Anesthesiology
1997; 86:825–35

A Technique for Population Pharmacodynamic Analysis of Concentration–Binary Response Data

James M. Bailey, M.D., Ph.D.,* Keith M. Gregg Ph.D.†

Background: **Pharmacodynamic data frequently consist of the binary assessment (a "yes" or "no" answer) of the response to a defined stimulus (verbal stimulus, intubation, skin incision, and so on) for multiple patients. The concentration–effect relation is usually reported in terms of C_{50}, the drug concentration associated with a 50% probability of drug effect, and a parameter the authors denote γ, which determines the shape of the concentration–probability of effect curve. Accurate estimation of γ, a parameter describing the entire curve, is as important as the estimation of C_{50}, a single point on this curve. Pharmacodynamic data usually are analyzed without accounting for interpatient variability. The authors postulated that accounting for interpatient variability would improve the accuracy of estimation of γ and allow the estimation of C_{50} variability.**

Methods: **A probit-based model for the individual concentration–response relation was assumed, characterized by two parameters, C_{50} and γ. This assumption was validated by comparing probit regression with the more commonly used logistic regression of data from individual patients found in the anesthesiology literature. The model was then extended to analysis of population data by assuming that C_{50} has a log-normal distribution. Population data were analyzed in terms of three parameters, $\langle C_{50} \rangle$, the mean value of C_{50} in the population; ω, the standard deviation of the distribution of the logarithm of C_{50}; and γ. The statistical characteristics of the technique were assessed using simulated data. The data were generated for a range of γ and ω values, assuming that C_{50} and γ had a log-normal distribution.**

Results: **The probit-based model describes data from individual patients and logistic regression does. Population analysis using the extended probit model accurately estimated $\langle C_{50} \rangle$, γ, and ω for a range of values, despite the fact that the technique accounts for C_{50} variability but not γ variability.**

Conclusion: **A probit-based method of pharmacodynamic analysis of pooled population data facilitates accurate estimation of the concentration–response curve. (Key words: Anesthetics: interpatient variability; C_{50}. Pharmacodynamics: concentration–response. Statistics: logistic regression; probit regression; random effects.)**

PHARMACODYNAMIC data are often recorded as binary variables; that is, the patient does or does not respond to a command, can or cannot maintain adequate spontaneous ventilation, does or does not have a hemodynamic or somatic response to surgical stimulus, and so forth. In this situation, the most common technique of data analysis is logistic regression, in which the probability of drug effect is evaluated as a function of C, the drug concentration in plasma (or at the effect site), with the equation

$$P = C^{\gamma}/(C50^{\gamma} + C^{\gamma}) \quad (1)$$

C_{50} is the concentration at which the probability of drug effect is 50% and γ is a measure of the steepness of the concentration-effect curve (throughout this article *P* refers to the probability of drug effect, such as the probability of ablating the response to some stimulus, such as skin incision). Logistic regression has been used for the analysis of the pharmacodynamics of inhaled and intravenous anesthetics,[1–10] usually with a primary focus of determining C_{50} values. However, logistic regression not only estimates the median point of the concentration-response curve (C_{50}), but it also provides information on the shape of this curve by estimating the steepness parameter γ. Note that equation 1, when applied to a single patient, implies that there is a fundamental element of intrapatient variability in the concentration-effect relation. There is a finite, albeit small, probability of drug effect at very low drug concentrations, and at high drug concentrations there is still some probability that a drug effect will not be observed. This reflects what many studies have shown: specifically that the drug concentration needed to block the response of an individual to a repetitive and unchanging stimulus varies randomly around some mean

* Associate Professor, Department of Anesthesiology, Emory University School of Medicine.

† Department of Anesthesiology, Stanford University School of Medicine.

Received from Emory University School of Medicine and Stanford University School of Medicine. Submitted for publication June 14, 1996. Accepted for publication December 3, 1996.

Address reprint requests to Dr. Bailey: Department of Anesthesiology, Emory University School of Medicine, 1364 Clifton Road, N.E., Atlanta, Georgia 30322. Address electronic mail to: James_Bailey@eushc.org

value during the study. This also is consistent with the clinical observation that anesthetic requirements vary during a procedure, even when the stimulus is seemingly unchanging. The parameter γ is a measure of this intrapatient variability and has substantial clinical significance. When γ is large (>6), the probability of effect is very close to zero when C is less than C_{50} (even if only slightly less) and very close to 1 when C is greater than C_{50} (even if only slightly greater). In other words, if γ is large there is a well-defined threshold for drug effect. In contrast, when γ is small, the drug effect-concentration curve is not as steep, and it is difficult to clearly define a threshold for drug effect. For example, if $\gamma = 1$, equation 1 implies that the concentration must increase by a factor of 9 to increase from a 10% chance to a 90% chance of drug effect. In contrast, if $\gamma = 10$, the concentration need only increase by a factor of 1.5 to change from a 10% to 90% chance of drug effect. Understanding the shape of the concentration-effect curve is important for clinical practice.[11,12] Although C_{50} is a single point of the concentration-effect curve, the parameter γ characterizes the entire relation. For example, given C_{50}, the drug concentration associated with a 90% or 95% probability of drug effect (C_{90} or C_{95}) depends on γ. Given that most anesthesiologists would prefer to maintain the anesthetic concentrations at C_{90} or C_{95} rather than C_{50}, it seems that accurate estimation of γ could be viewed as important as determining C_{50}.

Ideally, logistic regression would be applied to data taken from individual patients for evaluation of C_{50} and γ and then these individual values could be used to generate a measure of central tendency (mean, median, and so on) for the population. But in reality, many important stimuli, such as tracheal intubation, loss of consciousness, skin incision, or sternotomy, can only be evaluated once during the study of any one patient and logistic regression is most often applied to pooled data of multiple patients with one data point per patient. Analysis of pooled data should account for interpatient variability around the typical concentration-response curve. Evaluating interpatient variability has proved useful in pharmacokinetic investigations,[13-15] and we believe it is important to develop methods to evaluate pharmacodynamic interpatient variability.

Although quantifying interpatient variability is of interest in its own right, it is even more important to account for interpatient variability to evaluate more accurately the shape of the individual concentration-effect curves. Evaluation of the steepness of the individual concentration-effect relation may be confounded by the lack of distinction between intrapatient and interpatient variability. This is illustrated conceptually in figure 1, where the concentration-effect relations of nine hypothetical patients are presented. Although each curve is steep, the curve generated by one point from each patient is much flatter (and, in fact, the curve would have appeared even flatter if we had not selected low probability values from the leftward curves and high probability values from the rightward curves). Although this is a contrived example, it nevertheless illustrates how intrapatient variability may be overestimated if interpatient variability is significant. The resultant concentration-effect curve may appear artificially flat while the curve for any one patient may be steep. No readily accessible method is available to evaluate interpatient C_{50} variability or to distinguish interpatient C_{50} variability in pooled data analysis from flat individual concentration-effect curves. This article presents a technique to achieve these goals.

