J Pharmacokinet Pharmacodyn
DOI 10.1007/s10928-013-9305-5

ORIGINAL PAPER

Development of a placebo effect model combined with a dropout model for bipolar disorder

Wan Sun · Thomas P. Laughren · Hao Zhu · Guenther Hochhaus · Yaning Wang

Received: 11 January 2013 / Accepted: 12 February 2013

Abstract The aim of this study was to develop a placebo model for bipolar disorder to help optimize clinical trial designs for studies targeting manic episodes in bipolar disorder. A bipolar disease database was built based on individual longitudinal data collected from over 3,000 patients in 11 clinical trials for 5 approved bipolar drugs. An empirical placebo effect model with an exponential decay process plus a linear progression process was developed to quantify the time course of the Young Mania Rating Scale total score based on only placebo data from the database. In order to describe the dropout pattern during the trials, a parametric survival model was developed and the Weibull distribution was identified to be the best distribution to describe the data. Based on the likelihood ratio test, it was found that patients with higher baseline score, slower disease improvement and more rapid disease progression tended to dropout earlier, and the trial features such as trial starting year and trial site were also significant covariates for dropout. A combination of the placebo effect model and the dropout model was applied to simulate new clinical trials through Monte-Carlo simulation. Both the placebo effect model and dropout model described the observed data reasonably well based on various diagnostic plots. The joint placebo response and dropout models can serve as a tool to simulate the most likely level of placebo response with the expected dropout pattern to help design a new clinical trial.

Keywords Bipolar disorder · Placebo model · Dropout · Clinical trial simulation

W. Sun · Y. Wang (✉)
Division of Pharmacometrics, Office of Clinical Pharmacology, US Food and Drug Administration, 10903 New Hampshire Avenue, Silver Spring, MD 20993, USA
e-mail: yaning.wang@fda.hhs.gov

T. P. Laughren
Division of Psychiatry Products, US Food and Drug Administration, Silver Spring, MD, USA

H. Zhu
Division of Clinical Pharmacology 1, Office of Clinical Pharmacology, US Food and Drug Administration, Silver Spring, MD, USA

G. Hochhaus
Department of Pharmaceutics, University of Florida, 1600 SW Archer Road, Gainesville, FL 32610, USA

Introduction

Bipolar disorder is a treatable and recurrent psychiatric illness with different symptomatic episodes (mania, depression and mixed states) [1, 2]. Current drug treatments include atypical antipsychotics, anticonvulsants and mood stabilizers. The Young-Mania Rating Scale (YMRS) is the most commonly used rating scale to assess the severity of manic symptoms [3]. It consists of 11 items, each scored from 0 to either 4 or 8. The total score can range from 0 to 60, with a higher score implying more severe symptoms.

As is true for most neuropsychiatric disorders, clinical trials for bipolar disorder are generally considered to be challenging. One of the challenges is that placebo response varies widely, both within and across clinical trials. This variability in placebo response may be one explanation for the relatively modest drug-placebo differences observed in clinical trials targeting manic symptoms in bipolar patients, which significantly interferes signal detection for new medications and even results in failure of trials [4]. Thus it may be useful to develop a placebo model to quantify the

Published online: 02 March 2013

time course of placebo effect across different trials and to identify factors that can explain the heterogeneity in placebo response. Such a model could serve as a valuable drug development tool to help design future clinical trials by providing a quantitative assessment of the expected placebo response in a specific patient population. In the area of antipsychotic placebo effect modeling, various models, e.g., linear model, Weibull model, and inverse Bateman function, have been used to describe the placebo effect in schizophrenia [5, 6], depression [7], Alzheimer's disease [8] and Parkinson's disease [9]. A placebo effect model for bipolar disorder has not yet been described in the literature.

High dropout rate is another common problem in bipolar mania trials. Missing data due to dropout can introduce potential biases in the comparison of treatment groups, reduce the overall statistical power, and inflate type I error, depending on how the missing data are handled and the reasons for dropout. Last observation carried forward (LOCF) is one of the missing data imputation methods applied widely in practice. With this method, the last observed value is retained and carried forward to subsequent time points until the last time point with missing data under the assumption of a constant profile after dropout. Another similar imputation method is called baseline observation carried forward. These simple imputation methods are not recommended as the primary approach to the treatment of missing data unless the assumptions that underlie them are scientifically justified [10]. Dropout patterns in clinical trials can be quantified with parametric dropout models [11–13]. Certain patient characteristics and trial features can influence dropout in a systematic way. These factors can be incorporated into the dropout model to explain the between-patient or between-trial variability in dropout patterns. Such a model can be used together with the placebo effect model to simulate the possible placebo response with realistic dropout patterns to assist in the design of a new clinical trial. This approach has been implemented in other disease areas [12, 14] and is an essential component of model-based drug development [15, 16].

Methods

Disease database

A disease database was developed including patient-level data from 11 clinical trials targeting manic symptoms in bipolar patients for 5 drugs approved for the acute treatment of manic symptoms in patients with bipolar disorder. Trial selection was driven by availability of electronic datasets containing longitudinal YMRS score measurements. All the data were organized into a pooled database in a uniform structure and format. All of these clinical trials were randomized, double-blind, placebo-controlled, parallel group studies with the same primary efficacy endpoint (change from baseline in YMRS total score at day 21) and similar inclusion criteria. Some key differences in trial designs included trial duration (3 vs. 12 weeks), number of treatment groups (2 vs. 3), trial locations (with US centers vs. without US centers), and the time when the trials were conducted (before 2002 or after 2002). SAS 9.2 was used to organize the data and build the database.

