.00	0.12	5.85	1.77
pha	4.00	10.00	139793.25	0.26	186.57	5.23
candida	3.00	10.00	12960.33	0.27	17.31	2.85

for the treatment data and

$$\log(SI) = \text{Median(pos.controls)} - \text{Median(controls)} + \log(3)$$

for the positive control data.

- Calculate a resistant coefficient of variation (this is $\tilde{\phi}_L$) defined in Section 2.2.

Table 11 illustrates calculating the Stimulation Indices using the LAV method for patient AC234.

The calculations were carried out with the following S-PLUS function, named `m2`:

```
function(patient, db, looping = F)
{
#m2
#Example m2("AC153")
#
#patient is a patient id, his/her corresponding well counts are used
#in analysis looping is a logical set to true by the function m2.for.
#The variable patient in this case is a patients data. m2.for
#takes care of attributes, etc.
        if(!looping) {
                y <- db[patient, ]
                attributes(y) <- NULL
                y <- unlist(y)
        }
        else y <- patient
        z <- log(y)     #
#
#first calculate the medians
        m5c <- median(z[1:12], na.rm = T)
        m51 <- median(z[13:16], na.rm = T)
        m510 <- median(z[17:20], na.rm = T)
        m5100 <- median(z[21:24], na.rm = T)
        m7c <- median(z[25:36], na.rm = T)
        m71 <- median(z[37:40], na.rm = T)
        m710 <- median(z[41:44], na.rm = T)
        m7100 <- median(z[45:48], na.rm = T)
        mpha <- median(z[49:52], na.rm = T)
        mcand <- median(z[53:56], na.rm = T)    #
#
```

```
#now calculate the log(SIs)
        d51 <- m51 - m5c
        d510 <- m510 - m5c
        d5100 <- m5100 - m5c
        d71 <- m71 - m7c
        d710 <- m710 - m7c
        d7100 <- m7100 - m7c
        pha <- mpha - m5c + log(3)
        cand <- mcand - m5c + log(3)    #
#
#sample size and number of parameters
#
        N <- length(z[!is.na(z)])
        tmp <- c(d51, d510, d5100, d71, d710, d7100, m5c, m7c, pha
        , cand)
        npar <- length(tmp[!is.na(tmp)])        #
#
# scale estimators
#
        meds <- rep(c(m5c, m51, m510, m5100, m7c, m71, m710, m7100,
        mpha, mcand), c(12, 4, 4, 4, 12, 4, 4, 4, 4, 4))
        residuals <- z - meds
        phitil <- median(abs(residuals - median(residuals, na.rm =T)),
        na.rm = T)/0.6745
        phitil <- sqrt(N/(N - npar)) * phitil
        sclres <- residuals/phitil
        absres <- abs(sclres)
        weights <- ifelse(absres > 1.345, 1.345/absres, 1)
        phihat <- sqrt((sum(weights * (residuals)^2, na.rm = T))
        /(N - npar))    #
#
#calculate separate phitiles for day 5 and 7 control & treatments
#
        res5c <- residuals[1:12]
        N5c <- length(res5c[!is.na(res5c)])
        phitil5c <- median(abs(res5c - median(res5c, na.rm = T))
        , na.rm = T)/0.6745
        phitil5c <- phitil5c * sqrt(N5c/(N5c - 1))
        par5t <- c(m51, m510, m5100)
        Npar5t <- length(par5t[!is.na(par5t)])
        res5t <- residuals[13:24]
        N5t <- length(res5t[!is.na(res5t)])
        phitil5t <- median(abs(res5t - median(res5t, na.rm = T))
        , na.rm = T)/0.6745
        phitil5t <- phitil5t * sqrt(N5t/(N5t - Npar5t))
```

```
        res7c <- residuals[25:36]
        N7c <- length(res7c[!is.na(res7c)])
        phitil7c <- median(abs(res7c - median(res7c, na.rm = T))
        , na.rm = T)/0.6745
        phitil7c <- phitil7c * sqrt(N7c/(N7c - 1))
        par7t <- c(m71, m710, m7100)
        Npar7t <- length(par7t[!is.na(par7t)])
        res7t <- residuals[37:48]
        N7t <- length(res7t[!is.na(res7t)])
        phitil7t <- median(abs(res7t - median(res7t, na.rm = T))
        , na.rm = T)/0.6745
        phitil7t <- phitil7t * sqrt(N7t/(N7t - Npar7t)) #
#
#Other things of interest              #
#
        Nprime <- sum(weights, na.rm = T)
        tmp <- absres[!is.na(absres)]
        OO.005 <- length(tmp[tmp > qnorm(0.995)])
        OO.0005 <- length(tmp[tmp > qnorm(0.9995)])     #
#output results         #
#
        output <- c(d51, d510, d5100, d71, d710, d7100, m5c, m7c, pha,
        cand, phihat,  phitil, N, Nprime, OO.005, OO.0005, phitil5c,
        phitil5t, phitil7c, phitil7t)
        if(!looping)
                names(output) <- c("d5be1si", "d5be10si", "d5be100si",

                        "d7be1si", "d7be10si", "d7be100si", "day5",
                        "day7", "pc.pha", "pc.cand", "phihat",
                        "phitild", "N", "Nprime", "OO.995", "OO.9995",
                        "phitil5c", "phitil5t", "phitil7c", "phitil7t")
        output
}
```

Additional information included in the summary report (see Appendix E) for the LAV method are standardized residuals (which can be used to identify outliers) and the MAD estimates of the standard deviation for each set of culture conditions.

Table 11: Calculation of LAV SIs for Patient AC234

Data Subset	Sorted Data	Median	log(SI)	SI
Day 5 Controls	7.14 7.31 7.47 7.50 7.60 7.72 7.74 7.78 7.80 7.88 8.08 8.14	$(7.72 + 7.74)/2 = 7.73$		
Day 5 be1	7.54 7.77 7.84 8.06	$(7.77 + 7.84)/2 = 7.79$	$7.79 - 7.73 = 0.06$	1.06
Day 5 be10	8.14 8.30 8.43 8.45	$(8.30 + 8.43)/2 = 8.37$	$8.37 - 7.73 = 0.64$	1.90
Day 5 be100	8.90 8.99 9.22 9.34	$(8.99 + 9.22)/2 = 9.11$	$9.11 - 7.73 = 1.38$	3.97
Day 7 Controls	7.31 7.38 7.42 7.48 7.57 7.62 7.71 7.91 7.96 8.09 8.50 8.55	$(7.62 + 7.71)/2 = 7.67$		
Day 7 be1	6.98 7.54 7.55 7.72	$(7.54 + 7.55)/2 = 7.55$	$7.55 - 7.67 = -0.12$	0.87
Day 7 be10	8.06 8.75 8.84 8.86	$(8.75 + 8.84)/2 = 8.80$	$8.80 - 7.67 = 1.13$	3.10
Day 7 be100	9.25 9.49 9.50 9.51	$(9.49 + 9.50)/2 = 9.50$	$9.50 - 7.67 = 1.83$	6.23
pha	11.50 11.75 11.90 12.13	$(11.75 + 11.90)/2 = 11.83$	$11.83 - 7.73 + \log(3) = 5.20$	181.02
candida	9.27 9.33 9.74 10.05	$(9.33 + 9.74)/2 = 9.54$	$9.54 - 7.73 + \log(3) = 2.91$	18.33

