Research Article

Bayesian Quantitative Disease–Drug–Trial Models for Parkinson’s Disease to Guide Early Drug Development

Joo Yeon Lee^{1,2} and Jogarao V. S. Gobburu¹

Received 6 April 2011; accepted 6 July 2011; published online 27 July 2011

Abstract. The problem we have faced in drug development is in its efficiency. Almost a half of registration trials are reported to fail mainly because pharmaceutical companies employ one-size-fits-all development strategies. Our own experience at the regulatory agency suggests that failure to utilize prior experience or knowledge from previous trials also accounts for trial failure. Prior knowledge refers to both drug-specific and nonspecific information such as placebo effect and the disease course. The information generated across drug development can be systematically compiled to guide future drug development. Quantitative disease–drug–trial models are mathematical representations of the time course of biomarker and clinical outcomes, placebo effects, a drug’s pharmacologic effects, and trial execution characteristics for both the desired and undesired responses. Applying disease–drug–trial model paradigms to design a future trial has been proposed to overcome current problems in drug development. Parkinson’s disease is a progressive neurodegenerative disorder characterized by bradykinesia, rigidity, tremor, and postural instability. A symptomatic effect of drug treatments as well as natural rate of disease progression determines the rate of disease deterioration. Currently, there is no approved drug which claims disease modification. Regulatory agency has been asked to comment on the trial design and statistical analysis methodology. In this work, we aim to show how disease–drug–trial model paradigm can help in drug development and how prior knowledge from previous studies can be incorporated into a current trial using Parkinson’s disease model as an example. We took full Bayesian methodology which can allow one to translate prior information into probability distribution.

KEY WORDS: Bayesian method; drug development; prior knowledge; quantitative disease–drug–trial model.

INTRODUCTION

A recent survey (1) reported that pharmaceutical research and drug development is not efficient, for example, almost 50% of registration trials fail mainly because pharmaceutical companies employ one-size-fits-all development strategies. A majority of the trials are reported to fail due to lack of differentiation from placebo (i.e., effectiveness). For example, only 14% of 39 depression trials were successful in terms of effectiveness (2). Other causes of failure include unanticipated safety problems and commercial reasons. Our own experience at the Food and Drug Administration suggests that failure could be reduced by utilizing prior knowledge (3,4). Especially accrual of disease and trial knowledge from previous experiences can further allow more efficient planning during the early phase of drug development (5).

Quantitative disease–drug–trial models (6,7) allow learning from prior experience and summarize the knowledge in a ready-to-apply format. Employing these models to plan future development is proposed as a powerful solution to improve pharmaceutical R&D productivity. The ultimate goals of the disease–drug–trial model paradigms are to apply the disease and trial models to future development and regulatory decisions, and share them with the public.

Parkinson’s disease belongs to a group of conditions called movement disorders, and is principally the result of the loss of dopamine-producing brain cells in the midbrain. Parkinson’s disease can be characterized by disease status over time using a longitudinal model. A symptomatic effect of drug treatments as well as a natural rate of disease progression (disease modification) determines the rate of disease deterioration. Figure 1 illustrates the time course of Parkinson’s disease progression which consists of a symptomatic effect (short-term benefit) and a disease-modifying effect (long-term benefit). Treatments for symptomatic benefit may not affect the rate of disease progression whereas disease-modifying treatments can slow disease progression. Pharmaceutical companies are attempting to develop drugs that can slow the progression of Parkinson’s disease. How-

¹Division of Pharmacometrics, Office of Clinical Pharmacology, Office of Translational Science, Center for Drug Evaluation and Research, Food and Drug Administration, 10903 New Hampshire Ave., Room 3198, Building 51, Silver Spring, Maryland 20993-0002, USA.

²To whom correspondence should be addressed. (e-mail: jooyeon2.lee@fda.hhs.gov)

Fig. 1. Illustration of symptomatic and disease-modifying effects in Parkinson’s disease. The disease progression in the placebo group is illustrated to worsen linearly over time consistent with literature. A symptomatic drug is assumed to offset the symptoms, depicted as an initial dip, but does not change the rate of disease progression. A disease-modifying drug decelerates the disease progression. Also illustrated is the time course of UPDRS for a drug that is both symptomatic and disease modifying

ever, there is no approved drug which claims disease modification to date. Regulatory agencies have been asked to comment on trial design and statistical analysis methodology. Bhattaram *et al.* (8) published a Parkinson’s disease model to explore competing trial endpoints and analyses that could distinguish disease-modifying effects from symptomatic effects. The main objective of their research was to explore a variety of statistical tests which might be able to discern disease-modifying effects.

In our research work, we focus on how prior information, or knowledge from previous clinical trials, can be incorporated into a current clinical trial using a Parkinson’s disease model. Bayesian methodology was applied to incorporate prior knowledge of the parameters of interest in the model. Specifically, the power prior distribution introduced by Ibrahim and Chen (9) was employed to translate prior knowledge into probability distribution.