Placebo effect model

The longitudinal YMRS score data were analyzed using nonlinear mixed effect models. The models reported for other neuropsychiatric disorders were first explored as candidate models to describe the time course of YMRS scores. Although the time course for the mean YMRS total score in the placebo arm of clinical trials for bipolar disorder was generally similar to other neuropsychiatric disorders, none of reported models was able to capture the heterogeneous individual YMRS profiles. A model using a combination of an exponential decay process and a linear progression described the individual YMRS profiles reasonably well. The structure of the model is shown in Eq. 1:

$$\mathrm{E_i}(t) = S_{0i} \times e^{-K_i * t} + \mathrm{Slope_i} \times \mathrm{t} + \varepsilon_i(t) \quad (1)$$

where $\mathrm{E_i}$ is the observed score for *i*th subject at time t (day), $\mathrm{S_{0i}}$ is the observed baseline score for *i*th subject, $\mathrm{K_i}$ is the rate constant of exponential disease improvement for *i*th subject, $\mathrm{Slope_i}$ is the rate of linear disease progression for *i*th subject, and ε_i is the difference between the observed and expected scores and is assumed to follow normal distribution with a mean of 0 and variance of σ^2. Both Ki and $\mathrm{Slope_i}$ were assumed to follow log-normal distribution (Eq. 2).

$$P_i = \exp\left(LP_{pop} + \eta_i\right) \quad (2)$$

where $\mathrm{P_i}$ is the individual parameter, $\mathrm{LP_{pop}}$ is the log-transformed population geometric mean, and η_i is the difference between the individual and population parameters on a log scale and is assumed to follow a normal distribution with a mean of zero and variance of ω^2 (between-subject variability). Various patient characteristics such as age, gender, race, and baseline score were explored as potential covariates to explain the between-subject variability for the two parameters.

In order to constrain the simulated score within the YMRS range of 0–60, a logit function (Eq. 3) was used to transform both the data and the model (Eq. 1):

$$Logit(E) = \log\frac{E/60}{1 - E/60} \quad (3)$$

At the boundary of 0 or 60, 0.1 or 59.9 was used to avoid invalid logit transformation. A transform-both-sides approach was applied. NONMEM 7.2 was used to build the placebo effect model with the first-order conditional estimation (FOCE) method, and SAS 9.2 was applied to create the diagnostic plots.

Placebo dropout model

A parametric survival model (Eq. 4) was utilized to describe the dropout pattern during the trials.

$$T = \exp(x'\beta) * T_0 \quad (4)$$

where T is the dropout time, T_0 is the dropout time sampled from the baseline distribution corresponding to values of zero for the covariates, x' is the vector of covariate values and β is a vector of unknown parameters.

Various patient characteristics and trial features were explored as potential covariates to explain the between-patient or between-trial variability in dropout patterns. Patient characteristics included the observed baseline score, the estimated individual parameters K_i and $Slope_i$ from the placebo effect model, age, gender and race. Trial features were incorporated into the model as categorical variables, including trial starting year (0 for before 2002; 1 for after 2002), location of the trial (0 for trials with US sites; 1 for trials without US sites) and number of treatment arms (0 for 2 arms; 1 for 3 arms). Several commonly used distributions for time-to-event data, including exponential, log normal, Weibull, and gamma, were tested for T_0. Likelihood ratio tests and diagnostic plots were applied to select an appropriate distribution.

Evaluation of the dropout model was conducted through internal validation. One thousand sets of parameters were simulated based on the parameter estimates and their variance–covariance matrices from the final model. Then, 1,000 replicates of time-to-dropout data were simulated for each trial on the basis of the simulated parameters, the baseline YMRS score, and relevant covariates in the trials. Median survival curves and 95 % confidence intervals (CIs) were constructed across the 1,000 replicates. The simulated survival curves were compared to the observed survival curves (median and 95 % CI). SAS 9.2 was used to build the placebo dropout model (PROC LIFEREG) and evaluate the final model.

Model evaluation for joint placebo effect and dropout models

The placebo effect model and the dropout model were combined to simulate the longitudinal YMRS score in clinical trials through Monte-Carlo simulation. Each patient's parameters for placebo effect model were simulated by a bootstrap sampling from the individual parameter set in each trial. A full time course of YMRS scores was simulated for each patient based on the scheduled visit times in each trial. Dropout time was also simulated for each patient based on the final dropout model with the relevant covariates. The dropout time for each patient was used to truncate the simulated YMRS score. For example, if the scheduled visit times were baseline, day 7, 14 and 21, and the simulated dropout time was day 15 for a specific patient, the simulated YMRS score on day 21 for that patient was removed. This procedure was repeated 1,000 times for each trial. Median and 90 % CIs were summarized for the 5th, 50th and 95th percentiles at each visit time based on the simulated longitudinal YMRS score across the 1,000 replicates for each trial. These statistics were compared with the observed ones in each trial.

Results

Disease database

Patient-level longitudinal data from over 3,000 patients enrolled in 11 clinical trials for 5 drugs approved for the acute treatment of manic symptoms in bipolar disorder were included in the database (Table 1). Most of these trials were 3 weeks in duration, however, 2 trials for drug D had a duration of 12 weeks. For these latter trials, the additional 9 weeks were designed for assessment of the maintenance effect. All the trials with 2 treatment arms were conducted in the US, except for study B3; all of the 3-arm trials were carried out at international sites. Even though trial A1, B3, C1 and C2 included multiple international sites, over 50 % of the sites were in the US. All trials were conducted after 2000; studies C1, C2 and E1 were the most recently conducted trials. The demographic information and baseline YMRS total scores are summarized in Table 2 and compared across trials conducted at non-US sites, mainly US sites, and only US sites. The largest demographic differences across trial sites were in race. Higher baseline scores in non-US trials were due to the different enrollment criteria in these two trials.

Placebo effect model

The observed YMRS score ranged from 0 to 58. Less than 2 % of the data were at the boundary of 0. The choice of 0.1 to replace 0 for a valid logit transformation was supported by the insensitivity of the model parameters (<20 % change for any parameter) to this arbitrary small number (0.1, 0.001 or 0.0001). Even though more complex methods were proposed to handle bounded continuous data

[17, 18], the impact of the boundary issue on the intended utility of the model is minimal in our case and the relatively simple method of shifting the boundaries by a small margin is considered practically sufficient. No significant covariates were identified for the individual parameters, K_i and $Slope_i$. The final estimates of parameters as well as the precision measures of the estimates for the placebo effect model are listed in Table 3. Large between-subject variability was observed for K and slope, which should be expected given the highly variable YMRS score profiles across different patients. All parameters were estimated with good precision. The mean value of log-transformed baseline score, $\log(S_0)$ was 3.37 with a standard deviation of 0.202.