D. Quasi Likelihood Details

The quasi likelihood analysis uses the Gamma distribution to describe the error component of the model and a log link function. Calculations for the quasi analysis were carried out with the following functions. The first function, `m3`, carries out preliminary calculations and then makes a call to the `glm` function where the IWLS algorithm is implemented. On each iteration, the `glm` function calls the `ROBGAMMA` function (the second function shown). This function establishes the variance and link function and calculates the deviance and a set of robust weights on each iteration.

```
function(y)
{
#m3 - quasi likelihood analysis
#
#y is a vector of well counts from an LPT assay
#
#
#Transform data
    z <- log(y)    #
#
#first calculate the medians
    m5c <- median(z[1:12], na.rm = T)
    m51 <- median(z[13:16], na.rm = T)
    m510 <- median(z[17:20], na.rm = T)
    m5100 <- median(z[21:24], na.rm = T)
    m7c <- median(z[25:36], na.rm = T)
    m71 <- median(z[37:40], na.rm = T)
    m710 <- median(z[41:44], na.rm = T)
```

```
    m7100 <- median(z[45:48], na.rm = T)
    mpha <- median(z[49:52], na.rm = T)
    mcand <- median(z[53:56], na.rm = T)    #
#
#figure out number of parameters & sample size
    tmp <- c(m5c, m51, m510, m5100, m7c, m71, m710, m7100, mpha, mcand)
    npar <- length(tmp[!is.na(tmp)])
    N <- length(y[!is.na(y)])    #
#
#initial phitilde
#
    j <- rep(1:10, c(12, 4, 4, 4, 12, 4, 4, 4, 4, 4))
    yhat <- rep(tapply(y, j, median, na.rm = T), c(12, 4, 4, 4, 12, 4,

        4, 4, 4, 4))
    gamres <- (y - yhat)/yhat
    phitilde.init <- sqrt(N/(N - npar)) * mad(gamres, na.rm = T)    #
#
#let the ROBGAMMA function have access to the scale parameter
#
    assign("robscale", phitilde.init, fr = 1)    #
#
#fit model - XX is model matrix of indicator variables see Table 2
#            OFFSET is to account for differences in counting times
#
    OFFSET <- rep(c(0, log(10/30)), c(48, 8))
    logfit <- glm(y~ XX - 1 + offset(OFFSET), na.action =
na.omit,
        start= rep(tmp, c(12, 4, 4, 4, 12, 4, 4, 4, 4, 4)), family =
        ROBGAMMA)    #
#
#set up for output and do other calculations
#
    logySI <- logfit$coef
    phihat <- sqrt((sum(resid(logfit, type = "pearson")^2))
    /logfit$ df.residual)
    newres <- (y[!is.na(y)] - fitted(logfit))/fitted(logfit)
    phitilde.f <- sqrt(N/(N - npar)) * mad(newres)
    iters <- logfit$iter
    npr <- NPRIME
    c(logySI, phihat, phitilde.init, phitilde.f, N, npr, iters)}



$family:
```

```
           name           link       variance
 "Robust Gamma" "Log: log(mu)" "Square: mu^2"

$names:
[1] "Log: log(mu)"

$link:
function(mu)
log(mu)

$inverse:
function(eta)
care.exp(eta)

$deriv:
function(mu)
1/mu

$initialize:
expression({
    if(!is.null(dimy <- dim(y))) {
        if(dimy[2] > 1)
            stop("multiple responses not allowed")
        else y <- drop(y)
    }
    else y <- as.numeric(y)
    mu <- y + 0.167 * (y == 0)
}
, maxit <- 20)

$variance:
function(mu)
mu^2

$deviance:
function(mu, y, w, residuals = F, robust = T)
{
    old.deviance <- function(mu, y, w, residuals = F)
    {
        nz <- y > 0
        devi <- (y - mu)/mu
        devi[nz] <- devi[nz] - log(y[nz]/mu[nz])
        if(residuals)
            sign(y - mu) * sqrt(2 * abs(devi) * w)
        else 2 * sum(w * devi)
```

```
    }
    if(!robust)
        return(old.deviance(mu, y, w, residuals))
    a <- attr(w, "robust")
    if(is.null(a))
        return(old.deviance(mu, y, w, residuals))
    else {
        robust.scale <- a[1]
        k <- a[2] * robust.scale
        dev <- old.deviance(mu, y, w, T)    #
# remember if there are prior weights they are included here
        devtest <- abs(dev) <= k
        devsq <- dev^2 * devtest + (!devtest) * (2 * k * abs(dev)
         - k^2)
        if(residuals)
            sign(dev) * sqrt(devsq)
        else sum(devsq)
    }
}

$weight:
expression({
        robust.scale <- robscale
        attr(w, "robust") <- c(robust.scale, 1.345)
        robweight <- (1.345 * robust.scale)/abs(family$deviance(mu, y,
            w, T, F))
        robweight <- ifelse(robweight > 1, 1, robweight)#
#
#assigning this way allows other functions to pick up this variable
#without having to make it a component to be passed all the way back
#to the top expression
#
        assign("NPRIME", sum(robweight), fr = 1)    #
        w * robweight
}
)
```

E. Description of Summary Report

The S-PLUS output shown on the following pages is a concise summary of the three methods for analyzing the LPT data. Relevant statistics for each method are reported and a table of the SIs from each method is also shown. We have included three examples of the detailed output - AC153, AC147, and AC234 - at the end of this section.

The first section of the report, labeled `Method 1 - Current Method (ORISE Aug 93)` contains the summary statistics for the current analysis method. For each set of culture conditions, the following information is provided: The number of wells used to calculate the average and coefficient of variation (this column is labeled `Reps` in the output), the counting time (labeled `Time`), the mean (`Avg`), the coefficient of variation (`CV`), the log of the Stimulation Index (`log(SI)`), and the Stimulation Index (`SI`).

The second section of the report contains the results of the LAV analysis and is labeled `Method 2 - LAV Log(Well Counts)`. This section has three distinct subsections, the first contains summary statistics, the second shows the log of the Stimulation Indices and the Stimulation Indices, and the third contains the log of the well counts (labeled `Rep1 - Rep4`) along with the median (`Median`) and the resistant scale estimate (`S-MAD`) for each set of culture conditions, and the standardized residuals (`R1 - R4`).

The standardized residuals are defined as

$$\tilde{\epsilon}_{jk} = \frac{z_{jk} - \tilde{z}_j}{\tilde{\phi}_L}$$

and are the basis for several of the summary statistics shown in the first subsection. A set of weights are calculated from these residuals using Huber's loss function (See Section 2.3). Summing these weights provides us with quantity labeled `Nprime`. We also use these weights to calculate `Phihat` which is defined as

$$\hat{\phi} = \sqrt{\frac{\sum(w_{jk} \times \tilde{\epsilon}_{jk}^2)}{N - p}}$$

where w_{jk} is the calculated weight using Huber's loss function. This quantity is similar to $\tilde{\phi}_L$ which is defined in Section 2.2 and labeled `Phitilde` in the summary report. `N` is the number of available well counts used in the analysis, `N` $>$ `z(.995)` is the number of standardized residuals that exceed the 0.9995 percentile of the standard normal distribution (this number is 2.576), and `N` $>$ `z(.9995)` is the number of standardized residuals that exceed the 0.995 percentile of the standard normal distribution (3.291).

