The Use of Clinical Trial Simulation to Support Dose Selection: Application to Development of a New Treatment for Chronic Neuropathic Pain

Peter A. Lockwood,[1,4] Jack A. Cook,[1] Wayne E. Ewy,[2] and Jaap W. Mandema[3]

Received June 4, 2003; accepted July 9, 2003

Purpose. Pregabalin is being evaluated for the treatment of neuropathic pain. Two phase 2 studies were simulated to determine how precisely the dose that caused a one-point reduction in the pain score could be estimated. The likelihood of demonstrating at least a one-point change for each available dose strength was also calculated.

Methods. A pharmacokinetic-pharmacodynamic (PK/PD) model relating pain relief to gabapentin plasma concentrations was derived from a phase 3 study. The PK component of the model was modified to reflect pregabalin PK. The PD component was modified by scaling the gabapentin concentration–effect relationship to reflect pregabalin potency, which was based on preclincal data. Uncertainty about the potency difference and the steepness of the concentration–response slope necessitated simulating a distribution of outcomes for a series of PK/PD models.

Results. Analysis of the simulated data suggested that after accounting for the uncertainty, there was an 80% chance that the dose defining the clinical feature was within 45% of the true value. The likelihood of estimating a dose that was within an acceptable predefined precision range relative to a known value approximated 60%. The minimum dose that should be studied to have a reasonable chance of estimating the dose that caused a one-point change was 300 mg.

Conclusions. Doses that identify predefined response may be imprecisely estimated, suggesting that replication of a similar outcome may be elusive in a confirmatory study. Quantification of this precision provides a rationale for phase 2 trial design and dose selection for confirmatory studies.

KEY WORDS: clinical trial simulation; pharmacokinetic/pharmacodynamic model; dose estimation.

INTRODUCTION

Dose selection for phase 2 and phase 3 studies is a challenging issue in drug development. The phase 2 goal is to learn if a drug is safe and effective and how these endpoints vary with dose so that the performance of optimal doses can be confirmed in phase 3 studies. Often the highest dose studied in phase 2 is set by a maximally tolerated dose based on safety and tolerance results from phase 1. On the other hand, the criteria for setting the lowest and intermediate doses are often not explicitly defined. For the lowest dose, the target is often to use a dose below which the drug is unlikely to produce a clinically useful effect (no-effect dose) and the intermediate dose(s) to lie somewhere between the lowest and highest doses. However, such a strategy may not result in an optimally designed study to select doses for phase 3 trials and be predisposed toward a high risk of having to conduct multiple phase 2 studies in the event that the outcomes are ambiguous. For example, two different doses may be studied but not show a difference in effect. So does this imply that there truly was no difference in effect or that the study design was not sensitive enough to demonstrate a difference? Such ambiguity about the dose-response requires resolution before more expensive phase 3 trials are done to avoid the risk of studying a less than optimum dose, and this contributes to increased development costs and extended development time. On other occasions doses may be selected for a phase 3 study based on an apparent dose–response relationship derived from phase 2, which when studied in a confirmatory study demonstrates an effect of lesser magnitude. This may result in additional clinical studies to resolve the paradox, thereby delaying registration and escalating development costs. Development costs for a new chemical entity have recently been cited as being around US$800 million, with much of the cost attributable to rising clinical trial costs (1). Further failure to fully elucidate dose-response likely contributed to systematic overdoses of a number of marketed compounds, resulting in unnecessary harm to patients and subsequent labeling changes (2,3).

These pitfalls might be avoided if, at the time of phase 2 planning, there is a clear understanding of the desired clinical response(s) and how precisely the dose producing this response can be estimated from a trial design under different assumptions of drug pharmacokinetics (PK) and pharmacodynamics (PD). For example, if the desired clinical response was half of the maximum obtainable response possible, and it was shown that the corresponding dose (i.e., the ED_{50}) could not be estimated with any real precision, there would be less inclination to base dose selection for a confirmatory study on the results of such a phase 2 study. The likely consequence would be a modification of the study design, enabling the ED_{50} to be estimated more precisely. If on the other hand the ED_{50} was shown to be estimated with acceptable precision, then this would suggest that the design was well suited to elucidate dose-response and would substantiate dose selection for phase 3 based on the phase 2 study outcome.

A variation to this concept of evaluating a trial design by quantifying the precision of dose selection is the estimation of the likelihood that a trial design can correctly identify the dose that causes the response of interest. When planning clinical studies, some measure of how often a given design would correctly determine the "right" dose would be useful for guiding the design and building confidence in the study outcome. In this case, determination of the "right" dose necessitates calculating the probability that the dose falls within a prespecified range that corresponds to acceptable precision about a known or true value. Reverting back to the ED_{50} example, if there is a high likelihood (e.g., 80%) that the ED_{50} falls within a desired precision (e.g., 35% of the true value), then this suggests the study design is suitable to estimate the

[1] Clinical Pharmacokinetics and Pharmacodynamics, Pfizer Global Research and Development, Ann Arbor, MI.
[2] Department of Clinical Biostatistics, Pfizer Global Research and Development, Ann Arbor, MI.
[3] Pharsight Corporation, Mountain View, CA.
[4] To whom correspondence should be addressed. (email: peter.lockwood@pfizer.com)

0724-8741/03/1100-1752/0

dose producing the response of interest. In this context, the study could be described as adequately powered, and therefore, planning future studies on the outcome of the actual study would be reasonable. If this probability were lower, e.g., 40%, it would suggest the study was not well powered and that additional work or changes to the study design were necessary to increase the chances of characterizing the dose corresponding to the response of interest, so that confirmation would be optimized in future phase 3 studies.