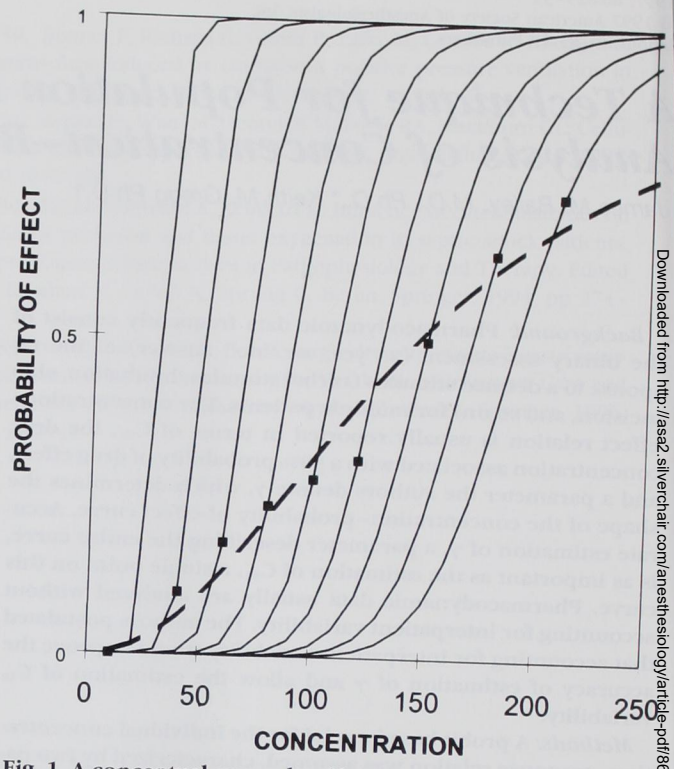

Fig. 1. A conceptual example of how the slope of the resultant concentration–effect curve (— ● —) may appear flat when generated by single points from multiple patients, each of whom have a steep concentration–effect curve (—) but with varying values of C_{50} (in this example, γ for each patient equals 10).

Methods

Appendix 1 presents the mathematical basis for this probit-based approach to pharmacodynamic analysis. The technique is based on the assumption that for individual patients the probability of a drug effect is given by

$$P = \Phi[\gamma \ln C - \gamma \ln C50] \quad (2)$$

where Φ denotes the standardized cumulative normal distribution; that is, $\Phi[\gamma \ln C - \gamma \ln C_{50}]$ is the area under a Gaussian curve (with standard deviation of 1) from minus infinity to $\gamma \ln C - \gamma \ln C_{50}$ (see appendix 1 for details). Thus equation 2 resembles equation 1 (logistic regression) in that it is characterized by two parameters, γ and C_{50}, and the probability increases with increasing C.

In two separate studies, Vuyk *et al.*[7,9] collected data during gynecologic surgery on the response to repetitive intra-abdominal stimuli (electric cautery) of individual patients as a function of alfentanil plasma concentration when supplemented by nitrous oxide or various concentrations of propofol. Because this is, to our knowledge, the largest collection of individual concentration-response data, we used it to assess the validity of probit analysis for individual patients (equation 2) by comparing the results to those of logistic regression (equation 1). Observed responses and drug concentrations were taken from the figures of these reports. The parameters (C_{50}, γ) in equations 1 and 2 were evaluated by maximum likelihood estimation. The logarithm of the likelihood of observed results

$$\log L = \sum_i L_i$$

$$= \sum_i \{R_i \log (P_i) + (1 - R_i) \log (1 - P_i)\} \quad (3)$$

(P is given by equation 1 or 2 and $R = 1$ if a drug effect is observed and $R = 0$ otherwise and I indexes patients) was maximized as a function of C_{50} and γ. The estimation procedure was easily implemented using an Excel spreadsheet (Microsoft, Redmond, WA), taking advantage of its built-in function that evaluates the error function and the Solver function for the maximization step.

Appendix 1 describes generalization of equation 2 to account for interpatient variability. The probability of drug effect when concentration-response data are collected from multiple patients with one observation per patient is given by

$$P = \Phi[(\gamma \ln C - \gamma \ln \langle C50 \rangle)/\sqrt{(1 + \omega^2\gamma^2)}] \quad (4)$$

where Φ again denotes the standardized cumulative normal distribution, but now the probability of drug effect is determined by three parameters, $\langle C_{50} \rangle$, the mean value of C_{50} in the population; γ; and ω^2, the variance of $\ln C_{50}$ values in the population. Derivation of equation 4 is contingent on the assumption that the probability of drug effect for individual patients is given by equation 2 and that the distribution of C_{50} values in the population has a log-normal distribution.

Direct validation of equation 4 was not possible because we are not aware of a sufficiently large data set comprised of single data points from multiple patients for which the interpatient variability of C_{50} (ω^2 in equation 4) is also known. Instead, computer simulation was used for indirect validation. Each simulated study consisted of 30 different participants (a number fairly typical of many anesthesiology pharmacodynamic studies). A C_{50} value and a γ value were assigned to each hypothetical "patient" using the Excel random-number generator and assuming that both C_{50} and γ had log-normal distributions. This is analogous to the enrollment of patients, during a true human participant study, from a population with randomly varying C_{50} and γ values. For all simulations we assumed that the mean C_{50} (denoted $\langle C_{50} \rangle$) was 100 units per milliliter, that the standard deviation of the distribution of log γ was 0.5 (see Discussion to follow), and the mean value of γ and the variance of C_{50} were varied, as will be described. Each simulated patient was also assigned a single drug concentration from one of 30 values distributed uniformly from 10 to 300 units/ml. This corresponds in a real study to the investigator assigning a drug dose to each patient enrolled in the study. At this point, if the concentration-effect relation were completely deterministic (no intrapatient variability) then a positive drug effect would be observed if the drug concentration, C, assigned to the patient exceeded the value of C_{50} for that patient (see appendix 1). However, as noted earlier, there is an element of randomness even in the individual concentration-effect relation. To account for this intrapatient variability, a normally distributed random variable (denoted ϵ), with mean value of 0 and standard deviation of 1, was generated for each "patient" and added to $\gamma \ln C - \gamma \ln C_{50}$ (see equation A-1 in the appendix 1). If this sum exceeded zero, the simulated patient was assumed to have a positive drug effect (the response variable R was given a value of 1).

