# Modeling and Simulation for Clinical Trial Design Involving a Categorical Response: A Phase II Case Study with Naratriptan

**Ivan Nestorov,[1] Gordon Graham,[1,5] Stephen Duffull,[3] Leon Aarons,[2] Eliane Fuseau,[4] and Peter Coates[4]**

*Received April 6, 2001; accepted May 4, 2001*

***Purpose.*** The overall aim of the present study was to investigate retrospectively the feasibility and utility of model-based clinical trial simulation as applied to the clinical development of naratriptan with effect measured on a categorical scale.

***Methods.*** A PK-PD model for naratriptan was developed by using information gathered from previous naratriptan and sumatriptan preclinical and clinical trials. The phase IIa naratriptan data were used to check the PK-PD model in its ability to describe future data. A further PK-PD model was developed by using the phase IIa naratriptan data, and a phase IIb trial was designed by simulation with the use of Matlab. The design resulting from clinical trial simulation was compared with that derived by using D-optimal design.

***Results.*** The PK-PD model showed reasonable agreement with the data observed in the phase IIa naratriptan clinical trial. Clinical trial simulation resulted in a design with four or five arms at 0 mg, 2.5 and/or 5 mg, 10 mg, and 20 mg, PD measurements to be taken at 0, 2, and 4 or 6 h and at least 150 patients per arm. A sub-D-optimal design resulted in two dosing arms at 0 and 10 mg and PD measurements to be taken at 1 and 2 h.

***Conclusions.*** Clinical trial simulation is a useful tool for the quantitative assessment of the influence of the controllable factors and is the only tool for the quantitative assessment of the uncontrollable factors on the power of a clinical trial.

**KEY WORDS:** simulation; design; clinical trials; naratriptan; pharmacokinetics; pharmacodynamics.

## INTRODUCTION

There is growing pressure from both within and outside the pharmaceutical industry to improve the efficiency of the clinical phases of drug development. In one report from the Food and Drug Administration (FDA) (1) for 12 NDAs evaluated in 1994–1995, the median number of clinical trials per NDA was 68, of which 27 were efficacy trials. A substantial fraction of the trials in these NDAs were adjudged by the FDA or the drug's sponsor to be seriously flawed or failed on scientific or performance grounds.

By analogy with large engineering projects, simulations can be used in situations in which the actual system or process is expensive or involves questions of safety and/or practices prescribed by law or regulatory bodies. Experiments performed on the simulation model can often provide valuable insights for the real system. There is recent and growing interest in using computer simulation of clinical trials to improve clinical trial design (2). Simulating clinical trials, applying techniques commonly used in other technology-based industries, such as the aerospace industry, can allow clinical trial designers to thoroughly test their designs and analyze the simulated results before actually conducting a clinical trial. Such careful scrutiny of a design can help expose weaknesses and identify reliance on assumptions in a plan. The result is that trials can be evaluated for robustness under various scenarios. The choices made by using this process should improve the chances of successful trials and of obtaining the necessary information for the product label.

Biologic systems are subject to many sources of variability. Consequently, clinical trial simulation involves Monte Carlo procedures addressing the typical behavior of the system and the departures from it. A recent joint report of two scientific conferences, organized by FDA, EUFEPS (European Federation for Pharmaceutical Sciences), ASCPT (American Society for Clinical Pharmacology and Therapeutics), AAPS (American Association of Pharmaceutical Scientists), and ACCP (American College of Clinical Pharmacy) (3), states that computer-assisted clinical trial modeling and simulation has achieved a sound conceptual basis by using formal pharmacokinetic and pharmacodynamic modeling principles, now that capable software has become available. It is our opinion, however, that more systematic research is needed both in the generic and methodologic issues, and the practical applications of clinical trial simulation and design, to gain evidence of the feasibility and utility of this approach.

It is recognized (4,5) that a clinical trial simulation and design (CTS&D) exercise is based on three major components: (i) a set of models, (ii) a clinical trial protocol, and (iii) adequate resources. Therefore, because simulation is a natural progression from the increased use of mathematical models (4) and the results can be no better than the models used, the modeling issues should be at the center of every generic research activity in CTS&D methodology. A class of models of special recent interest, commonly used in anesthesia and analgesia, have binary or categoric responses (6–9), where a number of methodologic issues, e.g., for optimal experimental design, remain to be addressed.

Early phase (I and II) clinical trials have been increasingly viewed as the most critical point at which to make assessments of the therapeutic effectiveness of a new drug and decide whether to proceed with large investments of time, money, and resources (1,2). There are a number of research questions, related to such trials, that still need to be answered (1,10–13). One of those, referring to parallel trials, is the rational or optimal selection of dosing regimens and sampling schedules.

Naratriptan is a novel $5HT_{1B/1D}$ agonist for the acute treatment of migraine. It has been developed for oral treatment and followed the development of the "first-in-line" triptan-sumatriptan (14,15). The overall aim of the present study

[1] Centre for Applied Pharmacokinetic Research, Manchester University, Oxford Road, Manchester M13 9PL, United Kingdom.
[2] School of Pharmacy, Manchester University, Oxford Road, Manchester M13 9PL, United Kingdom.
[3] School of Pharmacy, University of Queensland, Brisbane QLD 4072, Australia.
[4] Clinical Pharmacology, Glaxo Wellcome Research and Development Ltd., Greenford, Middlesex UB6 0HE, United Kingdom.
[5] To whom correspondence should be addressed. (e-mail: ggraham@fs1.pa.man.ac.uk)

0724-8741/01/0800-1210$19.50/0 

was to investigate the feasibility and utility of model-based CTS&D for the phase II development of naratriptan with effect measured on a categorical scale. The research was performed retrospectively and was not used to influence the development of naratriptan. Consequently, this research was used as a vehicle to explore various aspects of CTS&D. The work undertaken involved two separate projects. In the first project we had naratriptan pharmacokinetic data from a phase I clinical trial and sumatriptan pharmacodynamic data, together with preclinical data comparing the potency of naratriptan and sumatriptan. We used this information to develop a pharmacokinetic-pharmacodynamic model for naratriptan and applied this model to design a phase IIa study for naratriptan, with constraints set by logistics and regulation, with the principal objective of defining the dose/concentration-response relationship in an optimal fashion. The results of these simulations were compared with the actual phase IIa study that was performed. In the second project we used the phase IIa data to refine the naratriptan pharmacokinetic-pharmacodynamic model and used this model to explore some issues pertinent to the design of a phase IIb study. Principally, we investigated the application of optimal design methodology to the design of a pharmacodynamic study involving a categorical response variable. In particular, we evaluated the resources in terms of time, software, and experience needed to perform the CTS&D exercise. The current document is a detailed account of our experiences with CTS&D.