Figure 1 shows the comparison of individual observations, individual predictions and population predictions of time-course of YMRS total score for some representative patients. Even though the population profile is a simple monotone curve, the individual profiles could vary dramatically among patients. The final model was flexible enough to describe the various types of individual profiles. In Fig. 2, scatter plots of the individual observed and predicted YMRS total scores for each study is shown with the reference to the identity line to illustrate that the same structural model could fit all the studies reasonably well.

Table 1 Summary of drug and trials in the disease database for bipolar disorder

Drug	Trial	Duration (week)	No. of treatment Arms	Year of trial	Country
A	A1	3	3	2000–2001	Mainly-US
	A2	3	2	2000–2001	Only US
	A3	3	2	2002–2003	Only US
B	B1	3	2	2000–2001	Only US
	B2	3	2	2000–2001	Only US
	B3	3	2	2002–2003	Mainly-US
C	C1	3	3	2004–2006	Mainly-US
	C2	3	3	2004–2006	Mainly-US
D	D1	12	3	2001–2002	Non-US
	D2	12	3	2001–2002	Non-US
E	E1	3	2	2006–2007	Only US

Placebo dropout model

Weibull distribution was selected for the dropout model on the basis of likelihood ratio tests (Table 4). Among the various patient characteristics and trial features, five covariates were identified as significant predictors for time-to-dropout (Table 5). They are baseline score, individual parameter K_i and $slope_i$, trial location and starting time.

The estimated coefficients suggested that patients with higher baseline scores and steeper slopes (progression) for placebo effect tended to drop out earlier in the trial while patients with larger K (disease improvement) tended to drop out later, or not all. The trials without US sites and the more recent trials tend to have a longer time to dropout or a lower dropout rate at week 3.

The graphic comparison of survival curves with simulated data and those with observed data (Fig. 3) demonstrated that the dropout model could reasonably describe the observed dropout pattern under the placebo treatment in most trials except for trial A3. Trial A3 showed a relatively large discrepancy between the predicted data and the observed data. Efforts were taken to explore possible reasons for this discrepancy. No reasonable factors were identified to explain this observation. It could be due to between-trial variability. However, the number of trials is too limited to explicitly model between-trial variability.

Model evaluation for joint placebo effect and dropout models

The combination of the placebo effect model and the dropout model is necessary to simulate a new clinical trial. Figure 4 shows the graphic comparison of longitudinal YMRS score summaries at 5th, 50th, and 95th percentiles

Table 2 Demographic and baseline characteristics

Characteristic	Non-US ($n = 2$)	Mainly-US ($n = 4$)	Only-US ($n = 5$)	ALL ($n = 11$)
No. of patients (%)	599 (18.3)	1,584 (48.4)	1,087 (33.2)	3,270 (100)
Gender, Female (%)	52.8	44.3	49.4	47.5
% Race				
White (%)	62.9	59.8	66.2	62.5
Asian (%)	34.4	13.1	1.4	12.9
Black (%)	<1	16.0	26.7	16.6
Age, Median (range)	41 (18–79)	39 (18–76)	40 (18–76)	40 (18–79)
YMRS baseline				
Mean (SD)	33.2 (6.5)	28.6 (5.8)	28.4 (5.6)	29.4 (6.2)

between the simulated data and the observed data. The shaded area for each percentile represents the 90 % confidence interval based on the simulated data. Overall, the combination of the two models successfully reproduced the longitudinal YMRS score under the influence of dropout for each trial.

Table 3 Parameter estimates of placebo effect model for pooled dataset

Parameter	Estimate (SE)	ω^2 (SE)	Shrinkage %
Log(K)[a]	−3.25 (0.00373)	1.25 (0.0385)	23.5
Log(Slope)[a]	−1.49 (0.00149)	1.27 (0.0004)	15.5
δ	3.86 (0.00113)		

[a] K and slope were estimated to be 0.0388 and 0.225

Discussion

The aim of this study was to leverage existing clinical trial data under placebo treatment that can be quantified via parametric models to help optimize clinical trial designs for future bipolar mania studies. As part of this effort, a standard disease database was constructed, with common variables across multiple bipolar clinical trials to facilitate the model building process and future model updates with more accumulated data.