This simulation technique is best illustrated with an example. Each simulation was done on a spreadsheet and each row corresponded to a single "patient" and had entries for the assigned value of C and the randomly

generated values of C_{50}, γ, and ϵ and the sum $\gamma \ln C - \gamma \ln C_{50} + \epsilon$, which determined R. In one of the simulations, patient 6 was assigned a drug concentration of 80 units/ml. The values of C_{50}, γ, and ϵ returned by random-number generation were 95, 9.4, and 0.9, respectively. Thus $\gamma \ln C - \gamma \ln C_{50} + \epsilon$ was equal to -0.72, and this indicates that a drug effect was not observed (*i.e.*, R = 0). In contrast, patient 13 was assigned a drug concentration of 90 units/ml. The values of C_{50}, γ, and ϵ returned by random-number generation were 115, 5.4, and 1.4, respectively. For this patient, $\gamma \ln C - \gamma \ln C_{50} + \epsilon$ was equal to 0.08 and R = 1.

Six different simulations were conducted by assigning γ a mean value of either 10 (steep concentration-response curves), 5 (intermediate concentration-response curves), or 2 (flat concentration-response curves) and assigning ω^2 values of either 1 or 0.04. Twenty-five repetitions of each simulation were conducted to assess the statistical properties of this technique. Maximum likelihood estimates of $\langle C_{50} \rangle$, γ, and ω were calculated using equations 3 and 4 and implemented with the Excel spreadsheet. We also applied "naive" probit analysis (ignoring interpatient variability) to the simulated data sets by assigning ω the value of 0.

Appendix 2 provides a detailed description of how to use an Excel spreadsheet to implement this technique.

This technique was applied to real data presented by Ausems *et al.*[5] on the relation among alfentanil (when supplemented by nitrous oxide), plasma concentration, and the responses to intubation and skin incision, with data retrieved from the published graphs of Jacobs *et al.*[10] on the relation between midazolam plasma concentrations and loss of responsiveness, and of Bailey *et al.*[16] on the relation between sufentanil plasma concentrations and the responses to intubation, skin incision, and sternotomy in cardiac surgical patients, using the original data. For alfentanil, an indirect and approximate estimate of ω was made by equating it to the standard deviation of ln C_{50} values for the electroencephalographic effect reported by Egan *et al.*[17] For midazolam, the standard deviation of patient ages in the study of Jacobs *et al.*[10] was used as a measure of ω, because these investigators found that age was an important covariate of ln C_{50}. For sufentanil, ω was directly calculated from reported values of C_{50}.[16]

Results

Population pharmacodynamic analysis using equation (4) is based on the assumption that a probit model (equation 2) accurately describes the concentration-response relation for individual patients. We assessed this assumption using the data of Vuyk and colleagues, who have studied extensively the pharmacodynamics of alfentanil in combination with varying doses of propofol or nitrous oxide in healthy women undergoing lower abdominal gynecologic surgery.[7,9] They have presented concentration-response data from 38 patients. In seven cases the data were not amenable to either logistic or probit analysis (because a concentration was found above which there was always a positive drug effect and below which there was never a positive drug effect). For the remaining 31 patients, the log-likelihoods of the "best" (maximum likelihood) description of the data by either logistic regression or probit analysis are shown in table 1, along with C_{50} values. There was

Table 1. Comparison of Logistic and Probit Regression

Patient Number	Logistic Regression (equation 1)		Probit Regression (equation 4)	
	Log Likelihood	C_{50}	Log Likelihood	C_{50}
1	−1.06532	214.1	−1.03553	214.5
2	−.70989	322.0	−.80881	330.5
3	−1.60195	154.7	−1.59424	154.6
4	−2.58987	218.5	−2.56484	218.5
5	−2.87866	243.1	−2.87617	243.3
6	−5.77132	70.7	−5.76901	71.3
7	−1.75534	249.9	−1.73134	250.4
8	−3.12353	216.4	−3.11987	217.2
9	−1.17598	114.1	−1.1386	114.2
10	−1.25779	11.9	−1.24108	12.1
11	−1.98425	27.0	−2.4843	30.0
12	−1.31266	119.1	−1.29176	119.5
13	−1.3843	66.7	−1.3697	66.8
14	−2.07401	68.0	−2.03493	68.3
15	−4.3225	135.5	−4.30576	135.7
16	−1.17671	74.1	−1.14826	74.5
17	−2.83859	196.1	−2.80724	196.9
18	−3.70032	122.5	−3.69338	122.7
19	−3.75884	151.6	−3.753	152.0
20	−2.98351	101.6	−2.98554	101.0
21	−2.68345	171.6	−2.67342	172.4
22	−1.30887	33.5	−1.29635	33.6
23	−1.36657	27.8	−1.47043	26.5
24	−1.93444	21.2	−1.91961	21.3
25	−2.06559	45.5	−2.05672	45.7
26	−1.78894	33.3	−1.7653	33.4
27	−2.2905	58.7	−2.26178	58.6
28	−2.07936	21.3	−2.0662	21.5
29	−1.10739	78.0	−1.08837	78.0
30	−2.25199	37.5	−2.23398	37.5
31	−1.91965	29.2	−1.89681	29.2

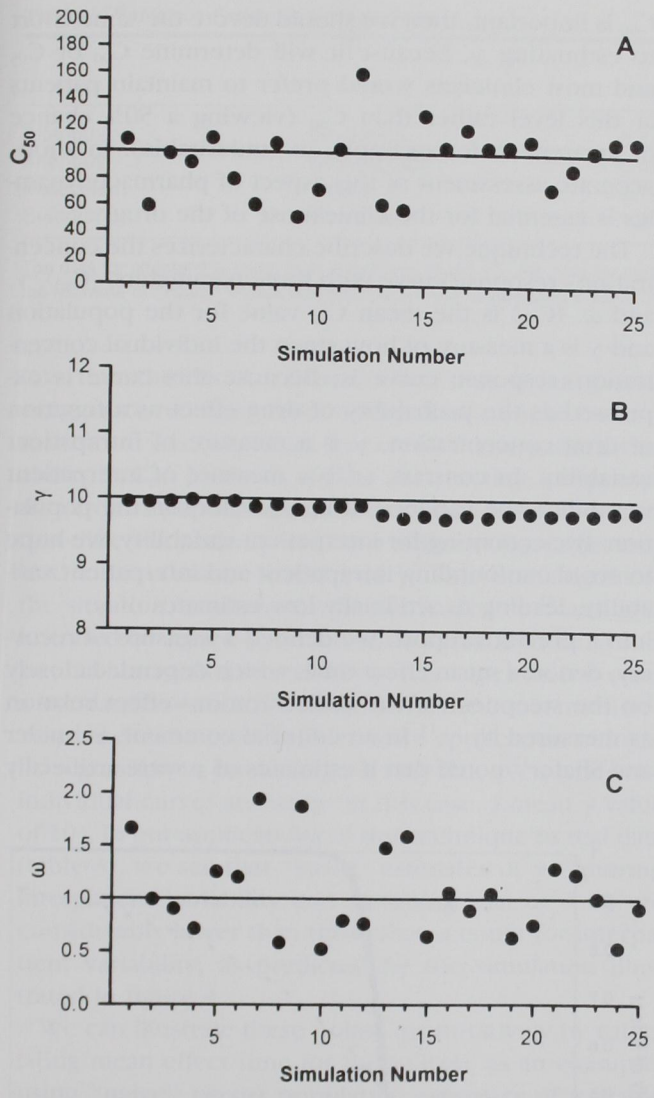

Fig. 2. Individual estimates of $\langle C_{50} \rangle$ (*A*), γ (*B*), and ω (*C*) for 25 repetitions of simulated data generated assuming $\langle C_{50} \rangle = 100$, mean $\gamma = 10$ (γ is log-normal distributed with a standard deviation of 0.5), and $\omega = 1$. These target values for the estimates are indicated by accented horizontal lines.

no difference between the two approaches in their ability to describe these data.