## MATERIALS AND METHODS

### Naratriptan and Sumatriptan Data

Pharmacodynamic models were based on the clinical end point of naratriptan, migraine pain relief/severity. Pain relief was measured on a five-point ordered categorical scale defined as follows: 0 = no pain relief, 1 = mild pain relief, 2 = moderate pain relief, 3 = considerable pain relief, and 4 = total pain relief. Pain severity was measured on a four-point ordinal scale with 0 = no pain, 1 = mild pain, 2 = moderate pain, and 3 = severe pain. Regulatory requirements define a measure of "success" of a drug as the ability to reduce pain severity or increase pain relief at 2 h "significantly." For pain severity, a success is a reduction from categories 3 or 2 to 1 or 0, and for pain relief, it is an increase from categories 0, 1, or 2 to 3 or 4. The simulations were based on dichotomizing the categorical responses into pain relief (success) and no pain relief (failure).

Information available for the analysis included preclinical and clinical data accumulated by Glaxo Wellcome up to 1993. The phase I studies included intravenous, subcutaneous, and oral (solution) routes of administration from a total of 26 healthy male volunteers. The phase IIa naratriptan data made available consisted of placebo or subcutaneous administration to about 400 patients (mostly female), including 33 patients on active treatment with both pharmacokinetic and headache score data. The phase IIa clinical trial that was performed was designed as a randomized, double-blind, placebo-controlled, dose-ranging, in-clinic study to evaluate the efficacy, safety, and tolerability of subcutaneous naratriptan (14). The subcutaneous doses of naratriptan were 0, 0.5, 1, 2.5, 5, and 10 mg, and 6 mg of sumatriptan was used as a positive control. There were 63, 60, 55, 42, 34, 34, and 47 patients in each dose group, respectively. The PK sampling times were 0, 0.167, 0.333, 0.5, 1, 1.5, 2, 3, 4, 6, and 8 h, whereas the pharmacodynamic variable was measured at the same times except for 6 and 8 h.

### Analyses

Two sets of analyses were performed. First, the integration of preclinical and phase I data for naratriptan and sumatriptan was used to construct a model of subcutaneous naratriptan to simulate the phase IIa clinical trial for naratriptan. This required the use of phase I naratriptan pharmacokinetic data, pharmacodynamic data from sumatriptan, and preclinical data comparing the pharmacodynamics of naratriptan and sumatriptan. In this analysis, the emphasis was on evaluating the predictive ability of the integrated model constructed for clinical trial simulations.

Second, the phase IIa subcutaneous naratriptan data were modeled, and the PK/PD models were used to simulate a phase IIb parallel clinical trial with oral naratriptan. In this analysis, the emphasis was on evaluating the sensitivity of the power of the study to various trial design factors (sample size per arm, model parameter uncertainties) and on selecting optimal dose sizes and effect sampling schedules.

### Models for Binary and Categoric Outcome Data

Longitudinal categorical data are often modeled by using proportional odds models (7). Such models have been applied to data from analgesic clinical trials in which pain relief/severity recorded on a categorical scale has been used as the pharmacodynamic measure (8,9). Proportional odds models exploit the ordinal nature of the categories to define log odds ratios that are then modeled. Therefore, rather than modeling the categories themselves, a transformation of the probabilities of being in a particular set of categories is modeled. The standard proportional odds model can be defined as

$$\text{logit}(\Pr(Y_i \le k)) = \log\left(\frac{\Pr(Y_i \le k)}{1-\Pr(Y_i \le k)}\right) = \sum_{h=1}^{k}\theta_h + f(\underline{\beta}, \underline{x}_i),\ k = 1,2,\ldots, K-1. \quad (1)$$

Pr(.) corresponds to a probability, $Y$ represents the categorical response data, the "logit" transformation is given by logit(Pr(Y)) = log[Pr(Y)/{1-Pr(Y)}], $f$(.,.) is the function used to model the logit transformed probabilities (or log odds ratios), $\underline{\beta}$ is the vector of model parameters, $\underline{x}_i$ is the $i$th vector of covariate values, $K$ is the number of categories and $\Sigma_{h=1}^{k}\theta_h$ is the cut point (probability of being in first $k$ categories for $\underline{x}_i$ = 0) up to category $k$. For a categorical response variable with $K$ categories, there are $K$-1 "parallel" lines on the log odds scale that represent the different sets of cumulative probabilities.

Proportional odds models can be defined with both fixed and random effects to enable the estimation of population mean and individual parameter values as well as the corresponding interindividual variance components. A common approach to developing a proportional odds model for

categorical pharmacodynamic data is to define a placebo/time component and a drug/concentration effect, which have an additive effect on the log odds scale (8).

### Modeling of Naratriptan Phase I Data and Preclinical Sumatriptan Data

A population pharmacokinetic analysis (NONMEM) of the phase I naratriptan data was undertaken. This analysis revealed that a two-compartment open model provided the best fit to the data. Intravenous infusion and subcutaneous and oral administration data were fitted simultaneously so that population estimates of rate constants of absorption and availability could be determined.

To allow scaling between sumatriptan and naratriptan, published preclinical sumatriptan data were used (15,16). The preclinical studies included 5HT receptor-binding assays in COSM6 cells that were transiently transfected with the $5HT_{1D}$ gene, isolated ring preparations of dog basilar and middle cerebral artery measuring isometric tension changes, measurements of carotid vascular resistance in anesthetized dogs and the effect of neurogenically mediated inflammation in the dura of anesthetized rats. Comparisons of naratriptan and sumatriptan results showed that naratriptan is two- to sixfold more potent than sumatriptan, but they have approximately the same maximum effect. A conservative value of 2 was used to scale the potency parameter in the sumatriptan pharmacodynamic model.

The model used to describe the phase II sumatriptan pharmacodynamic data is defined in Eq. (2) (17), which is expressed in two parts. The first line of Eq. (2) corresponds to a baseline pain severity score of 2 observed in 55.4% of migraineurs. If the patient reported a pain severity score of 3 after the baseline measurement, then this is recorded as 2 because only the decrease of pain severity is of interest. The second line corresponds to a baseline pain severity score of 3 observed in 44.6% of migraineurs. Therefore, the first part of the model has two cut points and the second line has three cut points. This is defined in Eq. (2) by conditioning on the baseline pain severity variable $score_0$. The placebo model is given by log(*time*), and the drug effect is defined as an "$E_{max}$"-type model where $C_e$ is the concentration of the drug at a hypothetical effect site (18) and a schematic of the model is given in Fig. 1. The random effects are denoted by $\eta_i$ (baseline score of 2) and $\zeta_i$ (baseline score of 3) and are normally distributed with mean zero and variance given by $\omega_\eta^2$ and $\omega_\zeta^2$, respectively.

$$\text{logit}(\Pr(Y_{ij} \le k | score_0)) = \begin{cases} \sum_{h=0}^{k} \theta_{h+1} + \theta_4 \log(time_{ij}) + \dfrac{\theta_5 C_{eij}}{\theta_6 + C_{eij}} + \eta_i, & score_0 = 2,\ k = 0,1 \\ \sum_{h=0}^{k} \phi_{h+1} + \phi_5 \log(time_{ij}) + \dfrac{\phi_6 C_{eij}}{\phi_7 + C_{eij}} + \xi_i, & score_0 = 3,\ k = 0,1,2 \end{cases} \quad (2)$$

Because of the regulatory requirement of significant pain relief defined earlier, only the cut point corresponding to $k = 1$ is necessary because this corresponds to dichotomizing the pain severity categorical data into pain relief (categories 0 and 1) and no pain relief (categories 2 and 3). Thus, the model becomes a binary logistic model for the dichotomization, pain relief/no pain relief as defined in Eq. (3).

$$\text{logit}(\Pr(Y_{ij} = 1 | score_0)) = \begin{cases} \theta_1 + \theta_4 \log(time_{ij}) + \dfrac{\theta_5 C_{eij}}{\theta_6 + C_{eij}} + \eta_i, & score_0 = 2 \\ \phi_1 + \phi_5 \log(time_{ij}) + \dfrac{\phi_6 C_{eij}}{\phi_7 + C_{eij}} + \xi_i, & score_0 = 3 \end{cases} \quad (3)$$

The naratriptan PD model is obtained from the sumatriptan PD model by scaling the affinity parameters $\theta_6$ and $\phi_7$ by the reported relative potency factor between naratriptan and sumatriptan (15,16).