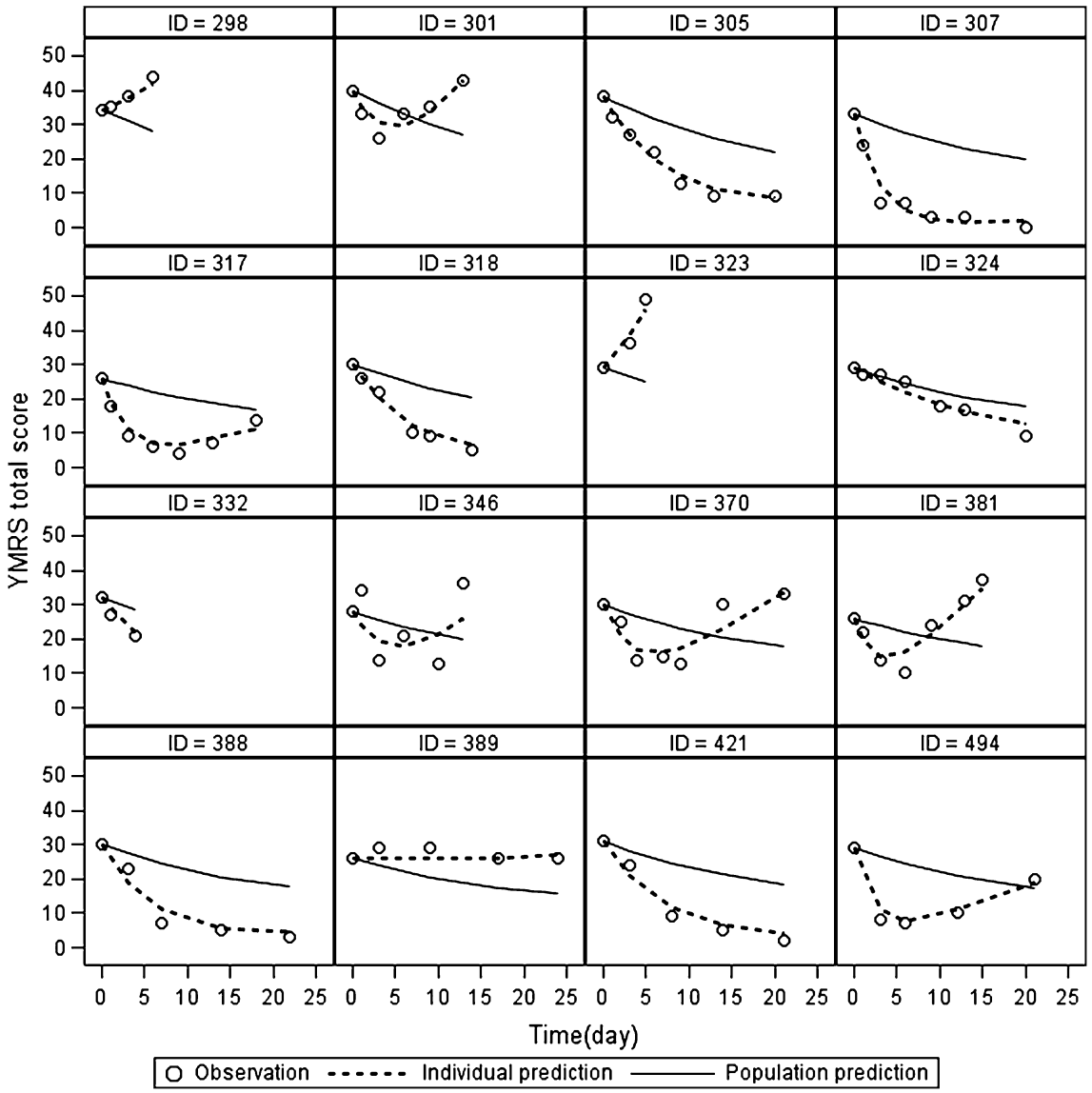

Fig. 1 Goodness of fit plots. Comparison of individual observation (*open circle*), individual prediction (*dotted line*) and population prediction (*solid line*) on time profile of YMRS total score for some representative patients

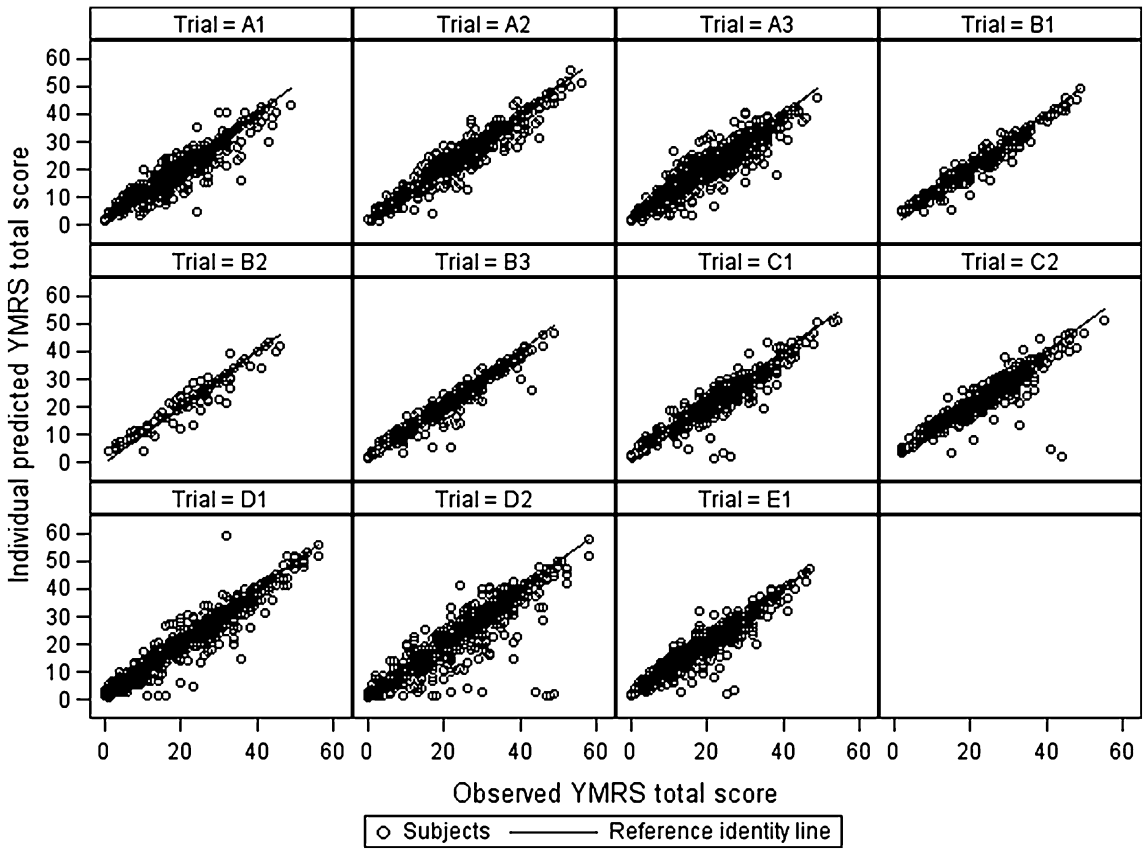

Fig. 2 Goodness of fit plots. Scatter plot of the individual observed and predicted YMRS total scores with the reference identity line for each study