In this paper, we start with a discussion of the conceptual framework of disease–drug–trial models. Next section presents a Bayesian exercise using a Parkinson’s disease model to show how prior knowledge can be translated to a probability scale, followed by a simple trial simulation which can be used as an example for designing an early phase trial

in Parkinson’s disease. Finally, we provide a detailed discussion of the findings.

QUANTITATIVE DISEASE–DRUG–TRIAL MODELS

Disease–drug–trial models are defined as mathematical representations of the time course of biomarker and clinical outcomes, placebo effects, a drug’s pharmacologic effects, and trial execution characteristics for both the desired and undesired responses (6). Table I summarizes the general concept of quantitative disease–drug–trial models.

Disease models quantify the relevant biological (physiologic) system in the absence of drug. The three major components of the disease model can be the relationship between biomarkers and clinical endpoints, natural disease progression, and placebo effect. In recent years, there have been considerable efforts in modeling natural disease progression, where the goal is to describe the change in the clinical outcome over time. For example, Holford and Peace (10) described the natural progression of Alzheimer’s disease as measured by the Alzheimer’s Disease Assessment Scale–Cognitive score using a linear model. The progression of Parkinson’s disease as reflected by total Unified Parkinson Disease Rating Scale (UPDRS) was modeled by Battaram *et al.* (1), Chan and Holford (11), and Holford *et al.* (12).

Drug models generally refer to dose or exposure response models (or pharmacokinetic–pharmacodynamic models). In drug development, better early dose-ranging studies are crucial in reducing the attrition rate in late-phase clinical trials. Thus, early phase trials that focus on bridging exposure response across patients, healthy subjects, animals, and *in vitro* results become very important. These models will also need to take into account the differences in the trial designs (e.g., dose–range and nonlinearity).

In addition to disease and drug models, other factors which determine the outcome of a clinical trial could be patient characteristics, inclusion/exclusion criteria, early discontinuation, and compliance to the protocol or to the prescribed regimen. Also, dropout rates are an important component in trial models. Reliable simulation of future trials is not feasible without incorporating such knowledge.

Actually, the disease–drug–trial models we described here are never complete, meaning that they are ever-evolving as new information accrues. Depending on the decision at hand, we may need only some of the aspects of the disease or drug models. In our work, we start with characterizing the time profile of Parkinson’s disease. Extending our work to

Table I. The Conceptual Depiction of Quantitative Drug–Disease–Trial Models

Disease model	Drug model	Trial model
Biology	Pharmacology	Patient population (demographics)
Biomarker–outcome relationship	Effectiveness	
Natural disease progression	Safety	
Placebo effect	Preclinical/healthy/patients	Dropout
	Product features	Compliance

include pharmacologic activity and/or preclinical information could be a subsequent step.

IMPLICATION OF DISEASE-TRIAL MODEL IN DEVELOPING DRUGS AGAINST PARKINSON'S DISEASE

We discussed the concept of quantitative disease–drug–trial models above; and in this section, we aim to show how to apply this concept to actual drug development. We adopted a Bayesian method for analyzing a Parkinson's disease model to show how to incorporate prior knowledge into current drug development using two clinical trials. The analyses were done by R2WinBUGS and SAS 9.2.

Data

Two randomized clinical trials were employed for the analyses: Earlier *versus* Later Levodopa Therapy in Parkinson Disease (ELLDOPA) study (13) is a multicenter, placebo-controlled, randomized dose-ranging, double-blind clinical trial where patients were randomized to placebo or carbidopa/levodopa at doses of 12.5/50, 25/100, and 50/200 mg three times daily; TVP-1012 in Early Monotherapy for Parkinson's disease Outpatients (TEMPO) study (14) is a double-blinded, randomized, fixed-dose parallel group study in early Parkinson's disease patients where patients were assigned to placebo, rasagiline 1 or 2 mg/day. Study durations for ELLDOPA and TEMPO studies were 24 and 26 weeks, respectively.

There was similarity between the two studies in terms of baseline demographics; mean ages were 61 and 64 years for ELLDOPA and TEMPO studies, most patients (more than 90%) were Caucasian, and more male patients were enrolled than female patients in both studies.

In the Bayesian analyses, TEMPO study was used for the current trial and ELLDOPA study was used as a historical study as levodopa has been approved earlier than rasagiline. The two dose levels of rasagiline (1 and 2 mg/day) in the TEMPO study appeared to be overlapping in terms of the disease progression time profile; so for the analyses, they were combined into one dose level.

Model

As shown in Fig. 1, characterization of disease progression can be challenging due to difficulty discerning disease-modifying effects from symptomatic effects. Most approaches published until now assume a linear relationship between time and disease status. To overcome this problem, Chan and Holford (11) modeled disease progression and symptomatic effects separately using data from multiple drugs. Guimaraes *et al.* (15) described nonlinearity of Parkinson's disease progression using a nonlinear mixed effect model. Bhattaram *et al.* (8) also applied a nonlinear mixed effect model but extended it to characterize treatment effect and placebo effect simultaneously.