### Modeling of Naratriptan Phase II Data

Two approaches to pharmacokinetic analysis of the naratriptan phase IIa data were undertaken. As a first step, the dose-binary response data were modeled by a stepwise logistic regression program (BMDPLR) to develop a PD model [Fig. 1a; Eq. (4)] with the most informative covariates–time (time), dose (dose), dose-time interaction, and type of migraine. The "Lag" in the PD model (Fig. 1a) represents the offset between the administration of the dose and the onset of the effect and, therefore, does not constitute a separate compartment. At the next step, a population sequential PK-PD model (Fig. 1b) was developed. The PK part of the PK/PD model is an ordinary two-compartment first-order absorption model. The PK and PD part are connected by a "link" model. This is a standard modeling approach (9,14) in which the effect is related to the concentration in an "effect" compartment and its rate constant *Keo* is used to describe the delay of the effect with respect to the plasma concentration. First, the

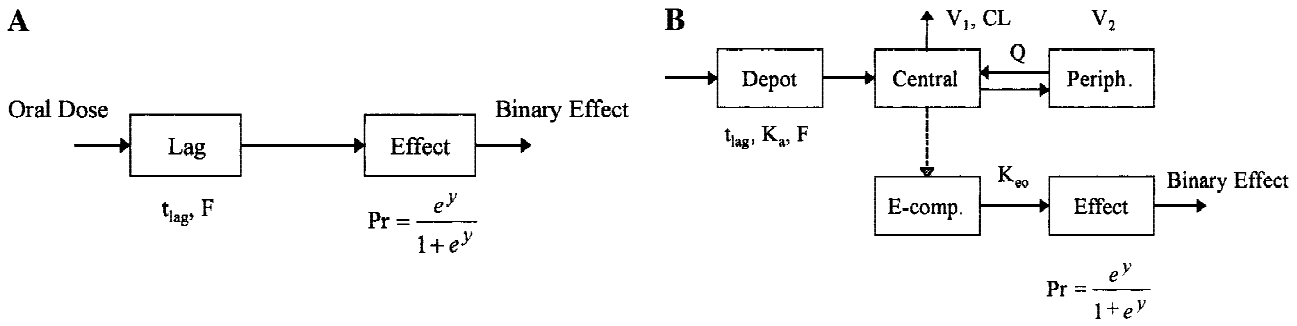


**Fig. 1.** (a) PD model for naratriptan defined by Equation (4). (b) PK-PD model for naratriptan defined by Eq. (5).

parameters of the PK model were estimated and post hoc estimates of the individual PK model parameters were derived. Then, the individual PK parameters were used to estimate the parameters of the PD model. The best fit for naratriptan was achieved by using time, concentration-effect site ($C_e$), a concentration-time interaction term, and type of migraine as covariates. More complex PK-PD models (involving time delays) were abandoned because the data were insufficient to support them.

$$\text{logit}(\Pr(Z=1)) = \theta + \beta_1 time + \beta_2 dose + \beta_3 time \times dose + \beta_4 type \quad (4)$$

$$\text{logit}(\Pr(Z=1)) = \theta + \beta_1 time + \beta_2 C_e + \beta_3 time \times C_e + \beta_4 type \quad (5)$$

$Z$ is the dichotomized pain relief variable (1 = pain relief, 0 = no pain relief) and is shown in Fig. 1. Both models were concurrently used in the simulations of the phase IIb oral naratriptan trial.

## Simulation for Clinical Trial Design

Clinical trial simulation (CTS) is essentially a Monte Carlo exercise addressing the typical behavior of the system and the departures from it. During this exercise, the influence of the trial controllable factors (dosing and sampling schedules, modes, sites, and number of subjects/arms) and uncontrollable factors (model parameters and structures, modeling and design assumptions, compliance, and limitations) on the trial outcomes is assessed. During the clinical trial design phase, the results of the simulations were analyzed by using various statistical techniques to determine the most favorable controllable components of the clinical trial that will lead to the achievement of the best outcomes of the trial with respect to the prespecified goals.

To determine the asymptotically "correct" values (nominal values) that will serve as a basis for detecting significant differences in the power studies, an intensive trial simulation was undertaken, with nine doses (0, 1, 2.5, 5, 10, 15, 20, 25, and 30 mg of naratriptan). One thousand trials were simulated with the PD model, and the statistical significance of the difference of the observed effect at 2, 4, and 6 h was assessed at a probability level of 0.05.

Concentrating on the power of the trial to detect statistically and clinically distinct doses, both the PD and the PK-PD model were used in extensive Monte Carlo simulations, investigating the sample size per arm, the number of subjects in the placebo arm, the response sampling times, and some of the modeling assumptions. Assumptions relating to absorption were of special interest because no experimental information was available for the oral formulation (the phase I oral administration data were obtained from a solution). In each of the above cases, 200 trials were simulated, each trial having 4 doses (arms)—placebo, 5, 10, and 20 mg naratriptan.

The sample size can be calculated alternatively by using the nominal output values of the models obtained after the initial asymptotic simulation, without doing any further simulations. For this purpose, statistical tables (19) using a normal approximation to the binomial distribution were applied.

## Optimal Experimental Design with Binary Data

The list of doses and the effect sampling times can also be determined solely on the basis of the models developed, without clinical trial simulation, by using design of experiments (DOE) techniques (20). The dose-ranging study can be redefined in a dose-effect model parameter estimation study. D-optimal experimental design theory can then be applied, which aims to minimize the uncertainty in the dose-effect model parameters. D-optimal designs maximize the determinant of Fisher information matrix (FIM) and, therefore, minimize the joint confidence regions of the parameter estimates (20).

If the type of the migraine in both the PD (Fig. 1a) and PK-PD (Fig. 1b) models is ignored, then the response model has four parameters to be estimated ($\theta$, $\beta_1$, $\beta_2$, and $\beta_3$) and two covariates—time and dose (PD model) or effect site concentration (PK-PD model). Because the number of design points of the optimal design is equal to the number of parameters estimated, the optimal design in our case includes two doses and two effect sampling times because it is analogous to a two-factor experiment, with each factor having two levels. The additional constraints to the sets of possible doses and effect sampling times are as follows: (i) they are positive, (ii) a placebo dose is required, and (iii) a 2-h effect sampling time is required by the regulator.