The third section, labeled `Method 3 - Quasi-Likelihood Resistant Regression (Well Counts)` shows the results of the quasi-likelihood analysis. The format of this section is similar to the section containing the LAV results. The first subsection contains summary statistics, the second shows the log of the Stimulation Indices and the Stimulation Indices, and the third shows the well counts (`Rep1 - Rep4`) on the original scale, the fitted values from the QL model (`Fitted`), a resistant scale estimate (`S-MAD`), and the standardized residuals `R1 - R4`). The standardized residuals for the QL analysis are defined in Section 2.3.

In the `Summary Statistics` subsection, the quantity `Phitilde.f` is the value $\tilde{\phi}$, defined in Section 2.3. The value `Phitilde.i` is an estimate of $\tilde{\phi}$ that uses residuals calculated from the initial estimates for the fitted values, i.e. the median for each set of culture conditions. The value of `Iterations` is simply the number of iterations required for convergence.

The final section provides a comparison of the log(SIs) and SIs as calculated by the three methods. This section is labeled `Summary Comparison for Three Methods`.

```
Method 1-Current Method (ORISE Aug 93)(Original Well Counts on next page)AC153
----------------------------------------
              Reps Time       Avg    CV log(SI)      SI
day5 controls   10   30  1220.000 0.281      NA      NA
     day5 be1    3   30   814.333 0.251  -0.404   0.667
    day5 be10    4   30  1744.750 0.217   0.358   1.430
   day5 be100    4   30  4647.750 0.222   1.338   3.810
day7 controls    8   30  2929.750 0.239      NA      NA
     day7 be1    3   30   982.000 0.237  -1.093   0.335
    day7 be10    3   30   761.333 0.130  -1.348   0.260
   day7 be100    4   30  7921.750 0.245   0.995   2.704
          pha    4   10 59633.750 0.255   4.988 146.640
      candida    4   10 25190.500 0.273   4.126  61.944
------------------------------------------------------
Method 2-LAV Log(Well Counts)
-------------------------------
Summary Statistics
------------------------------------------------------
Phitilde (Coef. of Variation)*
------------------------------
Overall:         0.367
Day 5 Control:   0.443  Day 5 Treated:   0.34
Day 7 Control:   0.563  Day 7 Treated:   0.276
------------------------------
N:               56     Nprime:          53.06
N > z(.995):     4      N > z(.9995):    3
-------------------------------------------------------
*Phitilde is MAD est. of the std. dev. on log scale (corresponds to CV
#on orig. scale)

Stimulation Indices
              Day 5               Day 7           Positive Controls
           be1  be10 be100    be1   be10 be100          pha candida
log(SI) -0.423 0.199 1.248 -1.122 -1.436 0.792        4.792   3.909
     SI  0.655 1.221 3.483  0.326  0.238 2.207      120.490  49.860
                    Log(Well Counts)                    Standardized Residuals
                Rep1   Rep2   Rep3   Rep4 Median S-MAD   R1   R2   R3   R4
day5 controls  6.872  7.067  6.719  6.759  7.182 0.443 -0.8 -0.3 -1.3 -1.2
day5 controls  7.296  8.887  6.929  6.883  7.182 0.443  0.3  4.6 -0.7 -0.8
day5 controls  7.313  7.455  7.422  7.597  7.182 0.443  0.4  0.7  0.7  1.1
     day5 be1  6.957  6.560  7.268  6.532  6.758 0.363  0.5 -0.5  1.4 -0.6
    day5 be10  7.347  7.290  7.415  7.741  7.381 0.107 -0.1 -0.2  0.1  1.0
   day5 be100  8.181  8.662  8.297  8.562  8.429 0.313 -0.7  0.6 -0.4  0.4
day7 controls  9.127  8.567  8.239  8.559  8.139 0.563  2.7  1.2  0.3  1.1
day7 controls  7.745  7.953  7.557  8.040  8.139 0.563 -1.1 -0.5 -1.6 -0.3
day7 controls  7.807  8.278  8.035  8.793  8.139 0.563 -0.9  0.4 -0.3  1.8
     day7 be1  6.571  7.034  8.713  7.000  7.017 0.397 -1.2  0.0  4.6  0.0
    day7 be10  6.667  6.741  7.922  6.480  6.704 0.223 -0.1  0.1  3.3 -0.6
   day7 be100  8.706  9.030  8.832  9.254  8.931 0.278 -0.6  0.3 -0.3  0.9
          pha 11.320 10.877 10.872 10.829 10.874 0.041  1.2  0.0  0.0 -0.1
      candida 10.477  9.982  9.978 10.003  9.992 0.021  1.3  0.0  0.0  0.0
```

Method 3 - Quasi-Likelihood Resistant Regression (Well Counts) AC153

Summary Statistics

Phihat:	0.509
Phitilde.i:	0.333
Phitilde.f:	0.295
N:	56
NPrime:	52.313
Iterations:	4

Stimulation Indices

	Day 5			Day 7			Positive Controls	
	be1	be10	be100	be1	be10	be100	pha	candida
log(SI)	-0.367	0.220	1.200	-1.028	-1.391	0.707	4.851	3.989
SI	0.692	1.247	3.321	0.358	0.249	2.028	127.815	53.992

	Well Counts						Standardized Residuals			
	Rep1	Rep2	Rep3	Rep4	Fitted	S-MAD	R1	R2	R3	R4
day5 controls	965	1173	828	862	1399.691	0.391	-1.1	-0.6	-1.4	-1.3
day5 controls	1474	7237	1021	976	1399.691	0.391	0.2	14.2	-0.9	-1.0
day5 controls	1500	1729	1672	1992	1399.691	0.391	0.2	0.8	0.7	1.4
day5 be1	1050	706	1434	687	969.250	0.321	0.3	-0.9	1.6	-1.0
day5 be10	1551	1466	1661	2301	1744.750	0.096	-0.4	-0.5	-0.2	1.1
day5 be100	3571	5780	4011	5229	4647.750	0.305	-0.8	0.8	-0.5	0.4
day7 controls	9202	5253	3786	5212	3906.641	0.420	4.6	1.2	-0.1	1.1
day7 controls	2310	2844	1915	3102	3906.641	0.420	-1.4	-0.9	-1.7	-0.7
day7 controls	2458	3936	3087	6588	3906.641	0.420	-1.3	0.0	-0.7	2.3
day7 be1	714	1135	6084	1097	1397.485	0.258	-1.7	-0.6	11.4	-0.7
day7 be10	786	846	2757	652	972.431	0.171	-0.7	-0.4	6.2	-1.1
day7 be100	6037	8349	6852	10449	7921.750	0.250	-0.8	0.2	-0.5	1.1
pha	82425	52954	52669	50487	59633.750	0.035	1.3	-0.4	-0.4	-0.5
candida	35501	21623	21551	22087	25190.500	0.018	1.4	-0.5	-0.5	-0.4