Table 4 AIC and −2 log likelihood (−2LL) values under various distributions (Exponential is nested within Weibull (df = 1), Weibull is nested within Gamma (df = 1), and log-normal is nested within Gamma (df = 1))

Distribution	AIC	−2LL
Exponential	2,360.5	2,346.5
Weibull	2,240.1	2,224.1[a]
Log-normal	2,305.1	2,289.1
Gamma	2,238.7	2,220.7[a]

[a] Reduction in −2LL is >3.8, corresponding to $P \leq 0.05$ for Chi square test with degree of freedom (df) of 1. The significance for Weibull is relative to Exponential. Relative to Log-normal, Gamma is significant. Between Gamma and Weibull, the reduction in −2LL is 3.4. Therefore, Weibull was selected

Table 5 Parameter estimates of placebo dropout model for pooled dataset

Parameter	Coefficient estimate	Std. error	Chi square	*P* Value
Intercept	6.24	0.54	131.71	<0.0001
Log(baseline)	−0.72	0.16	20.97	<0.0001
Log(k)	0.35	0.04	91.08	<0.0001
Log(slope)	−0.31	0.04	67.92	<0.0001
Country	0.86	0.11	63.06	<0.0001
Starting year	0.33	0.06	30.02	<0.0001

To quantify the longitudinal placebo effect, an empirical model was selected to describe the time course of the primary efficacy endpoint, YMRS total score. Empirical models are relatively simple and descriptive to fit the rating scale profile in neuropsychiatric disorders within the studied population and duration [4], while mechanism-based models or semi-mechanistic models may possess more predictive power to extrapolate into new scenarios that were not studied in the completed trials. In recent years, placebo responses have been analyzed with sophisticated biological tools at both biochemical and cellular levels as a biological phenomenon worthy of scientific investigation, and the increased understanding of the placebo-response mechanism provided potential opportunities to build mechanistic models for placebo response [19]. However, the underlying mechanisms for placebo response in bipolar clinical trials are still unclear and complicated due to both psychological effects produced by medical interventions and non-psychological factors, including trial design, protocol conduct, and demographic differences.

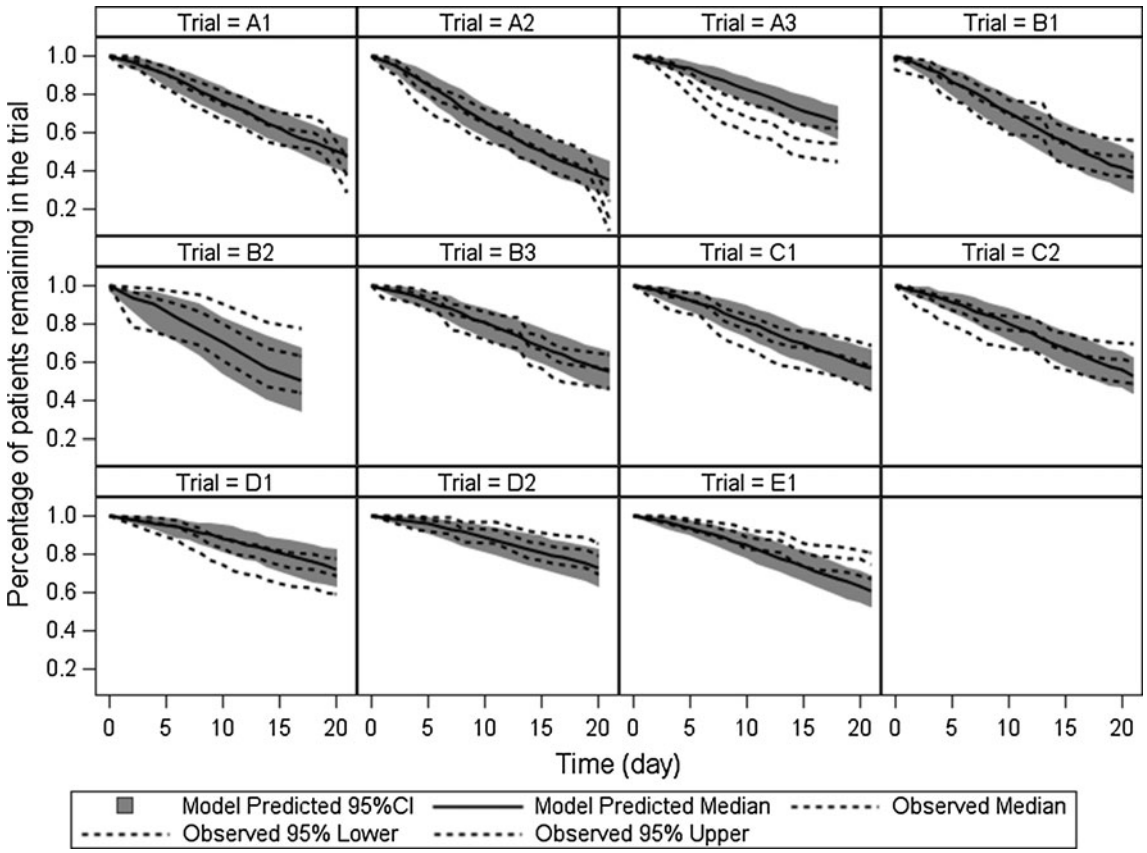

Fig. 3 Comparison of survival curves for simulated data and observed data. The *dotted lines* represent the median survival curves and its 90 % confidence intervals for the observed data; the *solid black lines* and *shaded area* represent the median survival curves and its 90 % confidence intervals for the simulated data

Empirical models were selected to quantify and describe the longitudinal placebo response in bipolar clinical trials.