Taking into account the above considerations, we extended the DOE theory to the case of binary response data with two covariates and an interaction term in the logistic function. The standard result for a linear logistic regression model with one covariate is to take measurements at the design points corresponding to the probabilities 0.176 and 0.824 (21). When there is a linear combination of covariates (including interactions), the FIM can be written as shown in Eq. (6), when the model is that given in Fig. 1a.

$$\pi_i = \frac{\exp(\theta + \beta_1 t_i + \beta_2 d_i + \beta_3 t_i \times d_i)}{1 + \exp(\theta + \beta_1 t_i + \beta_2 d_i + \beta_3 t_i \times d_i)},\ t = time,\ d = dose$$

$$w_i = \pi_i(1 - \pi_i)$$

$$M_F(\underline{\theta}, \underline{x}_i) = \begin{bmatrix} \sum_{i=1}^{n} w_i & \sum_{i=1}^{n} t_i w_i & \sum_{i=1}^{n} d_i w_i & \sum_{i=1}^{n} t_i d_i w_i \\ \sum_{i=1}^{n} t_i w_i & \sum_{i=1}^{n} t_i^2 w_i & \sum_{i=1}^{n} t_i d_i w_i & \sum_{i=1}^{n} t_i^2 d_i w_i \\ \sum_{i=1}^{n} d_i w_i & \sum_{i=1}^{n} t_i d_i w_i & \sum_{i=1}^{n} d_i^2 w_i & \sum_{i=1}^{n} t_i d_i^2 w_i \\ \sum_{i=1}^{n} t_i d_i w_i & \sum_{i=1}^{n} t_i^2 d_i w_i & \sum_{i=1}^{n} t_i d_i^2 w_i & \sum_{i=1}^{n} t_i^2 d_i^2 w_i \end{bmatrix} \quad (6)$$

Subsequently, a D-optimization procedure incorporating the simplex method, programmed in MATLAB, was implemented to determine the optimal design points. To compare alternative trial designs, which may be more logistically realistic, their relative efficiencies, defined as the ratio of the FIM for the new design with respect to the optimal design to the power of reciprocal number of parameters, were calculated.

**Table I.** Pharmacokinetic Parameters Obtained from Naratriptan Phase I Data

| Parameter | Mean | C.V. (%) |
|---|---|---|
| Cl (L $h^{-1}$) | 22.7 | 27.9 |
| $V_1$ (L) | 17.2 | 20.7 |
| $V_2$ (L) | 147 | 26.3 |
| Q (L $h^{-1}$) | 154 | 25.5 |
| ka ($h^{-1}$) | 4.01 | 64.0 |
| F | 0.96 | — |

*Note.* The parameters Cl, $V_1$, $V_2$ and Q were estimated from all three routes of administration simultaneously, whereas ka and F are the estimates for subcutantous naratriptan. C.V. (%) corresponds to the intersubject variability in terms of the coefficient of variation.

**Table II.** Sumatriptan Pharmacodynamic Model [Eq. (2)] Parameters (17)

| Proportional odds model | Model Parameter | Estimate (SE) |
|---|---|---|
| Model conditional on baseline score = 2 | | |
| Baseline score = 0 | $\theta_1$ | –6.79 (0.58) |
| Baseline score = 1 | $\theta_2$ | 4.73 (0.217) |
| Effect site equilibration rate constant keo ($h^{-1}$) | $\theta_3$ | 0.779 (0.49) |
| Placebo/time effect | $\theta_4$ | 1.93 (0.246) |
| Maximum effect | $\theta_5$ | 3.96 (1.2) |
| Ce50 (ng/mL) | $\theta_6$ | 3.51 (2.73) |
| Intersubject variability | $\omega_\eta^2$ | 17.7 (2.14) |
| Model conditional on baseline score = 3 | | |
| Baseline score = 0 | $\phi_1$ | –7.74 (0.677) |
| Baseline score = 1 | $\phi_2$ | 3.36 (0.2) |
| Baseline score = 2 | $\phi_3$ | 3.9 (0.238) |
| Effect site equilibration rate constant keo ($h^{-1}$) | $\phi_4$ | 2.04 (1.04) |
| Placebo/time effect | $\phi_5$ | 2.32 (0.234) |
| Maximum effect | $\phi_6$ | 9.85 (23) |
| Ce50 (ng/mL) | $\phi_7$ | 45.9 (162) |
| Intersubject variability | $\omega_\zeta^2$ | 19 (2.42) |

*Note.* There were 553 migraineurs with moderate pain at baseline and 435 with severe pain at baseline.

### Software Used

BMDP statistical package (procedures PL and LR) and NONMEM were used for the PD and PK/PD model parameter estimation. All simulations and DOE processing were performed by using MATLAB software. The power calculations and the analysis of the results were performed with Microsoft EXCEL.

## RESULTS AND DISCUSSION

### Modeling of the Naratriptan Data from Phase I Studies and Sumatriptan Preclinical Data

In the phase I data, there were 26 individuals available for analysis with intravenous, subcutaneous, and oral administration. A two-compartment model for the disposition was selected with a first-order absorption model for the subcutaneous and oral routes. The parameter values for the subcutaneous model are given in Table I. The parameter estimates for the pharmacodynamic model defined in Eq. (2) are given in Table II.

Figure 2a shows mean profiles of the naratriptan PK-PD model [Eq. (2)] predictions with respect to time and dose, which were based on combined naratriptan pharmacokinetic and sumatriptan pharmacodynamic data. Figure 2b shows that naratriptan at any particular time achieves greater pain relief than sumatriptan because of the lower value of the naratriptan Ce50 parameter.

We used the actual data from the phase IIa naratriptan clinical trial to check the performance of the integrated model. The phase I 95% prediction interval and the phase IIa pharmacokinetic data for naratriptan are shown in Fig. 3. It can be seen that the phase I model underpredicts the phase IIa data. The reason for this difference is unknown given that in sumatriptan clinical trials, there did not seem to be any difference between patients and healthy volunteers (22). The phase I data come from a healthy male population, and the phase IIa data are from a predominantly female population, but it is assumed there is no gender difference in the pharmacokinetics.

The predicted placebo model is in good agreement with the observed phase IIa placebo data. Figure 4a shows the comparison of the predicted and observed pain relief profile at 2 h. The general trend of the predicted profile is correct but systematically underpredicts the observed data. This could either be due to the pharmacokinetic model underpredicting the observed concentration-time data or the potency scaling factor being too small.

