Using Disease Progression Models as a Tool to Detect Drug Effect

DR Mould[1], NG Denman[2] and S Duffull[3]

Generally, information required for approval of new drugs is dichotomous in that the drug is either efficacious and safe or not. Consequently, the purpose of most confirmatory clinical trials is to test the null hypothesis. The primary reasons for designing hypothesis testing trials are to provide the information required for approval using analyses techniques that are relatively straightforward and free of apparent assumptions. However, the information required for approval is very different from that used by prescribers for decision making. In the clinic, decisions must be made about dose adjustment for individual patients in the presence of additional therapies and co-morbidities. Choice of drug and dosing regimen is therefore a classical risk to benefit decision that is often poorly informed from the results of confirmatory trials. Therefore, providing answers to the more difficult question of how to use the drug in a clinical setting is essential.

Sheiner[1,2] and others[3] suggested that greater statistical power can be achieved through the use of alternative hypotheses that test accepted scientific models of disease and response rather than simply testing the null hypothesis where the sole covariate in question is the presence of the study drug. Evaluating the dose-response surface involves developing disease progression and drug effects models, and is described by Sheiner as a part of the learning and confirming process in drug development.[4]

Modeling involves developing mathematical equations to describe quantitative relationships for individuals or populations. These equations can be used to predict the time course of disease and drug effects in settings other than those explicitly studied. NONMEM (NONlinear Mixed Effects Modeling) is a program that allows model building to be performed using a population approach.[5] One important feature of population analyses is the ability to describe between-subject variability, and to account quantitatively for covariate influences, such as weight, which can explain some of this variability.

The concept of evaluating disease progression through model-based evaluations is not new. In a work published in 1981 by Holford and Sheiner,[6] the authors proposed a new meaning for an old model:

$$E(t) = E_0 + \frac{E_{\max} Cp(t)}{EC_{50} + Cp(t)} \quad (1)$$

where $E(t)$ represents a measure of disease status at time t and E_0 represents the baseline disease status measurement. $E_{\max}$ and EC_{50} are parameters representing the maximal achievable drug effect and the drug concentration at half maximal response, respectively, and $C_p(t)$ is the predicted (plasma) drug concentration. In these terms, equation (1) represents a "zero progression" model for disease status such that over the course of evaluation, the observed disease status does not change except through therapeutic intervention.

However, there is ample evidence that disease status often does not remain stable over the course of a clinical trial. For example, Griggs *et al.*[7] described the results of a randomized, placebo controlled trial of prednisone in 99 boys aged 5–15 years with Duchenne muscular dystrophy, and reported that muscle strength decreased in the placebo group over 6 months of observation (**Figure 1**). In this study, the "zero progression" model was generalized to allow for disease status to change linearly over time (equation (2)):

$$S(t) = S_0 + \alpha t \quad (2)$$

where $S(t)$ is the disease status at some time t, S_0 represents the baseline disease status, and α represents the rate of change (slope) of disease status.

The linear model of disease progression was further modified by the addition of an "effect compartment",[8] in order to allow for a delay between the initiation of treatment and the time to observable response (equation (3)).

$$S(t) = S_0 + E_{\mathrm{OFF}}(Ce_A) + \alpha t \quad (3)$$

[1]Projections Research Inc., Phoenixville, Pennsylvania, USA; [2]School of Physical Sciences, University of Queensland, Brisbane, Queensland, Australia; [3]School of Pharmacy, University of Otago, Dunedin, New Zealand. Correspondence: DR Mould (drmould@attglobal.net)

Published online 16 May 2007. doi:10.1038/sj.clpt.6100228

where $E_{OFF}(Ce_A)$ represents a drug effect that provides symptomatic benefit as a transient improvement (*offset*) in the disease status during treatment, but which predictably reverses to the untreated status after treatment is discontinued. Conversely, the effect of treatment can be described as altering the progression of disease (equation (4))

$$S(t) = S_0 + (E_{PROG}(t) + \alpha)t \qquad (4)$$

where $E_{PROG}(t)$ is the effect of treatment on the *progress* of disease, modifying the slope of the disease progression. In this model, the disease status would not return to the pretreatment course when therapy was discontinued, but would be expected to result in a permanent improvement (**Figure 2**). Several publications give good overviews of disease progression models and associated parameterizations for describing treatment effects.[9,10]

