

How Modeling and Simulation Have Enhanced Decision Making in New Drug Development

Raymond Miller,^{1,*} Wayne Ewy,¹ Brian W. Corrigan,¹ Daniele Ouellet,¹
David Hermann,¹ Kenneth G. Kowalski,¹ Peter Lockwood,¹
Jeffrey R. Koup,¹ Sean Donevan,¹ Ayman El-Kattan,¹
Cheryl SW Li,¹ John L. Werth,¹ Douglas E. Feltner,¹
and Richard L. Lalonde¹

Received June 30, 2005—Final July 28, 2005

The idea of model-based drug development championed by Lewis Sheiner, in which pharmacostatistical models of drug efficacy and safety are developed from preclinical and available clinical data, offers a quantitative approach to improving drug development and development decision-making. Examples are presented that support this paradigm. The first example describes a preclinical model of behavioral activity to predict potency and time-course of response in humans and assess the potential for differentiation between compounds. This example illustrates how modeling procedures expounded by Lewis Sheiner provided the means to differentiate potency and the lag time between drug exposure and response and allow for rapid decision making and dose selection. The second example involves planning a Phase 2a dose-ranging and proof of concept trial in Alzheimer's disease (AD). The issue was how to proceed with the study and what criteria to use for a go/no go decision. The combined knowledge of AD disease progression, and preclinical and clinical information about the drug were used to simulate various clinical trial scenarios to identify an efficient and effective Phase 2 study. A design was selected and carried out resulting in a number of important learning experiences as well as extensive financial savings. The motivation for this case in point was the "Learn-Confirm" paradigm described by Lewis Sheiner. The final example describes the use of Pharmacokinetic and Pharmacodynamic (PK/PD) modeling and simulation to confirm efficacy across doses. In the New Drug Application for gabapentin, data from two adequate and well-controlled clinical trials was submitted to the Food and Drug Administration (FDA) in support of the approval of the indication for the treatment of post-herpetic neuralgia. The clinical trial data was not replicated for each of the sought dose levels in the drug application presenting a regulatory dilemma. Exposure response analysis submitted in the New Drug Application was applied to confirm the evidence of efficacy across these dose levels. Modeling and simulation analyses showed that the two studies corroborate each other with respect to

¹Pfizer Global Research and Development, Pfizer Inc, 2800 Plymouth Road, Ann Arbor, MI 48105, USA.

*To whom correspondence should be addressed. Telephone: +1-734-6222373; e-mail: raymond.miller@pfizer.com

the pain relief profiles. The use of PK/PD information confirmed evidence of efficacy across the three studied doses, eliminating the need for additional clinical trials and thus supporting the approval of the product. It can be speculated that the work by Lewis Sheiner reflected in the FDA document titled "Innovation or Stagnation: Challenge and Opportunity on the Critical Path to New Medical Products" made this scientific approach to the drug approval process possible.

KEY WORDS: drug development; pharmacokinetics; pharmacodynamics; PK/PD; simulation.

INTRODUCTION

Materia medica is a Latin term for any material or substance used in the composition of curative agents in medicine. The use of materia medica was historically based upon the idea that medicines possess magical curative powers. Folk healers among the peasantry believed in the law of signatures to identify which herbs would be effective against which diseases. By a process of trial and error, folk healers learned which herbs worked and which herbs did not work. Materia medica eventually developed into the science of pharmacology. The subject of pharmacology has grown rapidly in recent years and new drugs are being synthesized at an ever-increasing rate. Clinical pharmacology has been described as a bridging discipline that combines elements of pharmacology with clinical medicine. In the past few decades drug development has progressed from the relatively ad hoc trial and error process of materia medica to an advanced scientific process involving careful experimental design, valid data analysis, informative graphics, mathematical modeling, and computer science. Lewis Sheiner (1940–2004) was one of the pivotal individuals in this changing evolution. Through his scientific writings, teaching programs, and collaboration with scientists around the world, his impact on drug development has been immeasurable. His philosophy of “Learn, confirm” cycles throughout drug development has become the catchphrase in the field of clinical pharmacology (1). Regulatory bodies have come to realize that proof of efficacy is only one step in a long process from drug discovery to marketing. This is reflected in the most recent guideline that the Food and Drug Administration (FDA) has published:– “Innovation or Stagnation: Challenge and Opportunity on the Critical Path to New Medical Products” (2). The FDA has emphasized the importance of integrating pharmacokinetic and pharmacodynamic (PK/PD) information in drug development and its potential impact on decision-making. This manuscript describes three real examples where the application of modeling and simulation techniques, as suggested by Lewis Sheiner, aided drug development at critical decision points in preclinical development, phase 2 and regulatory review.