A comparison of the power of showing a significant difference between the active groups and placebo at 2 h for the

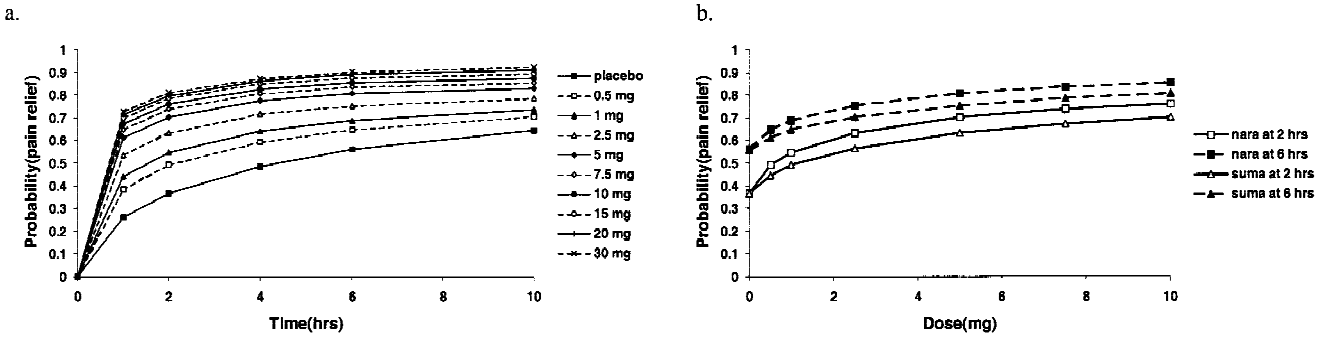


**Fig. 2.** (a) Naratriptan probability of pain relief profile vs. time using the naratriptan pharmacokinetic model, sumatriptan pharmacodynamic model, and the potency parameter scaled by a factor of 2. (b) Naratriptan and sumatriptan probability of pain relief profiles vs. dose at 2 and 6 h using the naratriptan pharmacokinetic model, sumatriptan pharmacodynamic model, and the potency parameter scaled by a factor of 2.

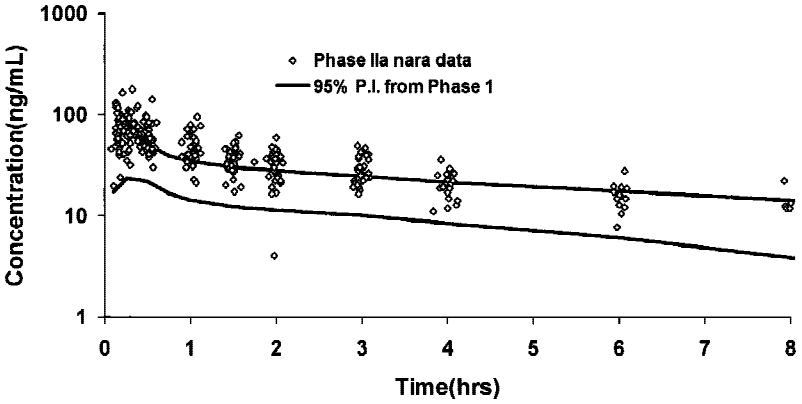


**Fig. 3.** Phase I 95% prediction interval (calculated pointwise) based on modeled phase I naratriptan pharmacokinetic data compared with observed phase IIa naratriptan pharmacokinetic data. The data have been jittered to show the amount of data at each time point.

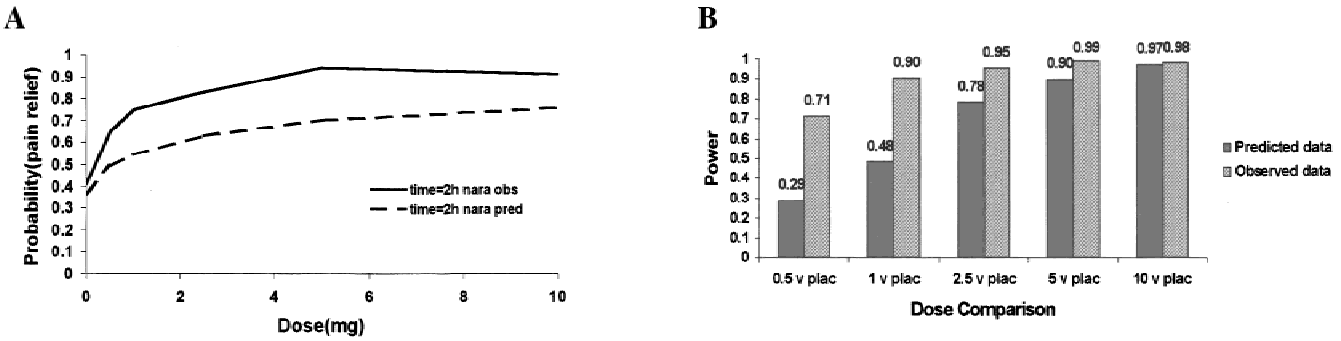


**Fig. 4.** (A) Observed naratriptan phase IIa pharmacodynamic data and predicted PK-PD pain relief profiles at 2 h. (B) Observed and simulated (from the PK-PD model) power at 2 h.

**Table III.** Parameters of the PD Model [Eq. (4)] Estimated from Naratriptan Dose-Response Data Using the Logistic Regression Procedure BMDPLR

| Parameter | Value | SD |
|---|---|---|
| tlag (h) | 0.9 | n.a.[a] |
| F (%) | 60 | 30 |
| θ | –1.264 | 0.630 |
| β1 | 0.543 | 0.270 |
| β2 | 0.042 | 0.021 |
| β3 | 0.164 | 0.082 |
| β4 | 0.4 | n.a. |

[a]n.a. = not applicable.

**Table IV.** Parameters of the PK-PD Model [Eq. (5)], Estimated by Using a Sequential Approach with NONMEM

| PK model parameters | | | Link & PD model parameters | | |
|---|---|---|---|---|---|
| Parameter | Value | SD | Parameter | Value | AD |
| tlag (h) | 0.35 | 0.35 | $K_{eo}$ (l/h) | 0.845 | n.a.[a] |
| F (%) | 48 | n.a. | θ | –0.11 | 0.06 |
| $K_a$ (l/h) | 0.85 | 0.28 | β1 | 0.482 | n.a. |
| CL (L/h) | 17.9 | 7.57 | β2 | 0.082 | n.a. |
| $V_1$ (L) | 23.00 | 20.1 | β3 | 0.083 | n.a. |
| $V_2$ (L) | 98.4 | 47.0 | β4 | 0.714 | n.a. |
| Q (L/h) | 101.4 | 36.8 | | | |

[a]n.a. = not applicable.