As an initial model development strategy, several empirical model structures from the literature [4] were tested to explore whether any of them could describe the time course of YMRS score under placebo treatment in bipolar clinical trials. The rationale was that the overall shape of mean YMRS total score time profile was similar to those of clinical scores used in other neuropsychiatric disorders. However, it was found that none of these models (Emax, Exponential, Weibull, inverse Bateman) could describe all possible individual profiles, such as monotone increase, monotone decrease, and initial decrease followed by increase. The final model that could provide flexible fittings to all these profiles was a combination of an exponential decay process and a linear progression. This model structure was also utilized as an empirical model to describe longitudinal HAMD-17 score for depression trials [20] and longitudinal tumor data for non-small cell lung cancer trials [21]. There are various approaches available for handling baseline responses in pharmacokinetics–pharmacodynamic analysis [22]. In our analysis, the observed instead of estimated baseline score was used as one of the predictors for the longitudinal placebo response. The baseline score was available for every patient in all the trials and was a standard patient recruitment criterion. Using observed baseline score is one of the methods discussed by Dansirikul et al. [22] and the authors concluded that this method was not significantly different from the other methods that estimated the baseline score. Moreover, using the observed baseline score can further simplify the model with fewer parameters to be estimated.

Despite the efforts to look for potential covariates to explain the between-subject variability in individual parameters related to placebo response, no significant covariate was identified from the commonly available demographic factors such as gender, age, and race. However, it might be possible that some other factors related to the disease and the treatment history may affect the change in YMRS total score over time. Lipkovich and colleagues [23] identified individual YMRS items, type of episode at baseline, treatment history, number of previous manic episodes, lack of disruptive-aggressive behavior and more prominent depressive symptoms at baseline as potential risk factors for relapse in patients who had responded to treatment. However, it was challenging to explore the

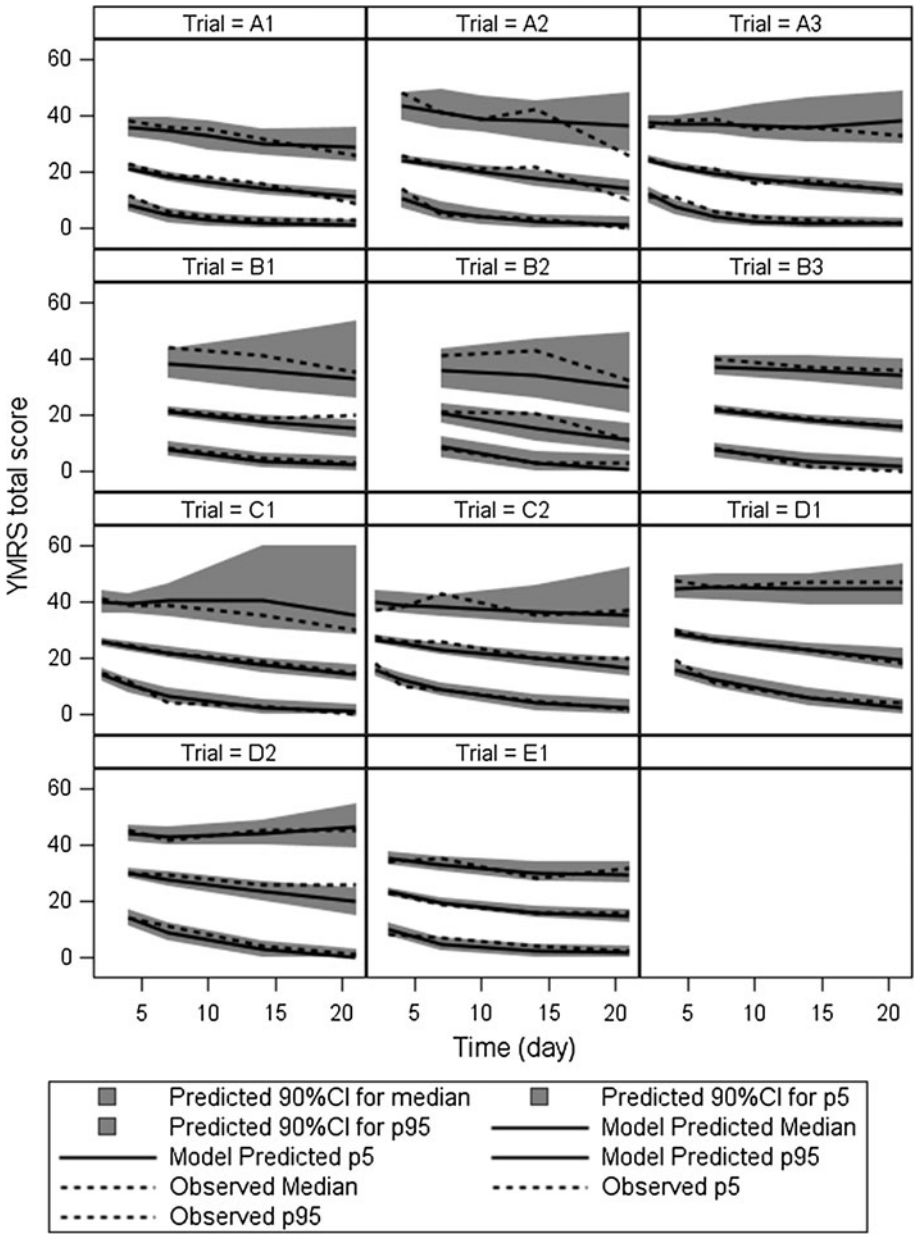

Fig. 4 Comparison of YMRS total score time profiles between simulation and observation for all the studies. The *dotted lines* represent the 5th, 50th, and 95th percentiles of YMRS time profile of the observed data; the *solid black lines* represent the 5th, 50th, and 95th percentiles of YMRS time profile of the simulated data; *shaded areas* represent the 90 % confidence intervals for the percentiles

relationship between those potential factors and the placebo effect for the pooled bipolar database, since different trials collected the information differently or some trials did not include the relevant information. Even though the two individual parameters, K_i and $Slope_i$, were quantified with statistical models without any covariates, these two parameters could be treated as a baseline patient characteristic and considered as "inherent" features of the patients. Researchers recently identified a potential gene to predict placebo response in irritable bowel syndrome [24]. Therefore, they could be included as potential predictors, together with other patient factors or trial features, to explain the between-subject and between-trial variability in dropout patterns.