The precision and bias of population analysis using equation (4) was evaluated by computer simulations. Figures 2 and 3 show the distribution of population estimates of $\langle C_{50} \rangle$, γ, and ω for two of these simulations. For figure 2, the simulated data were generated assuming that $\langle C_{50} \rangle = 100$, $\gamma = 10$, and $\omega = 1$. Qualitatively, this represents steep individual concentration-response curves with substantial interpatient variability in the C_{50} value. For figure 3, the target values are $\langle C_{50} \rangle = 100$, $\gamma = 2$, and $\omega = 0.2$, which qualitatively represents flat individual concentration-response curves (large intrapatient variability) with relatively insignificant interpatient variability. The statistical fidelity of the method is indicated by how close the individual estimates are to these target values (shown as horizontal lines in the figures). For quantitative appraisal, the mean estimates of $\langle C_{50} \rangle$, γ, and ω for 25 repetitions of six different simulations (varying in the values of γ and ω)

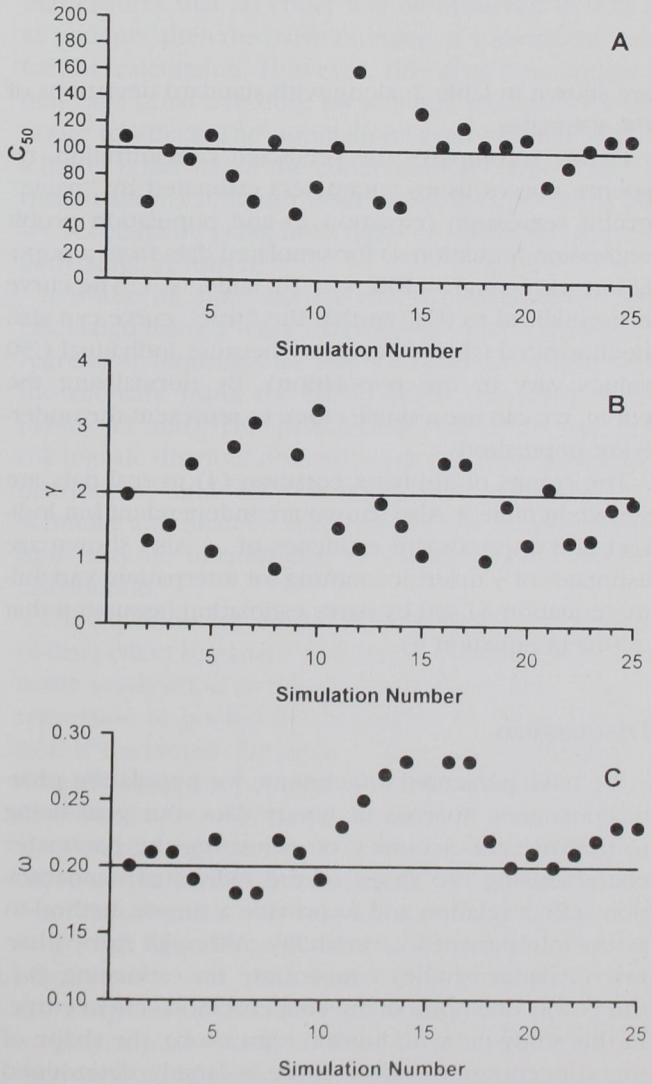

Fig. 3. Individual estimates of $\langle C_{50} \rangle$ (*A*), γ (*B*), and ω (*C*) for 25 repetitions of simulated data generated assuming $\langle C_{50} \rangle = 100$, mean $\gamma = 2$ (γ is log-normal distributed with a standard deviation of 0.5), and $\omega = 0.2$. These target values for the estimates are indicated by accented horizontal lines.

Table 2. Results of Computer Simulations

Target Parameters			Estimated Parameters		
C_{50}	γ	ω	C_{50}	γ	ω
100	10	1	94.9 (25.6)	9.8 (0.9)	1.05 (0.17)
100	10	0.2	102.0 (10.9)	10.5 (0.8)	0.25 (0.08)
100	5	1	96.4 (30.2)	4.9 (0.4)	0.97 (0.24)
100	5	0.2	103.2 (11.6)	4.9 (1.1)	0.21 (0.06)
100	2	1	102.5 (35.6)	2.1 0.3)	1.0 (0.34)
100	2	0.2	98.7 (18.4)	1.8 (0.7)	0.23 (0.04)

Estimated parameters are shown with standard deviation of estimate (for 25 repetitions) in parentheses. Target parameters are those used to generate the simulated patient data and thus are the targets for ideal estimation.

are shown in table 2, along with standard deviations of the estimates.

Figure 4 compares the predicted concentration-response curves using parameters estimated by "naive" probit regression (equation 2) and population probit regression (equation 4) for simulated data from a population with $\langle C_{509} \rangle = 100$, $\gamma = 10$, and $\omega = 1$. The curve is normalized to $\langle C_{50} \rangle$ so that the "true" curve can also be illustrated (this is necessary because individual C50 values vary in the population). By normalizing the curve, we can use a single curve to represent the underlying population).

The results of applying equation (4) to real data are shown in table 3. Also shown are independent but indirect and approximate estimates of ω. Also shown are estimates of γ made accounting for interpatient variability (equation 4) and by naive estimation (assuming that $\omega = 0$ in equation 4).

Discussion

We have presented a technique for population pharmacodynamic analysis of binary data, our goal being to improve the accuracy of estimating the parameter characterizing the shape of the individual concentration-effect relation and to provide a simple method to assess interpatient C_{50} variability. Although many pharmacodynamic studies concentrate on estimating C_{50}, this is only one point of the concentration-effect curve. In this study (as with logistic regression), the shape of the concentration-effect curve is largely determined by a single parameter, which we denoted γ. Drugs with a relatively large value of γ have a steep concentration-effect relation with well-defined thresholds for effect. Furthermore, if we believe that accurate estimation of C_{50} is important, then we should devote the same effort to estimating γ, because it will determine C_{90} or C_{95} and most clinicians would prefer to maintain patients at this level rather than C_{50} (viewing a 50% chance of awareness, for example, as undesirable). In short, accurate assessment of this aspect of pharmacodynamics is essential for the clinical use of the drug.