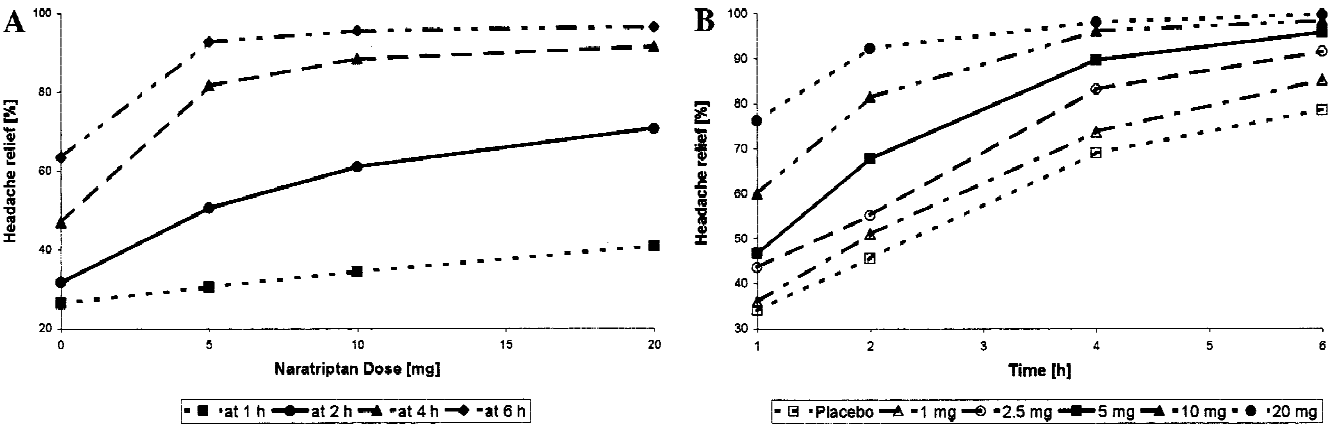


**Fig. 5.** (A) Dose-effect profile of the PK-PD model, Eq. (5). (B) Time-effect profile of the PD model, Eq. (4).

observed probabilities of pain severity and those predicted by the simulation model was made. The power calculations for the simulation model were made with the same number of patients as in the actual clinical trial, which was simulated 1000 times. Figure 4b shows that for doses > 2.5 mg, there was >80% power for both the observed and predicted data.

### Modeling of the Naratriptan Phase IIa Data

The parameters of the PD model [Eq. (4)] and PK-PD model [Eq. (5)], identified from the phase IIa naratriptan data are given in Tables III and IV, respectively. The parameters related to the absorption of the oral form were chosen on the basis of previous information with sumatriptan and literature

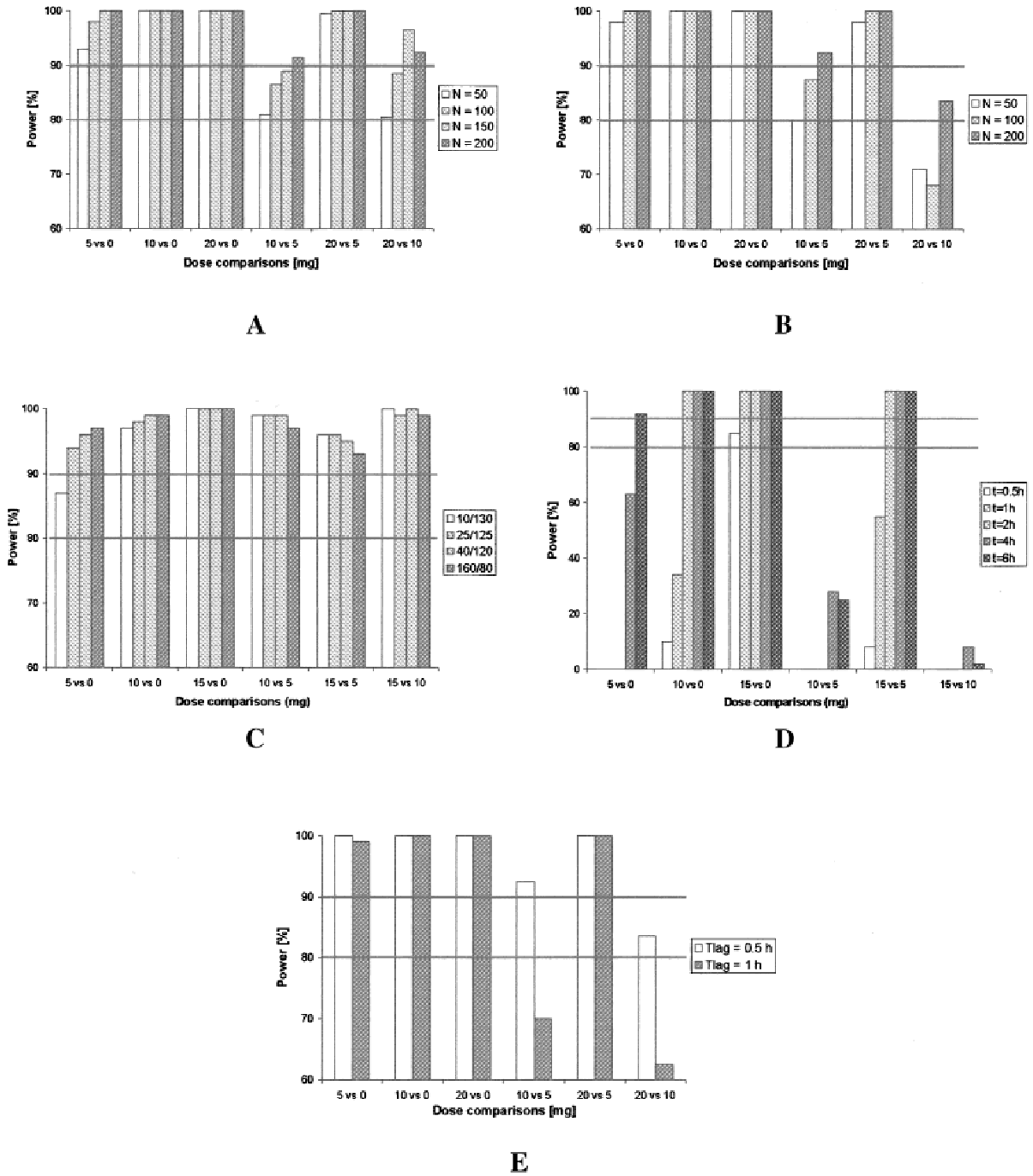


**Fig. 6.** Results of the simulation of various clinical trial designs. (A) Influence of the sample size (number of subjects, N) per arm (dose) on the power of the trial [PD model, Eq. (4)]. (B) Influence of the sample size (number of subjects, N) per arm (dose) on the power of the trial [PK-PD model, Eq. (5)]. (C) Effect of the number of subjects in the placebo arm [PD model, Eq. (4)]. The total number of subjects per trial (with 4 doses) was kept at 400. The number of placebo arm subjects was varied at 10, 25, 40, and 160; the number of subjects on the active treatment arms is adjusted accordingly. (D) Effect of the headache relief sampling time on the power of the trial [PD model, Eq. (4)]. (E) Effect of the PK lag-time on the power of the trial [PK-PD model, Eq. (5)].

**Table V.** Sample Size Calculation [Using Statistical Tables (19)] to Give 80% Power of Showing a 95% Significant Difference

| | Effect sampling times (h) | | | |
|---|---|---|---|---|
| Dose comparisons | 1 | 2[a] | 4 | 6 |
| 5 mg *vs.* 0* mg | 1414 | 82 | 22 | 23 |
| 10 mg *vs.* 0* mg | 400 | 35 | 15 | 18 |
| 20 mg *vs.* 0* mg | 130 | 19 | 12 | 17 |
| 10 mg *vs.* 5 mg | 1822 | 280 | 337 | 860 |
| 20 mg *vs.* 5 mg | 267 | 71 | 144 | 430 |
| 20 mg *vs.* 10 mg | 700 | 286 | 1148 | 4693 |

[a] Required by the regulator.

data. Because the stepwise logistic regression procedure of BMDPLR does not provide estimates for the standard deviation (SD) of the PD parameters evaluated, for the purposes of simulation, a conservative CV of 50% has been assumed and given in the last column of Table III. The estimates of SD in Table IV are given by NONMEM. NONMEM seemed to assign all the variability to the first parameter of the model–θ. This is not surprising, taking into account the low information contents of the categorical data.