Linear equations have been used to describe the progression of several diseases, including Alzheimer's disease and schizophrenia.[11,12] In these models, the disease progression component included a placebo response model, incorporating both the trajectory of disease as well as a transient change in disease status attributed to placebo response. It is important to note that there are differences between a true placebo effect and a perceived placebo effect.[13] The perceived placebo effect is a mixture of several factors, including a natural tendency for individuals to regress to their mean status, introduction of unidentified parallel interventions, and a true placebo effect. This distinction coincides with the approach taken with disease progression modeling, which attempts to describe the natural time course of the disease as well as the effect of placebo.[9]

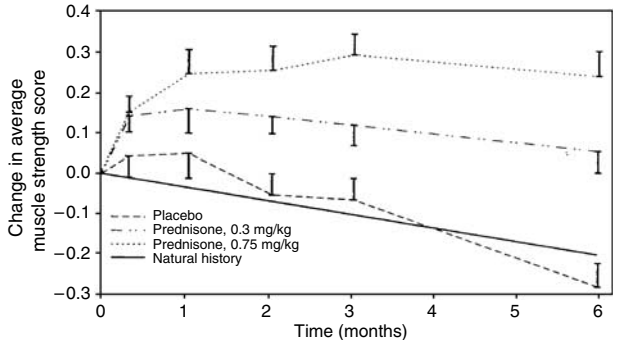

Figure 1 Change in average muscle strength score over time in patients with Duchenne muscular dystrophy. The solid line represents the trajectory of the disease; the dashed line the placebo response; the dashed and dotted line prednisone at 0.3 mg/kg, and the dotted line prednisone at 0.75 mg/kg. The solid bars are standard error bars.

As mentioned previously, models for disease progression characteristically have high unexplained between-subject variability. This is consistent with the observation that between-subject variability is greater for pharmacodynamic than pharmacokinetic evaluations.[14,15] In some cases, measures of disease status are based on summary scores, such as the Hamilton depression score (HAM-D), a composite score measuring the severity of depressive symptoms. The HAM-D score consists of either 17 or 21 items for which an interviewer provides ratings for selected symptoms, including overall depression, guilt, suicide, insomnia, and other symptoms. Perhaps as a consequence of the multiple ratings used to generate the summary score, evaluations of depression scores tend to have high variability, making evaluation of data from clinical trials in depression difficult using traditional empirical methods. However, model-based evaluations should provide a better tool to detect drug effect because these evaluations naturally allow disease severity to change over time.

Another potential benefit of modeling is that more informative clinical trial designs can be developed using optimal design techniques. These techniques are based on evaluating the population Fisher information matrix for any model, given some set of design variables. Maximizing its determinant, a summary measure of the overall information matrix, is called D-optimal design, which is based on the seminal work of Mentre *et al.*,[16] and updated by Retout *et al.*[17]

When performing an empirical evaluation of clinical trials of antidepressants, the considerable variability in HAM-D and the pronounced placebo response, which may itself be

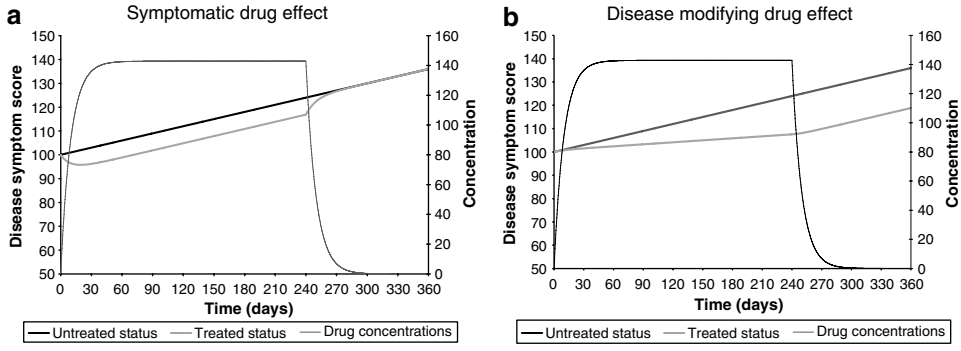

Figure 2 Symptomatic and disease modifying drug effect on the progression of disease. (**a**) A drug with symptomatic effect and (**b**) a disease modifying effect. For both panels, the solid grey line represents active treatment; the solid black line the trajectory of the disease; and the gray line the disease progression with treatment effect.