EARLY DRUG DEVELOPMENT

The major objectives of early drug development are to select promising compounds and to identify potentially safe and effective doses and dosing regimens. Integration of PK/PD in early development helps with compound selection and guides creation of an efficient clinical development strategy.

This first example illustrates the application of PK/PD modeling in predicting response in a Phase 1 study and assessing the potential for differentiation between compounds. Three compounds of the same class (A, B, and C) were being investigated for the same indication and were at different stages in their development. These compounds were centrally active and speed of onset was a crucial determinant of efficacy. Thus, differences in both potency and time course of effect were important in identifying the most efficacious dose. The drug development strategy was to accelerate development of back-up compounds by capitalizing on previous information from compounds further along in the development and to advance the lead compound unless backups were clearly differentially better. Compound C was the last in the series, but had the potential to be a superior candidate.

Dose-ranging preclinical pharmacology studies using similar designs were conducted assessing response at scheduled times post administration of each compound. In addition to measuring pharmacologic response, an objective of these studies was to understand the pharmacokinetic (PK) profile. Thus, PK samples were obtained in a parallel group of animals. Compounds A, B, and C differ only slightly in terms of PK properties. However, they had different potency and quite different time courses of pharmacologic response after a single dose, with a delay in onset of pharmacologic activity relative to the PK profile. Figure 1 illustrates the time course of

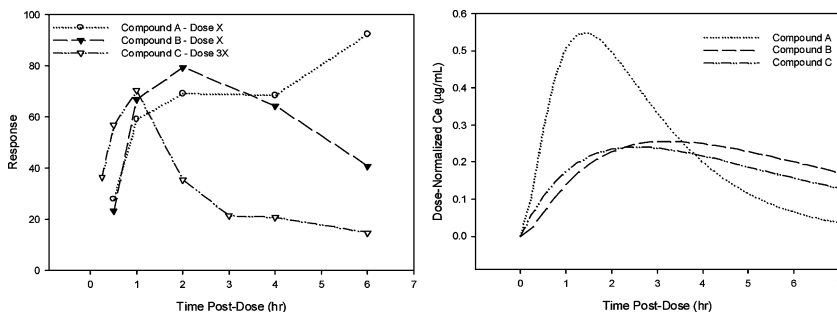


Fig. 1. Time course of pharmacologic response in preclinical studies (Left Panel) and Dose-Normalized concentrations in the effect site (Right Panel) for compounds A, B and C.

pharmacodynamic response for compounds A, B and C. Note that the mean response at 6 hr for compound A was not consistent with the rest of the data, however, this pattern was not observed at other dose levels, and did not adversely impact pharmacodynamic parameter estimates obtained by simultaneously fitting data from all dose levels. The relationship between concentration and response was assessed using the classical approach described by Sheiner *et al.* (3) popularly known as the “effect compartment model”. This seminal work made it practical to quantify the lag between the exposure to a drug and the onset and offset of the pharmacologic response. The PK/PD modeling was conducted in NONMEM version V (4) using individualized PK parameters to determine potency of each compound (EC50) and the delay between peak concentration and effect as described using effect site equilibration rate constant (ke_0). Results of the PK/PD modeling are summarized in Table I and predicted effect-concentration-time profiles illustrated in Fig. 1. The temporal delay observed between systemic concentrations and pharmacodynamic response may be related to slow distribution to the site of action, post-receptor mechanisms or other physiological processes involved in generating the measured response. At this stage of development quantifying the delay was adequate for comparing the compounds.