The final dropout model included baseline YMRS total score, $\log(K_i)$, $\log(\text{Slope})$, year when the trials were conducted and trial site as predictors for time to dropout. Steeper slope (progression) was associated with earlier dropout while larger K (disease improvement) was associated with late dropout. This finding is consistent with the documented most common reason for dropout in the placebo group: lack of efficacy. Higher baseline score was found to be associated with earlier dropout. William-Faltaos et al. [14] identified a similar association for Alzheimer's disease. It is widely observed that patients with higher baseline scores tend to experience greater improvement on active treatment, which may support an enrichment design including only those patients with more severe symptoms at baseline. The

increased dropout for these patients under placebo treatment should be considered in such a trial to avoid missing data due to a high dropout rate.

Several trial features were explored as potential factors to explain the different dropout patterns in various trials. Trial starting year was included with the expectation that a more recent trial may have fewer dropouts due to improved medical care and facilities over time. The result did support this expectation. The number of treatment arms is related to the probability of patients being assigned to the active treatment groups in a placebo controlled trial, and has been reported to affect the placebo response for depression [25, 26], migraine [27], and schizophrenia [28]. It was shown that the higher the likelihood of receiving active treatment, the higher the response to placebo. However, the opposite was reported for irritable bowel syndrome [29]. A similar analysis was conducted to explore whether the number of treatment arms affected the placebo response and dropout rate for bipolar trials. Given the narrow range of this variable (2 or 3 arms), the number of treatment arms was not identified as a significant factor for either the placebo response or the dropout pattern.

Trial location has been reported as an important contributing factor for the heterogeneous results between or within trials [30, 31]. A general observation has been that non-US trials tend to have lower dropout rates. Our results also suggest that non-US sites were associated with a lower dropout rate. However, this finding needs to be explained with caution. There may be other confounding factors, one of which is trial duration. In this database, trial location is highly confounded by trial duration, making it difficult to conclude which factor was the true driver for the differential dropout rates. The lower dropout rate at week 3 observed in the non-US sites could be due to the different culture or clinical practice in those countries. It could also result from other incentives that may have been offered to the patients to induce them to stay in the trials longer given the planned longer duration to assess the maintenance effect.

Simulation based on the combination of the placebo effect model and the dropout model indicated good predictive capability of the joint models. Even though bootstrap re-sampling could lead to an under-representation of the between-subject variability due to the shrinkage of the individual parameters, this was not evident based on the comparison between the simulated and observed results. Since no between-trial variability was included in the model due to the limited number of trials, the estimated between-subject variability was inflated to a certain degree by the between-trial variability. Therefore, the shrinkage effect was compensated by the over-estimated between-subject variability. For the sake of simulating a new clinical trial, such a model should represent a more conservative scenario to avoid an over-optimistic view of the between-subject variability and a potentially underpowered study. In addition, bootstrap re-sampling was applied to maintain the trial specific distribution of the individual parameters and covariates in order to compare the simulated and the observed results for each trial. Simulating from the estimated parameter distributions would have generated almost identical simulated results for each trial because it is impossible to match the trial specific distribution of the individual parameters. Such a tool can be used to generate expected placebo response for the general bipolar population with the expected dropout pattern. This expected placebo response profile can serve as the reference for a new compound that may become an effective treatment. When the new compound is studied in a phase 2 placebo-controlled trial, the observed placebo response can be compared with the expected placebo response simulated from the tool. If consistent placebo responses are observed, the effect size for the new compound estimated from the phase 2 trial is reliable and can be used to design the phase 3 trial with confidence. Otherwise, the unusual placebo response observed in the phase 2 trial should be explained with caution and, more importantly, the effect size estimated for the new compound may not be reliable and could lead to an underpowered phase 3 trial. The same tool can also be applied to simulate the expected placebo response for phase 3 trials. In addition, an exposure–response model and a dropout model for the new compound should be developed based on the compound specific data. The drug specific models can be combined with the placebo models to simulate a complete phase 3 trial.

Conclusion

This study compiled a bipolar disease database with individual data from over 3,000 patients. This database can serve as a standard data template and a convenient data source for meta-analysis to address many challenging questions that require across trial comparisons. When data from future trials with more arms become available, these new data can be easily integrated into the database with the standard format. Ultimately, it may be possible to reach a more definitive conclusion on the impact of the likelihood of receiving the active treatments on the placebo response and dropout probability for bipolar trials. The joint placebo response and dropout models properly characterized the observed longitudinal placebo response by taking dropout into account. The developed models can be applied to simulate the most likely level of placebo response to help design a new clinical trial.

Disclaimer The views expressed in this paper are those of the authors and do not necessarily represent those of the FDA.