affected by the trial design, yields a high failure rate. This failure is not necessarily because the antidepressant is ineffective, but often results from an inability to extricate drug effect from the placebo response.[18,19] Fifty-two randomized, double-blind, placebo-controlled pivotal clinical trials were evaluated to correlate placebo response as the percentage mean change from baseline in HAM-D score with trial outcome.[20] It was found that in trials with a greater than 30% mean change from baseline in the placebo group, only 21.1% of the trials found active treatment superior to placebo. However, in trials with a smaller placebo response, more trials (74.2%) showed superiority of active treatment. Thus, the magnitude of placebo response in depression trials is variable and has a considerable impact on the power of antidepressant trials to evaluate the effect of active treatment.

The choice of statistical methods has also been cited as a factor in identifying the effect of active treatment,[21] and exploiting advances in statistical techniques has been recommended.[22] The use of disease progression modeling therefore constitutes an important tool in the evaluation of drug effect, particularly when the placebo response is substantial and/or the data are highly variable.

EXAMPLE CASE STUDY

Clinical trials evaluating the effect of novel antidepressants make a good example for comparing the power of model-based evaluation to traditional evaluation of drug effect. Several models have been proposed for evaluating the time course of depression and the associated placebo response, including the inverse Bateman model[23,24] and the K-PD model.[25–27] Shang *et al.*[28] evaluated these models and also a two-transit-compartment model (**Figure 3**)[29] for their ability to describe the time course of placebo response. This evaluation suggested that the transit model had better characteristics than the inverse Bateman function or the K-PD model. The influence of placebo was handled by creating a driver for the placebo effect in the form of a one-compartment model with bolus input and first-order elimination. The parameter estimates for the placebo model are presented in **Table 1**. Placebo was dosed at a daily unit dose to mimic placebo treatment schedules. The parameters used for the pharmacokinetic model (equation (5)) were selected to ensure that concentrations of placebo were close to zero at the end of each day.

$$\frac{\mathrm{d}Cplacebo}{\mathrm{d}Time} = -\frac{CLplacebo\ Cplacebo}{Vplacebo} \tag{5}$$

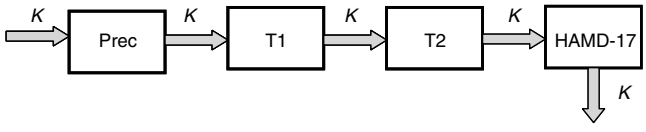

Figure 3 Schematic of two transit compartment model used to describe the change of HAM-D scores over time. "Prec" represents a precursor status; "T1" and "T2" transit compartments; and HAMD the observed depression score. "*K*" is the transit rate such that $K = \mathrm{MTT}/3$.

where *Cplacebo* is the predicted placebo "concentration", *CLplacebo* and *Vplacebo* are the "clearance" and "volume" of placebo, respectively. The equations describing the transit model are shown in equation (6).

$$\begin{aligned}
\frac{\mathrm{d}Prec}{\mathrm{d}Time} &= KS_0(1 - Slope\ Cplacebo) - K\ Prec \\
\frac{\mathrm{d}T_1}{\mathrm{d}Time} &= K\ Prec - KT_1 \\
\frac{\mathrm{d}T_2}{\mathrm{d}Time} &= KT_1 - KT_2 \\
\frac{\mathrm{d}HAMD}{\mathrm{d}Time} &= KT_2 - K\ HAMD \\
HAMD &= HAMD(t) + \varepsilon_{\mathrm{Additive}}
\end{aligned} \tag{6}$$

Determination of optimal study design

Denman *et al.*[30] proposed informative study designs for the three structural models of placebo effect in depression. The optimization procedure was performed using WinPOPT.[31]

Commonly implemented study designs for clinical depression involve a short run-in period with 6–8 weeks treatment and no washout. HAM-D scores are assessed at screening and then weekly thereafter. The optimal study design for estimating the disease progression model parameters included a short run-in period, a 4-week treatment period and 2 weeks washout. HAM-D scores are collected over the entire 6-week period at 1, 2, 14, 25.3, 40, and 56 days. The last HAM-D observation time is therefore 14 days after the last treatment. This design involves the same study duration as the traditional design but provides better information for model-based evaluation, and would not be expected to negatively impact traditional empirical evaluation. The present design was not optimized to estimate the pharmacokinetics of an active drug.