Summary Comparison for Three Methods

	Day 5			Day 7		
	be1	be10	be100	be1	be10	be100
Log(SIs)						
Method 1	-0.404	0.358	1.338	-1.093	-1.348	0.995
Method 2	-0.423	0.199	1.248	-1.122	-1.436	0.792
Method 3	-0.367	0.22	1.2	-1.028	-1.391	0.707
SIs						
Method 1	0.667	1.43	3.81	0.335	0.26	2.704
Method 2	0.655	1.221	3.483	0.326	0.238	2.207
Method 3	0.692	1.247	3.321	0.358	0.249	2.028

Method 1-Current Method (ORISE Aug 93)(Original Well Counts on next page)AC147

	Reps	Time	Avg	CV	log(SI)	SI
day5 controls	9	30	1806.667	0.243	NA	NA
day5 be1	4	30	2509.250	0.148	0.329	1.389
day5 be10	4	30	11295.500	0.240	1.833	6.252
day5 be100	3	30	7415.333	0.132	1.412	4.104
day7 controls	11	30	2804.455	0.249	NA	NA
day7 be1	3	30	2366.667	0.362	-0.170	0.844
day7 be10	4	30	17929.000	0.233	1.855	6.393
day7 be100	4	30	17146.000	0.219	1.811	6.114
pha	3	10	86739.000	0.191	4.970	144.032
candida	4	10	17866.000	0.278	3.390	29.667

Method 2-LAV Log(Well Counts)

Summary Statistics

Phitilde (Coef. of Variation)*

Overall:	0.264		
Day 5 Control:	0.363	Day 5 Treated:	0.13
Day 7 Control:	0.39	Day 7 Treated:	0.218
N:	56	Nprime:	51.514
N > z(.995):	5	N > z(.9995):	2

---

* Phitilde is MAD est. of the std. dev. on log scale (corresponds to #CV on orig. scale)

Stimulation Indices

	Day 5			Day 7			Positive Controls	
	be1	be10	be100	be1	be10	be100	pha	candida
log(SI)	0.260	1.856	1.450	-0.013	1.805	1.813	5.040	3.315
SI	1.297	6.398	4.264	0.987	6.082	6.130	154.543	27.519

	Log(Well Counts)						Standardized Residuals			
	Rep1	Rep2	Rep3	Rep4	Median	S-MAD	R1	R2	R3	R4
day5 controls	8.422	8.097	7.928	7.444	7.533	0.363	3.4	2.1	1.5	-0.3
day5 controls	7.621	7.351	7.482	8.223	7.533	0.363	0.3	-0.7	-0.2	2.6
day5 controls	7.246	7.137	7.566	7.501	7.533	0.363	-1.1	-1.5	0.1	-0.1
day5 be1	7.758	7.828	7.678	8.015	7.793	0.129	-0.1	0.1	-0.4	0.8
day5 be10	9.426	9.531	8.919	9.352	9.389	0.153	0.1	0.5	-1.8	-0.1
day5 be100	8.964	9.526	8.749	9.003	8.983	0.218	-0.1	2.1	-0.9	0.1
day7 controls	7.624	7.929	7.966	8.237	7.964	0.390	-1.3	-0.1	0.0	1.0
day7 controls	7.640	7.574	8.316	8.066	7.964	0.390	-1.2	-1.5	1.3	0.4
day7 controls	7.962	11.129	7.734	7.978	7.964	0.390	0.0	12.0	-0.9	0.1
day7 be1	8.025	7.250	7.877	8.781	7.951	0.663	0.3	-2.7	-0.3	3.1
day7 be10	10.060	9.497	9.735	9.804	9.769	0.262	1.1	-1.0	-0.1	0.1
day7 be100	9.819	9.735	9.407	9.958	9.777	0.191	0.2	-0.2	-1.4	0.7
pha	11.889	11.485	11.122	11.465	11.475	0.311	1.6	0.0	-1.3	0.0
candida	9.853	9.646	9.454	10.095	9.749	0.342	0.4	-0.4	-1.1	1.3

Method 3-Quasi-Likelihood Resistant Regression (Well Counts) AC147

Summary Statistics

Phihat:	0.787
Phitilde.i:	0.26
Phitilde.f:	0.337
N:	56
NPrime:	51.168
Iterations:	4

Stimulation Indices

	Day 5			Day 7			Positive Controls	
	be1	be10	be100	be1	be10	be100	pha	candida
log(SI)	0.155	1.669	1.386	-0.029	1.718	1.673	4.945	3.217
SI	1.168	5.308	3.998	0.971	5.573	5.329	140.421	24.941

	Well Counts						Standardized Residuals			
	Rep1	Rep2	Rep3	Rep4	Fitted	S-MAD	R1	R2	R3	R4
day5 controls	4547	3285	2774	1710	2148.980	0.281	3.3	1.6	0.9	-0.6
day5 controls	2041	1558	1776	3725	2148.980	0.281	-0.1	-0.8	-0.5	2.2
day5 controls	1403	1258	1931	1809	2148.980	0.281	-1.0	-1.2	-0.3	-0.5
day5 be1	2341	2511	2160	3025	2509.250	0.120	-0.2	0.0	-0.4	0.6
day5 be10	12405	13775	7476	11526	11407.678	0.169	0.3	0.6	-1.0	0.0
day5 be100	7814	13715	6303	8129	8591.519	0.182	-0.3	1.8	-0.8	-0.2
day7 controls	2047	2776	2882	3779	3217.398	0.334	-1.1	-0.4	-0.3	0.5
day7 controls	2079	1946	4088	3183	3217.398	0.334	-1.0	-1.2	0.8	0.0
day7 controls	2869	68132	2285	2915	3217.398	0.334	-0.3	59.9	-0.9	-0.3
day7 be1	3056	1408	2636	6509	3124.210	0.452	-0.1	-1.6	-0.5	3.2
day7 be10	23397	13323	16897	18099	17929.000	0.228	0.9	-0.8	-0.2	0.0
day7 be100	18387	16898	12176	21123	17146.000	0.211	0.2	0.0	-0.9	0.7
pha	145721	97238	67646	95333	100586.974	0.252	1.3	-0.1	-1.0	-0.2
candida	19021	15453	12758	24232	17866.000	0.300	0.2	-0.4	-0.8	1.1

Summary Comparison for Three Methods

	Day 5			Day 7		
	be1	be10	be100	be1	be10	be100
Log(SIs)						
Method 1	0.329	1.833	1.412	-0.17	1.855	1.811
Method 2	0.26	1.856	1.45	-0.013	1.805	1.813
Method 3	0.155	1.669	1.386	-0.029	1.718	1.673
SIs						
Method 1	1.389	6.252	4.104	0.844	6.393	6.114
Method 2	1.297	6.398	4.264	0.987	6.082	6.13
Method 3	1.168	5.308	3.998	0.971	5.573	5.329