Quantification of the precision for estimating a dose that causes a desired clinical effect and estimation of the likelihood that a dose falls within a prespecified precision range (i.e., determining the likelihood that a design will correctly estimate the right dose) are measures of trial performance and provide a means of comparing the efficiency of different study designs to satisfy the objectives. These analyses can be accomplished before commencement of the phase 2 study paradigm through the implementation of a clinical trial simulation (CTS) strategy that takes into account the variability associated with patient population kinetics, dynamics, and compliance. When a particular design is simulated on multiple occasions, a probability distribution of trial outcomes is generated, and the subsequent evaluation of these distributions can then provide some insight into the trial performance and thus offer a rational basis for making decisions about different aspects of a clinical study design given the different realms of uncertainty (4,5). Without the aid of a simulation strategy to evaluate the impact of uncertainty in key assumptions or to determine the elements of trial designs that are critical to a successful outcome, scientifically and commercially efficient drug development programs will continue to be elusive.

In this paper, a clinical trial simulation study is presented that evaluated how well two 8-week parallel group phase 2 studies could identify doses that corresponded to a response of clinical interest for pregabalin [CI-1008, (S)-3-isobutyl GABA, (S)-(+)-3-(aminomethyl)-5-methylhexanoic acid], a new drug under development for multiple indications, which include pain relief in diabetic neuropathic patients. Diabetic neuropathy is defined as peripheral somatic or autonomic nerve damage attributable solely to diabetes mellitus, and the symptoms of the disease may include severe painful and persistent burning, thermal hyperalgesia, prickling sensations, insomnia, weight loss, anxiety, and depression. The response of interest was a one-point reduction in pain intensity on an 11-point numerical pain rating scale. The dose that produced this effect was defined as the minimum effective dose (MED), as this change was considered the minimum clinically significant change (6).

There was a body of preclinical information about the potency of pregabalin compared to gabapentin, another drug shown to be clinically effective in diabetic neuropathy, which enabled a conservative, best, and optimistic guess at the clinical effectiveness of pregabalin. The information was linked through a series of mathematical models, and the potential inaccuracies of the model predictions resulting from the uncertainty of the "true" model parameter values were incorporated into the decision framework for evaluating the trial performance. As the simulation study was completed after protocol finalization, the outcome of the simulation study failed to impact the development program.

The objective of this simulation study was to quantify how precisely the MED could be estimated from the proposed phase 2 plan and the likelihood that the MED was correctly identified from the plan (i.e., how often the estimated MED was within a prespecified range of a known true value).

METHODS

Derivation of the Simulation Models

The models used in the simulation study were based on a precursor pharmacokinetic/pharmacodynamic (PK/PD) model for gabapentin. The gabapentin PK/PD model was derived from a placebo-controlled phase 3 study of gabapentin efficacy in neuropathic pain (7) and incorporated components that described the drug PK and PD, a placebo effect, and subject dropout rate. There were insufficient data to precisely estimate the slope of the concentration–response relationship, and this led to the derivation of three different precursor models. The models reflected a conservative, a most likely (best guess), and an optimistic description of the concentration–response data. The final set of simulation models was obtained by modifying the precursor models to reflect the pregabalin drug effect. Replacing the gabapentin PK component with a PK model of pregabalin and scaling the potency parameter (EC_{50}) according to the relative difference in potency between the two compounds accomplished this. The scaling of the potency was based on preclinical models of the drug's effects.

Precursor PKPD Models Based on Gabapentin

PK Model

Average gabapentin plasma concentrations were estimated for each subject using a single-compartment model with zero-order input [Eq. (1)] and individual measures of weight and creatinine clearance, collected at the outset and during the course of the study. Pharmacokinetic studies in humans have shown that the apparent volume of distribution is proportional to weight (9) and that gabapentin, which is cleared only by renal excretion (8,9), is dependent on renal function. Gabapentin renal clearance was shown to be linearly related to creatinine clearance in a single-dose PK study in 60 subjects with various degrees of renal failure (10). Therefore, renal clearance was derived from creatinine clearance (ClCr) based on individual serum creatinine measurements and application of the Cockcroft and Gault equation (11) and was subsequently used in the estimation of average gabapentin plasma levels. Previous studies also demonstrated that plasma gabapentin concentrations increased in a less than proportionate manner with increasing oral doses, and this resulted in bioavailability estimates ranging from 60% to 30% over a dose range of 300 to 1200 mg t.i.d. (9,12). Therefore, the relationship between amount of gabapentin administered and that reaching the systemic circulation was described by Eq. (4).

