0022-1767/81/1264-1614$02.00/0
THE JOURNAL OF IMMUNOLOGY

Vol. 126, No. 4, April 1981
Printed in U.S.A.

LIMITING DILUTION ASSAYS FOR THE DETERMINATION OF IMMUNOCOMPETENT CELL FREQUENCIES

I. Data Analysis

CARL TASWELL[1]

From the Unit of Human Cancer Immunology, Lausanne Branch, Ludwig Institute for Cancer Research, CH-1066 Epalinges, Switzerland

A statistical method was developed for the analysis of experimental data from limiting dilution assays. Formulas for the estimation of the frequency of immunocompetent cells within a test population were derived by the statistical methods of weighted averaging, likelihood maximization, and χ^2 minimization. Equations for the latter 2 were solved by Newton's method of iterative approximation. Estimates obtained by these methods were found to be more valid than those obtained by least squares (LS) fitting as judged by the χ^2 test and as established by Monte Carlo experiments. χ^2 minimization was chosen as the preferable estimation method with maximum accuracy and precision (minimum bias and variance) for the standard determination of frequencies; likelihood maximization was used only for the confirmation of results. When data from previously published experiments were reanalyzed, both results and conclusions were found to differ significantly from those originally obtained by LS fitting, thus demonstrating the importance of using proper data analysis methods. In conjunction with the use of available calculators or microcomputers, the method presented here provides a simple and rapid procedure for the valid determination of immunocompetent cell frequencies.

Limiting dilution assays (LDA)[2] incorporating sensitive *in vitro* microculture techniques can be used to determine the frequencies of immunocompetent cells (IC) possessing a wide variety of functionally defined activities (1–7). LDA are quantal dose-response bioassays that detect an all-or-nothing (positive or negative) immune response in each individual culture within groups of replicate cultures that vary in the dose of cells tested. Dose-response data analysis is then used to determine the frequency of IC within the test cell population. For this analysis, it is necessary to assume 1 of 4 possible models for the generation of a positive response: 1) only 1 cell of only 1 cell type is necessary for a positive response (single-hit), 2) a given number (≥2) of cells of only 1 cell type is necessary (multi-hit), 3) only 1 cell of each of a given number (≥2) of cell types (multi-target), and 4) a combination of the multi-hit and multi-target models. The theory of limiting dilution analysis including these models and the experimental conditions necessary for their validity were discussed extensively by Miller *et al.* (8) and by Lefkovits and Waldmann (9). This article presents data analysis methods for the single-hit model only. This model assumes that the IC are diluted to limiting doses, that every single IC generates a detectable response, and that all other cell types and culture factors remain at saturating (nonlimiting) doses.

If all dilutions of the test cell population are homogeneous suspensions, then the Poisson probability distribution (10) describes the occurrence of the number of test cells and consequently the number of IC sampled and aliquoted into each replicate culture. These numbers vary between individual replicate cultures; however, on the average, they equal the intended dose of test cells and IC. For the single-hit model, a positive immune response, which is quantal and not quantitative, cannot be used to distinguish between the presence of 1 or more IC. However, in the absence of an IC, no immune response can occur. Therefore, in order to apply Poisson probability theory to the single-hit model, the zero term of the Poisson equation must be used to describe the relation between the average number of cells tested per replicate culture and the number of negatively responding cultures per group (see *Appendix*). This single-hit Poisson model (8, 9, 11) provides the theoretical basis for a wide variety of data reduction methods for the estimation of the frequency of IC from LDA dose-response data. These estimation methods range from the use of simple hand-drawn plots with eye-estimated frequencies to the use of more complex statistical estimators with computer-calculated frequencies. These estimators include the following: least squares (LS), weighted mean (WM), maximum likelihood (ML), and minimum χ^2 (MC). However, these methods vary not only in complexity but also in accuracy and precision.

Although data from some of the earliest LDA for the determination of IC frequencies (12, 13) were analyzed by a valid method (11, 14), the majority of data from experiments published since then has been analyzed by simpler but less valid methods. The advantages of increased accuracy and precision of results were not considered to be adequate compensation for the disadvantages of increased time and effort to obtain the results. Some immunologists thus resorted to standard LS fitting as a compromise method despite its invalidity for the analysis of LDA data (15). Yet the more widespread the use of limiting dilution analysis becomes, the more necessary valid data reduction methods become. It would therefore be desirable to develop an accurate and precise statistical method for the analysis of LDA data that would be usable by all investigators. In this article, I present such a method, demonstrate its merit, and provide formulas that can be readily used on any calculator or microcomputer.

MATERIALS AND METHODS

Terminology and notation. Frequency determinations are described here by terms and symbols adapted from Finney (15) and Halsall and Makinodan (16).

Given assay design parameters: a) Let l be the number of groups of replicate cultures and each group be labeled dose i for $i = 1, 2, 3, \ldots l$. b) Let n_i be the number of replicate cultures of each dose i and N be the total number of cultures ($N = \sum n_i$). c) Let x_i be the number of cells tested in each replicate culture of each dose i. Actually, x_i is the mean of the distribution of the Poisson variable x_{ij} for constant i and variable j where x_{ij} is the true number of cells tested in the j^{th} culture of the i^{th} dose (see *Appendix*). d) Let $P_i = e^{-\phi x_i}$ be the probability of a negatively responding culture for each dose i (see *Appendix*).

Received for publication June 4, 1980.
Accepted for publication November 26, 1980.
The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] Present address: Medical Scientist Training Program, New York University School of Medicine, 550 First Avenue, New York, NY 10016.

[2] Abbreviations used in this paper: LDA, limiting dilution assay(s); IC, immunocompetent cell(s); f, frequency estimate; LS, least squares; WM, weighted mean; ML, maximum likelihood; MC, minimum χ^2; V, variance; CI, confidence interval; df, degrees of freedom; CTL-P, cytolytic T lymphocyte precursor; all other mathematical and statistical notation as explained in the text.

Observed data: Let r_i be the number of negatively responding cultures of each dose i.