The predicted dose-effect curves for naratriptan using the PK-PD models are shown in Figure 5a. The PD model showed similar mean dose-effect curves.

### Simulation for Clinical Trial Design

From the "nominal" mean time-effect curve of all 1000 trials, shown in Figure 5b, four dose bands, which give significantly different effects (>10% at 0.05 significance level at 2 h) can be determined: {placebo, 1 mg, and 2.5 mg}, {5 mg}, {10 mg}, and {20 mg}. For the subsequent clinical trial simulations power evaluation, the following naratriptan doses (arms) were selected: placebo, 5 mg, 10 mg, and 20 mg.

A comparison of the power of the trial with different sample sizes (50, 100, 150, and 200 subjects per dose) at 2 h is shown in Figures 6a (PD model) and 6b (PK-PD model). It can be seen that, even with the PK-PD model, which seems to be more conservative, a sample size of 200 subjects per arm (dose) would guarantee a power of the trial (i.e., probability to detect a true difference) > 80%.

It is interesting to note that fewer subjects may be needed in the placebo arm. Figure 6c shows the influence of the number of subjects in the placebo arm (the total number of subjects in all four arms was kept constant at 400) on the power of the trial. Even with only 25 subjects in the placebo arm (and 125 in each of the other three) 90% power is ensured.

The effect of the headache relief sampling time on the power of the trial is shown in Figure 6d. Because the 2-h time point is required by the regulator, the power of the trial will be maximized if 4 and/or 6 h are included in the sampling schedule.

It is known that migraine is usually accompanied by delayed gastric emptying, which would increase the PK lag-time (23). The data available following oral administration to healthy subjects did not provide information about the magnitude of the PK lag. Therefore, the influence of the lag-time on the power of the trial was explored. The results of the simulations with lag-times of 0.5 and 1 h are shown in Figure 6e. Based on these simulations it is realistic to expect that for most doses the proposed trial would perform well, with a power >80%, if the PK lag-time of naratriptan is in this range.

The experimental data also show evidence of significant variability in the bioavailability of naratriptan, the likely CV being in the region of 40–60%. Yet, the population model used for the PK-PD model was unable to estimate this variability, possibly assigning it to other parameters. The same consideration applies to the model rate constant *Keo,* which represents the PD delay in the system. To study the influence of the uncertainties on the power of the trial, clinical trials with different CVs (0, 30%, and 60%) for the parameters of interest were simulated. In both cases, even with a CV > 30%, the power of the clinical trial was adequate.

Table V shows the sample size calculation, using statistical tables (19) with a normal approximation to the binomial distribution. The tables give a conservative estimate of the sample size needed, compared with the simulation results, due to the approximation used.

Based on the above simulations, the following clinical trial design is recommended:

1. Four or five arms with doses: placebo, 2.5 mg and/or 5 mg, 10 mg, or eventually 20 mg. Because the simulations show that the minimum effective dose is in the range of 1 mg, the latter dose may be included in the design, if the regulator requires it. The simulations show that the maximum no-effect dose is in the range of 0.5–0.75 mg.
2. At least 150 subjects per arm with possibly less for placebo if the required power of the trial is 80%.
3. The headache relief should be monitored at 2 h, 4 and/or 6 h.

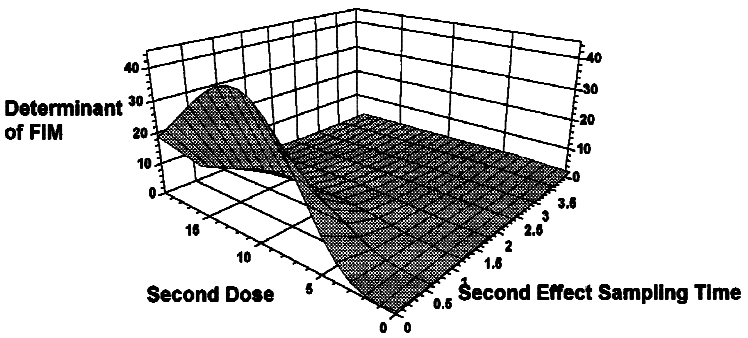


**Fig. 7.** Determinant of the Fisher information matrix surface for naratriptan with the first dose fixed at 0 mg and the first sampling time fixed at 2 h.

**Table VI.** Relative Efficiency of Several Possible Clinical Trial Designs for Naratriptan

| Doses (mg) | Effect sampling times (h) | Replicates of each dose[a] | Number of points | FIM | Relative efficiency (%) | Note |
|---|---|---|---|---|---|---|
| 0[b], 10 | 1, 2[b] | 4 | 16 | 637 | 100 | Suboptimal design |
| 0[b], 2.5, 10, 20 | 2[b], 4 | 2 | 16 | 33.3 | 47.8 | Proposed design 1 |
| 0[b], 1, 2.5, 10 | 2[b], 4 | 2 | 16 | 75.0 | 58.6 | Proposed design 2 |
| 0[b], 1, 2.5, 5 | 2[b], 4 | 2 | 16 | 115 | 65.2 | Proposed design 3 |
| 0[b], 0.1, 0.25, 1, 2.5, 5, 7.5, 10 | 2[b], 4 | 1 | 16 | 107 | 63.9 | Applied design (14) |

[a] The replicates were introduced for the purpose of the relative efficiency comparison, so that each of the compared designs had the same number of doses (8 as in the last proposed design).
[b] Required by the regulator.

### D-Optimal Experimental Design for Naratriptan

The FIM surface, which was computed with two fixed points of the design—a placebo dose and an effect sampling point at 2 h—is shown in Figure 7. It can be seen that the maximum of the FIM is attained at early sampling times (close to 0 h) and at a second dose of approximately 10 mg. For the sake of the balance between logistics and optimality, we selected as a suboptimal design the one with a second dose of 10 mg and a second sampling time of 1 h because of the lag in the drug being absorbed.

To make a quantitative comparison between alternative trial designs, which are more logistically realistic, their relative efficiencies, with respect to the suboptimal design, were calculated from the respective FIMs and are shown in Table VI. It can be seen that, in terms of D-optimality, the proposed designs have close or even better efficiencies than the design that was actually used and published (14). At the same time, the proposed designs include only half of the number of the doses (arms) and, consequently, half of the resources needed for the execution of the trial. The latter results show the need for a careful balance between the design optimality and the logistics to achieve the best trial results.

## CONCLUSION

The studies described in this report were performed retrospectively and did not influence the naratriptan drug development program. The design of clinical trials for a "second-in-line" drug with a similar mechanism of action are more straightforward than a drug with a novel mode of action, in that considerable insight can be gleaned from the experience of the development of the "first-in-line" drug. Nevertheless, modeling is a powerful vehicle for carrying information forward and designing studies in an optimal fashion. Although clinical trial simulation was not used in the design of the naratriptan trials, modeling did play an important part in trial development, which can be seen from the fact that the actual phase IIa study was not very dissimilar to the "optimal" design that we ended up with. In terms of "real-time" development, the major limitation of optimal clinical trial design is gaining speedy access to relevant information within the timelines of the development program.