$$C_{avg} = (Dose_{app}/\tau)/Cl[1 - e^{-\left(\frac{Cl}{V}\right)\cdot t}] \quad (1)$$

Where

$$Cl = (ClCr \cdot 1.06 + 0.84) \quad (2)$$

$$V = 0.8 \cdot wt \quad (3)$$

$$Dose_{app} = 823 \cdot Dose/(Dose + 1120) \quad (4)$$

In Eqs. (1)–(4), C_{avg} = average plasma concentration between peak and trough levels (μg/ml), CL = oral clearance (L/day), ClCr = creatinine clearance (L/day), D_{app} = apparent dose (mg), V = volume of distrubution (L), wt = weight (kg), τ = dosing interval (h), t = time (day). [Equations (2) and (3) are from references 9 and 10.]

PD Model

The gabapentin PD model was derived from a double-blind placebo-controlled trial of gabapentin for the treatment of painful diabetic peripheral neuropathy (7). Following a 1-week screening phase, patients were randomized to receive gabapentin or placebo according to an 8-week, double-blind, parallel-group, multicenter study design. Gabapentin (or matching placebo) was titrated from 900 to 3600 mg/day divided into three daily doses (t.i.d.) during the first 4 weeks of the double-blind phase. In the second 4 weeks, dosage was maintained at the maximum dose that was tolerated by each patient. A total of 165 patients were randomized to treatment: 84 received gabapentin, and 81 received placebo. The primary efficacy measure was daily pain severity as measured on an 11-point Likert scale (0 = no pain; 10 = worst possible pain). A random-effects model was used to characterize the relationship between daily pain score and gabapentin exposure in individual patients taking into account placebo, time, and baseline effects. The modeling of the clinical trial data suggested a drug and placebo effect [Eq. (5)] that were dependent on the baseline pain response. Drug effect was also dependent on average drug concentration, and this relationship was described by a sigmoid E_{max} function (13). Random effects were associated with the placebo onset rate constant (k_{pl}), the baseline score *(Base),* and a scaling factor describing the magnitude of the placebo effect *(PLM)* using a proportional error model. Random effects were assumed to follow a multivariate normal distribution with mean 0 and diagonal variance-covariance matrix Ω, with diagonal elements ($\omega_1^2 \ldots \omega_m^2$) (15). An additive residual error component was included in the model.

$$E = Base\left(1 + (PLM) \cdot (1 - e^{-k_{pl} \cdot t}) + \frac{E_{max} \cdot C_{avg}^n}{C_{avg}^n + EC_{50}^n}\right) + \varepsilon \quad (5)$$

where E = pain score (0-10), Base = pretreatment baseline score (0-10), PLM = the magnitude of the placebo effect, k_{pl} = the first order rate constants describing the onset of the placebo effect ($days^{-1}$), t = time (days), E_{max} = the maximal drug effect, EC_{50} = the concentration at which the effect due to drug is 50% of E_{max} (μg/ml), C_{avg} = the average gabapentin concentration based on dose, and estimated clearance (μg/ml), n = Hill Coefficient or slope of dose–response relationship, and ε = the within-subject random effect. Between-subject random effects were associated with Base, PLM, and k_{pl}.

The models were fitted to the data with the nonlinear regression program NONMEM (version V) using the first-order method (14). Initial fitting of the data with unconstrained parameters resulted in a maximal drug effect (E_{max}) that approximated a 20% reduction in pain score and a slope parameter equal to 10 that suggested a near-vertical concentration–response slope (Model 1). Constraining the slope parameter to reflect a nonlinear monotonically increasing concentration–response profile with an asymptote (slope parameter = 1) yielded a maximal drug effect that approximated a 50% reduction in pain score (Model 2). The objective function value (OFV) for both of these models was equivalent (OFV = 13018), which suggested that the models characterized the gabapentin neuropathy trial data equally well. Diagnostic plots indicated that the average response in the placebo-treatment group was well characterized but was underpredicted in the treatment group with Model 1 and overpredicted with Model 2. Therefore, an additional set of parameters characterizing the drug effect (Model 3) was derived through a series of manual iterations by assigning values of E_{max}, EC_{50}, and the slope to reflect a concentration–response profile intermediate to those obtained from Models 1 and 2, and to more accurately describe the average pain score over time for the subjects who received gabapentin (Fig. 1, right panel). The final values for Model 3 were E_{max} = −0.25, EC_{50} = 5.35 μg/ml, and n = 2. Model 1 was considered the most conservative PKPD model, Model 2 the most optimistic, and Model 3 the most likely or best guess. The relationship between the drug effect characterization and the concentration range over which the response was evaluated is displayed in Fig. 2. There was a paucity of data outside of this concentration range that would explain why the data were equally well described by Models 1 and 2. Figure 1 compares the average response over time for the placebo and drug-treated groups with the model predictions.