Calculated results: a) Let $p_i = r_i/n_i$ be the fraction of negatively responding cultures of each dose i; p_i is an estimate of P_i. b) Let f be the estimate of the true frequency ϕ of the defined subpopulation of IC within the test cell population.

Note that the term *dose* may refer either to the number I of doses or to the value x_i of an individual dose. $\sum$ and $\prod$ refer, respectively, to summations and products over i (i.e., for $i = 1$ to $i = I$) unless specified otherwise. The meaning of all other symbols should be clear from the text and/or standard to basic mathematics and statistics. Finally, the term *estimator* refers to an estimation method, whereas the term *estimate* refers to an estimated frequency.

Determination of ϕ by LS estimation. A formula for the estimation of ϕ by the method of LS fitting was derived by simple algebraic substitution in the standard formula for unweighted LS fitting (17) of values for the equation of the line

$$y = bx + a \tag{1}$$

with the corresponding values for the equation of the line

$$\ln p_i = -\phi x_i + a \tag{2}$$

derived from a logarithmic transformation of the zero term of the Poisson equation (see *Appendix*). Thus, the LS frequency estimate f_{LS} was determined as the value of ϕ that minimizes

$$\sum (\ln p_i + \phi x_i)^2 \tag{3}$$

and was calculated as

$$f_{LS} = \frac{(\sum x_i)(\sum \ln p_i) - I \sum (x_i \ln p_i)}{I \sum (x_i^2) - (\sum x_i)^2}. \tag{4}$$

Since doses with $r_i = 0$ result in the indeterminate form ln 0, they were excluded from the calculations.

Determination of ϕ by WM estimation. A formula for the estimation of ϕ by the method of weighted averaging was derived (see *Appendix*) by constructing a weighted arithmetic mean

$$\frac{\sum f_i w_i}{\sum w_i} \tag{5}$$

of the estimates f_i of ϕ obtained independently from each individual dose i. The weight w_i for each estimate was chosen as the reciprocal of its variance divided by the sum of the weights (18). Thus, the WM frequency estimate f_{WM} was calculated as

$$f_{WM} = \frac{\sum \left(\frac{-\ln p_i}{x_i}\right)\left(\frac{x_i^2 r_i}{1 - p_i}\right)}{\sum \left(\frac{x_i^2 r_i}{1 - p_i}\right)} \tag{6}$$

and its variance V as

$$V(f_{WM}) = \frac{1}{\sum \left(\frac{x_i^2 r_i}{1 - p_i}\right)}. \tag{7}$$

Since doses with $r_i = 0$ result in the indeterminate form ln 0, they were excluded from the calculations.

Determination of ϕ by ML estimation. The derivation (see *Appendix*) of a formula for the estimation of ϕ by the method of likelihood maximization was adapted from Finney's analysis of dilution series in quantal response assays (11). The ML frequency estimate f_{ML} was determined as the value of ϕ that maximizes

$$\ln L = \sum \left[\ln \binom{n_i}{r_i} - r_i \phi x_i + (n_i - r_i) \ln (1 - e^{-\phi x_i}) \right] \tag{8}$$

where ln L is the natural logarithm of the likelihood function L (see *Appendix*). The f_{ML} was calculated by Newton's method of iterative approximation (19) as

$$f_{j+1} = f_j - \left. \frac{(\partial \ln L/\partial \phi)}{(\partial^2 \ln L/\partial \phi^2)} \right|_{f_j} \tag{9}$$

where f_j is the j^{th} iterative f_{ML} estimate of ϕ, and

$$\left. \frac{\partial \ln L}{\partial \phi} \right|_{f_j} = \sum \left[-r_i x_i + \frac{(n_i - r_i) x_i e^{-\phi x_i}}{(1 - e^{-\phi x_i})} \right] \Bigg|_{f_j} \tag{10}$$

and

$$\left. \frac{\partial^2 \ln L}{\partial \phi^2} \right|_{f_j} = \sum \left[\frac{-(n_i - r_i) x_i^2 e^{-\phi x_i}}{(1 - e^{-\phi x_i})^2} \right] \Bigg|_{f_j} \tag{11}$$

are respectively the 1st and 2nd partial derivatives of ln L with respect to ϕ evaluated at f_j. $V(f_{ML})$ was calculated as the negative reciprocal of the 2nd derivative of ln L

$$V(f_{ML}) = \left. \frac{-1}{(\partial^2 \ln L/\partial \phi^2)} \right|_{f_{ML}} \tag{12}$$

evaluated at f_{ML} (20). Note that $V(f_{ML})$ is the reciprocal of a sum and not the sum of reciprocals. All experimental data were included in the calculations.

Calculation of χ^2 and determination of ϕ by MC estimation. A formula for the calculation of the χ^2 statistic and for the estimation of ϕ by the method of χ^2 minimization was derived (see *Appendix*) by simple algebraic substitution of observed and expected class frequencies in the standard formula for the Pearson χ^2 (21). Thus, the MC frequency estimate f_{MC} was determined as the value of ϕ that minimizes

$$\chi^2_{[I-1]} = \sum \left[\frac{(r_i - n_i e^{-\phi x_i})^2}{n_i e^{-\phi x_i}(1 - e^{-\phi x_i})} \right] \tag{13}$$

where the subscript $I-1$ indicates the number of degrees of freedom (df). The f_{MC} was calculated by Newton's method of iterative approximation as

$$f_{j+1} = f_j - \left. \frac{(\partial \chi^2/\partial \phi)}{(\partial^2 \chi^2/\partial \phi^2)} \right|_{f_j} \tag{14}$$

where f_j is the j^{th} iterative f_{MC} estimate of ϕ, and

$$\left. \frac{\partial \chi^2}{\partial \phi} \right|_{f_j} = \sum \left[\frac{n_i x_i e^{-\phi x_i}(2r_i - n_i) + r_i^2 x_i (e^{\phi x_i} - 2)}{n_i (1 - e^{-\phi x_i})^2} \right] \Bigg|_{f_j} \tag{15}$$