For our analysis, change from the baseline in total UPDRS score (ΔUPDRS , hereafter) was modeled using a nonlinear random effects model as below;

$$\Delta\text{UPDRS}_{it} \sim N(\mu_{it}, \sigma^2), i = 1, \dots, N, t = 1, T$$

$$\mu_{it} = \text{slope}_i \times \text{time} - \text{symeff}_i \times (1 - \exp(-ke0 \times \text{time})) \quad (1)$$

where,

$$\begin{aligned} \text{slope}_i &= \beta_0 + \beta_1 \times \text{trt} + b_{1i} \\ \text{symeff}_i &= \gamma_0 + \gamma_1 \times \text{trt} + b_{2i} \\ \begin{pmatrix} b_{1i} \\ b_{2i} \end{pmatrix} &\sim \text{MVN} \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \sum_{2 \times 2} \right] \end{aligned}$$

ΔUPDRS_{it} is the change from the baseline in total UPDRS score at each visit (time), which was assumed to follow a normal distribution with mean μ_{it} and variance residual error (σ^2). $\text{Trt}=(0,1,\dots)$ refers to treatment group (where 0 means placebo). The parameters of slope_i and symeff_i represent individual (patient)-specific slope in natural disease progression and symptomatic effect, respectively. β_0 and β_1 are population-averaged effects for disease progression for patients given placebo and study drug, respectively. γ_0 and γ_1 are population-averaged symptomatic effects for short-term benefit for the patients given placebo and study drug, respectively. The parameter $ke0$ measures the speed with which the initial symptomatic effect is achieved, which was assumed to be the same for both placebo and drug arms. Lastly, b_{1i} and b_{2i} are random effects which are assumed to follow a multivariate normal distribution with mean (0,0) and covariance Σ . For simplicity, b_{1i} and b_{2i} were assumed to be independent, that is, $\Sigma = \begin{pmatrix} 0 & \omega_1^2 \\ \omega_2^2 & 0 \end{pmatrix}$.

In this model, the ability of the drug to slow disease progression can be shown by testing whether $\beta_1=0$ or not, and $\gamma_1=0$ implies no detectable short-term benefit provided by a study drug.

Prior Elicitation

The Bayesian paradigm has been widely used in recent years since new computational procedures have made Bayesian methods a viable approach (16–21). The Bayesian approach treats the parameter of interest as a probability rather than a fixed unknown value. Inference is based on the posterior distribution of the parameter, which is proportional to the product of the likelihood of observed data and the prior distribution. The prior distribution is the expected distribution before observing data. Knowledge of the parameter of interest is updated using observed data. The new probability distribution, the posterior distribution, is computed using Bayes' rule. Suppose that $Y=(y_1, y_2, \dots, y_n)$ is the observed data drawn from a normal distribution with mean μ and variance σ^2 . Assume that the variance, σ^2 is known so that the parameter of interest is the mean, μ , and we have prior knowledge on μ that it follows normal distribution with mean μ_0 and variance τ^2 . In this simple setting, it is not hard to show the posterior distribution of μ given data Y becomes a normal distribution. Let $L(\mu, \sigma^2; y_1, y_2, \dots, y_n)$ and $f(\mu)$ represent the likelihood function from observed data and prior distribution of μ , respectively. Then, the posterior distribution of μ can be computed as the product of likelihood function and the prior distribution

as follows,

$$\begin{aligned} f(\mu|Y, \sigma^2) &\propto L(\mu, \sigma^2|Y) \times f(\mu) \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\sum \frac{(y_i - \mu)^2}{2\sigma^2}\right) \times \frac{1}{\sqrt{2\pi\tau^2}} \exp\left(-\frac{(\mu - \mu_0)^2}{2\tau^2}\right) \\ &\propto \exp\left(-\frac{(\mu - \mu^*)^2}{2\delta^*}\right) + \text{constant} \end{aligned} \quad (2)$$

where,

$$\begin{aligned} \mu^* &= \delta^* \left(\frac{\mu_0}{\tau^2} + \frac{\bar{y}}{\sigma^2/n} \right), \\ \delta^* &= \left(\frac{1}{\tau^2} + \frac{n}{\sigma^2} \right)^{-1} \end{aligned}$$

As shown in the simple example, one advantage of the Bayesian approach is the ability to combine prior information with the observed data from a current trial through the use of informative priors. However, one cannot fully avoid subjectivity in determining the location or precision parameters, μ_0 and τ^2 in the example. Thus, the Bayesian method has been criticized due to the subjectivity in the selection of the prior distribution, and therefore prior elicitation becomes an important task in Bayesian analysis.

The impact of prior distribution on posterior inference can be easily illustrated using the example above (2); if there is little prior knowledge on the parameter of interest (μ), then we may put relatively flat prior distribution, i.e., $\tau^2 \gg \sigma^2$. In this case, the mean and variance in the posterior distribution (2) become $\mu^* \rightarrow \bar{y}$, $\delta^* \rightarrow \sigma^2/n$, implying that the posterior distribution of μ will be dominated by the observed data. On the other hand, it would not be hard to show that the posterior distribution of μ will be governed by prior distribution when there is little data or strong belief on the parameter is expressed by informative prior distribution. In this case, the caution should be given in constructing prior distribution and one may perform a sensitivity analysis by assigning different prior distributions to the parameter of interest and examining the stability of posterior distribution accordingly.