Dropout Model

A dropout model describing a 3% weekly (0.43% day) dropout rate was derived based on a review of dropout data from the gabapentin study. Dropout rates were approximately the same in both groups. This was implemented in the simulation model as a survival model according to the following equation:

$$S(t) = \exp[-\int_0^t h(x)dt] \quad (6)$$

where

$$\int_0^t h(x)dt = 0.0043 \quad (7)$$

For each patient, a uniform (0,1) random variate was generated for each study day, and the patient was dropped from the study on the first day that a random variable less than 0.0043 was obtained.

Final Models

The precursor PK/PD models were modified to reflect pregabalin PK and the likely PD. Mean and variance estimates of pregabalin clearance and volume were obtained from an oral rising single dose tolerance and pharmacokinetic study in healthy volunteers (15) and replaced those of gabapentin. Individual plasma concentrations were simulated using a one-compartment model and measures of pregabalin clearance and volume of distribution that were derived by random sampling with replacement from log normal distributions with a mean (standard deviation) of 4.4 (0.14) L/h and

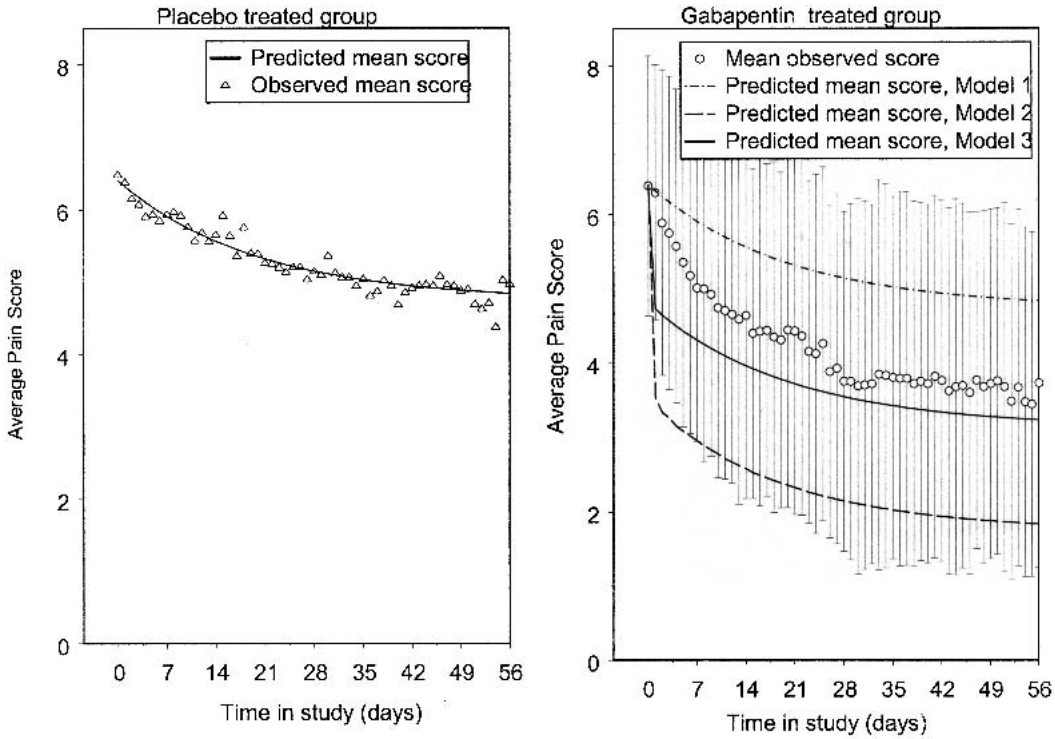

Fig. 1. Average pain score observations and model predictions for the placebo and gabapentin treatment groups. Error bars represent 1 standard deviation. Simulation was run 500 times.

36.6 (0.18) L, respectively. $Dose_{app}$ was set to equal the actual dose because pregabalin is essentially completely absorbed across the entire dose range. The gabapentin potency parameter (EC_{50}) was multiplied by a constant based on the pregabalin/gabapentin potency ratios derived from preclinical models of pain relief in inflammatory (formalin footpad test, carageenan-induced thermal and mechanical hyperalgesia model), surgical (plantaris muscle incision model), or neuropathic pain (streptozocin model of neuropathic pain) (16–21). Preclinical pain relief models indicated that the maximum drug effect of pregabalin was the same as that of gabapentin, and the average potency difference based on the EC_{50}s was approximately 3, but estimates from different models ranged from 2 to 4. Thus, the gabapentin PK/PD models were scaled based on a conservative (scaling factor = 2), most likely (scaling factor = 3), and optimistic (scaling factor = 4) estimate of the difference in potency of the two compounds.