and

$$\left. \frac{\partial^2 \chi^2}{\partial \phi^2} \right|_{f_j} = \sum \left[\frac{n_i^2 x_i^2 (e^{-\phi x_i} - e^{-3\phi x_i}) + n_i r_i x_i^2 (-2e^{-\phi x_i} + 2e^{-3\phi x_i})}{n_i (1 - e^{-\phi x_i})^4} \cdots \right.$$

$$\left. \cdots \frac{r_i^2 x_i^2 (e^{\phi x_i} - 4 + 7e^{-\phi x_i} - 4e^{-2\phi x_i})}{n_i (1 - e^{-\phi x_i})^4} \right] \Bigg|_{f_j} \tag{16}$$

are, respectively, the 1st and 2nd partial derivatives of χ^2 with respect to ϕ evaluated at f_j. $V(f_{MC})$ was calculated as 2 times the reciprocal of the 2nd derivative of χ^2

$$V(f_{MC}) = \left. \frac{2}{(\partial^2 \chi^2/\partial \phi^2)} \right|_{f_{MC}} \tag{17}$$

evaluated at f_{MC} (20). Note that $V(f_{MC})$ is the reciprocal of a sum and not the sum of reciprocals. All experimental data were included in the calculations.

Calculation of f, its confidence interval (CI) and probability value, and the CI for the ratio of 2 independent f. The noniterative estimates f_{LS} and f_{WM} were calculated explicitly by Equations 4 and 6, respectively. The iterative estimates f_{ML} and f_{MC} were approximated by Equations 9 and 14, respectively. Iteration was terminated when the incremental value was less than or equal to 0.005% of the estimate (4-digit accuracy). The f_{WM} was used as the starting value for iterative approximation to f_{ML}, and f_{ML} was then used as the starting value for f_{MC}. The 95% CI (17) for f was calculated as

$$95\%\ \mathrm{CI}(f) = f \pm 1.96\ \mathrm{SE}(f) \tag{18}$$

where the standard error (SE) of f is

$$\mathrm{SE}(f) = \sqrt{V(f)}. \tag{19}$$

Probability values corresponding to $\chi^2_{[df]}$ values as a measure of the goodness of fit of f were obtained by using a Monroe 1930 calculator with a built-in function for that purpose. The 95% CI for the ratio m of 2 independent frequency estimates f_a and f_b where

$$m = \frac{f_a}{f_b} \tag{20}$$

was calculated according to Fieller's theorem assuming the normal approximation (18) as

$$95\%\ \mathrm{CI}(m) = \frac{m}{h} \pm \frac{1.96}{h f_b} \sqrt{h V(f_a) + m^2 V(f_b)} \tag{21}$$

where

$$h = \frac{f_b^2 - 1.96^2 V(f_b)}{f_b^2}. \tag{22}$$

All calculations (except those for probability values) were performed with a Hewlett-Packard 9845A microcomputer using a program written in BASIC. Calculation and printout of all results for each frequency determination were completed in seconds.

Use of previously published experimental data. All 4 estimation methods (LS, WM, ML, and MC) were performed on experimental data previously published by Taswell *et al.* (22). Some LS estimates reported here differ from those published originally (22). These discrepancies are due to rounding errors. The original LS estimates were first calculated and then recorded as whole number reciprocals. They were later reconverted to fractions for publication in Reference 22. Since the LS estimates presented here were recalculated directly from the data, they are the correct LS estimates. For these LDA, specific cytolysis was the activity assayed with a positive response of a culture defined as any ^{51}Cr release value greater than 3 standard deviations (SD) above the mean background level. Thus, cytolytic T lymphocyte precursors (CTL-P) were the functionally defined IC for which frequencies were determined. The minimum positive level of 3 SD above the mean has been arbitrarily accepted as a standard by most investigators working with LDA for CTL-P. Furthermore, it has been shown that generally the distribution of ^{51}Cr release values for positively responding cultures within each group of replicate cultures is independent of the test cell dose. A description of culture methods and a review of experiments with LDA for CTL-P has been recently published by MacDonald *et al.* (23).

Monte Carlo experiments. In order to test empirically the relative efficiency of the 4 different estimators, a series of Monte Carlo experiments (24) were performed in a manner analogous to that of Berkson (25, 26). In each experiment, 1000 sets of values for the numbers r_i of negatively responding cultures were generated randomly in a manner that directly simulated an LDA of a given assay design with known doses x_i and numbers n_i of replicate cultures for a given test cell population with known IC frequency ϕ. For each of the 1000 sets of data, f_{LS}, f_{WM}, f_{ML}, and f_{MC} were calculated; and subsequently for each of the 4 sets of 1000 estimates, the means, variances, and mean square errors were calculated as a measure of the bias and variance of each of the 4 estimators. These experiments were performed on the Hewlett-Packard microcomputer. Complete details of the methods will be published separately (C. Taswell and W. F. Taylor, manuscript in preparation).

RESULTS

Summary of results. The validity of LS fitting, weighted averaging, likelihood maximization, and χ^2 minimization as methods for the estimation of the IC frequency ϕ was investigated for the single-hit Poisson model (see *Introduction* and *Appendix*) by several different procedures. The results can be summarized as follows: 1) When LS, WM, ML, and MC estimates (see *Materials and Methods*) from experimental data were subjected to χ^2 testing with probability values as a measure of the goodness of fit, MC, ML, and WM estimates were found to be better than LS estimates (see Tables II, III, and IV). 2) When LS, WM, ML, and MC estimates from artificial data generated by Monte Carlo methods were subjected to Monte Carlo experimentation with the mean square error as a measure of the efficiency (see *Materials and Methods*), the MC estimator was found to be the best (minimum bias and variance, maximum accuracy and precision), whereas the ML estimator was 3 to 15% worse, the WM estimator 20 to 40%, and the LS estimator 300 to 1800% worse than the MC estimator (see below). 3) When various estimators for the combination of data from multiple determinations were subjected to χ^2 testing and Monte Carlo experimentation, the MC estimator for data pooled from the multiple determinations was found to be the best (see Table IV and below). 4) Different estimators can produce results leading to different conclusions from the same data (see Table III).