Thus, in terms of potency, the compounds rank from highest to lowest as follows: $B > A > C$ with potency relative to A of 0.3 for B and 4 for C. The compounds ranked as $C > A > B$, starting with the most rapid onset, with ke_0 relative to A of 12 for C and 0.7 for B. From these pre-clinical results, a model was developed to simulate the response profile in humans for compound C. The PK predictions were made based on allometric scaling of the PK parameters of systemic clearance and volume of distribution, estimates of absorption rate constant and assuming a similar lag time of absorption. For the pharmacodynamic parameters, ke_0 was held constant from the preclinical data, while ED50 was adjusted based on the

Table I. Parameter Estimates of the PK-PD Modeling of Preclinical and Clinical Data

PD Parameter	Compound A	Compound B	Compound C
Preclinical Data			
ke_0 (hr^{-1})	0.32	0.23	3.84
EC50 ($\mu g/mL$)	2.00	0.634	7.52
Clinical Data			
ke_0 (hr^{-1})	1.05	0.45	5.63
EC50 ($\mu g/mL$)	1.21	0.405	29.5

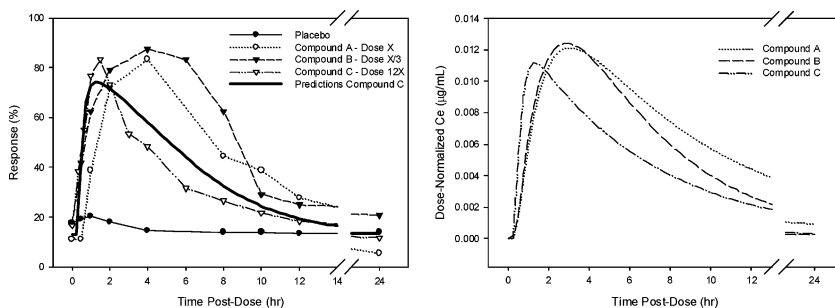


Fig. 2. Time course of pharmacologic response in Phase 1 study (Left Panel) and Dose-Normalized concentrations in the effect site (Right Panel) for compounds A, B and C.

potency ratios in Table I. A sigmoid E_{max} model has been developed based on data obtained following administration of compound A in healthy subjects (Fig. 2). Parameters of the pharmacodynamic model such as baseline, E_{max}, and sigmoidicity factor were assumed to be constant in making predictions. Although compound B was more potent than C, development was discontinued because of its slower onset and offset of action.

Results obtained in the single rising dose tolerance study in the subsequent clinical trial confirmed the simulation results (Fig. 2), showing that the relative potency and relative rate of onset determined in the preclinical model were consistent with the clinical data (Table I). Compound C had lower potency than predicted, but the rank order for potency and onset rate was maintained. These parameters were adjusted based on Phase 1 data and further used to predict doses in proof-of-concept studies.

The preclinical and clinical PK/PD models developed for compounds A and B were used together with preclinical data for compound C to predict the time course of pharmacologic activity in humans for compound C. This allowed development of go/no go decision criteria based on onset time observed in the first in human (FIH) study of compound C. The PK/PD predictions then provided a rationale for greater investment in compound C to move it to the FIH decision point more quickly. Finally, PK/PD modeling of relative potency was useful for dose selection for the FIH study, and later, using also phase 2a information from compounds A and B, for a phase 2a study of compound C.