Simulation Methods

The uncertainty in both the concentration–response slope and potency scaling parameter resulted in the use of five distinct models for the simulations (SM). In models SM1, SM2, and SM3, the scaling factor was set to equal 3 (most likely scaling factor) while the slope of the concentration–response relationship was varied to span the range of uncertainty (i.e., 1–10). In models SM4 and SM5, the concentration–response slope was assumed to equal 2 (most likely slope), while the scaling factor was set to 2 and 4, respectively. Overall, SM3 was considered the most likely representation of the pregabalin PK/PD (best guess model) because the concentration-effect shape was intermediate to those defined by SM1 and SM2, and the scaling factor was the average value obtained from preclinical data. Changing parameter estimates used in the simulation models are listed in Table I, and fixed parameter estimates are listed in Table II.

For each of the two phase 2 study designs listed in Table III, daily pain scores were simulated in 5000 patients for the five pregabalin PK/PD models using the Pharsight Trial Simulator, version 1 (22). The simulated outcome measure was converted to a discrete value on a 0 to 10 scale by rounding to the nearest integer. The primary measure of drug effectiveness was the difference in the mean daily pain score at the last week of treatment relative to the mean daily baseline score (baseline-adjusted pain score). The two sets of simulated data were pooled for each set of model assumptions, and the MED was derived by fitting the mean baseline adjusted pain score for each treatment group with a sigmoid E_{max} model (S-Plus ver 4.0) (23) and then calculating the MED on the basis of the E_{max} model parameter values. These estimates represented the true value of the dose that corresponded with the desired predefined clinical response in the virtual patient population.

For each study, 80 patients/dose-group were sampled with replacement from the virtual population on 500 occasions, each replicate reflecting a simulated clinical study. For each replicate, the mean baseline adjusted pain score was determined at each dose, and the means from both studies were then pooled. The pooled means were fitted with a sigmoid E_{max} model (S-Plus ver 4.0), and the MED was subsequently calculated from the model parameters. Trial performance metrics were calculated on the basis of the distribution of estimated MED values. To estimate the precision with which the MED could be estimated with 80% confidence, the 10th and 90th percentile values were obtained from the dis-

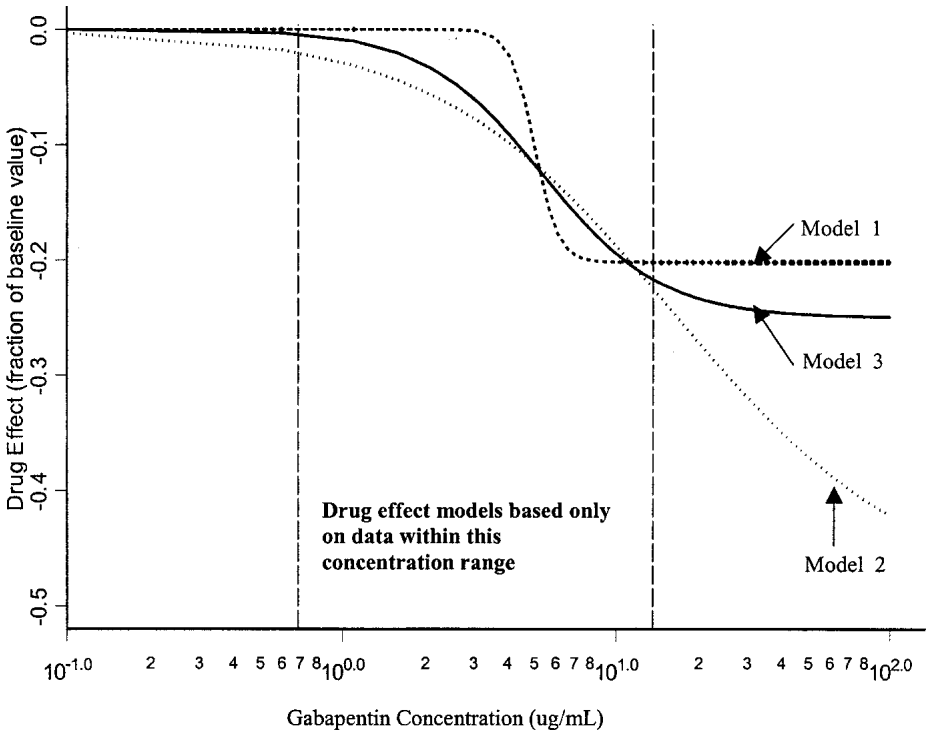

Fig. 2. Concentration–response profiles for the drug effect component of the pharmacodynamic models used in the simulation studies. Models of drug effect (Model 1 and 2) were derived from a clinical study of gabapentin and represented the data equally well. Model 3 was considered the most likely representation of the drug effect. The figure displays the concentration range containing the response data on which the drug effect models were based. The paucity of response data outside of this region resulted in models 1 and 2 characterizing the data equally well.

tribution for each set of model assumptions. The bias of the outcome relative to the true value was determined from the 50th pecentile. The percentage of studies that estimated the MED to be within 35% of the true value was also calculated, with the rationale for the cutoff value being that there was approximately a twofold difference between the upper and lower limits. This difference was arbitrarily selected for the purposes of the simulation study.