The technique we describe characterizes the concentration-response curve with three parameters, $\langle C_{50} \rangle$, γ, and ω. $\langle C_{50} \rangle$ is the mean C_{50} value for the population and γ is a measure of how steep the individual concentration-response curve is. Because this curve is expressed as the probability of drug effect as a function of drug concentration, γ is a measure of intrapatient variability. In contrast, ω^2 is a measure of interpatient variability, the variance of ln C_{50} values in the population. By accounting for interpatient variability, we hope to avoid confounding intrapatient and interpatient variability, leading to artificially low estimates of γ.

In a previous report, we defined a measure of recovery, denoted mean effect time, which depended closely on the steepness of the concentration-effect relation as measured by γ.[11] In an editorial comment, Schnider and Shafer[12] noted that if estimates of γ were artificially

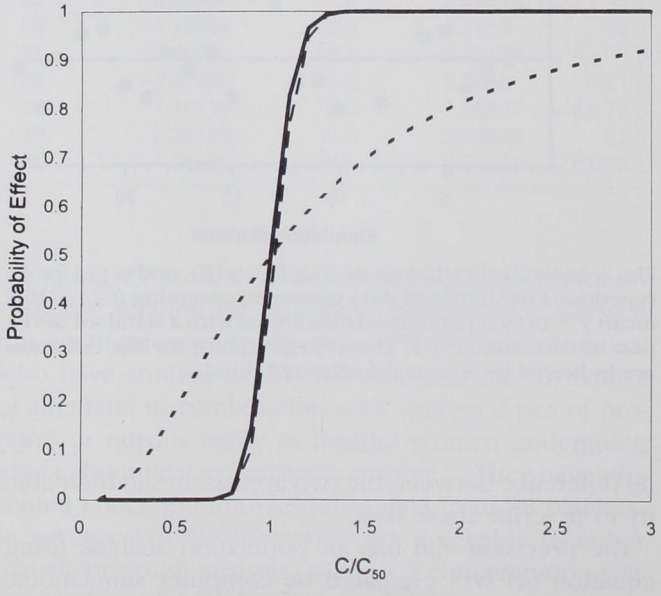

Fig. 4. The probability of response as a function of drug concentration, normalized to $\langle C_{50} \rangle$. Simulated data were generated assuming that the mean value of $\gamma = 10$ (γ is log-normal distributed with a standard deviation of 0.5) and that $\omega = 1$. Shown are the concentration–response curves for the typical patient with $\gamma = 10$ (––), the curve estimated using population probit analysis (equation 4) (– – –), and the curve estimated with naive probit analysis (equation 2) (. . . . . .).

Table 3. Variance Estimate Comparison

Drug	Stimulus	ω Probit	ω Indirect	γ	γ Naive
Alfentanil	Intubation	0.37	0.65	8.05	2.56
Alfentanil	Skin incision	0.45	0.65	3.18	1.81
Midazolam	Verbal	0.74	0.38	5.05	1.31
Sufentanil	Combined	1.27	0.67	3.02	0.76

The data for alfentanil is taken from references 5 and 16. The data for midazolam is taken from reference 10. The estimate of γ uses the full equation (4), and the estimate of "naive" γ uses equation (4) but with ω assigned a value of zero.

low, then the calculated mean effect time would grossly exaggerate recovery time. A source of error in the estimation of γ is the failure to consider interpatient variability in C_{50}. Methods of data analysis that use pooled data from a population and do not account for the fact that different patients will have different C_{50} values may lead to inaccurate estimates of γ. This is analogous to the situation in pharmacokinetic analysis[13-15] and is illustrated by figure 4, which shows that if there is substantial interpatient variability of C_{50} values (in this case, a standard deviation for ln C_{50} of 1), the concentration-response curve estimated by naive application of logistic regression to pooled data is very flat, even when the individual curves are steep (in this case, a mean γ value of 10). In our applications of this technique to real data (table 3), we see that "naive" estimates of γ (ignoring interpatient variability and assuming that $\omega = 0$) are considerably lower than those that account for interpatient variability, as predicted by the simulation illustrated in figure 4.

We can illustrate these points quantitatively by calculating mean effect time for midazolam, as an example, using "naive" *versus* population estimates of γ (table 3). The technique for calculating the mean effect time is described by Bailey,[11] with the substitution of equation 2, rather than equation 1, for the probability of drug effect, and we consider the specific case of a duration of administration of 100 min. After discontinuing administration, the mean effect time based on the naive estimate of γ (1.31) is 42 min. Using the population-based estimate of γ (5.05), the estimate is 14 min. The naive estimate exaggerates the estimate of recovery time. However, we must also emphasize that accurate estimation of γ is important for understanding the pharmacology of the drug only if the clinician titrates the drug to effect. If drugs are administered without end points for titration, the naive estimate of γ may actually be more informative, although this would be grossly misleading about the concentration-effect curve of the typical patient. For example, if we seek the drug concentration that ensures that an effect will be observed in 90% of all patients, then the naive estimate of γ should be used for the calculation. However, this drug concentration may be a gross overdose for any one patient if we can titrate to effect. The population-based estimate of γ is a better reflection of the concentration-effect curve in this situation, although both estimates are useful and we recommend using both types of analysis in pharmacodynamic research.

It is also of some interest that the mean effect time calculated using a naive estimate of γ based on logistic regression (equation 1) was considerably larger than the estimate using the naive probit (equation 2) (89 *versus* 42 min). This presumably reflects the fact that the logistic distribution has longer tails than the normal distribution and may confirm the speculation by Schnider and Shafer[12] that calculations of mean effect time may be inflated due to the tails of the logistic distribution.

The equation we derive to describe the probability of drug effect (equation 4) is not as straightforward and easily understood as the one used for "naive" logistic regression of pooled data (equation 1). Thus a discussion is warranted. Equation 4 relates the probability of drug effect for the population to the cumulative normal distribution, by convention denoted $\Phi(x)$. If we plot a standardized normal curve (a Gaussian or bell-shaped curve with a standard deviation of 1) on the x-y axis, centered around $x = 0$, $\Phi(x)$ is the area under this curve to the left of x. Note that x can be either positive or negative; if $x = 0$, then $\Phi(x) = 0.5$, because $x = 0$ is the midpoint of the curve. In equation 4, $x = (\gamma \ln C - \gamma \ln\langle C_{50}\rangle)/\sqrt{(1 + \omega^2\gamma^2)}$. So as C increases, $\Phi(x)$ also increases, which is logical because the probability of drug effect should increase as the drug concentration increases. The probability of drug effect is 50% when $C = \langle C_{50}\rangle$ because $x = (\gamma \ln C - \gamma \ln\langle C_{50}\rangle)/\sqrt{(1 + \omega^2\gamma^2)} = 0$ at this point and the normal curve is centered

around $x = 0$. Also, note that as γ becomes larger, $x = (\gamma \ln C - \gamma \ln\langle C_{50}\rangle)/\sqrt{(1 + \omega^2\gamma^2)}$ increases more per unit increase in concentration; that is, the concentration-effect curve is steeper. Finally, as ω becomes larger (greater interpatient variability), the increase in x and $\Phi(x)$ per unit increase in drug concentration decreases; that is, the concentration-curve for the population is flatter. This reflects the fact that with greater interpatient variability there will be more patients with extremes of C_{50}, both low and high, so that for the population the probability of drug effect at low drug concentrations is increased while it is decreased at high concentrations, flattening the curve.