Comparison of LS, WM, ML, and MC estimators by χ^2 testing. The quality of LS, WM, ML and MC estimators was investigated by performing χ^2 tests on estimates calculated from previously published data (22). Table I presents the experimental data (adapted from Fig. 1 of Reference 22) for determinations of CTL-P frequencies in normal murine thymus and spleen cell populations (C57BL/6 responding cells cultured with DBA/2 stimulating cells and assayed against P815 target cells). Table II then presents the results for each of the 4 estimators as the estimate f of the frequency ϕ, the standard error (SE) of f, and the χ^2 and p values from the goodness of fit test for f. For both thymus and spleen frequency determinations, the χ^2 and probability values were lower and higher, respectively, for the WM, ML, and MC estimates than those for the LS estimates. These lower χ^2 and higher probability values indicate that the WM, ML, and MC estimates provided better agreement between the observed dose-response results and those expected assuming that the single-hit Poisson model applies to these LDA. In the case of the spleen frequency determination, the decrease in the χ^2 value was large enough to increase the probability value from 0.07 for the LS estimate (barely acceptable) to 0.56 to 0.57 for the WM, ML, and MC estimates (clearly acceptable). Although no significant differences were observed between WM, ML, and MC estimators for either f or SE or χ^2 in the 2 examples presented here, differences as large as 10 to 50% were observed in some cases (unpublished results).

TABLE I

Experimental data from limiting dilution assays for the determination of murine CTL-P frequencies[a]

Cell Population	Dose (i)	Cultures (n_i)	Cells per Culture (x_i)	Negative Cultures (r_i)	Fraction Negative Cultures (p_i)
Thymus	1	33	100	32	0.970
	2	33	500	28	0.848
	3	33	1000	24	0.727
	4	33	1500	21	0.636
	5	33	2000	14	0.424
Spleen	1	24	250	17	0.708
	2	24	500	13	0.542
	3	24	750	6	0.250

[a] Adapted from Figure 1 of Reference 22. C57BL/6 responding cells were incubated with DBA/2 stimulating cells, assayed for specific cytolysis against P815 target cells, and considered to be positive for CTL-P if the ^{51}Cr release value was greater than 3 SD above the mean background level.

TABLE II

Comparison of LS, WM, ML, and MC estimates of murine CTL-P frequencies[a]

Cell Population	Estimator	$f \times 10^{-3}$	$SE(f) \times 10^{-3}$	$\chi^2(f)$ [df]	$P(\chi^2)$
Thymus	LS	0.4060		$2.048_{[4]}$	0.727
	WM	0.3423	0.0519	$1.111_{[4]}$	0.893
	ML	0.3506	0.0523	$1.082_{[4]}$	0.897
	MC	0.3511	0.0529	$1.082_{[4]}$	0.897
Spleen	LS	2.083		$5.278_{[2]}$	0.071
	WM	1.447	0.256	$1.152_{[2]}$	0.562
	ML	1.501	0.259	$1.122_{[2]}$	0.571
	MC	1.492	0.262	$1.121_{[2]}$	0.571

[a] Data from Table I were analyzed by the LS, WM, ML, and MC estimators as explained in *Materials and Methods*.

Comparison of LS, WM, ML, and MC estimators by Monte Carlo experimentation. The relative efficiency of the LS, WM, ML, and MC estimators was investigated by performing a series of Monte Carlo experiments for a wide range of values of ϕ ($\phi = 1$ to $\phi = 1 \times 10^{-6}$) and a wide variety of symmetric and asymmetric assay designs with $I = 2$ to $I = 5$ and $N = 48$ to $N = 192$. (Symmetric and asymmetric assay designs are designs in which the doses x_i are spaced evenly or unevenly, respectively, about a dose x_i corresponding to $P_i = 0.5$.) The efficiency of each estimator was observed to be independent of the value of ϕ but dependent upon the values n_i and x_i of the assay design. χ^2 minimization was clearly found to be the minimum-bias minimum-variance estimator of ϕ as measured by the mean square error of the estimates. Depending upon the assay design, likelihood maximization was found to have 3 to 15% greater bias and variance, weighted averaging, 20 to 40%, and LS fitting, 300 to 1800%. Complete details of these results will be published separately (C. Taswell and W. F. Taylor, manuscript in preparation). The MC estimator was therefore considered to be the best statistical method for accurately estimating ϕ. The MC estimates of Table II together with the experimental data of Table I are displayed in Figure 1 as a plot of the logarithm of the fraction p_i of negatively responding cultures versus the number x_i of cells tested. The results for each frequency determination are displayed as 3 lines where the central one represents the estimate f and the other 2 represent the 95% confidence limits of f. The MC estimator was used to determine all other frequencies reported in this article. All MC estimates allowed for satisfactory acceptance of the single-hit Poisson model. The hypothesis was not rejected at the 5% level of significance, i.e., $p > 0.05$ and in most cases $p \gg 0.05$ unless indicated otherwise.

Valid vs invalid conclusions. Proper data analysis may be particularly important for those experiments in which frequencies are determined only once for each set of conditions investigated. In order to examine this possibility, data from experiments on the effect of cyclophosphamide treatment on the CTL-P frequency in

the thymus and spleen were reanalyzed by the MC estimator and the results were compared with those originally obtained by the LS estimator. Table III presents the frequency estimates and cell recoveries (adapted from Table III of Reference 22). According to the LS estimates, cyclophosphamide significantly altered the CTL-P frequency only in the spleen when it was used at a dose of 300 mg/kg, whereas according to the MC estimates, cyclophosphamide significantly altered the CTL-P frequency in both the spleen and thymus when used at that dose. Reanalysis of the data thus produced results different from those obtained originally. This difference was all the more striking because it could not have been predicted merely by comparing the original LS estimates. For example, according to the LS estimates, cyclophosphamide at both 100 and 300 mg/kg reduced the CTL-P frequency in the thymus by approximately 25%, whereas according to the MC estimates, cyclophosphamide at 100 mg/kg decreased the frequency by 3% and at 300 mg/kg by 58%.The consideration that some of the LS estimates are clearly unacceptable with $p < 0.05$ whereas the MC estimates are acceptable with $P > 0.10$ further emphasizes the difference. Since the MC estimator is the more valid statistical method (see above), conclusions derived from MC estimates provide a better interpretation of experiments than do conclusions derived from LS estimates. This statement applies especially to experiments performed only once as opposed to experiments performed repeatedly (see below).