RESULTS

The true MED values for all five models and the 10th, 50th (median), and 90th percentiles of the MED distribution are displayed in Table IV. Percentiles are expressed as a fraction of the true value for the different assumptions about the steepness of the concentration–response relationship and the size of the scaling factor. The predicted MEDs were, on average, slightly negatively biased, with the MED being within 10% of the true value, depending on the model assumptions. The precision of the MED relative to the true value varied among models. Across all models (SM1–SM5), the precision of the MED estimate was within 45% of the true value 80% of the time.

The power to estimate the MED to be within 35% of the true value for the different models is displayed in Table V. The results indicate that the power increases with increasing slope under the conditions of the most likely scaling factor (comparing SM1, SM2, and SM3). The power to estimate the MED assuming the most likely estimates of scaling factor and concentration effect slope was 61% across the range of assumption uncertainty.

Table I. Values of Varied Parameters for Each Simulation Model

Model no.	Slope factor (n)	Scaling factor	Pregabalin EC_{50}	E_{max}
1	10	3	1.67	–0.202
2	1	3	5.33	–0.489
3	2	3	5.35	–0.25
4	2	2	8.03	–0.25
5	2	4	4.01	–0.25

E_{max} = the maximal drug effect.
EC_{50} = the concentration at which the effect of the drug is 50% of E_{max} (μg/ml). The EC_{50} for pregabalin was derived by dividing the gabapentin EC_{50} by the scaling factor.
n = Hill coefficient or slope of dose–response relationship.
Model 3 was considered the most likely model to predict the pregabalin outcome.

Table II. Mean (Variance) of Fixed Parameters in Simulation Study Models

Baseline	PLM	k_{pl}	σ^2
6.4 (0.05)	–0.26 (0.12)	0.05 (1.8)	1.2

Base = pretreatment baseline score (0–10).
PLM = the magnitude of the placebo effect.
k_{pl} = the first-order rate constants describing the onset of the placebo effect ($days^{-1}$).
σ^2 = residual variance.

Table III. Dose, Treatment Duration, and Outcome Variable of Simulated Studies

Protocol	Baseline duration (weeks)	Titration duration (weeks)	Stable treatment duration (weeks)	Daily dose (mg)	Primary outcome
Study 1	1	2	4	0,150,600	Weekly mean pain score
Study 2	1	1	4	0,75,300,600	Weekly mean pain score

DISCUSSION

This simulation project quantified the reliability of a phase 2 program to identify the dose that caused a predetermined response. The reliability was expressed in terms of a confidence interval for quantifying how often the characterizing dose was estimated to be within a certain range of the true value. The results indicated that over a range of assumptions in drug potency and concentration–response slope, the dose defining the features of interest was within 45% of the true value 80% of the time, and the likelihood of the program to estimate the dose with acceptable precision approximated 50% and ranged from 40% to 70% depending on the key assumptions.

This suggests that the identification of the selected dose-response feature with any real precision from the trial design paradigm is borderline. Therefore, if the objective was to confirm the outcome in a future phase 3 study, selecting a dose based on this single outcome might be "risky." Borderline precision should promote consideration of the rationale underlying what is acceptable and how to improve the precision for identifying a dose that defines a feature of interest. Some options might include changing the number of treatments, the number of efficacy measurements, the desired clinical response, or the acceptance limits. Further options might be the gathering of additional information to reduce the uncertainty associated with key assumptions that influence how reliably a dose can be estimated. Investigating the impact of these options or combinations thereof may assist in identifying key failure points in a development program that, if remedied, may optimize decision making and enable the selection of effective doses for confirmatory studies with a high level of confidence.

The marginal precision raises the question as to what is the best dose to study to ensure a clinical outcome of at least a one-point change in pain score, given the dosing options available. This question was elucidated by evaluating the MED distribution derived from the simulations that characterized the probability of detecting a one-point change over a range of doses, with the true value having a probability of 0.5. In this simulation study the dosing options were placebo, 75 mg, 150 mg, 300 mg, 450 mg, and 600 mg, and the percentage of simulated studies that estimated the MED to lie within these dose ranges are listed in Table VI. This analysis indicates that there is a 70% likelihood of detecting at least a one-point change at 300 mg, with the probability increasing to 90% at 450 mg. Thus, it would be prudent to study at least a dose of 300 mg to confirm the drug effect of at least a one-point improvement in pain score.

The process described addressed different aspects needed to apply CTS to optimize clinical trial design and development strategies. First, the concept of building a model based on a variety of data sources was described. The model provided the framework to incorporate knowledge from different sources about the drug effect, the placebo effect, patient dropout, and the variability associated with the different model components. The dropout model was included to reflect the assumption that patients would drop out completely at random, independent of the treatment, and that the dropout rate would be similar to what was observed in the gabapentin study. If the dropout rate were greater than what was assumed, then the precision of the MED would be overestimated.