Clinical trial simulation and evaluation

Two basic study designs were investigated: the traditional study design for 6 weeks treatment with no washout evaluations, and the design optimized for model evaluation. These designs were evaluated using an empirical analysis. The optimal design was also evaluated using model-based analysis.

Empirical analysis. A series of studies were simulated to assess the power of the optimal and empirical designs to detect a true drug effect. The simulated studies included 200 subjects with a 1:1 placebo to active treatment allocation.

Table 1 Pharmacokinetic parameter values for placebo function

Parameter (units)	Parameter name	Value
Clearance (l/h)	*CLplacebo*	0.125
Volume of distribution	*Vplacebo*	1.00
Placebo effect	*Slope*	0.046

Subjects received daily placebo doses regardless of treatment assignment. The active drug was administered daily. The following two different drug effects were assessed: (1) a weakly active drug that resulted in a 2 HAM-D unit reduction over placebo and (2) a strongly active drug that resulted in a 5 HAM-D unit reduction over placebo. The effect of drug was implemented as a simple offset model (equation (7)), and hence the drug had no effect on the progression of the depressive illness.

$$HAMD = HAMD(t) - Active + \varepsilon_{\text{Additive}} \quad (7)$$

The parameters for the transit model and the active drug are presented in **Table 2**. Two hundred replicates of the clinical study were simulated and evaluated using the traditional statistical evaluation involving calculation of the change from baseline HAM-D (ΔHAM-D). Comparison of ΔHAM-D between the placebo and active cohorts was undertaken using a non-paired *t*-test ($\alpha = 0.05$). The drug effect was considered clinically significant if the mean ΔHAM-D was at least 2 U greater than placebo. All patients in the simulated clinical trials had HAM-D scores greater than or equal to 20 during the 2-week run-in period.

Model-based analysis. A model-based evaluation was conducted by simulating, and then fitting the data using the transit model and evaluating the precision of the estimate of the active drug effect. Only the optimal design using the weakly active drug was evaluated. The pharmacokinetics of the active drug were not considered in this evaluation. Because the 25.3-day sampling time for the active drug is at steady state, it was assumed that the dose would be a sufficient statistic for the concentration. Furthermore, the between-subject variability in this value (49%) should account for variability in clearance, as well as any additional residual variability on the HAM-D score associated with the active drug over and above placebo.

The simulation and estimations were repeated 100 times to provide summary statistics about the precision with which the active drug effect could be estimated. All simulations were performed using MATLAB Version 2006a (The Mathworks, Natick, MA). NONMEM Version V Level 1.1 (Globomax, Hanover, MD), compiled using Compaq Digital Visual Fortran Version 6.6.3C (Hewlett Packard, Palo Alto, CA), was used for model-based evaluations. The first-order conditional estimation method was used for model evaluations.

RESULTS

Empirical analysis. The power to detect a difference between the treatment and placebo using the standard empirical analysis is shown in **Table 3**. For the weakly active drug, the power is poor for both the empirical and the design optimized for model development (41 and 46.5%, respectively). For the strongly active drug, the power was 100% for both study designs. These results suggest that using a design that is optimized for model development does not negatively impact on the ability to detect drug effect when using standard empirical evaluation.

Model-based analysis. Using a model-based evaluation of the data, the mean drug effect (±SE) was estimated to be −1.97 (±0.402), and the 95% confidence interval of the mean drug effect ranged from −2.75 to −1.18 HAM-D units. Frequency histograms for all of the parameters are shown in **Figure 4**. The distributions for all parameters are narrow, indicating that the drug effect was estimated with high precision and sufficient confidence to show the weak drug effect under the optimally designed conditions.

DISCUSSION

The stated purpose for most clinical trials is to test the null hypothesis, which provides sufficient information for drug approval. However, such analyses may not provide useful information for use of the drug in a clinical setting.