Translating prior information into probability distribution is not an easy task so prior elicitation becomes one of the most important tasks in Bayesian inference (22–26). One can assign a non-informative prior when there is no prior information on the parameters, which will provide equivalent parameter estimates to maximum likelihood estimates in a frequentist setting. Alternatively, one can use field expert opinions for certain parameters of interest. Chaloner (22), Chaloner and Rhame (23), and O'Hagan (25) described how to use expert opinion to elicit prior distributions.

Historical data from previous similar studies can be useful in eliciting prior distributions for a current study. While historical priors provide a useful tool in Bayesian analysis, investigators may want to control the impact of historical data on the analysis using data from current trial depending on the similarity between historical and current studies. Ibrahim and Chen (9) introduced the power prior and described robust

statistical properties in various kinds of regression models (27,28). In general, the power prior has a form as follows,

$$f(\theta|D_0, \alpha_0) \propto L(\theta|D_0)^{\alpha_0} f_0(\theta) \quad (3)$$

where

θ $\{\beta_0, \beta_1, \gamma_0, \gamma_1, \text{ke0}\}$, the regression parameters in the model
 D_0 (n_0, Y_0) denote historical data from previous studies
 $L(\theta|D_0)$ is the likelihood of observed data from historical data
 $f_0(\theta)$ is a prior distribution on θ before D_0 is observed, initial prior

Lastly, α_0 is a scalar prior parameter that weights the historical data relative to the likelihood of the current study, where $0 \leq \alpha_0 \leq 1$.

Basically, a power prior on parameter θ is just the historical posterior distribution which is raised to the power α_0 . When a flat prior is assigned to an initial prior, i.e., $f_0(\theta) \propto 1$, the posterior distribution of θ for the current data becomes

$$f(\theta|D, D_0, \alpha_0) \propto L(\theta|D) \times L(\theta|D_0)^{\alpha_0} \quad (4)$$

where, $D=(n, Y)$ is the data from the current study. In our example, $L(\theta|D)$ and $L(\theta|D_0)$ denote the likelihood from TEMPO and ELLDOPA studies.

Hence, the parameter α_0 controls the influence of the historical data on the current data and it can be interpreted as a relative precision parameter for the historical data. By assigning a point mass prior of 1 to α_0 , the power prior distribution becomes the posterior distribution of historical data, implying full borrowing from historical data into the current trial. On the other hand, putting a point mass prior of 0 to α_0 means no borrowing of information, which is identical to using a non-informative prior. For generalized mixed effect models, Ibrahim and Chen (28) recommended constructing the power prior on historical likelihood given random effects, rather than on the marginal likelihood which can be formed by integrating over random effects. In general, power prior specification can be completed by assigning another prior distribution such as a beta distribution, $\text{Be}(a, b)$, to weight parameter α_0 as we do not have full confidence in how much similarity between studies exists. Neelon and O'Malley (29) recommended use of a range of fixed values to α_0 based on expert opinion depending on relevance of historical data to the analyses using current data. Hence, we assigned a range of point mass priors to the parameter, α_0 , and examined the impact of point mass priors on posterior estimates for each parameter as a sensitivity analysis.

First, we begin with separate analyses of ELLDOPA and TEMPO studies by assigning non-informative priors to the parameters in the model. The same nonlinear mixed effect model (1) was applied to both studies. The dispersed two chains with 10,000 iterations were run with Monte Carlo Markov Chain (MCMC) after a 5,000 burn-in period and every tenth sample was retained (thinning=10) to improve convergence and also reduce autocorrelation between samples. Trace plots and Gelman and Rubin's diagnostics (16) were used as convergence checks. For a model diagnostics, posterior predictive distribution (30) is a commonly used tool in the Bayesian paradigm. If the model fits the data

reasonably well, we should expect the observed mean to be on the mode of posterior distribution. As a more formal approach, posterior predictive p value (17,31) was calculated, which is defined as the probability that the replicated data could be more extreme than the observed data as measured by the test quantity, $T(\cdot)$, that is,

$$P(T(y^{\text{rep}}, \theta) \geq T(y, \theta) | y) \quad (5)$$

The test quantity, $T(\cdot)$ is commonly chosen to measure a feature of the data not directly addressed by the model. For our analysis, we chose residual sum of squares (17,30).

To complete our analyses, we attempted to construct power prior distributions for the parameters $\{\beta_0, \beta_1, \gamma_0, \gamma_1, \text{ke0}\}$ in the model using ELLDOPA study, and non-informative priors were assigned to σ^2 and between-subject variability in two random effects (ω_1^2, ω_2^2).

RESULTS

Table II summarizes the posterior means with standard deviations using non-informative priors for all parameters in the model from independent analysis for two studies. A normal distribution with mean 0 and variance 100 was used as a prior for parameters $\theta = \{\beta_0, \beta_1, \gamma_0, \gamma_1, \text{ke0}\}$, and an inverse gamma distribution, $\text{IG}(1, 0.1)$ was assigned to variance in the two random effects (ω_1^2, ω_2^2) and σ^2 . There are lots of similarities shown in the parameter estimates; both placebo (β_0) and drug effects (β_1) on slope show the same direction; the speed to reach maximum symptomatic effect seems to be similar between the two studies also. However, the symptomatic effect of the drug appears to be more apparent in TEMPO study than in ELLDOPA study (1.61 vs. 0.36). Also notice that there appears to be large between-subject variability in symptomatic effect.