Within the framework of the model and its assumptions, the mathematical derivation of equation 4 is exact. Thus the utility of this technique is determined solely by the validity of the assumptions of the model. The critical assumptions are (1) the concentration-response relation for individual patients is described by equation 2, a probit relation, rather than by the more commonly used logistic equation (equation 1); (2) the distribution of C_{50} values in the population conforms to a log-normal relation; and (3) we can ignore the interpatient variability of γ without substantial loss of accuracy.

The first assumption, that individual concentration-response curves are appropriately described by a probit model, was directly evaluated using data for individual patients and was compared with logistic regression. Although logistic regression has been commonly used for binary response pharmacodynamic analysis, we know of no previous effort to determine whether intrapatient variability is well described by a logistic distribution. In our analysis of the data of Vuyk *et al.*,[7,9] we found that the probit model (which uses a normal distribution to describe intrapatient variability) fit the data as well as the logistic model. This reflects the fact that the logistic and normal distributions are not that dissimilar, both being symmetric and sigmoid. We cannot conclude that the probit model is superior to the logistic model for individual patients, but it seems apparent that it is just as appropriate.

The second assumption of the model is that the distribution of C_{50} values conforms to a log-normal distribution. This type of assumption is commonly used in both pharmacokinetic and pharmacodynamic modeling and conforms to observed distributions of pharmacokinetic and pharmacodynamic parameters.[13] However, if the actual distribution of C_{50} values are multimodal, this assumption is clearly incorrect.

The final critical assumption of this model is that it is "γ naive"; that is, we have not accounted for variability in γ values. Although we have referred to γ as the parameter that determines the steepness of the concentration-response relation, the slope of this curve is also influenced by C_{50}. Thus ignoring γ variability is not equivalent to assuming that the shape of the concentration-response curve is the same for all patients. In addition, interpatient γ variability could be incorporated into the model if the distribution of γ were normal. As we noted previously, a log-normal distribution seems more plausible. However, if the variance is not too large, a normal distribution could approximate a log-normal, as a "first-order" approximation similar to that used in NONMEM,[13,14] and the model could be expanded to consider interpatient variability in γ. However, we have not extended the model in this manner because our primary focus is the variability of C_{50}. Accounting for γ variability would add a parameter to the model (two additional parameters would be required if covariance between C_{50} and γ is considered). Few anesthesiology data sets of this type are large enough to support accurate estimation of four (or five) parameters. In our simulations, we have assumed a log-normal distribution of γ values in the population, with a standard deviation equal to 50% of the mean value, a substantial variability. We chose this value because the mean standard deviation of ln γ values among the separate patient subsets of the data from Vuyk *et al.*[7,9] (after "trimming" the highest and lowest γ value in each group) was 0.477. Thus a ln γ standard deviation of 0.5 seems realistic. We found that our estimates of the mean value of γ were accurate despite our "γ naive" estimating equation (equation 4). This seems to parallel the observation that the "ω^2 naive" approach, logistic regression, is a good estimator of $\langle C_{50}\rangle$ (see figure 4) but may fail in the estimation of γ, a measure of variability (intrapatient). Similarly, our "γ naive" technique estimates $\langle\gamma\rangle$ reasonably accurately but provides no information about its variability. Although our model does not fully account for interpatient variability, we believe it is an improvement over the fully "naive" logistic regression approach.

A possible concern with this project is that we cannot validate our technique directly. There are simply no data sets available consisting of single concentration-response observations from multiple patients for whom the variance of C_{50} in the population is known. Consequently, we evaluated the statistical properties of the technique using computer simulation. We "created" 30

patients for each simulation by randomly generating values of C_{50} and γ, analogous to the experimental investigator enrolling patients from a population with log-normal distributions of these variables, and we assigned a single drug concentration to each patient, analogous to the experimental investigator selecting a drug dose for each patient in the study. We then accounted for intrapatient variability by letting the response variable, R, be determined by the sum of $\gamma \ln C - \gamma \ln C_{50}$ and a term representing random intrapatient variability, generated by a random-number generator. Thus the simulations emulated a typical pharmacodynamic study. We wish to emphasize that computer simulation is a commonly used validation procedure for statistical methods, and our approach parallels the validation of widely used population pharmacokinetic techniques.[14] We believe our simulations are realistic. They are based on a probit model (equation 2) for the individual concentration-response relations, which we have shown is a reasonable assumption, and on the assumption of log-normal distributions for C_{50} and γ. If these assumptions are plausible, our simulations indicate that this technique can accurately estimate $\langle C_{50} \rangle$, γ, and ω.

We illustrated this technique by applying it to three sets of actual patient data, that reported by Ausems *et al.*[5] for the relation between alfentanil (supplemented with nitrous oxide) concentration and the response to intubation or skin incision, that reported by Jacobs *et al.*[10] for the loss of responsiveness and midazolam plasma concentration, and data for sufentanil reported by Bailey *et al.*[16] In each of these cases we have approximate, indirect measures of ω, and the estimates with our technique and these indirect measures are in reasonable agreement. But this observation must be tempered by the approximate nature of the indirect measures. In the case of alfentanil, the indirect measure is the standard deviation of electroencephalographic measures of ln C_{50} reported by Egan *et al.*[17] Whether this has an important relation to the somatic or sympathetic response to intubation or incision is uncertain. In the case of midazolam, the investigators found that patient age was a significant covariate in the estimation of ln C_{50}; that is, ln C_{50} was proportional to age.[10] The independent estimate of ω in the table is simply the standard deviation of the patient ages in the study (multiplied by the coefficient of proportionality reported by the authors) and so should probably be viewed as a lower bound to the true standard deviation of C_{50} in the population. For the sufentanil data, the estimate of ω is actually direct but must be viewed as crude because the number of patients used for the estimate was small (n = 5) and the responses to different stimuli (intubation, incision, and sternotomy) were pooled for analysis.