Method 1-Current Method (ORISE Aug 93)(Original Well Counts on next page)AC234

	Reps	Time	Avg	CV	log(SI)	SI
day5 controls	12	30	2247.750	0.287	NA	NA
day5 be1	4	30	2486.000	0.214	0.101	1.106
day5 be10	4	30	4177.750	0.139	0.620	1.859
day5 be100	4	30	9181.250	0.200	1.407	4.085
day7 controls	10	30	2158.300	0.277	NA	NA
day7 be1	4	30	1779.750	0.279	-0.193	0.825
day7 be10	3	30	6777.667	0.057	1.144	3.140
day7 be100	4	30	12628.000	0.118	1.767	5.851
pha	4	10	139793.250	0.259	5.229	186.578
candida	3	10	12960.333	0.271	2.851	17.298

Method 2-LAV Log(Well Counts)

Summary Statistics

Phitilde (Coef. of Variation)*

Overall:	0.315		
Day 5 Control:	0.29	Day 5 Treated:	0.196
Day 7 Control:	0.418	Day 7 Treated:	0.075

N:	56	Nprime:	53.861
N > z(.995):	2	N > z(.9995):	0

----

* Phitilde is MAD est. of the std. dev. on log scale (corresponds to #CV on orig. scale)

Stimulation Indices

	Day 5			Day 7			Positive Controls	
	be1	be10	be100	be1	be10	be100	pha	candida
log(SI)	0.074	0.636	1.371	-0.122	1.130	1.829	5.197	2.906
SI	1.077	1.889	3.940	0.885	3.097	6.228	180.708	18.280

	Log(Well Counts)						Standardized Residuals			
	Rep1	Rep2	Rep3	Rep4	Median	S-MAD	R1	R2	R3	R4
day5 controls	7.717	7.136	7.782	7.742	7.729	0.290	0.0	-1.9	0.2	0.0
day5 controls	7.878	7.469	8.079	7.796	7.729	0.290	0.5	-0.8	1.1	0.2
day5 controls	8.135	7.604	7.503	7.306	7.729	0.290	1.3	-0.4	-0.7	-1.3
day5 be1	8.059	7.766	7.538	7.842	7.804	0.251	0.8	-0.1	-0.8	0.1
day5 be10	8.136	8.452	8.304	8.427	8.366	0.127	-0.7	0.3	-0.2	0.2
day5 be100	8.986	9.215	8.903	9.336	9.101	0.268	-0.4	0.4	-0.6	0.7
day7 controls	7.481	7.622	7.714	8.503	7.668	0.418	-0.6	-0.1	0.1	2.7
day7 controls	7.307	8.548	7.378	8.088	7.668	0.418	-1.1	2.8	-0.9	1.3
day7 controls	7.418	7.960	7.568	7.907	7.668	0.418	-0.8	0.9	-0.3	0.8
day7 be1	7.543	7.549	6.984	7.720	7.546	0.151	0.0	0.0	-1.8	0.6
day7 be10	8.755	8.065	8.842	8.864	8.798	0.094	-0.1	-2.3	0.1	0.2
day7 be100	9.503	9.491	9.249	9.509	9.497	0.015	0.0	0.0	-0.8	0.0
pha	12.130	11.505	11.755	11.901	11.828	0.321	1.0	-1.0	-0.2	0.2
candida	9.267	9.741	10.049	9.332	9.537	0.406	-0.9	0.6	1.6	-0.6

Method 3-Quasi-Likelihood Resistant Regression (Well Counts) AC234

Summary Statistics

Phihat:	0.296
Phitilde.i:	0.307
Phitilde.f:	0.32
N:	56
NPrime:	54.11
Iterations:	3

Stimulation Indices

	Day 5			Day 7			Positive Controls	
	be1	be10	be100	be1	be10	be100	pha	candida
log(SI)	0.095	0.614	1.402	-0.296	0.923	1.651	5.223	3.015
SI	1.100	1.848	4.062	0.744	2.516	5.211	185.527	20.384

	Well Counts						Standardized Residuals			
	Rep1	Rep2	Rep3	Rep4	Fitted	S-MAD	R1	R2	R3	R4
day5 controls	2247	1257	2397	2302	2260.482	0.283	0.0	-1.4	0.2	0.1
day5 controls	2639	1753	3225	2432	2260.482	0.283	0.5	-0.7	1.3	0.2
day5 controls	3412	2006	1814	1489	2260.482	0.283	1.6	-0.4	-0.6	-1.1
day5 be1	3162	2358	1878	2546	2486.000	0.230	0.8	-0.2	-0.8	0.1
day5 be10	3416	4684	4040	4571	4177.750	0.132	-0.6	0.4	-0.1	0.3
day5 be100	7990	10050	7351	11334	9181.250	0.252	-0.4	0.3	-0.6	0.7
day7 controls	1774	2043	2239	4929	2423.313	0.356	-0.8	-0.5	-0.2	3.2
day7 controls	1491	5155	1601	3254	2423.313	0.356	-1.2	3.5	-1.1	1.1
day7 controls	1666	2864	1935	2716	2423.313	0.356	-1.0	0.6	-0.6	0.4
day7 be1	1888	1899	1079	2253	1803.309	0.173	0.1	0.2	-1.3	0.8
day7 be10	6340	3181	6919	7074	6096.200	0.103	0.1	-1.5	0.4	0.5
day7 be100	13397	13242	10397	13476	12628.000	0.016	0.2	0.2	-0.6	0.2
pha	185261	99187	127343	147382	139793.250	0.295	1.0	-0.9	-0.3	0.2
candida	10584	16998	23131	11299	15359.339	0.357	-1.0	0.3	1.6	-0.8

Summary Comparison for Three Methods

	Day 5			Day 7		
	be1	be10	be100	be1	be10	be100
Log(SIs)						
Method 1	0.101	0.62	1.407	-0.193	1.144	1.767
Method 2	0.074	0.636	1.371	-0.122	1.13	1.829
Method 3	0.095	0.614	1.402	-0.296	0.923	1.651
SIs						
Method 1	1.106	1.859	4.085	0.825	3.14	5.851
Method 2	1.077	1.889	3.94	0.885	3.097	6.228
Method 3	1.1	1.848	4.062	0.744	2.516	5.211

F. Detailed Protocol - Lymphocyte Proliferation Assay

The ORISE protocol for performing Lymphocyte Proliferation assays essentially adheres to the recommendations of the expert panel (i.e., Committee to Accredit Beryllium Sensitivity Testing [CABST]) convened jointly by the U.S. DOE Office of Health and the Beryllium Industry Scientific Advisory Committee (BISAC) at a meeting held in Washington, DC, on February 3- 4, 1992. We collect approximately 30 ml of venous whole blood in sterile vacutainers containing sodium heparin for each assay (Figure 14). Tubes are inverted to mix blood with the anticoagulant and transported to the laboratory for processing. Cells are maintained at room temperature overnight. Within 24 hours after blood collection, mononuclear cells are separated using Ficoll-hypaque density gradient centrifugation, carried through three sequential washes, and counted in triplicate on an automated cell counter. Lymphocytes are cultured in RPMI 1640 culture medium (GIBCO) buffered with Hepes salts, and supplemented with 2mM/1-glutamine, 100 units per ml penicillin, and 100 μg per ml streptomycin. Pooled human serum is added at a final concentration of 10 percent. We are using 96 well flat-bottom microtiter plates and a final cell concentration of 2.5×10^5 cells per well contained in 0.2 ml volume of medium.