MCMC estimated all parameters without any convergence problems. The Gelman–Rubin diagnostic was estimated as 1 for all parameters and trace plots shows good mixing (Fig. 2).

Figure 3 shows the posterior predictive distribution of ΔUPDRS (change in UPDRS from baseline) score at week 26 for TEMPO study to see how well the model describes the data. Histograms represent posterior predictive distributions of placebo and treatment groups and dotted lines are observed mean of ΔUPDRS score at week 26. The posterior

predictive p -value was also calculated as described in the previous section. Ideally, it should be close to 0.5. If this probability is close to 0 or 1 such as less than 0.01 or more than 0.99, a model may be suspect. For our calculation, posterior predictive p -value was computed as 0.74 which shows no reason to question the adequacy of the model. Combining the graphical and p -value approaches, we find no compelling evidence for lack of fit to observed data.

Table III summarizes the posterior means with posterior standard deviation for the parameters in the model for TEMPO study using power priors by different point mass prior to weight parameter, α_0 . It should be noted that a value close to 0 means less borrowing information from ELLDOPA study ($\alpha_0=0$ is identical to the result using non-informative priors, which is presented in Table I) and weight of 1.0 indicates full borrowing of information from historical study (ELLDOPA study). To assess the impact of different point mass priors for weight parameters, we varied α_0 over the set $\{0.1, 0.5, 1.0\}$. It is evident that uncertainty around parameter estimates was reduced as more prior information was incorporated, which is also shown in the posterior distribution (Figs. 4 and 5). The posterior mean of β_1 , was estimated to be a negative value regardless of different point mass priors of α_0 , which magnitude gets larger with more weight to the historical data (ELLDOPA study). In terms of symptomatic effect the posterior mean of γ_1 implies that drug offers short-term benefit. However, it seems to diminish as more information is borrowed from the ELLDOPA study.

Figures 4 and 5 present how the posterior distribution of each parameter for TEMPO study can be affected by different weights in power prior. Black solid lines are the posterior distributions of each parameter in the model using power priors with $\alpha_0=0.1, 0.5$, and 1.0, and dotted and broken lines are the posterior distributions of the same parameters in ELLDOPA and TEMPO studies with non-informative prior. As expected, the posterior distribution of each parameter with power priors falls between those of ELLDOPA and TEMPO studies using non-informative priors. However, it can be also noticed that the posterior distributions of parameters for TEMPO study using power priors with bigger weight gravitates toward the posterior distribution of parameters for historical data. The distribution after incorporating prior knowledge from the historical study is narrower, implying that posterior estimates can be estimated with better precision, which is also shown in Table III.

Table II. The Posterior Mean with Posterior Standard Deviation (SD) of Parameters for ELLDOPA and TEMPO Studies

Parameter	Study	
	ELLDOPA	TEMPO
β_0 (placebo effect on slope, ΔUPDRS score/week)	0.99 (0.13)	0.73 (0.09)
β_1 (drug effect on slope, ΔUPDRS score/week)	−0.52 (0.07)	−0.30 (0.11)
γ_0 (placebo on symptomatic effect, ΔUPDRS /week)	1.93 (0.70)	1.12 (0.48)
γ_1 (drug on symptomatic effect, ΔUPDRS /week)	0.36 (0.37)	1.61 (0.60)
Ke_0 (speed to reach max symptomatic effect)	1.87 (0.19)	1.46 (0.61)
σ^2 (residual error)	8.68 (0.49)	8.53 (0.31)
ω_1^2 (between subject variability in slope)	1.13 (0.15)	0.47 (0.07)
ω_2^2 (between subject variability in symptomatic effect)	44.0 (4.65)	18.61 (2.31)

Non-informative priors were applied to all parameters in the model

Fig. 2. Trace plot for each parameter in the model from 5,000 samples after 5,000 burn-in period from two chains. The samples were collected with thinning=10

Also, we predicted Δ UPDRS score at week 26 for TEMPO study (data not shown) based on the posterior estimates of the parameters. Posterior means for the placebo group in the

TEMPO study were predicted to be slightly increased from 3.62 with non-informative priors to 3.64, 3.73, and 3.78 with weight parameter (α_0) of 0.1, 0.5, and 1.0 in the power priors, respectively. However, the change in Δ UPDRS score in rasagiline group by different weight parameter appears to be minimal (0.3, 0.0, -0.02, and -0.01 with $\alpha_0=0, 0.1, 0.5$, and 1.0.)