The application of the so called “effect compartment model” approach that Lewis Sheiner described to quantify the discontinuity frequently seen between exposure to a drug and the observed clinical response has had enormous impact on the drug development discipline. Determination of relative potency, as measured by EC₅₀, and relative rate of onset of pharmacologic action, as assessed using the parameter ke_0 , can be scaled from

animal models to human and can be used to discriminate between compounds, saving considerable time and money that would otherwise be needed for human trials.

PHASE 2 DRUG DEVELOPMENT

Phase 2 proof of concept (POC) and dose-ranging trials in a select group of patients are essential to verify that the compound demonstrates sufficient efficacy and safety. At this stage it is important to characterize the relationship between drug exposure and response, be it with a relevant biomarker or some clinical endpoint. Information that is gathered during this phase provides grounds for the decision of whether the balance between efficacy and safety justifies investing in the much more expensive phase 3 program. Lewis Sheiner's influential publication describing the learn-confirm cycles that developers need to go through brilliantly describes this information gathering process (1). The procedure can be thought of as progressive cycles starting with learning in phase 1 about optimal conditions to initiate phase 2, and learning from phase 2 to decide whether or not to initiate phase 3 and the procedure that is most likely to meet with success (i.e. correctly answer the questions). During phase 2 a key feature of this approach is to use previous clinical trial information in the same disease with prior PK/PD information on the new compound to better predict the most probable trial outcome. The following example applies some of the principles championed by Lewis Sheiner and illustrates the use of the lessons learned during the preclinical phase and phase 1, together with literature data to design a phase 2 study.

CI-1017 is an M1-muscarinic acid agonist for which we had PK/PD information from Phase 1 and preclinical studies (5). The question facing the drug development team was how to proceed with the Phase 2 study and what criteria should be applied for a go/no-go decision based on its results. Traditionally, Alzheimer's Disease (AD) trials used parallel group, placebo-controlled designs, assessing the Alzheimer's Disease Assessment Scale-Cognitive Subscale (ADASCog) at 4, 6, or 8-week intervals over a trial duration of 12–30 weeks. In this instance, the availability of toxicology study results limited the total exposure to study drug to 12 weeks.

The drug development team decided to use the combined knowledge about AD and CI-1017 to simulate various clinical trial scenarios to aid the design of the first Phase 2 study. The trial was to answer several questions, three of which were most important. First, does the drug have benefit? Second, since preclinical data implied a possible inverse U-shape dose-response, is more drug better than less? The shape of the exposure-response curve was an important issue since this would have a major

impact on the dose selection and design to be used in the phase 3 study. In fact, if the exposure-response analysis did exhibit an inverse U-shape, this product might well be unmarketable. Finally, if the drug proved beneficial, an objective was to identify the target dose range for subsequent trials. The design of the phase 2 trial, therefore, needed to focus on the exposure-response profile not just the question of whether the drug gives a beneficial effect.

At this stage a useful benchmark was tacrine (an approved drug for the treatment of AD symptoms). Extensive in-house clinical trial data for tacrine was available as well as literature data describing the exposure-response relationship (6–8). The tacrine data together with phase 1 PK and preclinical data for CI-1017 were used to simulate various clinical trial scenarios to determine the optimal clinical trial design that would have the best likelihood of answering the above questions.

Prospective designs included a conventional parallel group design and three Latin Square (crossover) designs with different numbers of treatments and treatment durations. Each study was constrained to be about the same size, with sixty patients and a total of about six ADASCog measurements per patient, since one design objective was to choose the best design conditional on a specified level of investment. The intrinsic merit of the different designs was evaluated using two key metrics: (i) the percentage of simulated trials that correctly detected a drug effect, using the appropriate analysis method and decision rule (this is essentially the conventional statistical power) and (ii) the percentage of simulated trials that correctly identified the shape of the dose-response profile (monotonic or inverse U-shaped). Accuracy of target dose estimation was also assessed.