Building a family of models enabled the impact of key assumption uncertainty to be investigated with the resultant conclusion that neither the uncertainty in potency ratio or dose–response slope significantly impacted the precision of

Table IV. Distribution of Minimum Effective Dose (MED) Across the Range of Uncertainty in Concentration–Response Slope Value and Scaling Factor

Model	Scale factor	Slope value	MED true value (mg)	Percentiles of MED distribution 10%	50%	90%
SM1	3	10	215	0.82	1.03	1.48
SM2	3	1	285	0.58	0.90	1.38
SM3*	3	2	275	0.59	0.88	1.47
Average				0.66	0.94	1.44
SM4	2	2	418	0.55	0.88	1.38
SM3*	3	2	275	0.59	0.88	1.47
SM5	4	2	202	0.59	0.92	1.49
Average				0.58	0.89	1.45

Tabular values are expressed as fraction of the true value and were derived from analysis of two proposed phase 2 study designs, which were simulated 500 times. Asterisk (*) represents most likely model.

Table V. Power of the Study Designs to Estimate the Minimum Effective Dose (MED) to Be Within 35% of the True Value Across the Range of Uncertainty in the Scaling Factor and Concentration–Effect Slope

Model	Percentage of studies with MED estimate less than 35% of true value	Percentage of studies with MED estimate greater than 35% of true value	Total percentage of studies with MED within 35% of true value
SM1	8	24	68
SM2	56	1	43
SM3*	6	33	61
SM4	12	41	47
SM5	47	3	50
Average			54

Tabular values are expressed as a percentage of the total number of MED estimates derived from the analysis of two proposed phase 2 studies, which were simulated 500 times. Asterisk (*) represents most likely model.

Table VI. Power to Estimate at Least a One-Point Change for the Available Dose Strengths

Model	<75 mg	75 to <150 mg	150 to <300 mg	300 to <450 mg	450 to <600 mg	>600 mg
SM1	8 (8)	1 (9)	76 (85)	13 (98)	1 (99)	1 (100)
SM2	1 (1)	7 (8)	58 (66)	29 (95)	5 (100)	0 (100)
SM3	2 (2)	6 (8)	61 (69)	25 (94)	5 (99)	1 (100)
SM4	10 (10)	1 (11)	24 (35)	39 (74)	18 (92)	8 (100)
SM5	2 (2)	26 (28)	62 (90)	8 (98)	1 (99)	1 (100)
Average	5 (5)	18 (13)	56 (69)	23 (92)	6 (98)	2 (100)

Power is based on the percentage of MED estimates that were within the listed dose range. The MED distribution was derived from the analysis of two proposed phase 2 study designs, which were simulated 500 times. Values in parentheses indicate the cumulative probability as the dose range increases.

the phase 2 program under consideration to estimate doses that characterized dose-response features. Had uncertainty in one key assumption influenced the outcome to a greater extent than the other, additional preclinical studies could have been conducted or the data reanalyzed in an attempt to reduce the uncertainty associated with that particular assumption. This notion reflects how CTS might, therefore, influence preclinical development of new drugs and promote greater liaison between preclinical and clinical development teams. Gomeni *et al.* (24) provide another example of this integration of PK and PD knowledge and the use of CTS to design and proof concept trials for the first time in humans. Also, in retrospect the authors acknowledge that a posterior predictive check (PPC) would have been useful before conducting the simulations. The PPC quantifies how well the model predicted the data on which it was based, thus providing some measure of the predictive power of the model. Although a PPC was not conducted for this simulation study, the author's recommend this be conducted in future simulation studies.

Another aspect discussed was the use of metrics to quantify how well the study design could identify clinically relevant endpoints. This simulation study assisted in the identification of an important clinical endpoint and prompted consideration of how precisely the dose corresponding to this clinical feature should be estimated relative to the known true value. This defined what constituted "acceptable precision," which then served as a marker of trial performance; i.e., how many of the simulated trials estimated the dose that corresponded to the clinical feature to fall within the acceptable precision range. Thus, defining the desired dose-response features and acceptable precision and confidence estimates at the outset of a clinical development program is an integral part of a clinical trial simulation program and leads to decision making based more on information and less on empiricism, thereby optimizing the overall development program.

The simulation program provided some key findings. Noteworthy was that CTS can be complex and time consuming. Therefore, if this technology is to become effective in drug development, model development should begin as early as possible and best practices implemented to facilitate this. Furthermore, this work served to educate by demonstrating across disciplines that the power of the simulation was not just in simulating potential outcomes but in focusing on what specifically the phase 2 program was designed to accomplish and subsequently examining the impact of assumption uncertainty on these objectives, thus removing some of the empiricism in dose selection for future studies.