Fig. 3. The posterior predictive distribution of Δ UPDRS score at week 26 from TEMPO study with non-informative priors to the parameters. The dotted line indicates the observed mean of Δ UPDRS score at week 26

Trial Simulation

In this section, we describe how the model can be used to simulate future trials. We intend to illustrate how the concept of the trial model can be applied in practice and therefore, we did not consider the distribution of covariates for this simulation to be consistent with the model we used for our analyses. One hundred replicates of disease progression time profiles were simulated. The posterior means with variability around mean for placebo effect in the model were taken from the results using power priors with $\alpha_0=0.5$. The parameters ke0 and the two random effects were also derived from the posterior estimates using power priors. We assumed that a drug provides only symptomatic effect by varying the γ_1 parameter in the model. Four dose levels of 10, 20, 30, and 40 mg with placebo were considered for the simulation. Different study durations—12, 16,

Table III. The Posterior Mean with Posterior Standard Deviation of Parameters for TEMPO Study

Parameter	Weight parameter (α_0) in the power prior distribution		
	0.1	0.5	1.0
β_0 (placebo effect on slope, Δ UPDRS score/week)	0.77 (0.08)	0.82 (0.07)	0.85 (0.06)
β_1 (drug effect on slope, Δ UPDRS score/week)	-0.37 (0.10)	-0.46 (0.07)	-0.49 (0.06)
γ_0 (placebo on symptomatic effect, Δ UPDRS score/week)	1.36 (0.44)	1.63 (0.33)	1.72 (0.29)
γ_1 (drug on symptomatic effect, Δ UPDRS score/week)	1.22 (0.51)	0.76 (0.35)	0.59 (0.28)
Ke_0 (speed to reach max symptomatic effect)	1.49 (0.24)	1.62 (0.21)	1.69 (0.17)
σ^2 (residual error)	8.5 (0.30)	8.5 (0.31)	8.5 (0.30)
σ_{11}^2 (between-subject variability in slope)	0.48 (0.07)	0.49 (0.07)	0.50 (0.07)
σ_{22}^2 (between-subject variability in symptomatic effect)	18.3 (2.21)	17.7 (2.03)	17.3 (1.91)

The power prior distributions (constructed using data from ELLDOPA study as described in the text) were applied to the set of parameter $\{\beta_0, \beta_1, \gamma_0, \gamma_1, \text{ and } ke_0\}$, by different point mass priors for the weight parameter, α_0 . Non-informative priors, $IG(1, 0.1)$ were assigned to variance in the two random effects (ω_1^2, ω_2^2) and residual error (σ^2)

20, and 24 weeks were simulated. A linear mixed model was employed to analyze the simulated data. The null hypothesis was that the slope of dose-response is zero at an $\alpha=0.05$. A dropout model was also considered for the simulation. Following the definition by Little and Rubin (32) and also Diggle and

Kenward (33), dropout is defined as missing at random (MAR) if dropout at time t depends only on observed data up to time $t-1$, and it is missing completely at random (MCAR) if dropout does not depend on any observed data. Also, it is missing not at random (MNAR) if dropout at time t depends on unobserved

Fig. 4. Placebo effect: the comparison of posterior distributions of parameters, β_0 (slope, left) and γ_0 (symptomatic effect, right). The dotted (ELLDOPA study) and broken (TEMPO study) lines represent the posterior distribution of β_0 and γ_0 with no prior information (non-informative prior). The solid lines indicate the posterior distribution of β_0 and γ_0 for TEMPO study with borrowing prior information from ELLDOPA study (power prior distribution) by different weight (α_0)

Fig. 5. Drug effect (rasagiline): the comparison of posterior distributions of parameters, β_1 (slope, *left*) and γ_1 (symptomatic effect, *right*). The *dotted* (ELLDOPA study) and *broken* (TEMPO study) lines represent the posterior distribution of β_1 and γ_1 with no prior information (non-informative prior). The *solid* lines indicate the posterior distribution of β_1 and γ_1 for TEMPO study with borrowing prior information from ELLDOPA study (power prior distribution) by different weight (α_0)

data at time t . For our simulation, dropout is considered to follow a MAR mechanism and missing values were generated using the dropout model (34) as follows;

$$\text{Logit } P(R_{it} = 1) = \delta_0 - \delta_1 \cdot Y_{it-1} \quad (6)$$

where, $R_{it}=1$ means a patient i remains in the study at visit, and 0 otherwise. Y_{it-1} is the change from baseline in total UPDRS score at visit $_{t-1}$. Also, $P(R_{it}=1)$ indicates the probability of remaining in the study at visit $_t$. This model implies that the probability of dropout at visit t depends only on the previous ΔUPDRS score. More specifically, those who had higher ΔUPDRS score at the previous visit are less likely to remain in the study at the current visit. Without the term of $\delta_1 \cdot Y_{it-1}$, the model assumes MCAR whereas it is MNAR with the addition of a term such as $\delta_2 \cdot Y_{it}$. Dropout rate was considered approximately 20% at the end of study for a 20 or 24 weeks study. This rate was observed in both studies. The rate was assumed to decrease as the study duration shortened; 13% for a 16-week study and 10% for a 12-week study. To meet this assumed dropout rate, the parameters of δ_0 and δ_1 were manipulated depending on the study duration.