Because we were unsure of the form of the true exposure-response relationship four different models were simulated: three monotonic models (Emax [super-linear], linear, and sigmoid Emax [sublinear]) and one inverse U-shaped model. Assuming a three-point effect as the minimally acceptable profile (based on tacrine's effect), each of these response models were constructed to reach a maximum of three points in the tested dose range: at the highest dose for the monotonic models and at the "best dose" for the inverted U-shaped model. The four drug effect models are depicted in Fig. 3.

For each treatment sequence, a population of patients was created and sampled with replacement to generate individual clinical trials (from 100 to 2000 clinical trials depending upon the precision needed for the particular objective). One objective of the simulations was to determine the relative robustness of the designs to detect a significant ($\alpha = 0.05$) treatment effect for each of the assumed dose-response relationships. Table II reports the power or percentage of trials in which a statistically significant treatment

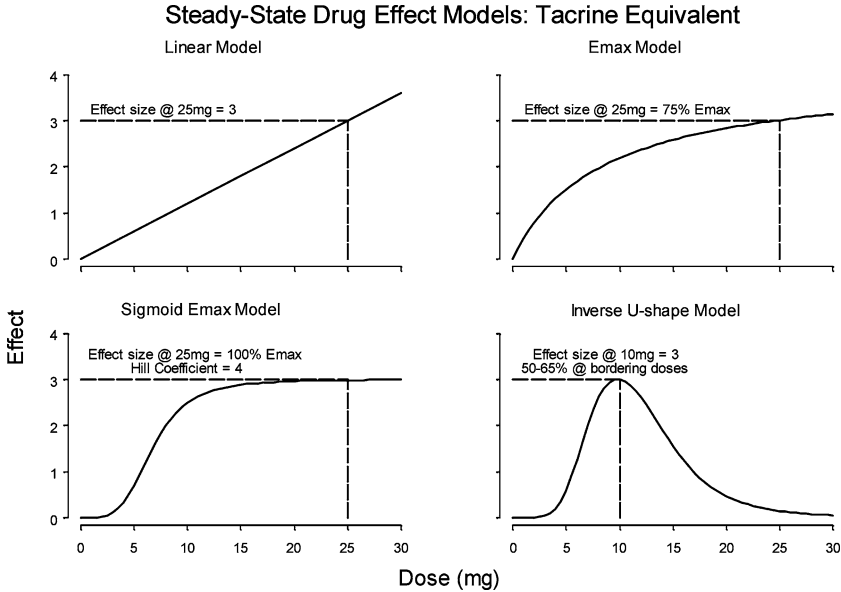


Fig. 3. Hypothetical drug effect models used to assess robustness of the designs to various assumptions for the CI-1017 dose-response relationship (from ref. 5).

Table II. Estimated Power to Detect a Significant ($\alpha = 0.05$) Treatment Effect (ref. 5)

Trial design	Parallel	4 × 4	4 × 4	6 × 6
Period length (weeks)	12	4	3	2
Dose response shape				
Linear	29	84	51	41
E _{max}	28	88	67	43
Sigmoid E _{max}	43	96	85	68
Inverse				
U-Shape	21	57	49	39
Average	30	81	63	48

effect was detected. The 4 × 4 Latin Square with four-week periods was the most robust design and performed well (> 80% power) for all three monotonic dose-response relationships examined. Power was lower to detect the inverse U-shaped dose-response. The parallel group design ($N = 60$) did not perform well, confirming the original power calculations that hundreds of patients would be needed. Based on these results, a 4 × 4 design with four-week periods was recommended and ultimately conducted saving significant resources (in excess of two million dollars compared to a parallel group

study with similar power), and the total elapsed conduct time was perhaps eight months shorter because of the reduced enrollment burden. The statistical models were ready for a rapid analysis of the data after the completion of the study. In this case, the results failed to meet the predefined criteria, and the drug project was cleanly terminated without the need for endless rework of the data looking for an elusive signal, as is often the case with negative studies.