There is a considerable statistical literature on the analysis of binary or ordered categorical data (see McCullagh and Nelder[18] and Zeger and Liang[19] for overviews) and, in particular, Sheiner[20] has described the analysis of categorical data within the framework of the NONMEM program. We have not examined the analysis of binary data with NONMEM, primarily because the probit technique is so easily implemented with an Excel spreadsheet (as detailed in appendix 2). However, in the analysis of binary data it is necessary to evaluate complicated integrals. In our probit model, these integrals (the cumulative normal distribution) are provided by a built-in function with a high degree of accuracy. With NONMEM, the integrals are evaluated by a technique known as the Laplace approximation.[20,21] To our knowledge, application of the Laplace approximation to the analysis of sparse binary data (one data point per patient) has not been validated in simulations such as those we used here.

We have presented a method for population analysis of binary, "yes/no," pharmacodynamic data. The technique estimates the mean value of C_{50}, a parameter γ that influences the steepness of the concentration-response curve, and ω^2, the variance of ln C_{50} in the population. The model assumes a log-normal distribution of C_{50} values. The underlying model for the individual concentration-response relations was validated by analysis of individual patient data. The statistical properties of the technique were evaluated by computer simulation.

References

1. Waud DR: On biological assays involving quantal responses. J Pharmacol Exp Ther 1972; 183:577-607

2. De Jong RE, Eger EI II: MAC expanded: AD_{50} and AD_{90} values of common inhalation anesthetics in man. ANESTHESIOLOGY 1975; 42:384-9

3. Fisher DM, Zwass MS: MAC desflurane in 60% nitrous oxide in infants and children. ANESTHESIOLOGY 1992; 76:354-6

4. Chortkoff BS, Bennett HL, Eger EI II: Subanesthetic concentrations of isoflurane suppress learning as defined by the category-example task. ANESTHESIOLOGY 1993; 79:16-22

5. Ausems ME, Hug CC Jr, Stanski DR, Burm AGL: Plasma concentrations of alfentanil required to supplement nitrous oxide anesthesia for general surgery. ANESTHESIOLOGY 1986; 65:362-73

6. Wessen A, Persson PM, Nilsson A, Hartvig P: Concentration-effect relationship of propofol after total intravenous anesthesia. Anesth Analg 1993; 77:1000-7

7. Vuyk J, Lim T, Engbers FHM, Burm AGL, Vletter AA, Bovill JG: Pharmacodynamics of alfentanil as a supplement to propofol or nitrous oxide for lower abdominal surgery in female patients. ANESTHESIOLOGY 1993; 78:1036–45

8. Glass PSA, Doherty M, Jacobs JR, Goodman D, Smith LR: Plasma concentration of fentanyl, with 70% nitrous oxide, to prevent movement at skin incision. ANESTHESIOLOGY 1993; 78:842–7

9. Vuyk J, Lim T, Engbers FHM, Burm AGL, Vletter AA, Bovill JG: The pharmacodynamic interaction of propofol and alfentanil during lower abdominal surgery in women. ANESTHESIOLOGY 1995; 83:8–22

10. Jacobs JR, Reves JG, Marty J, White WD, Bai SA, Smith LR: Aging increases pharmacodynamic sensitivity to the hypnotic effects of midazolam. Anesth Analg 1995; 80:143–8

11. Bailey JM: A technique for quantifying the duration of intravenous anesthetic effect. ANESTHESIOLOGY 1995; 83:1095–1103

12. Schnider TW, Shafer SL: Evolving clinically useful predictors of recovery from intravenous anesthetics. ANESTHESIOLOGY 1995; 83:902–5

13. Sheiner LB: The population approach to pharmacokinetic data analysis: Rationale and standard data analysis methods. Drug Metab Rev 1984; 15:153–71

14. Sheiner LB, Beal SL: Evaluation of methods for estimating population pharmacokinetic parameters. II. Biexponential model and experimental pharmacokinetic data. J Pharmacokinet Biopharm 1981; 9:635–51

15. Fisher DM: Propofol in pediatrics: Lessons in pharmacokinetic modeling. ANESTHESIOLOGY 1994; 80:2–5

16. Bailey JM, Schwieger IM, Hug CC Jr: Evaluation of sufentanil anesthesia obtained by a computer-controlled infusion for cardiac surgery. Anesth Analg 1993; 76:247–52

17. Egan TD, Minto CF, Hermann DJ, Barr J, Muir KT, Shafer SL: Remifentanil versus alfentanil. Comparative pharmacokinetics and pharmacodynamics in healthy adult male volunteers. ANESTHESIOLOGY 1996; 84:821–33

18. McCullagh P, Nelder JA: Generalized linear models, 2nd ed. London, Chapman and Hall, 1989

19. Zeger SL, Liang KY: An overview of methods for the analysis of longitudinal data. Stat Med 1992; 11:1825–39

20. Sheiner LB: A new approach to the analysis of analgesic drug trials, illustrated with bromfenac data. Clin Pharm Ther 1994; 56: 309–22

21. De Bruijn NG: Asymptotic methods in analysis. Amsterdam, North Holland, 1961

Appendix 1

Although logistic regression is commonly used in anesthesiology research, the underlying statistical model is seldom described. A useful general reference to these types of statistical models is the textbook by McCullagh and Nelder.[18] This model assumes that anesthetic effect is measured by an underlying, continuous variable, denoted y, that is related to the drug concentration by the equation

$$y = \gamma \ln C - \gamma \ln C50 + \epsilon \quad (5)$$

where ϵ is a random variable that is assumed to have a logistic distribution. Because we describe the concentration-effect curve by two parameters (C_{50} and γ), we can arbitrarily assume that the logistic distribution of ϵ has a mean of zero and a standard deviation of $\pi^2/3$. The model is applicable to binary, "yes/no" data because we assume that a positive drug effect is observed only if $y \geq 0$ or equivalently if $\epsilon \geq -(\gamma \ln C - \gamma \ln C_{50})$. Thus the probability of drug effect is

$$P = \int_{-(\gamma \ln C - \gamma \ln C50)}^{\infty} \frac{e^{-x}}{(1 + e^{-x})^2} dx = C^{\gamma}/(C^{\gamma} + C50^{\gamma}) \quad (6)$$

(The integrand is the logistic density function.)

As an alternative, we propose using probit analysis. The basic model is again

$$y = \gamma \ln C - \gamma \ln C50 + \epsilon \quad (A\text{-}1)$$

but ϵ now has a normal distribution with a mean of zero and a variance of 1. Again we assume that a drug effect is observed only if $y \geq 0$. Also note that again the concentration-effect curve is characterized by two parameters, γ and C_{50} (which is why the normal distribution may be arbitrarily assumed to have a variance of 1). With this model the probability of observing a drug effect is

$$P = \frac{1}{\sqrt{2\pi}} \int_{-(\gamma \ln C - \gamma \ln C50)}^{\infty} e^{-x^2/2} dx = \Phi[\gamma \ln C - \gamma \ln C50] \quad (A\text{-}2)$$

where Φ denotes the cumulative standardized normal distribution.