Table 3 Comparison of expected results using a standard evaluation of drug effect based on change from baseline HAM-D at end of treatment

	Placebo[a] (%)	Weakly active drug (%)	Strongly active drug (%)
Optimized design	0	41	100
Empirical design	0	46.5	100

HAM-D, Hamilton depression score. [a]No difference expected therefore equivalent to alpha error.

Table 2 Parameter values for transit model and active drug pharmacokinetics

Parameter (units)	Typical parameter value	Parameter name	Between-subject variability
HAM-D transit model parameters			
S_0	23	S_0	7%
Mean transit time	11.2 days	MTT	—
HAM-D residual error	—	$\varepsilon_{\text{Additive}}$	3.2
Active drug pharmacokinetic model parameters			
Offset effect	−2	Active	49%

HAM-D, Hamilton depression score. S_0 was simulated using a truncated distribution. The minimum value of S_0 was 20 HAM-D units and the maximum value of S_0 was 30 HAM-D units.

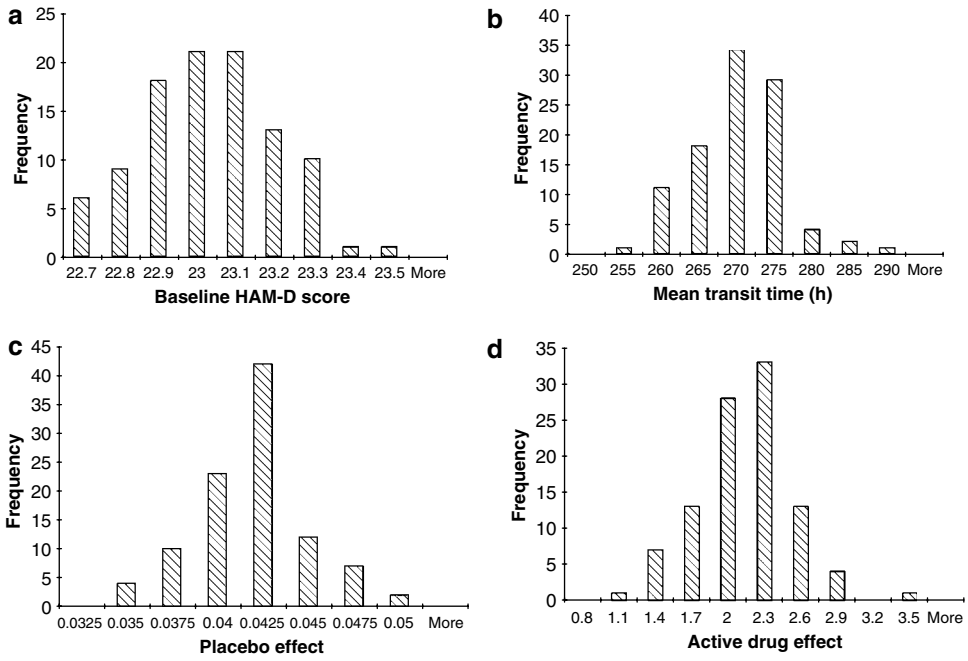

Figure 4 Frequency histograms of the distribution of the population parameter estimates from the case study. (**a**) The baseline HAMD-D score (S_0); (**b**) the mean transit time (MTT); (**c**) the placebo effect (slope); and (**d**) the active drug effect.

Furthermore, when the end point being evaluated is highly variable or when there is a substantial placebo response, the power to detect a drug effect using traditional techniques is often dramatically reduced. In the example illustrated here, there was an approximately twofold decrease in the power of traditional empirical analysis compared to a model-based evaluation.

Disease progression modeling is an additional evaluation tool that has an improved ability to detect a drug effect and provides useful dosing information for prescribers. Modeling is also an important component in the regulatory requirements for a new drug application.[32,33] Models of disease progression have been used for a wide variety of clinical indications. For each indication, the functions used to describe the progression must be based on the marker of disease activity being evaluated. In some cases, an indicator of drug activity that is considered acceptable for drug approval may be the avoidance of an event (*e.g.*, osteoporotic fracture) rather than alteration in a biomarker of disease, in which case disease progression modeling of the important predictor(s) (*e.g.*, bone mineral density) will yield information about the probability of the event, given the drug exposure and information regarding future monitoring of treatment in the clinic. In some cases, the information contained in the study design is also an important factor in determining an appropriate model for describing disease progression. Therefore, study conditions must be adequately evaluated to ensure that physiologically relevant models can be developed.