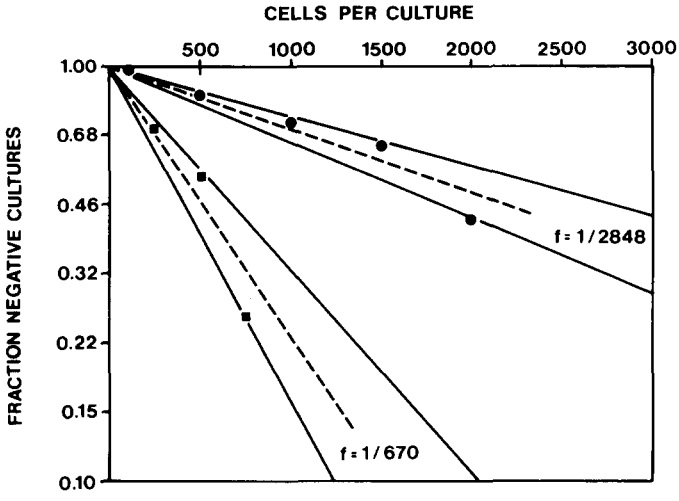

Figure 1. MC estimates of murine CTL-P frequencies. MC frequency estimates for the data of Table I were determined as explained in *Materials and Methods*. For each frequency determination, the lower line represents the upper 95% confidence limit of f_{MC}, the middle line represents the estimate f_{MC}, and the upper line represents the lower 95% confidence limit of f_{MC}. ●, Thymus; ■, spleen.

Combination of frequency estimates from multiple determinations. Although several statistical methods are available for the combination of frequency estimates from multiple determinations (15), the simple procedure recommended by Porter and Berry (14) was considered to be the best method for the LDA discussed here. Thus, combined frequency estimates can be obtained by analyzing with the MC estimator all data pooled from all determinations to be combined. Data from invalid frequency determinations (those with $p < 0.05$, see also *Discussion*) should not be included in the pooling of data. To test this method, data from experiments on the effect of cortisone treatment on the CTL-P frequency in the thymus and the spleen were pooled and analyzed by the MC estimator, and the results were compared with those originally obtained as the means of the LS estimates. Table IV presents the frequency estimates and cell recoveries (adapted from Table II of Reference 22). Despite absolute differences between the MC estimates for the pooled data and the means of the LS estimates, the relative results and conclusions derived from them were similar for both methods of statistical estimation. MC estimates for the treated populations were found to be consistent when combined ($p > 0.05$), whereas those for the untreated populations were inconsistent when combined ($p < 0.05$). Since the single-hit Poisson model was accepted as being valid for each of the MC estimates

TABLE III

Comparison of results obtained by LS and MC methods of data analysis for single experiments on the effect of cyclophosphamide treatment of mice[a]

Cyclophosphamide Dose	Cell Population	Method	Untreated			Treated			Recovery as % Untreated		
			CTL-P f[b]	95% CI(f)[b]	P(f)[c]	CTL-P f	95% CI(f)	P(f)	Total cells	CTL-P[d]	95% CI[d]
100 mg/kg	Thymus	LS	0.81		0.136	0.60		0.002	82	61	
		MC	1.04	0.74–1.33	0.477	1.01	0.75–1.27	0.346	82	80	54–120
	Spleen	LS	1.39		0.002	1.87		0.886	69	93	
		MC	2.15	1.54–2.77	0.120	1.93	1.29–2.58	0.893	69	62	38–96
300 mg/kg	Thymus	LS	0.84		0.360	0.63		0.369	20	15	
		MC	1.01	0.66–1.35	0.606	0.42	0.23–0.62	0.845	20	8	4–15
	Spleen	LS	2.08		0.071	0.16		0.345	10	1	
		MC	1.49	0.98–2.01	0.571	0.24	0.14–0.34	0.748	10	2	1–3

[a] Adapted from Table III of Reference 22. Each mouse received a single i.p. injection of cyclophosphamide and was sacrificed 2 days later. Pools of 3 organs were assayed for each determination as described in Table I. A single experiment was performed separately at each dose indicated. The data were then analyzed by the LS and MC estimators as explained in *Materials and Methods*.

[b] Expressed as the number of CTL-P per 10^3 cells.

[c] The p value corresponding to the χ^2 value of f.

[d] CTL-P recovery was calculated as the total cell recovery times the ratio of the CTL-P frequencies of the treated to untreated populations.

TABLE IV

Comparison of results obtained by LS and MC methods of data analysis for multiple experiments on the effect of cortisone treatment of mice[a]

Cell Population	Method	Untreated			Treated			Recovery as % Untreated		
		CTL-P f[b]	95% CI(f)[b]	P(f)[c]	CTL-P f	95% CI(f)	P(f)	Total cells	CTL-P[d]	95% CI[d]
Thymus	Mean of LS[e]	0.47		0.042	10.4		0.020	4	89	
	MC of pooled data[f]	0.47	0.38–0.56	0.042	7.10	5.38–8.82	0.377	4	60	43–82
Spleen	Mean of LS	2.43		0.000	2.59		0.211	37	39	
	MC of pooled data	1.56	1.13–2.00	0.038	1.90	1.36–2.45	0.688	37	45	30–68

[a] Adapted from Table II of Reference 22. Cortisone treatment experiments were performed analogously to the cyclophosphamide treatment experiments described in Table III. Multiple experiments (3 for thymus and 2 for spleen) were performed separately all at a cortisone dose of 100 mg/kg. Frequency estimates for each experiment were then combined as explained below.