References

1. Swann AC (2006) What is bipolar disorder? Am J Psychiatry 163(2):177–179
2. Angst J (2006) Do many patients with depression suffer from bipolar disorder? Canadian journal of psychiatry Revue canadienne de psychiatrie 51(1):3–5
3. Young RC, Biggs JT, Ziegler VE, Meyer DA (1978) A rating scale for mania: reliability, validity and sensitivity. Br J Psychiatry 133: 429–435
4. Pilla Reddy V, Kozielska M, Johnson M, Vermeulen A, de Greef R, Liu J, Groothuis GM, Danhof M, Proost JH (2011) Structural models describing placebo treatment effects in schizophrenia and other neuropsychiatric disorders. Clin Pharmacokinet 50(7): 429–450. doi:10.2165/11590590-000000000-00000
5. Kimko HC, Reele SS, Holford NH, Peck CC (2000) 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(5):568–577
6. Friberg LE, de Greef R, Kerbusch T, Karlsson MO (2009) Modeling and simulation of the time course of asenapine exposure response and dropout patterns in acute schizophrenia. Clin Pharmacol Ther 86(1):84–91
7. Shang EY, Gibbs MA, Landen JW, Krams M, Russell T, Denman NG, Mould DR (2009) Evaluation of structural models to describe the effect of placebo upon the time course of major depressive disorder. J Pharmacokinet Pharmacodyn 36(1):63–80
8. Holford NH, Peace KE (1992) Results and validation of a population pharmacodynamic model for cognitive effects in Alzheimer patients treated with tacrine. Proc Natl Acad Sci USA 89(23):11471–11475
9. Bhattaram VA, Siddiqui O, Kapcala LP, Gobburu JV (2009) Endpoints and analyses to discern disease-modifying drug effects in early Parkinson's disease. AAPS J 11(3):456–464
10. Panel on Handing Missing Data in Clinical Trials, National Research Council (2010) The prevention and treatment of missing data in clinical trials. The National Academies Press, Washington, DC
11. Gomeni R, Lavergne A, Merlo-Pich E (2009) Modelling placebo response in depression trials using a longitudinal model with informative dropout. Eur J pharm Sci 36(1):4–10. doi:10.1016/j.ejps.2008.10.025
12. Dutta S, Diderichsen P, Liu W (2011) Clinical trial simulation of outcome after accounting for high dropout typical of acute pain studies. Paper presented at the American College of Clinical Pharmacology 40th Annual Meeting, Chicago
13. Hu C, Sale ME (2003) A joint model for nonlinear longitudinal data with informative dropout. J Pharmacokinet Pharmacodyn 30(1):83–103
14. William-Faltaos D, Chen Y, Wang Y, Gobburu J, Zhu H (2012) Quantification of disease progression and dropout for Alzheimer's disease. Int J Clin Pharmacol Ther. doi:10.5414/CP201787
15. U.S. Food and Drug Administration (2009) The critical path initiative: transforming the way FDA-regulated products are developed, evaluated, manufactured, and used
16. Gobburu JV, Lesko LJ (2009) Quantitative disease, drug, and trial models. Annu Rev Pharmacol Toxicol 49:291–301. doi:10.1146/annurev.pharmtox.011008.145613
17. Hutmacher MM, French JL, Krishnaswami S, Menon S (2011) Estimating transformations for repeated measures modeling of continuous bounded outcome data. Stat Med. doi:10.1002/sim.4155
18. Hu C, Yeilding N, Davis HM, Zhou H (2011) Bounded outcome score modeling: application to treating psoriasis with ustekinumab. J Pharmacokinet Pharmacodyn 38(4):497–517. doi:10.1007/s10928-011-9205-5
19. Benedetti F (2008) Mechanisms of placebo and placebo-related effects across diseases and treatments. Annu Rev Pharmacol Toxicol 48:33–60. doi:10.1146/annurev.pharmtox.48.113006.094711
20. Gomeni R, Merlo-Pich E (2007) Bayesian modelling and ROC analysis to predict placebo responders using clinical score measured in the initial weeks of treatment in depression trials. Br J Clin Pharmacol 63(5):595–613. doi:10.1111/j.1365-2125.2006.02815.x
21. Wang Y, Sung C, Dartois C, Ramchandani R, Booth BP, Rock E, Gobburu J (2009) Elucidation of relationship between tumor size and survival in non-small-cell lung cancer patients can aid early decision making in clinical drug development. Clin Pharmacol Ther 86(2):167–174
22. Dansirikul C, Silber HE, Karlsson MO (2008) Approaches to handling pharmacodynamic baseline responses. J Pharmacokinet Pharmacodyn 35(3):269–283. doi:10.1007/s10928-008-9088-2
23. Lipkovich IA, Houston JP, Ahl J (2008) Identifying patterns in treatment response profiles in acute bipolar mania: a cluster analysis approach. BMC Psychiatry 8:65
24. Hall KT, Lembo AJ, Kirsch I, Ziogas DC, Douaiher J, Jensen KB, Conboy LA, Kelley JM, Kokkotou E, Kaptchuk TJ (2012) Catechol-*O*-methyltransferase val158met polymorphism predicts placebo effect in irritable bowel syndrome. PLoS One 7(10):e48135. doi:10.1371/journal.pone.0048135
25. Papakostas GI, Fava M (2009) Does the probability of receiving placebo influence clinical trial outcome? A meta-regression of double-blind, randomized clinical trials in MDD. Eur Neuropsychopharmacol 19(1):34–40. doi:10.1016/j.euroneuro.2008.08.009
26. Sinyor M, Levitt AJ, Cheung AH, Schaffer A, Kiss A, Dowlati Y, Lanctot KL (2010) Does inclusion of a placebo arm influence response to active antidepressant treatment in randomized controlled trials? Results from pooled and meta-analyses. J Clin psychiatry 71(3):270–279. doi:10.4088/JCP.08r04516blu
27. Diener HC, Dowson AJ, Ferrari M, Nappi G, Tfelt-Hansen P (1999) Unbalanced randomization influences placebo response: scientific versus ethical issues around the use of placebo in migraine trials. Cephalalgia 19(8):699–700
28. Mallinckrodt CH, Zhang L, Prucka WR, Millen BA (2010) Signal detection and placebo response in schizophrenia: parallels with depression. Psychopharmacol Bull 43(1):53–72
29. Enck P, Klosterhalfen S, Weimer K, Horing B, Zipfel S (2011) The placebo response in clinical trials: more questions than answers. Philos Trans R Soc Lond B Biol Sci 366(1572): 1889–1895. doi:10.1098/rstb.2010.0384
30. Khin NA, Chen YF, Yang Y, Yang P, Laughren TP (2011) Exploratory analyses of efficacy data from major depressive disorder trials submitted to the US Food and Drug Administration in support of new drug applications. J Clin Psychiatry 72(4):464–472. doi:10.4088/JCP.10m06191
31. Khin NA, Chen YF, Yang Y, Yang P, Laughren TP (2012) Exploratory analyses of efficacy data from schizophrenia trials in support of new drug applications submitted to the US Food and Drug Administration. J Clin Psychiatry 73(6):856–864. doi:10.4088/JCP.11r07539