Phase 3 and Beyond

Phase 3 clinical trials confirm the efficacy of new therapies as the final step prior to submission to a regulatory body for approval. Based on the 1962 amendment to the Federal Food, Drug and Cosmetic Act of 1938 (9), regulators interpretation was that substantial evidence of effectiveness required at least two adequate and well-controlled (AWC) investigations, presumably to ensure that the findings are repeatable. The FDA Modernization Act of 1997 and the FDA “effectiveness” guidance of 1998 (10,11) opened the door to using exposure-response information in combination with a single pivotal clinical trial as sufficient evidence of effectiveness. Although this situation may be relatively rare, there are much more common situations for which exposure-response data can support registration decisions and labeling. The recent gabapentin sNDA approval for neuropathic pain is a good example. Gabapentin (Neurontin®) was originally approved in the U.S. in 1993 as adjunctive therapy in the treatment of partial seizures in patients with epilepsy. Post marketing, anecdotal evidence suggested that patients suffering with post herpetic neuralgia (a syndrome of neuropathic pain following herpes zoster and occurring in approximately 10–15% of all herpes zoster patients) may receive pain relief with gabapentin.

To support the marketing approval of gabapentin for the indication of PHN, the sNDA package contained efficacy data from two trials conducted in patients with PHN. Unfortunately, these trials did not study the same dose levels—patients in study 1 were randomized to a final dose of 3600 mg/day while patients in study 2 were randomized either to 1800 or 2400 mg/day. Thus there was not replicate data for efficacy at these doses. Understanding of the exposure-response for gabapentin in PHN was further complicated by the saturable absorption of gabapentin, leading to less than dose proportional changes in exposure. The lack of replicated findings at a single dose, coupled with the saturable absorption and nonlinear exposure, presented a unique regulatory hurdle to approval for the PHN indication.

An exposure-response (ER) analysis was completed and included with the submission. Although ER concepts are not new, their value as part of

regulatory submissions has changed. ER evaluation took on a new role with the advent of FDAMA 1997 (10,11), which stated in Section 115 that “... based on relevant science, that data from one adequate and well-controlled clinical investigation and confirmatory evidence (obtained prior to or after such investigation) are sufficient to establish effectiveness...”. This means that instead of the conventional requirement of two AWC clinical trials to support a new drug application, the sponsor could rely on the results of a single AWC clinical trial combined with confirmatory evidence (such as ER data). Since 1977, prior to the gabapentin case no approved drug labeling has made use of this concept.

The gabapentin sNDA represented the first application of ER information to establish a linkage across two pivotal clinical studies to provide confirmatory evidence of dose response under the FDAMA ruling. Population PK/PD analyses were submitted showing that gabapentin exhibits exposure-dependent decreases in daily pain scores (Figs. 4 and 5) (12). In this analysis, the decrease in daily pain score was adequately described by an E_{max} model expressed as change from baseline minus effect of drug (gabapentin dose corrected for estimated bioavailability) and time-dependent placebo effect, with gender and disease duration having minimal effect on the ER relationship. Data from the titration period from each study were also successfully incorporated into the model to provide more information about within-subject dose response. The models and the supporting data were all made readily available to the FDA reviewers.

During deliberation among the FDA review staff which included clinicians, clinical pharmacologists, pharmacometricians and biostatisticians, it was proposed by the clinical pharmacologists and agreed to by others to explore whether the PK/PD analyses could provide the confirmatory