[b] Expressed as the number of CTL-P per 10^3 cells.

[c] The p value corresponding to the χ^2 value of f.

[d] CTL-P recovery was calculated as the total cell recovery times the ratio of the CTL-P frequencies of the treated to untreated populations.

[e] LS estimates for each frequency determination were averaged.

[f] Data from each frequency determination to be combined were pooled and analyzed by the MC estimator as explained in *Results*.

from the frequency determinations that were combined, these inconsistencies were interpreted to represent heterogeneity between the estimates combined rather than invalidity of the single-hit Poisson model for the final pooled estimate. Monte Carlo experiments (see above) were performed to compare the efficiency of 8 different estimators: 2 versions of each of the 4 estimators, LS, WM, ML, and MC. The 1st version of each of the 4 estimators was the mean of the estimates from the determinations combined, and the 2nd version was the estimate for the pooled data from the determinations combined. Of the 8 estimators, the one presented in Table IV as the MC pooled data method was found to be the most valid (C. Taswell and W. F. Taylor, manuscript in preparation).

DISCUSSION

I compared the validity of 4 statistical methods for the estimation of IC frequencies by χ^2 testing (see Table II) in a manner analogous to that of Fisher (27) and by Monte Carlo experimentation (see *Results*) in a manner analogous to that of Berkson (25, 26). I found that the 4 estimators can be ranked in the following order by decreasing bias and variance (increasing accuracy and precision): LS, WM, ML, and MC. This empirical result was expected by statistical theory. LS (see Equation 4) minimizes the sum of the squares of the differences between the logarithms of observed p_i and expected P_i values (see Equation 3) from a transformation (see Equation 2) of the single-hit Poisson-model equation (see *Appendix*). It is a function of only part of the experimental data: x_i and p_i from all doses i except those with $r_i = 0$. It should be noted that the use of p_i by itself without either r_i or n_i represents a loss of information. For example, a value of $p_i = r_i/n_i = 0.8$ can be obtained from an LDA with either $r_i/n_i = 16/20$ or $r_i/n_i = 80/100$ or any other ratio of numbers equal to 0.8. Furthermore, the LS function falsely assumes that the variance $V(p_i)$ is constant for all values of p_i. WM (see Equation 6) resolves some of these problems by weighting individual dose estimates (see Equation 29) with terms related to their respective variances (see Equation 31) and then taking their mean. However, it remains another version of a logarithmic transformation (see Equation 28) of the single-hit Poisson-model equation and is likewise a function of only part (albeit more than LS) of the experimental data: x_i, p_i, and r_i (and therefore n_i implicitly) from all doses i except those with $r_i = 0$. ML (see Equations 9, 10, and 11) maximizes a quantity called the likelihood, which is defined to be proportional to the probability that the dose-response data observed should be observed if in fact the true frequency ϕ equals the frequency estimate f. It retains the single-hit Poisson-model equation in its exponential form (see Equations 26 and 27) and is a function of all of the experimental data: x_i, r_i, and n_i from all doses i. Like LS, MC (see Equations 14, 15, and 16) also minimizes a version of the sums of the squares of the differences between observed and expected values; but like ML, it is a version that retains the single-hit Poisson-model equation in its exponential form and is a function of all of the experimental data. The estimators thus incur a loss of information to varying degrees, both as predicted theoretically and as observed experimentally.

Since results (f, SE(f), and $\chi^2(f)$) obtained independently from both ML and MC estimators tend to equality as the number of cultures increases assuming the validity of the single-hit Poisson model (15, 20), those cases in which discrepancies occur may represent LDA to which the Poisson model does not apply and for which the data and results cannot be considered acceptable. Thus, 2 independent validity tests can be performed for each estimate: the standard χ^2 test and an ML/MC divergence test. Failure to pass either one or both tests may invalidate both the frequency estimate and the single-hit Poisson model. In our laboratory we have arbitrarily defined $p < 0.05$ and/or ML/MC divergence $>$ 10% as the critical region for rejection of the estimate and the model. This close convergence of ML and MC results continues to encourage debate between statisticians about the relative efficiencies of these 2 estimators. Berkson (28, 29) has provided convincing arguments for preference of the MC estimator. It is possible to conceive of ML as a function that maximizes gain of information and MC as a function that minimizes loss of information. If it is assumed that the best estimator from a practical point of view is the one with the least error or loss of information, then MC should, in theory, be the best estimator. When these intuitive thoughts are tested by Monte Carlo experimentation for the particular model in question, it is possible to discover which estimator does, in fact, incur the least error. As reported in this article, the MC estimator is the best estimation method with the least error for the determination of IC frequencies based upon the single-hit Poisson model (see *Results*).

Since MC is the most accurate and precise estimator, it should be used whenever possible. If programmable calculators or microcomputers with sufficient capacity to obtain f_{MC} by Newton's iterative method (see Equation 14) are not available, the MC estimator can still be used with a standard calculator by performing iterations manually (solving Equation 13 for successive values of f until $\chi^2(f)$ is minimized and f_{MC} is obtained). The ML estimator should then be used (with iterations performed either automatically or manually) either to confirm or as a substitute for the results obtained by the MC estimator. If practicality nevertheless necessitates use of a noniterative method, then the WM estimator should be used and the LS estimator avoided. (The WM estimator was found to be much better than even a weighted LS estimator [C. Taswell and W. F. Taylor, manuscript in preparation].) Finally, it should be noted that regardless of the efficiency of a statistical estimator, the accuracy of an estimate depends upon the accuracy with which the minimum positive detection level discriminates between positive and negative responses. If the level is defined to be lower than its true value, then frequency estimates will be higher than their true values and *vice versa*. However, even if this level is thought to be more or less inaccurate, the resultant inaccuracy of an estimate should not be compounded by use of an estimator that is known to be inaccurate.