Equation (A-2) is applicable to data from a single patient. For the analysis of pooled population data with one observation per individual, this expression must be generalized. We rewrite equation A-1 as

$$y_i = \gamma \ln C_i - \gamma \ln C50_i + \epsilon_i \quad (A\text{-}3)$$

where the subscript I indexes patients. Note that we assume that γ is the same for all patients. We also assume that C_{50} has a log-normal distribution; that is, $C_{50_i} = \langle C_{50} \rangle \exp(\delta_i)$ where $\langle C_{50} \rangle$ is the mean C_{50} value and δ_i is normally distributed with a mean of zero and in unknown variance ω^2 and is independent of ϵ_i. With these assumptions we can write

$$y_i = \gamma \ln C_i - \gamma \ln \langle C50 \rangle + (\gamma \delta_i + \epsilon_i) \quad (A\text{-}4)$$

Because the sum of two independent normally distributed variables is itself a normally distributed variable whose variance is the sum of the variances of its individual components, in this model y_i is a normally distributed variable with variance equal to $1 + (\gamma\omega)^2$. Because we observe a drug effect if $y_i \geq 0$, the probability of a drug effect in the ith individual for population analysis is

$$p = \frac{1}{\sqrt{2\pi}} \int_{-\frac{\gamma \ln C - \gamma \ln \langle C50 \rangle}{\sqrt{1 + (\gamma\omega)^2}}}^{\infty} e^{-x^2/2} dx = \Phi\left[\frac{\gamma \ln C - \gamma \ln \langle C50 \rangle}{\sqrt{1 + (\gamma\omega)^2}}\right] \quad (A\text{-}5)$$

(For convenience we have dropped the index I). Note that the probability of drug effect for this population model is now characterized by three parameters: $\langle C_{50} \rangle$, the mean value of C_{50} for individual patients, γ a measure of the steepness of the concentration-response curve, and ω^2, the variance of C_{50} values in the population.

Appendix 2

In this appendix we describe the implementation of population probit analysis using an Excel spreadsheet. This implementation is illustrated in figure 5. These instructions assume a computer with a mouse For personal computers, cells are highlighted by using the

	A	B	C	D	E	F
1	64	0	250	-1.33612	0.090755	-0.04132
2	85	0	5	-1.05786	0.14506	-0.06806
3	123	1	1	-0.6955	0.24337	-0.61373
4	139	0		-0.57559	0.282447	-0.14415
5	153	0		-0.48149	0.315085	-0.16436
6	203	0		-0.20421	0.419094	-0.23589
7	206	1		-0.18983	0.424723	-0.37189
8	217	0		-0.13881	0.444798	-0.25555
9	221	0		-0.1209	0.451884	-0.26113
10	226	0		-0.09897	0.460583	-0.26808
11	237	0		-0.05236	0.479119	-0.28326
12	241	1		-0.03595	0.48566	-0.31367
13	251	0		0.003914	0.501562	-0.30239
14	270	1		0.075467	0.530078	-0.27566
15	271	1		0.079092	0.53152	-0.27448
16	278	0		0.104099	0.541455	-0.33862
17	279	1		0.10762	0.542851	-0.26532
18	291	0		0.148913	0.559189	-0.35575
19	347	1		0.321497	0.626083	-0.20337
20	374	1		0.394973	0.653568	-0.18471
21	384	1		0.420847	0.663067	-0.17844
22	388	1		0.431009	0.666769	-0.17602
23	406	1		0.475476	0.682776	-0.16572
24	432	0		0.536343	0.704139	-0.52891
25	451	1		0.578549	0.718553	-0.14354
26	527	1		0.731258	0.767689	-0.11481
27	571	1		0.809889	0.790998	-0.10182
28	577	1		0.820139	0.793932	-0.10022
29	674	0		0.97251	0.834601	-0.78147
30	700	1		1.009625	0.843662	-0.07383
31	738	1		1.061462	0.85576	-0.06765
32	794	1		1.133181	0.871431	-0.05977
33	796	1		1.135648	0.871948	-0.05951
34						-7.77311

Fig. 5. A typical Excel spreadsheet for population probit analysis. The entries in this example are explained in appendix 2.

mouse to align the "✚" over a cell and striking the mouse button. Options from the menu above the spreadsheet itself are selected by using the mouse to align the "↗" over the option and clicking the mouse button.

1. Enter data in columns A and B. The concentrations at which observations of drug effect were made are entered in column A (cells A1:A33 in this example). The response variables are entered in column B (enter a "1" if a drug effect was observed and "0" otherwise).
2. Enter initial estimates, or guesses, of $\langle C_{50} \rangle$ in cell C1, of γ in cell C2, and of γ in cell C3.
3. Highlight cell D1. Type "= (c2*ln(a1) − c2*ln(c1))/((1 + (c2*c3)^2)^.5)". (The quotation marks are not included in the expression to be typed). Strike the Enter key. Now position the mouse over the lower right corner of cell D1 until the "+" symbol appears (distinguish this from the "✚" symbol). While depressing the mouse button, drag the mouse down column D, in this example to row 33. This will fill in cells D2:D33 with the above expression except that the drug concentrations in cells A2:A33 will be used (in place of the value in cell A1) for cells D2:D33.
4. Highlight cell E1. Select the "function wizard," which is the box labeled "f_x" in the menu above the spreadsheet. A menu of categories of functions will appear on the screen. Select "Statistical." A menu of statistical functions will appear. Scroll down to "NORMSDIST" and select this option. Then select the "Next" option. A line for inserting the argument of the function will appear. Type "d1" and strike the Enter key. This will cause the value of the cumulative standardized normal distribution of the value in cell D1 to appear in cell E1. Place the mouse over the lower right-hand corner of cell E1 until the "+" symbol appears. Keeping the mouse button depressed, drag the mouse down column E to fill in cells E2:E33 with the values of the normal distribution corresponding to cells C2:C33 and D2:D33.
5. Highlight cell F1. Type "= b1*log(e1) + (1 − b1)*log(1 − e1)" and strike the Enter key. Place the mouse over the lower right-hand corner of cell F1 until the "+" symbol appears, and while keeping the mouse button depressed drag the mouse down column F to fill in cells F2:F33.
6. Highlight the cell below the last cell entered in column F. In this example, this is cell F34. Select the icon labeled "Σ" in the menu above the spreadsheet and strike the Enter key. This will enter the sum of the values in column F (the log-likelihood values).
7. Select the "Tools" option in the menu at the top of the spreadsheet and then select the "Solver" option. Select "Set Target Cell" and designate the cell containing the sum of the log-likelihood values, in this case typing "f34". Select "Max" and then select "By Changing Cells" and type "c1,c2,c3". Strike the Enter key. This will result in maximum likelihood estimation by varying the values of $\langle C_{50} \rangle$, γ, and ω in cells C1:C3. This process should be repeated using multiple starting values in cells C1:C3. Convergence of estimation can be improved by constraining $\langle C_{50} \rangle$, γ, and ω to have positive values. This option is available in the Solver menu.