Beryllium sulfate ($BeSO_4$, Aldrich Chemicals, 99.9% purity) in concentrations of 1, 10, and 100 μM is being used to evaluate donor lymphocyte hypersensitivity to Be metals. As positive controls we are using concanavalin-A (10 μg/ml) and phytohemagglutinin (30 μg/ml). For each set of exposures, quadruplicate wells are being evaluated to obtain estimates of lymphocyte proliferation response. Unstimulated control wells are run in replicates of 12 because other laboratories have observed considerable variability in rates of tritiated thymidine incorporated in the control series, and extra replicates are needed to achieve the required levels of statistical confidence. All cells are incubated at $37 \pm 0.5°C$ in an atmosphere of 5% $C0_2$ in air. Cells assayed for response to Be are harvested at five and seven days with a terminal six-eight hour pulse of 1.0 μCi of tritiated thymidine (sp. act. 6.7 mCi/mM). We are using a Packard 96 well cell harvester which deposits lymphocytes from each individual well on a standard glass filter paper which can be counted intact on the Packard Matrix 96 gas ionization counter, or punched for assay using a liquid scintillation counter. The Matrix 96 unit is less efficient in detecting beta decays than scintillation counters, but has the great advantage of simultaneously detecting beta radiation emissions from all 96 wells. Statistical accuracy can be achieved quite readily by increasing counting time using this instrument.

ORISE LPT CULTURE ASSAY

I. Culture Method

- Heparinized blood (~15ml)
- Ficoli-hypaque centrifugation
- Separated lymphocytes

RPMI 1640 Medium
10% pooled human serum
antibiotics

II. Beryllium Challenge

- 2.5 x 10^5 lymphocytes per well
- 96 well flat-bottomed microtiter plates

Beryllium Sulfate(µM)	#Replicate Wells	Day of Harvest
0	12	5, 7
1	4	5, 7
10	4	5, 7
100	4	5, 7
PHA(30µg/ml)	4	5
CON A(10 µM)	4	5

III. Harvest Method (day 5 and day7)

- Add tritiated thymidine (-8:00 A.M.) (1µCi/well sp. act. 5-7 mCi/mMol)
- Freeze plates at -20°C (-4:00 P.M.)
- Perform 30 min counts on Packard Matrix 96 gas ionization counter

IV. Data Reduction

- CONTROL WELLS
12 replicates - drop outlines
calculate mean +/- cv

- Be TREATMENTS
4 Replicates drop 1 outlier
calculate mean +/- cv

- STIMULATION INDEX = (SI)

$$SI = \frac{\text{mean Be treated}}{\text{mean control}}$$

Figure 14: ORISE LPT Culture Assay

F.1. Quality Control

Excess variability in counts between replicate wells within a treatment, i.e., "outliers" could result from technical errors in initiating the tests, or possibly from intrinsic biological variables associated with the characteristics of lymphocyte proliferation response in certain cell donors. Sources of technical error might include mistakes in pipetting, such as failures to add appropriate numbers of cells to individual wells, lack of addition or double addition of tritiated thymidine to specific wells, or improper washing of filters resulting in residual counts of unincorporated thymidine, or smearing of radiolabel across the filter paper.

Stringent methods for quality control are used routinely to guard against inadvertent technical errors. To minimize the risk of pipetting errors, all media and other test reagents are delivered to complete rows or columns of the test plate using electronic micropipetters that deliver up to 8 or 12 aliquots simultaneously. Thus, it is not likely that the operator could "loose her place" in adding reagents. Cells are harvested onto the surface of filter paper using a Packard 96 Well Harvester that simultaneously aspirates the cellular contents from each well. To ensure complete washings of culture plates, a wash volume of approximately 10 times that recommended by the manufacturer is used. For all tests, we routinely leave all wells in rows A and H empty as a quality control measure to allow evaluation of background counts on both the top and bottom of the filter paper. Erratic or high counts in these empty wells would signal incomplete washing of plates or "smearing" or radioactivity from one well to another.

Filter papers are counted intact on the Matrix 96 gas ionization counter, which simultaneously records counts and counts-per-minute with attendant errors for each well. Because the Matrix Counter is a gas ionization unit, only those beta decays that are emitted at right angles to the surface filter pad are detected and recorded. Thus, the sensitivity of the instrument in detecting counts is considerably less than that of a liquid scintillation counter (about 20% of emissions are detected using the gas ionization unit). For this reason, all plates are counted for longer periods of time to accumulate enough counts for statistical accuracy. Routinely, all plates containing control wells and wells challenged by beryllium salts are counted for 30 min, whereas mitogen-stimulated positive controls are counted for 10 min each.

To allow direct comparisons of lymphocyte proliferation response between different blood donors, we routinely initiate 5-day and 7-day tests on lymphocytes from 3 separate donors on a single test plate. The plate map that we routinely employ and examples of new data are displayed below.

Patient #1 Patient #2 Patient #3

	1	2	3	4	5	6	7	8	9	10	11	12
A												
B	C	C	C	C	C	C	C	C	C	C	C	C
C	C	C	C	C	C	C	C	C	C	C	C	C
D	C	C	C	C	C	C	C	C	C	C	C	C
E	*Be1	Be1	Be1	Be1	Be1	Be1	Be1	Be1	Be1	Be1	Be1	Be1
F	Be10	Be10	Be10	Be10	Be10	Be10	Be10	Be10	Be10	Be10	Be10	Be10
G	Be100	Be100	Be100	Be100	Be100	Be100	Be100	Be100	Be100	Be100	Be100	Be100
H												

	P#1	P#2	P#3									
	1	2	3	4	5	6	7	8	9	10	11	12
A	PHA	PHA	PHA									
B	PHA	PHA	PHA									
C	PHA	PHA	PHA									
D	PHA	PHA	PHA									
E	ConA	ConA	ConA									
F	ConA	ConA	ConA									
G	ConA	ConA	ConA									
H	ConA	ConA	ConA									

*Beryllium Sulfate 1,10,100 µ Molar Solutions

Figure 15: ORISE Plate Maps for LPT Assay.

This figure displays the platemap for initiating LPT for Oak Ridge Beryllium Workers. Cells from 3 persons are cultured on the same microtiter plate. Cells from Patient 1 are pipetted into columns 1-4; cells from Patient 2, into columns 5-8; and cells from Patient 3, into columns 9-12. Rows A and H are left blank to monitor background counts in the culture system. Rows B, C, and D are replicate sets of control wells, whereas rows E, F, and G contain beryllium concentrations of 1, 10, and 100 μ M respectively. The lower half of the figure demonstrates the platemap for initiating cultures with phytohemagglutinin or ConA.