Having established χ^2 minimization as the best estimation method for the determination of IC frequencies, I then reanalyzed LDA data from a variety of previously published experiments (see Tables III and IV) and found that not only did the quantitative results differ significantly but also the qualitative conclusions as well. The importance of using a proper statistical method for the analysis of LDA data cannot be overemphasized, as the examples presented here demonstrate. Although it is more difficult to standardize experimental conditions, it is relatively easy to standardize data analysis methods, including the adoption of conventions for the minimum positive detection level. The validity of this level should be tested and established for each IC investigated, whether a CTL-P used as an example here or any other B or T lymphocyte or cell of the immune system. Proper data analysis methods, if standardized, would then enable all investigators to compare and interpret results and conclusions with greater accuracy and precision.

Acknowledgments. I thank D. J. Finney (University of Edinburgh), C. B. Begg (Harvard University), and W. F. Taylor (Mayo Medical School) for helpful advice throughout the development of this work.

APPENDIX

Derivation of the single-hit Poisson-model equation. The probability P_{neg} that a culture responds negatively can be derived as follows. If a culture receives a fixed dose of x cells from a population such that the probability is ϕ that any 1 cell is an IC and responds positively, then $P_{neg}(\text{constant } x)$ is the probability that there are not any IC among x and

$$P_{neg}(\text{constant } x) = (1 - \phi)^x. \tag{23}$$

Since the dose of x cells is sampled from a homogeneously suspended population, x is not a known constant but an unknown Poisson variable. If x is a value from a Poisson distribution with mean X, then the average value of $P_{neg}(\text{variable } x)$ equals $P_{neg}(\text{constant } x)$ averaged over the Poisson distribution of x. Thus,

$$\begin{aligned} P_{neg}(\text{variable } x) &= \sum_{x=0}^{\infty} \frac{X^x e^{-X}}{x!}(1 - \phi)^x \\ &= e^{-X} \sum_{x=0}^{\infty} \frac{(X - X\phi)^x}{x!} \\ &= e^{-X} e^{X - X\phi} \\ P_{neg}(\text{variable } x) &= e^{-\phi X}. \end{aligned} \tag{24}$$

Alternatively, P_{neg} can be derived more simply as follows. If a culture receives on the average a dose of X cells, then the Poisson distribution gives the probability $P_c(X)$ that the culture receives any number c of IC as

$$P_c(X) = \frac{(\phi X)^c e^{-\phi X}}{c!}. \quad (25$$

$P_{neg}(X)$ is then the probability that the culture does not receive any IC and equals the zero term of the Poisson Equation 25. Thus,

$$P_{neg}(X) = P_0(X) = e^{-\phi X}. \quad (26$$

Equation 26 is the mathematical formulation of the single-hit Poisson model.

Derivation of the f_{WM} Equation 6. If X and $P_{neg}(X)$ of Equation 26 are defined to be x_i and P_i of dose i with a number of n_i of replicate cultures sufficiently large, then P_i is approximated by its experimentally observed value p_i equal to the fraction r_i/n_i of negatively responding cultures. Thus,

$$r_i/n_i = p_i \simeq P_i = e^{-\phi x_i} \quad (27$$

and a logarithmic transformation of Equation 27 is

$$\ln p_i \simeq -\phi x_i. \quad (28$$

The individual dose estimate f_i is then

$$f_i = \frac{-\ln p_i}{x_i} \simeq \phi. \quad (29$$

Assuming that r_i is a variable and that n_i and x_i are constants and applying the general rules for the variance of a function, the variance of f_i can be derived as

$$V(f_i) \simeq \frac{n_i P_i(1 - P_i)}{x_i^2 r_i^2}. \quad (30$$

When P_i is approximated by p_i, the variance of f_i becomes

$$V(f_i) \simeq \frac{(1 - p_i)}{x_i^2 r_i}. \quad (31$$

By definition of f_{WM} as

$$\frac{\sum f_i w_i}{\sum w_i} \quad (5$$

and substitution of f_i by its value from Equation 29 and of w_i by the reciprocal of the value of $V(f_i)$ from Equation 31, f_{WM} can be derived as

$$f_{WM} = \frac{\sum_{i=1}^{I} \left(\frac{-\ln p_i}{x_i}\right)\left(\frac{x_i^2 r_i}{1 - p_i}\right)}{\sum_{i=1}^{I} \left(\frac{x_i^2 r_i}{1 - p_i}\right)}. \quad (6$$

Derivation of the ln L Equation 8. If P_i from Equation 27 is the probability that a culture responds negatively, then the probability $P(r_i)$ that r_i out of n_i cultures respond negatively is

$$P(r_i) = \binom{n_i}{r_i} P_i^{r_i}(1 - P_i)^{n_i - r_i} \quad (32$$

by the binomial probability distribution. The likelihood L that r_i out of n_i cultures respond negatively in each of I doses is then

$$L = \prod_{i=1}^{I} P(r_i) \quad (33$$

by the multiplicative property of independent probabilities. Now, since a logarithmic transformation is a strictly 1-to-1 monotonic transformation, the value of ϕ that maximizes L equals the value that maximizes the natural logarithm of L. Thus, ln L can be derived as

$$\ln L = \sum_{i=1}^{I} \left[\ln\binom{n_i}{r_i} - r_i \phi x_i + (n_i - r_i) \ln (1 - e^{-\phi x_i})\right]. \quad (8$$

Derivation of the χ^2 Equation 13. For a value of the IC frequency ϕ with LDA dose-response data, χ^2 is calculated from class frequencies as

$$\chi^2_{(I \cdot 1)} = \sum_{i=1}^{I} \sum_{j=1}^{2} \left[\frac{(O_{ij} - E_{ij})^2}{E_{ij}}\right] \quad (34$$

where O_{ij} and E_{ij} are the observed and expected frequencies for the ij^{th} class. There are 2 I class frequencies with a positive ($j = 1$) and negative ($j = 2$) response class for each of the I dose classes. If for each class frequency n_{ij} is the number of trials and P_{ij} is the probability of an event, then it can be shown that

$$\chi^2_{(I \cdot 1)} = \sum_{i=1}^{I} \left[\frac{(O_{i2} - n_{i2}P_{i2})^2}{n_{i2}P_{i2}(1 - P_{i2})}\right]. \quad (35$$