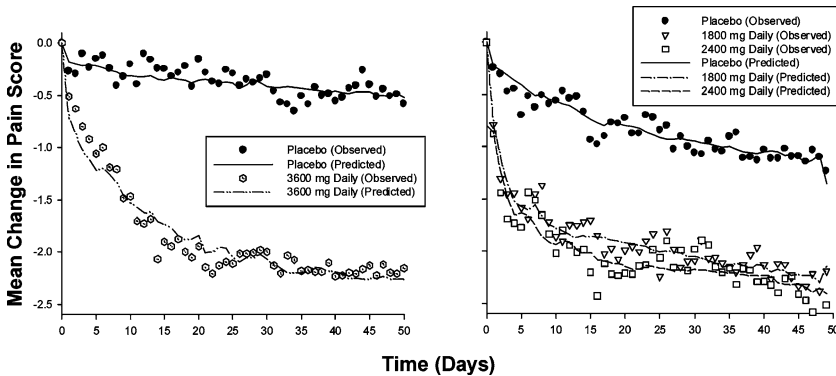


Fig. 4. Change in pain score from baseline over time.

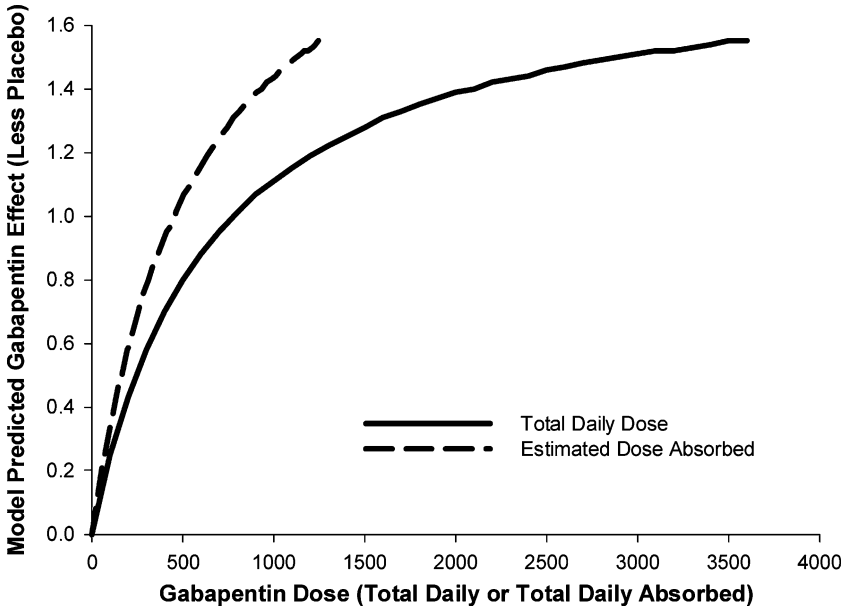


Fig. 5. Model predicted gabapentin effect by total daily dose and estimated dose absorbed.

evidence of efficacy. To replace a replicate clinical trial the PK/PD analysis would have to withstand the same qualitative and quantitative review that data from an AWC would, so the analyses were further tested and reviewed by the FDA, who reached similar conclusions:– The PK/PD results confirmed evidence of efficacy across the three studied doses, so that additional clinical trials would not be needed for approval of the product.

In the Package Insert of approved drugs, the establishment of effectiveness of drugs is described in the clinical studies section, where the design, patient populations, and statistical considerations of the studies are explained. In the case of gabapentin, the clinical trials section of the Package Insert states “PK/PD modeling provided confirmatory evidence of efficacy across all doses” to explain the basis of establishing the effectiveness of Neurontin® for the PHN indication (13).

In a recent publication (14) Lewis Sheiner and co-authors contend that “convincing evidence of the pharmacologic mechanism of the clinical effect of a drug serves the same purpose as—and can be derived more directly from sources other than—a second large and expensive clinical trial”. They proposed that ‘contrary to naïve fears of the consequences of “lowering standards” of effectiveness, adopting the now legally

sanctioned new standard of a single clinical trial plus confirmatory evidence (SCT-CE) would provide a scientific and regulatory framework for encouraging more rational, more efficient, and more informative clinical drug development.’ The rational drug development paradigm that he so enthusiastically encouraged seems to be gaining traction.