Clinical trial simulation is a useful tool for the quantitative assessment of the influence of controllable clinical trial variables, such as sample size, sampling schedule, and administration strategy, on the power of the trial to determine a particular outcome. In some cases, such as the power of the trial to determine the headache relief at 2 h in the case of naratriptan, classical power calculations may be used. However, for complex end points, particularly those based on PK-PD modeling, simulation is required. Simulation is the only method that can be used to assess the influence of uncontrollable variables such as parameter uncertainties and modeling assumptions.

The number of simulations that need to be performed increases exponentially with the number of trial factors and levels of those factors. Thus, for three factors, such as dose, time, and formulation, each at three levels, 27 sets of simulations are required, and to obtain reliability, each of these sets need to be run several hundred times. Optimal design theory may help to reduce the simulation burden by defining sensitive regions of the design space. Pure optimal designs tend to be unrealistic, often involving replicated experiments. Consequently, a combination of optimal design theory and simulation, taking into account questions of logistics and ethics, offers a very pragmatic approach to clinical trial design.

The actual resources, in software and time, to perform the simulations and analyses reported here were not demanding. However, clinical trial design should be viewed as a modeling exercise that brings together diverse disciplines such as pharmacology and statistics. Consequently, the required interdisciplinary expertise is considerable. Therefore, the main impediment to the widespread application of CTS&D in the pharmaceutical industry is the lack of suitably trained personnel.

CTS may also be used as an educational tool. Thus, by means of simulation, the purpose and conduct of a clinical trial can be explained to the clinical trial staff. For example, the importance of accurate timekeeping can be explored by using a sensitivity analysis. It is also possible to assign resources and develop costing on the basis of a CTS.

Clinical trial simulation and design raises many interesting and fundamental research issues in the areas of optimal design of experiments, mixed effects modeling and sensitivity analysis. These issues are complex and will be the subject of future research. Presently, the CTS&D tools that are available would seem to be sufficient for CTS to make a significant contribution to the drug development process.

## REFERENCES

1. C. C. Peck and R. E. Desjardins. Simulation of clinical trials: Encouragement and cautions. *Appl. Clin. Trials* **5**:30–32 (1996).
2. M. Hale, W. R. Gillespie, S. K. Gupta, B. Tuk, and N. H. Holford. Clinical trial simulation: Streamlining your drug development process. *Appl. Clin. Trials* **5**:35–40 (1996).
3. L. J. Lesko, M. Rowland, C. C. Peck, and T. F. Blaschke. Opti-

mizing the science of drug development: Opportunities for better candidate selection and accelerated evaluation in humans. *Eur. J. Pharm. Biopharm.* **10**:iv–xiv (2000).
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. N. H. G. Holford, M. Hale, H.C. Ko, J.-L. Steimer, L. B. Sheiner, and C.C. Peck (eds.). Simulation in drug development: good practices. Publication of CDDS, http://www.dml.georgetown.edu/cdds/sddgp723.html (1999).
6. L. P. Balant and M. Gex-Fabry. Modeling during drug development. *Eur. J. Pharm. Biopharm.* **50**:13–26 (2000).
7. L. B. Sheiner and S. L. Beal. Integrated pharmacokinetic-pharmacodynamic model for different kinds of pharmacodynamic data: continuous, categoricalal and time-to-event. In M. Danhof and J. L. Steimer (eds.), *Measurement and Kinetics of In Vivo Drug Effects: Advances in Simultaneous Pharmacokinetic / Pharmacodynamic Modelling*, Leiden/Amsterdam Center for Drug Research, Leiden, 1998 pp. 3–6.
8. L. B. Sheiner. A new approach to the analysis of analgesic drug trials, illustrated with bromfenac data. *Clin. Pharmacol. Ther.* **56**:309–322 (1994).
9. C. Y. Liu and N. Sambol. Pharmacodynamic analysis of analgesic clinical trials using empirical methods. *Pharm. Res.* **12**:438–445 (1995).
10. L. B. Sheiner, S. L. Beal, N. C. Sambol. Study designs for dose ranging. *Clin. Pharmacol. Ther.* **46**:63–78 (1989).
11. L. B. Sheiner and D. B. Rubin. Intention-to-treat analysis and the goals of clinical trials. *Clin. Pharmacol. Ther.* **57**:6–15 (1995).
12. J.-P. Boissel, I. Durieu, P. Girard, P. Nony, F. Chauvin, and M. Haugh. Dose-ranging trials: Guidelines for data collection and standardized descriptions. *Control. Clin. Trials* **16**:319–330 (1995).
13. P.-H. Hsu. An integrated design for dose comparison trials. *J. Clin. Pharmacol.* **33**:762–766 (1993).
14. E. Fuseau, R. Kempsford, P. Winter, M. Asgharnejad, N. Sambol, and C. Y. Liu. The integration of the population approach into drug development: a case study, Naratriptan. In L. Aarons *et al.* (eds.), *The Population Approach: Measuring and Managing Variability in Response, Concentration and Dose*, COST B1 Medicine, European Commission, Brussels, 1997 pp. 203–214.
15. C. Dahlof, L. Hogenhuis, J. Olesen, H. Petit, J. Ribbat, J. Schoenen, D. Boswell, E. Fuseau, H. Hassani, and P. Winter. Early clinical experience with subcutaneous naratriptan in the acute treatment of migraine: a dose-ranging study. *Eur. J. Neurol.* **5**:469–477 (1998).
16. H. E. Connor, W. Feniuk, D. T. Beattie, P. C. North, A. W. Oxford, D. A. Saynor, and P. P. A. Humphrey. Naratriptan: Biological profile in animal models relevant to migraine. *Cephalalgia* **17**:145–152 (1997).
17. O. Petricoul and E. Fuseau. Meta-analysis of the exposure/efficacy relationship for sumatriptan nasal spray. Poster at Population Approach Group in Europe, Saintes, France (1999).
18. N. Holford and L. B. Sheiner. Kinetics of pharmacological response. In M. Rowland and G. Tucker (eds.), *Pharmacokinetics: Theory and Methodology*, Pergamon Press, Oxford, 1986 pp. 189–212.
19. S. Lemeshow, D. W. Hosmer, J. Klar, and S. K. Lwanga. *Adequacy of sample size in health studies*, WHO, New York, 1990.
20. E. Walter and L. Pronzato. Qualitative and quantitative experiment design for phenomenological models—A survey. *Automatica* **26**:195–213 (1990).
21. S. Minkin. Optimal designs for binary data. *J. Am. Stat. Assoc.* **82**:1098–1103 (1987).
22. V. F. Cosson and E. Fuseau. Mixed effect modeling of sumatriptan pharmacokinetics during drug development. II. From healthy subjects to phase II dose ranging in patients. *J. Biopharm. Pharmacokinet.* **26**:149–171 (1998).
23. N. R. Cutler, E. K. Hussey, J. J. Sramek, B. D. Clements, L. A. Paulsgrove, M. A. Busch, and K. H. Donn. Oral sumatriptan in pharmacokinetics in the migrainous state. *Cephalalgia* **11**:222–223 (1991).