First, the relationship between power and sample size by different study durations was assessed. The maximum treatment effect (ΔUPDRS score) at the end of study was assumed to be 2.2 ($\gamma_1=0.3$). Figure 6 displays the result. Based on our simulation results, shorter studies required a larger sample size to achieve 80% power; a total of 500 patients (100 patients for each arm) are needed for a 20-week long study, whereas a total of 400 patients are necessary for a 24-week study. However, the simulation also showed that more than 600 patients are needed for a 12-week study to achieve 80% power. Across all sample sizes considered, power appears to be only slightly increasing with full data (no dropout) as the dropout rate considered for the simulation is not substantial.

The relationship between power and treatment effect was assessed by different sample sizes assuming a study duration of 24 weeks (Fig. 7). The treatment effect was defined as difference in ΔUPDRS score between placebo and 40 mg at week 24, which was determined by parameter γ_1 in the model. When the drug provides little symptomatic effect, even a sample size of 800 patients would not detect the

Fig. 6. The simulated power vs. different sample size by different study duration. The *left graph* is from full data (no dropout) and the *right graph* is when a certain portion of patients are assumed to be dropped out by the end of study (20% for 20 and 24 weeks, 13% for 16 weeks, and 10% for 12 weeks)

treatment effect whereas only 40 patients per arm will be needed to detect the drug effect of a 2.0 point difference. False-positive rate was approximately 3% with no dropouts and 2% with 15% of dropout rate.

DISCUSSION

One of the principal challenges in drug development is testing for proof-of-concept and selecting an appropriate trial design for late phase clinical trials. Screening molecules

against Parkinson's disease in the clinic is even more challenging due to the lack of useful biomarkers. An innovative solution would be to perform "smarter" trials. Using prior knowledge with respect to the disease and trials across multiple programs, which are independent of drug properties, can enhance the efficiency of early patient trials. Efficiency constitutes of optimizing the trial duration and size while preserving the informativeness.

We discussed a conceptual framework for quantifying disease, drug, and trial information that can guide future

Fig. 7. The simulated power vs. different additional symptomatic effect by drug by different sample size assuming the study duration is 24 weeks. The *left graph* is from full data (no dropout) and the *right graph* is from assumption of 20% dropout

development plans. This concept can be applied in practice using a Parkinson's disease model as the example.

Most pharmacometric analyses utilize maximum likelihood methods. Models and parameter estimates from each analysis are reported independently. Such an approach is not readily amenable for another researcher to build directly on the previous work. That is, the prior parameter estimates can at best be used as initial estimates in the model development using a new data set. Some newer approaches allow specifying prior distributions more formally (35). Conceptually, researchers approach drug development in a Bayesian fashion, albeit informally. For this reason, the Bayesian methodology was applied in the analyses. Bayesian methodology provides a number of advantages in drug development; when one wants to incorporate prior knowledge, it can be done in a flexible manner by applying different priors to the parameters; informative priors can improve the precision of estimates and increase the ability to detect treatment effects if any. Computational advances allow implementing Bayesian modeling.

The nonlinear mixed effect model was able to characterize Parkinson's disease progression reasonably. MCMC methods were able to estimate all parameters without convergence problems. The parameters in the model were estimated with better precision when informative priors were used. This advantage will be more prominent when sample size is relatively small or event rate is low—which is a property of early patient trials. When there is little observed information from a current trial as is often the case for rare diseases, accurately elicited prior distribution should be critical.

There are certain technical aspects of the modeling that need attention. The weight parameter in the power prior distribution we applied for our analyses is a subjective choice. To apply our method in future drug development, the use of specific weight should be justified based on clinical relevance. This is not much different from real-world decision making. The Bayesian analysis allows such assumptions to be transparent.

CONCLUSION

The goal of our work was to show how to utilize prior information in future drug development. Full Bayesian analyses were employed. Bayesian method allowed us to incorporate prior information from previous study into current trial by informative priors. Our model described the observed data reasonably well as assessed by the posterior predictive distribution and posterior predictive p -value.

The present work is focused on symptomatic effects and can be easily extended to disease-modifying effects. To the best of our knowledge, this is first attempt to build a Bayesian disease model, and we hope that future researchers will develop the model further to enhance its utility.

ACKNOWLEDGMENT

We would like to acknowledge Parkinson's Study Group and NIH Exploratory Trials in Parkinson's Disease (NET-PD) Group for providing access to clinical trial data. We also thank the Division of Pharmacometrics, Office of Clinical

Pharmacology for helpful discussions over the years. The views expressed in this article are those of the authors and do not necessarily reflect the official views of the Food and Drug Administration.