CONCLUSION

In the publication by the FDA titled “Innovation or Stagnation: Challenge and Opportunity on the Critical Path to New Medical Products” (2) the opportunities offered by the concept of model-based drug development are mentioned, and the reference cited is the seminal manuscript by Lewis Sheiner regarding the “learn-confirm” paradigm (1). This present review gives some examples of work based on the concepts championed by Lewis Sheiner, in which pharmaco-statistical models of drug efficacy and safety are developed from preclinical and available clinical data to improve drug development knowledge management and development decision-making. With the publication of the Innovation or Stagnation document, the FDA also has acknowledged the value of these concepts and will collaborate with others in the use of modeling and simulation to improve trial design and drug development decision-making.

REFERENCES

1. L. B. Sheiner. Learning versus confirming in clinical drug development. *Clin. Pharmacol. Ther.* **61**:275–291 (1997).
2. Food and Drug Administration. Innovation; Stagnation: Challenge and Opportunity on the Critical Path to New Medical Products. www.fda.gov/oc/initiatives/critical-path/whitepaper.html. 2004.
3. L. B. Sheiner, D. R. Stanski, S. Vozeh, R. D. Miller, and J. Ham. Simultaneous modeling of pharmacokinetics and pharmacodynamics: application to d-tubocurarine. *Clin. Pharmacol. Ther.* **25**:358–371 (1979).
4. S. L. Beal and L. B. Sheiner. *NONMEM Users Guide*. NONMEM project group. University of California, San Francisco, 1998.
5. R. Miller, J. D. Hermann, W. Ewy, and K. G. Kowalski. The strategic role and application of pharmacokinetic/pharmacodynamic modeling in drug development. In: *Pharmacokinetics in Drug Development: Clinical Study Design and Analysis*, Vol. 1, P. L. Bonate and D. R. Howard (eds.), AAPS Press, Arlington VA, 2004, pp. 551–581.
6. N. H. G. Holford and K. E. Peace. Methodologic aspects of a population pharmacodynamic model for cognitive effects in Alzheimer’s patients treated with tacrine. *Proc. Natl. Acad. Sci. USA.* **89**:11466–11470 (1992a).
7. N. H. G. Holford and K. E. Peace. Results and validation of a population pharmacodynamic model for cognitive effects in Alzheimer’s patients treated with tacrine. *Proc. Natl. Acad. Sci. USA.* **89**:11471–11475 (1992b).
8. N. H. G. Holford and K. E. Peace. The effect of tacrine and lecithin in Alzheimer’s disease. A population pharmacokinetic analysis of five clinical trials. *Eur. J. Clin. Pharm.* **47**:17–23 (1994).

9. Federal Food, Drug and Cosmetic Act of 1938, Drug Amendments of 1962. Pub. L. No. 87-781, 76 Stat. 780 (1962).
10. Food and Drug Administration Modernization Act of 1997, Pub. L. No. 105-115, 111 Stat. 2295 (1997).
11. Guidance for Industry: Exposure-response relationships, study design, data analysis and regulatory applications. Food and Drug Administration, May 2003, Guidance for Industry: Providing Clinical Evidence of Effectiveness for Human Drugs and Biological Products. Food and Drug Administration, May 1998.
12. B. Frame, P. Burger, R. Miller, J. Cook, E. Garafolo, L. Knapp, H. N. Bockbrader, and B. W. Corrigan, *An Analysis of Exposure Response data for Gabapentin for the Treatment of Neuropathic Pain*. Neuropathic Pain Meeting, San Francisco, 2001.
13. Neurontin® (Parke-Davis) Capsules. Physicians Desk Reference Electronic Library™ Online: ©2002. Thomson MICROMEDEX.
14. C. C. Peck, D. B. Rubin, and L. B. Sheiner. Hypothesis: A single clinical trial plus causal evidence of effectiveness is sufficient for drug approval. *Clin. Pharmacol. Ther.* **73**:481-490 (2003).