PROTOCOL #: 1 NAME: Ac.coh 10-JUN-93 12:01

TIME 30:00 ELAPSED TIME 30:00

	1	2	3	4	5	6	7	8	9	10	11	12
1-A:	57	47	48	52	126	68	99	69		27	37	36
1-B:	515	881	489	303	191	260	673	382	1300	1451	3353	127
1-C:	535	742	1602	676	310	420	251	669	2850	1368	634	1478
1-D:	923	570	510	568	253	550	333	439	540	1654	1487	1330
1-E:	17700	10749	19080	18855	696	372	270	434	1236	1991	1173	1743
1-F:	19197	27501	27280	31033	286	383	758	1369	1175	1591	1617	1877
1-G:	21083	38090	45938	29685	454	428	366	654	1772	2415	2766	3737
1-H:	41	63	52	75	66	83	91	43	49	44	31	24

Figure 16: Typical Printout Sheets of Data from Three Different Individuals.

This figure displays a typical printout sheet of data from three different individuals. The test is a 5-day plate, counted for 30 min. Data are shown as total counts. Patient 1 displays a pronounced response to all 3 levels for beryllium challenged wells, whereas Patient 3 demonstrates higher levels of counts in control wells, but also demonstrates no response to beryllium. Direct comparisons of data between the three persons can be readily made from a single printout sheet. This approach readily allows comparisons of counts within replicate treatments for lymphocytes from the same donor, as well as comparisons of inter-individual variability in counts between different subjects.

08-JUN-93 10:48
TIME 30:00 ELAPSED TIME 30:00

	1	2	3	4	5	6	7	8	9	10	11	12
1-A:	36	30	25	33	55	40	51	44	60	56	60	54
1-B:	1734	1603	758	1501	34	25	37	33	832	6931	1121	670
1-C:	1036	607	2710	523	35	33	25	36	2248	5274	726	1230
1-D:	1018	597	1861	2372	34	27	244	23	1854	896	1566	5718
1-E:	7336	3841	4513	4343	52	33	42	40	28604	3092	16175	20088
1-F:	3145	6285	1628	1850	40	37	40	46	31880	36893	6556	48059
1-G:	3730	9293	18045	6304	49	37	43	38	22735	49557	18859	31056
1-H:	90	97	68	86	54	45	42	46	25	66	64	37

10-JUN-93 12:01
TIME 30:00 ELAPSED TIME 30:00

	1	2	3	4	5	6	7	8	9	10	11	12
1-A:	37	33	33	31	31	32	43	37	46	35	56	37
1-B:	1910	3867	1433	1371	36	41	26	44	3007	968	2460	4952
1-C:	1360	3474	14034	1659	47	40	35	35	3531	16753	1131	5535
1-D:	1336	1391	1096	6996	25	29	37	26	3002	20512	3036	5312
1-E:	10028	1784	13968	3306	39	32	39	50	52721	91752	32199	1115
1-F:	15032	5155	2105	5255	60	39	34	48	23931	21429	15385	12076
1-G:	14231	4622	8587	7924	50	35	38	50	17279	43695	25714	13323
1-H:	55	81	57	50	73	27	28	51	72	78	65	82

Figure 17: Repeat Cultures for a Person.

This figure represents repeat cultures for a person who initially showed a positive response to beryllium. For this set of data, the patient's cells were cultured in columns 1-4 and in 9-12, with the middle for columns being blank. The patient's cells were cultured in two types of serum; i.e., heat inactivated at $56°C$ for 30 min, and in the same lot of sera which had not been heat inactivated. Note pronounced difference in the level of response between the two types of sera.

ORNL/TM-6818

INTERNAL DISTRIBUTION

1. D. E. Conrad
2. T. S. Darland
3. D. J. Downing
4–8. E. L. Frome
9. S. Garrett
10. F. Hartman
11. O. W. Jones
12. G. D. Kerr
13. C. A. Little
14. G. L. Love
15. M. D. Morris
16. C. E. Oliver
17. G. Ostrouchov
18–22. Stephanie Raby
23. D. E. Reichle
24. P. S. Rohwer
25–29. R. F. Sincovec
30. C. C. Travis
31. R. C. Ward
32. J. K. Williams
33. Central Research Library
34. K-25 Applied Technology Library
35–36. Laboratory Records Dep.
37. Laboratory Records – RC
38. Patent Office
39. Y-12 Technical Library

EXTERNAL DISTRIBUTION

40. Barbara P. Barna, Cleveland Clinic Foundation, Immunopathology, One Clinic Center, 9500 Euclid Avenue, Cleveland, OH 44195-5131
41. J. R. Beall, Office of Health and Environmental Research, ER-72, U. S. Department of Energy, Washington, DC 20585
42. N. Breslow, Department of Biostatistics, University of Washington, Seattle, WA 98195
43. Roger W. Brockett, EPMD Advisory Committee , Pierce Hall, 29 Oxford Street, Harvard University, Cambridge, MA 02138
44. B. G. Brooks, Epidemiology and Health Surveillance, EH-42 CC 270, U. S. Department of Energy, 19901 Germantown Road, Germantown, MD 20874
45. E. Cardis, International Agency for Research on Cancer, 150 Cours Albert Thomas, 69372 Lyon Cedex 08, France
46. Harvey Checkoway, Department of Environmental Health, University of Washington, SC34, Seattle, WA 98195
47–51. S. P. Colyer, Medical Division, Oak Ridge Institute for Science and Education, P. O. Box 117, Oak Ridge, TN 37831
52. D. L. Cragle, Medical Division, Oak Ridge Institute for Science and Education, P. O. Box 117, Oak Ridge, TN 37831
53. K. W. Crase, Westinghouse Savannah River Co., Savannah River Site, Bldg. 735A, Aiken SC 29802
54. D. Crawford-Brown, Department of Environmental Science, University of North Carolina, Chapel Hill, NC 27599-7400

55. Donald J. Dudziak, EPMD Advisory Commitee, Department of Nuclear Engineering, 110B Burlington Engineering Labs, North Carolina State University, Raleigh, NC 27695-7909
56. Mark Durst, Lawrence Berkeley Laboratory, 1 Cyclotron Road, Building 50, Room 3238, Berkeley, CA 94720
57. Merril Eisenbud, 711 Bayberry Drive, Chapel Hill, NC 27514
58. J. Estéve, International Agency for Research on Cancer, 150 Cours Albert Thomas, 69372 Lyon Cedex 08, France
59. Roger B. Falk, EG&G, Rocky Flats, P. O. Box 464, Golden, CO 80401
60. D. Flanders, Department of Epidemiology, School of Public Health, Emory University, Atlanta, GA 30322
61. Jerome Friedman, Department of Statistics, Sequoia Hall, Stanford University, Stanford, CA 94305
62. R. J. M. Fry, Radiation Research, Biology Division, ORNL, P.O.Box 2009, Oak Ridge, TN 37831-8077
63. S. A. Fry, Medical Division, Oak Ridge Institute for Science and Education, P. O. Box 117, Oak Ridge, TN 37831
64. Warren Galke, Occupational Medicine Group, Los Alamos National Laboratory HSE-2, MS: K404, Los Alamos, New Mexico 87545
65. G. R. Gebus, Occupational Medicine, EH-43, U. S. Department of Energy, 19901 Germantown Road, Germantown, MD 20874
66. E. S. Gilbert, Battelle Pacific Northwest Laboratories, Battelle Boulevard, P.O. Box 999, Richland, WA 99352
67. M. Ginevan, Step 5 Corporation, 1101 17th Street NW, Suite 501, Washington, DC 20036
68. Ronald Harbek, National Jewish Center for Immunology and Respiratory Medicine, 1400 Jackson Street, Denver, CO 80206
69. R. W. Hornung, National Institute for Occupational Safety and Health, 4676 Columbia Parkway, Cincinnati, OH 45226-1998
70. G. R. Howe, NCIC Epidemiology Unit, University of Toronto, McMurrich Building, 12 Queens Park Crescent W, Toronto, Ontario, Canada M5S 1A8
71. G. W. Kneale, Department of Public Health and Epidemiology, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom
72. Richard Kradin, Massachusetts General Hospital, Pulminary Immunology and Molecular Biology, Boston, MA 02114
73. Kathleen Kreiss, OccupationalEnvironmental Medicine Division, National Jewish Center for Immunology and Respiratory Medicine, 1400 Jackson Street, Denver, CO 80206
74. M. Kutner, Cleveland Clinic Foundation, One Clinic Center, 9500 Euclid Avenue, Cleveland, OH 44195
75. James E. Leiss, Route #2, Box 142C, Broadway, VA 22815