Since n_{i2} is the number n_i of replicate cultures, P_{i2} is the probability P_i of a negative response from Equation 27, and O_{i2} is the number r_i of negatively responding cultures, χ^2 can be derived as

$$\chi^2_{(I \cdot 1)} = \sum_{i=1}^{I} \left[\frac{(r_i - n_i e^{-\phi x_i})^2}{n_i e^{-\phi x_i}(1 - e^{-\phi x_i})}\right]. \quad (13$$

REFERENCES

1. Jonard, J., and J. Panijel. 1973. *In vitro* primary response in microcultures containing less than fifty thousand lymphoid cells. Eur. J. Immunol. 3:245.
2. Quintans, J., and I. Lefkovits. 1973. Precursor cells specific to sheep red cells in nude mice. Estimation of frequency in the micro-culture system. Eur. J. Immunol. 3:392.
3. Skinner, M. A., and J. Marbrook. 1976. An estimation of the frequency of precursor cells which generate cytotoxic lymphocytes. J. Exp. Med. 143: 1562.
4. Teh, H.-S., E. Harley, R. A. Phillips, and R. G. Miller. 1977. Quantitative studies on the precursors of cytotoxic lymphocytes. I. Characterization of a clonal assay and determination of the size of clones derived from single precursors. J. Immunol. 118:1049.
5. Lindahl, K. F., and D. B. Wilson. 1977. Histocompatibility antigen activated cytotoxic T lymphocytes. II. Estimates of the frequency and specificity of precursors. J. Exp. Med. 145:508.
6. Ryser, J.-E., and H. R. MacDonald. 1979. Limiting dilution analysis of alloantigen-reactive T lymphocytes. I. Comparison of precursor frequencies for proliferative and cytolytic responses. J. Immunol. 122:1691.
7. Swain, S. L., P. R. Panfili, R. W. Dutton, and I. Lefkovits. 1979. Frequency of allogeneic helper T cells responding to whole H-2 differences and to an H-2K difference alone. J. Immunol. 123:1062.
8. Miller, R. G., H.-S. Teh, E. Harley, and R. A. Phillips. 1977. Quantitative studies of the activation of cytotoxic lymphocyte precursor cells. Immunol. Rev. 35:38.
9. Lefkovits, I., and H. Waldmann. 1979. Limiting Dilution Analysis of Cells in the Immune System. Cambridge University Press, Cambridge.
10. Haight, F. A. 1967. Handbook of the Poisson Distribution. John Wiley & Sons, Inc., New York.
11. Finney, D. J. 1952. Statistical Method in Biological Assay. 1st Ed. Charles Griffin & Company, London.
12. Kennedy, J. C., J. E. Till, L. Siminovitch, and E. A. McCulloch. 1966. The proliferative capacity of antigen-sensitive precursors of hemolytic plaque-forming cells. J. Immunol. 96:973.
13. Groves, D. L., W. E. Lever, and T. Makinodan. 1970. A model for the interaction of cell types in the generation of hemolytic plaque-forming cells. J. Immunol. 104:148.
14. Porter, E. H., and R. J. Berry. 1963. The efficient design of transplantable tumour assays. Br. J. Cancer 17:583.
15. Finney, D. J. 1978. Statistical Method in Biological Assay. 3rd ed. Charles Griffin & Company, London.
16. Halsall, M. K., and T. Makinodan. 1974. Analysis of the limiting-dilution assay used for estimating frequencies of immunocompetent units. Cell. Immunol. 11:456.
17. Remington, R. D., and M. A. Schork. 1970. Statistics with Applications to the Biological and Health Sciences. Prentice-Hall, Inc., Englewood Cliffs, NJ.
18. Colquhoun, D. 1971. Lectures on Biostatistics. An Introduction to Statistics with Applications in Biology and Medicine. Clarendon Press, Oxford.
19. Johnson, L. W., and R. D. Riess. 1977. Numerical Analysis. Addison-Wesley, Inc., Reading, MA.
20. Fisher, R. A. 1922. On the mathematical foundations of theoretical statistics. Philos. Trans. R. Soc. London; A, 222:309.
21. Lancaster, H. O. 1969. The Chi-Squared Distribution. John Wiley & Sons, Inc., New York.
22. Taswell, C., H. R. MacDonald, and J.-C. Cerottini. 1979. Limiting dilution analysis of alloantigen-reactive T lymphocytes. II. Effect of cortisone and cyclophosphamide on cytolytic T lymphocyte precursor frequencies in the thymus. Thymus 1:119.
23. MacDonald, H. R., J.-C. Cerottini, J.-E. Ryser, J. L. Maryanski, C. Taswell, M. B. Widmer, and K. T. Brunner. 1980. Quantitation and cloning of cytolytic T lymphocytes and their precursors. Immunol. Rev. 51:93.
24. Hammersley, J. M., and D. C. Handscomb. 1964. Monte Carlo Methods. Methuen & Co., Ltd., London.
25. Berkson, J. 1955. Maximum likelihood and minimum χ^2 estimates of the logistic function. J. Am. Statistical Assoc. 50:130.
26. Berkson, J. 1956. Estimation by least squares and by maximum likelihood. Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability. Vol. 1. Edited by Jerzy Neyman. University of California Press, Berkeley and Los Angeles, Pp. 1–13.
27. Fisher, R. A. 1970. Statistical Methods for Research Workers. 14th Ed. Oliver & Boyd, Edinburgh.
28. Berkson, J. 1961. The purpose of estimation. Some remarks on Mr. C. R. Rao's paper "Apparent Anomalies and Irregularities in Maximum Likelihood Estimation." Bull. Int. Stat. Inst. 38:1.
29. Berkson, J. 1980. Minimum chi-square, not maximum likelihood! Ann. Stat. 8:457.