REFERENCES

1. Tufts Center for the Study of Drug Development. Tufts CSSD new drug development cost study. *www.tufts.edu/med/cssd/Nov30CostStudy/PressRelease.html* (2001).
2. J. Cross, H. Lee, A. Westelinck, J. Nelson, C. Grudzinskas, and C. Peck. Postmarketing drug dosage changes of 499 FDA-approved new molecular entities. *Pharmacoepidemiol. Drug Safety* **11**:439–446 (2002).
3. E. R. Heerdink, J. Urquhart, and H. G. Leufkens. Changes in prescribed drug doses after market introduction. *Pharmacoepidemiol. Drug Safety* **11**:447–453 (2002).
4. N. H. G. Holford, H. C. Kimko, J. P. R. Monteleone, and C. C. Peck. Simulation of clinical trials. *Annu. Rev. Pharmacol. Toxicol.* **40**:209–234 (2000).
5. P. L Bonate. Clinical trial simulation in drug development. *Pharm. Res.* **17**:3252–3256 (2000).
6. J. T. Farrar, J. P. Young, Jr., L. LaMoreaux, J. L. Werth, and R. M. Poole. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain* **94**: 149–158 (2001).
7. M. Backonja, A. Beydoun, K. R. Edwards, S. L. Schwartz, V. Fonseca, M. Hes, L. LaMoreaux, and E. Garofalo. Gabapentin for the symptomatic treatment of painful neuropathy in patients with diabetes mellitus: a randomized controlled trial. *JAMA* **280**: 1831–1836 (1998).
8. B. Schmidt. Potential antiepileptic drugs; gabapentin. In: R. Levy, R. Mattson, B. Meldrum, J. K. Penry, and F. E. Dreifuss (eds). *Antiepileptic Drugs,* 3rd ed. Raven Press, New York, 1989, pp. 925–935.
9. H. N. Bockbrader. Clinical pharmacokinetics of gabapentin. *Drugs Today* **31**:613–619 (1995).
10. R. A. Blum, J. Comstock, D. A. Sica, R. W. Schultz, E. Keller, P. Reetze, H. Bockbrader, D. Tuerck, J. Busch, P. A. Reece, and A. J. Sedman. Pharmacokinetics of gabapentin in subjects with varying degrees of renal function. *Clin. Pharmacol. Ther.* **56**:154–159 (1994).
11. D. R. Luke, C. E. Halstenson, J. A. Opsahl, and G. R. Matzke. Validity of creatinine clearance estimates in the assessment of renal function. *Clin. Pharmacol. Ther.* **48**:503–508 (1990).
12. B. H. Stewart, A. R. Kugler, P. R. Thompson, and H. N. Bockbrader. A saturable transport mechanism in the intestinal absorption of gabapentin is the underlying cause of the lack of proportionality between increasing dose and drug levels in plasma. *Pharm. Res.* **10**:276–281 (1993).
13. N. H. G. Holford and L. B. Sheiner. Understanding the dose–effect relationship: clinical application of pharmacokinetic and pharmacodynamic models. *Clin. Pharmacokinet.* **6**:429–453 (1981).
14. S. L. Beal and L. B. Sheiner. *NONMEM users guides, part I–VIII.* University of California, San Francisco, 1996.
15. J. A. Busch, J. A. Strand, E. L. Posvar, H. N. Bockbrader, and L. L. Radulovic. Pregabalin (CI-1008) single-dose pharmacokinetics and safety/tolerance in healthy subjects after oral administration of pregabalin solution or capsule doses. *Epilepsia* **39**(Suppl 6):58 (1998).
16. D. Dubuisson and S. G. Dennis. The formalin test:A quantitative study of the analgesic effects of morphine, meperidine, and brain stem stimulation in rats and cats. *Pain* **4**:161–174 (1977).

17. H. Wheeler-Aceto and A. Cowan. Standardization of the rat paw formalin test for the evaluation of analgesics. *Psychpharmacol.* **104**:35–44 (1991).
18. K. Hargreaves, R. Dubner, F. Brown, C. Flores, and J. Joris. A new and sensitive method for measuring thermal nociception in cutaneous hyperalgesia. *Pain* **32**:77–88 (1988).
19. L. O. Randall and J. J. Selitto. A method for measurement of analgesic activity on inflamed tissue. *Arch. Int. Pharmacodyn.* **111**:409–419 (1957).
20. T. J. Brennan, E. P. Vandermeulen, and G. F. Gebhart. Characterization of a rat model of incisional pain. *Pain* **64**:493–501 (1996).
21. M. J. Field, S. McCleary, J. Hughes, and L. Singh. Gabapentin and pregabalin, but not morphine and amitriptyline, block both static and dynamic components of mechanical allodynia induced by streptozocin in the rat. *Pain* **80**:391–398 (1999).
22. *Pharsight Trial Simulator Version 1.* Pharsight Corporation, Mountain View, California, 1998.
23. *S-Plus 4 for Windows.* Insightful Corporation, Seattle, Washington, 1998.
24. R. Gomeni, C. Falcoz, C. D'Angeli, and A. Bye. *In-silico* prediction of drug properties in man using preclinical data and computer assisted drug development. *Drug Inf. J.* **35**:1047–1063 (2001).