REFERENCES

1. Elias T, Gordian M, Singh N, Zimmel R. Why products fail in phase III. *In Vivo*. 2006;24:49–54.
2. Hooper M, Amsterdam JD. Do clinical trials reflect drug potential? A review of FDA evaluation of new antidepressants. Presented at NCDEU 39th Annual Meeting, Boca Raton, FL. 11–14 June, 1998.
3. Bhattaram VA, Booth BP, Ramchandani RP, Beasley BN, Wang Y, Tandon V, *et al*. Impact of pharmacometrics on drug approval and labeling decisions: a survey of 42 new drug applications. *AAPS J*. 2005;7:E503–12.
4. Bhattaram VA, Bonapace C, Chilukuri DM, Duan JZ, Garnett C, Gobburu JV, *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*. 2007;81:213–21.
5. Gobburu J. Disease model. *Clin Adv Hematol Oncol*. 2008;6:241–42.
6. Gobburu J, Lesko LJ. Quantitative disease, drug and trial models. *Annu Rev Pharmacol Toxicol*. 2009;49:291–301.
7. Zhang L, Sinha V, Forgue ST, Callies S, Ni L, Peck R, *et al*. Model-based drug development: the road to quantitative pharmacology. *J PKPD*. 2006;33(3):369–93.
8. Bhattaram VA, Siddiqui O, Kapcala L, Gobburu J. Endpoints and analyses to discern disease-modifying drug effects in early Parkinson's disease. *AAPS J*. 2009;11:456–64.
9. Ibrahim J, Chen M. Power prior distributions for regression models. *Stat Sci*. 2000;15:46–60.
10. Holford NH, Peace KE. Methodologic aspects of a population pharmacodynamic model for cognitive effects in Alzheimer patients treated with tacrine. *Proc Natl Acad Sci USA*. 1992;89:11466–70.
11. Chan PL, Holford NH. Drug treatment effects on disease progression. *Annu Rev Pharmacol Toxicol*. 2001;41:625–59.
12. Holford NH, Chan PL, Nutt JG, Kiebertz K, Shoulson I. Disease progression and pharmacodynamics in Parkinson disease—evidence for functional protection with levodopa and other treatments. *J Pharmacokinet Pharmacodyn*. 2006;33:281–311.
13. Parkinson Study Group. Levodopa and the progression of Parkinson's disease. *NEJM*. 2004;351:2498–508.
14. Parkinson Study Group. A controlled trial of rasagiline in early Parkinson disease: the TEMPO study. *Arch Neurol*. 2002;59(12):1937–43.
15. Guimaraes P, Kiebertz K, Goetz CG, Elm J, Palesch Y, Huang P, *et al*. Nonlinearity of Parkinson's disease progression: implication for sample size calculation in clinical trials. *Clin Trials*. 2005;2:509–18.
16. Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences (with discussion). *Stat Sci*. 1992;7:457–511.
17. Gelman A, Carlin J, Stern H, Rubin D. *Bayesian data analysis*. 1st ed. London: Chapman & Hall/CRC; 1995.
18. Wakefield JC. The Bayesian analysis of population pharmacokinetic models. *J Am Stat Ass*. 1996;91:61–76.
19. Wakefield JC. Bayesian individualisation via sampling-based methods. *J Pharmacokinet Biopharm*. 1996;24:103–31.
20. Spiegelhalter DJ, Freedman L, Parmar MKB. Bayesian approaches to randomized clinical trials (with discussion). *J Roy Stat Soc Ser A*. 1994;157:357–415.
21. Gilks WR, Richardson S, Spiegelhalter DJ. *Markov Chain Monte Carlo in practice*. London: Chapman and Hall; 1996.
22. Chaloner K. Elicitation of prior distributions. In: Berry DA, Stangl DK, editors. *Bayesian biostatistics*. New York: Marcel Dekker; 1996. p. 141–56.
23. Chaloner K, Rhome FS. Quantifying and documenting prior beliefs in clinical trials. *Stat Med*. 2001;20:581–600.

24. Kadange JB, Wolfson LJ. Experiences in elicitation. *The Statistician*. 1998;47:3–19.
25. O'Hagan A. Eliciting expert beliefs in substantial practical applications. *The Statistician*. 1998;47:21–35.
26. Press SJ. *Subjective and objective Bayesian statistics*. 2nd ed. New Jersey: Wiley; 2003.
27. Chen M, Ibrahim J. The relationship between the power prior and hierarchical models. *Bayesian Anal*. 2006;1:551–74.
28. Chen H, Ibrahim J, Shao Q, Weiss RE. Prior elicitation for model selection and estimation in generalized linear mixed models. *J Stat Plan Inf*. 2003;111:57–76.
29. Neelon B, O'Malley AJ. Bayesian analysis using power priors with application to pediatric quality of care. *J Biomet Biostat*. 2010;1:1–9.
30. Gelman A, Meng A, Stern H. Posterior predictive assessment of model fitness via realized discrepancies. *Stat Sin*. 1996;6: 733–807.
31. Meng X. Posterior predictive p values. *Ann Stat*. 1994;22:1142–60.
32. Little RJA, Rubin DB. *Statistical analysis with missing data*. 2nd ed. New Jersey: Wiley; 2002.
33. Diggle P, Kenward MG. Informative drop-out in longitudinal data analysis. *Applied Statistics*. 1994;43:49–73.
34. Little RJA. Modeling the drop-out mechanism in repeated-measures studies. *J Am Stat Assoc*. 1995;90:1112–21.
35. NONMEM: nonlinear-mixed effects modelling, Icon Development Solutions, Ellicott City, MD, USA.