Information about the onset and offset of drug action is likely to be important to understanding exposure response, especially when the onset of drug activity may be overshadowed by placebo response. As active drug effect includes the placebo effect, it is difficult to delineate these processes without the aid of modeling. Informative study designs (*e.g.*, those using formal optimal design techniques such as D-optimality) improve the amount of information obtained from a clinical trial.

In conclusion, model-based evaluation of disease progress provides insight into drug activity. With appropriately informative study designs, greater insight into the mechanism and extent of drug effect on disease progression can be evaluated.

CONFLICT OF INTEREST

The authors declared no conflict of interest.

1. Jonsson, E.N. & Sheiner, L.B. More efficient clinical trials through use of scientific model-based statistical tests. *Clin. Pharmacol. Ther.* **72,** 603–614 (2002).
2. Sheiner, L.B. & Rubin, D.B. Intention-to-treat analysis and the goals of clinical trials. *Clin. Pharmacol. Ther.* **57,** 6–15 (1995).
3. Little, R.J. & Rubin, D.B. Causal effects in clinical and epidemiological studies via potential outcomes: concepts and analytical approaches. *Annu. Rev. Public Health* **21,** 121–245 (2000).
4. Sheiner, L.B. Learning versus confirming in clinical drug development. *Clin. Pharmacol. Ther.* **61,** 275–291 (1997).
5. Beal, S.L. & Sheiner, L.B. The NONMEM system. *Am. Stat.* **34,** 118–119 (1980).
6. Holford, N.H. & Sheiner, L.B. Understanding the dose–effect relationship: clinical application of pharmacokinetic–pharmacodynamic models. *Clin. Pharmacokinet.* **6,** 429–453 (1981).
7. Griggs, R.C. *et al.* Prednisone in Duchenne dystrophy. A randomized, controlled trial defining the time course and dose-response. Clinical