76–80. L. G. Littlefield, Medical Division, Oak Ridge Institute for Science and Education, P. O. Box 117, Oak Ridge, TN 37831

81. A. C. Lucas, Victoreen, Inc., 6000 Cochran Road, Cleveland, OH 44139-3395

82. B. H. Margolin, Department of Biostatistics, University of North Carolina, School of Public Health, Chapel Hill, NC 27599-7400

83. Thomas Markham, Brush Wellman Inc., 1200 Hanna Building, Cleveland, OH 44115

84. G. Matanoski, Johns Hopkins University, 624 North Broadway, Baltimore, MD 21205

85. C. B. Meinhold, Radiological Sciences Division, Bldg. 703 M, Brookhaven National Laboratory, Long Island, NY 11937

86. Ted Meinhardt, Deputy Directory, Division of Surveillance, Hazards Evaluation & Field Studies, R-12, National Institute of Occupational Health, 4676 Columbia Parkway, Cincinnati, OH 45226

87. Frederick Miller, Department of Pathlogy, SUNY Stony Brook, T-140, HSC, Stony Brook, New York 11794-8691

88. Neville Moray, EPMD Advisory Committee, Department of Mechanical and Industrial Engineering, University of Illinois, 1206 West Green Street, Urbana, IL 61801

89. Margaret M. Mroz, OccupationalEnvironmental Medicine Division, National Jewish Center for Immunology and Respiratory Medicine, 1400 Jackson Street, Denver, CO 80206

90. C. R. Muirhead, National Radiological Protection Board, Chilton, Didcot, Oxon OX11 0RQ, United Kingdom

91. William E. Murray, National Institute for Occupational Safety and Health 4676 Columbia Parkway Cincinnati, Ohio 45226

92. Lee S. Newman, OccupationalEnvironmental Medicine Division, National Jewish Center for Immunology and Respiratory Medicine, 1400 Jackson Street, Denver, CO 80206

93. Emmanuel Ojo-Amaize, Speciality Laboratories, Inc., 2211 Michigan Avenue, Santa Monico, CA 90404

94. L. Christine Oliver, Massichusetts General Hospital, Occupational Medicine, 50 Staniford St., Sixth Floor, Boston, MA 02114

95. D. A. Pierce, Department of Statistics, Radiation Effects Research Foundation, 5-2 Hijiyama Park, Minami-ku, Hiroshima, 732 Japan

96. D. L. Preston, Department of Statistics, Radiation Effects Research Foundation, 5-2 Hijiyama Park, Minami-ku, Hiroshima, 732 Japan

97. Carl Quong, Lawrence Berkeley Laboratory, 1 Cyclotron Road, Building 50, Berkeley, California 94720

98. Robert Rinsky, National Institute for Occupational Safety and Health, 4676 Columbia Parkway, Cincinnati, Ohio 45226

99. Milton D. Rossman, Hospital of the University of Pennsylvania, 812 E. Gates Building, 3400 Spruce Street, Philadelphia, PA 19104
100. Jerome Sacks, National Institute of Statistical Sciences, P.O. Box 14162, Research Triangle Park, NC 27709-4162
101. Scott Schneider, Occupational Health Foundation, 1126 16th Street NW, Washington, DC 20036
102. W. J. Schull, Epidemiological Research Center, School of Public Health, University of Texas Health Science Center, Houston, TX 77225
103. P. Seligman, Environmental Safety and Health, U. S. Department of Energy, 19901 Germantown Road, Germantown, MD 20874
104. R. E. Shore, Institute of Environmental Medicine, New York University, 341 East 25th Street, 2nd Floor, New York, NY 10010-2598
105. W. K. Sinclair, National Council on Radiation Protection, 7910 Woodmont Avenue, Suite 800, Bethesda, MD 20814
106. J. Smith, Radiation Studies Branch, Center for Disease Control, 1600 Clifton Road, NE, Mailstop F-28, Atlanta, GA 30333
107. Mathew H. Smith, Klemm Analysis Group, Inc., 2971 Flowers Road, South, Oglethorpe Building, Suite 120, Atlanta, GA 30341
108. Daniel L. Solomon, Department of Statistics, North Carolina State University, P.O. Box 5457, Raleigh, NC 27650
109. A. M. Stewart, Department of Public Health and Epidemiology, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom
110. H. Stockwell, Epidemiology and Health Surveillance, EH-42 CC 270, U. S. Department of Energy, 19901 Germantown Road, Germantown, MD 20874
111. C. H. Strader, Epidemiology and Health Surveillance, EH-42 CC 270, U. S. Department of Energy, 19901 Germantown Road, Germantown, MD 20874
112. W. G. Tankersley, Oak Ridge Institute of Science and Education, Center for Epidemiological Research, P.O. Box 117, Oak Ridge, TN 37831-0117
113. D. Thomas, Department of Preventive Medicine, University of South California, 1420 San Pablo Street, PMB B201, Los Angeles, CA 90033-9987
114. Ray A. Waller, S-1, Statistics, Los Alamos National Laboratory, P.O. Box 1663, Los Alamos, NM 87545
115. Paul F. Wambach, Occupational Medicine, EH-43, U. S. Department of Energy, 19901 Germantown Road, Germantown, MD 20874
116. J. Watkins, Medical Division, Oak Ridge Institute for Science and Education, P. O. Box 117, Oak Ridge, TN 37831
117. J. E. Watson, School of Public Health, University of North Carolina, Chapel Hill, NC 27599-7400
118. Mary F. Wheeler, EPMD Advisory Committee, Rice University, Department of Mathematics, P. O. Box 1892, Houston, TX 77251
119. S. Wing, Department of Epidemiology, School of Public Health, University of North Carolina, Chapel Hill, NC 27599-7400

120–121. Office of Scientific and Technical Information, P.O. Box 62, Oak Ridge, TN 37830