Investigation of Duchenne Dystrophy Group. *Arch. Neurol.* **48,** 383–388 (1991).
8. Holford, N.H.G. & Sheiner, L.B. Kinetics of pharmacologic response. *Pharmacol. Ther.* **16,** 143–166 (1982).
9. Holford, N.H.G., Mould, D.R. & Peck, C. Disease Progression Models. In *Principles of Clinical Pharmacology* 2nd edn. (eds. Atkinson, A.J. Jr, Abernethy, D.R., Daniels, C.E., Dedrick, R.L. & Markey, S.P.) pp 313–322 (Academic Press, New York, NY, 2006).
10. Mould, D.R. Development of Models of Disease Progression. In *Pharmacometrics: The Science of Quantitative Pharmacology* (eds. Ette, E. & Williams, P.) pp 547–581 (John Wiley and Sons, Hoboken, NJ, 2007).
11. Holford, N.H. & Peace, K.E. Methodologic aspects of a population pharmacodynamic model for cognitive effects in Alzheimer patients treated with tacrine. *Proc. Natl. Acad. Sci. USA* **89,** 11466–11470 (1992).
12. Kimko, H.C., Reele, S.S., Holford, N.H. & Peck, C.C. Prediction of the outcome of a phase 3 clinical trial of an antischizophrenic agent (quetiapine fumarate) by simulation with a population pharmacokinetic and pharmacodynamic model. *Clin. Pharmacol. Ther.* **68,** 568–577 (2000).
13. Ernst, E. & Resch, K.L. Concept of true and perceived placebo effects. *BMJ* **311,** 551–553 (1995).
14. Hattis, D. & Lynch, M.K. Empirically observed distributions of pharmacokinetic and pharmacodynamic variability in humans – implications for the derivation of single point component uncertainty factors providing equivalent protection as existing RfDs, www2.clarku.edu/faculty/dhattis/Lipscomb%20EPA%20report.doc.
15. Levy, G. Predicting effective drug concentrations for individual patients: determinants of pharmacodynamic variability. *Clin. Pharmacokinet.* **34,** 323–333 (1998).
16. Mentre, F., Mallet, A. & Baccar, D. Optimal design in random-effects regression models. *Biometrika* **84,** 429–442 (1997).
17. Retout, S., Duffull, S. & Mentre, F. Development and implementation of the population Fisher information matrix for the evaluation of population pharmacokinetic designs. *Comput. Methods Programs Biomed.* **65,** 141–151 (2001).
18. Khan, A. & Schwartz, K. Study designs and outcomes in antidepressant clinical trials. *Essent. Psychopharmacol.* **6,** 221–226 (2005).
19. Dworkin, R.H., Katz, J. & Gitlin, M.J. Placebo response in clinical trials of depression and its implications for research on chronic neuropathic pain. *Neurology* **65**(Suppl 4): S7–S19 (2005).
20. Khan, A., Detke, M., Khan, S.R.F. & Mallinckrodt, C. Placebo response and antidepressant clinical trial outcome. *J. Nerv. Ment. Dis.* **191,** 211–218 (2003).
21. Klein, D.F. A checklist for developing and reviewing comparative treatment evaluations in the areas of psychotherapy and pharmacotherapy. In *Psychotherapy Indications and Outcomes* (ed. Janowsky, D.S.) pp 367–381 (American Psychiatric Press Inc., Washington, DC, 1999).
22. Greist, J.H., Mundt, J.C. & Kobak, K. Factors contributing to failed trials of new agents: can technology prevent some problems? *J. Clin. Psychiatry* **63,** 8–13 (2002).
23. Holford, N., Li, J., Benincosa, L. & Birath, M. Population disease progress models for the time course of HAMD score in depressed patients receiving placebo in anti-depressant clinical trials. PAGE 11 (2002) Abstr. 311 [www.page-meeting.org/?abstract=311].
24. Holford, N. The time course of placebo response in clinical trials – do antidepressants really take two weeks to work? *AAPS Annual Meeting* (2005) [http://www.aapspharmaceutica.com/inside/focus_groups/ModelSim/imagespdfs/Holford05.pdf].
25. Jacqmin, P. *et al.* Modeling drug induced changes in biomarkers without using drug concentrations: introducing the K-PD model. PAGE 10 (2001) Abstr. 232 [www.page-meeting.org/?abstract=232].
26. Pillai, G., Gieschke, R., Goggin, T., Jacqumin, P., Schimmer, R.C. & Steimer, J-L A semimechanistic and mechanistic population PK–PD model for biomarker response to ibrandronate, a new bisphosponate for the treatment of osteoporosis. *Br. J. Clin. Pharmacol.* **58,** 618–631 (2004).
27. Cosson, V. & Gomeni, R. Modelling placebo response in depression using a mechanistic longitudinal model approach. PAGE 14 (2005) Abstr. 818 [www.page-meeting.org/?abstract=818].
28. Shang, E. *et al. Model-based Analysis of Placebo Response in Major Depression* (ACCP, Cambridge, MA, 2006).
29. Friberg, L.E., Freijs, A., Sandstrom, M. & Karlsson, M.O. Semiphysiological model for the time course of leukocytes after varying schedules of 5-fluorouracil in rats. *J. Pharmacol. Exp. Ther.* **295,** 734–740 (2000).
30. Denman, N., Mould, D. & Duffull, S. *Study Designs to Assess the Placebo Response (PBO-R) for Hamilton Depression (HAMD) Scores for Depression* (AAPS, San Antonio, TX, 2006).
31. Duffull, S.B., Eccleston, J., Kimko, H. & Denman, N.G. *WinPOPT User Guide*, (University of Otago, Dunedin, 2006) [www.winpopt.com].
32. Bhattaram, V.A. *et al.* Impact of pharmacometrics on drug approval and labeling decisions: a survey of 42 new drug applications. *AAPS J.* **7,** E503–E512 (2005).
33. Bhattaram, V.A. *et al.* Impact of pharmacometric reviews on new drug approval and labeling decisions – a survey of 31 new drug applications submitted between 2005 and 2006. *Clin. Pharmacol. Ther.* **81,** 213–221 (2007).