Pharm Res (2014) 31:2829–2843
DOI 10.1007/s11095-014-1380-9

RESEARCH PAPER

Handling Missing Data in a Duloxetine Population Pharmacokinetic/Pharmacodynamic Model – Imputation Methods and Selection Models

Eunice Yuen • Ivelina Gueorguieva • Leon Aarons

Received: 21 November 2013 / Accepted: 7 April 2014 / Published online: 3 May 2014

ABSTRACT

Purpose In pharmacokinetic (PK)/pharmacodynamic (PD) modelling and simulations (M&S), omitting dropouts can cause inaccuracies in parameter estimation and clinical trial simulations (CTS). This study examines the impact of different imputation methods for missing data on the interpretation of model results, as well as develops a selection model (where dropout and efficacy are jointly modelled) for use in CTS.

Methods Missing data were imputed using single and multiple imputation and pattern mixtures methods for a previously reported duloxetine PK/PD model. The probability of dropout was described in the selection model and CTS was conducted with a hypothetical drug to examine the impact of dropout on trial results.

Results The study completion rate was 75% and dropouts were not random. Model parameters obtained with different imputation methods were mostly within 40% (range 0 to 63%) compared to the model without dropouts. CTS showed 0.3 points lower median pain scores and 3% lower coefficient of variation over the 12-week simulations when dropout was included.

Conclusions Missing data had little impact on the original population PK/PD analyses. Sensitivity analyses for dropouts should be conducted in M&S exercises. The utility of selection models in CTS was explored via a hypothetical case study.

KEY WORDS Bayesian methods · dropout · imputation methods · population pharmacokinetics/pharmacodynamics · selection model

E. Yuen (✉) • I. Gueorguieva
Global PK/PD and Pharmacometrics, Eli Lilly and Co.
Erl Wood Manor, Windlesham Surrey, GU20 6PH, UK
e-mail: yuenes@lilly.com

L. Aarons
Manchester Pharmacy School, University of Manchester Manchester, UK

ABBREVIATIONS

ACMV	Available case missing value
b.i.d.	Twice daily
bgr	Gelman-Rubin convergence
BOCF	Baseline observation carried forward
CCMV	Complete case missing value
CRD	Completely random dropout
C_{ss}	Concentration at steady state
DPNP	Diabetic peripheral neuropathic pain
EC_{50}	Concentration required to achieve 50% of maximal effect
EMA	European Medicines Agency
E_{max}	Maximum effect
FDA	Food and Drug Administration
FOCE	First order conditional estimation
ID	Informative dropout
LOCF	Last observation carried forward
MC	Monte Carlo
NCMV	Neighbouring case missing value
NONMEM	Non-linear mixed effects modelling
NRS	Numerical rating scale
PD	Pharmacodynamics
PK	Pharmacokinetics
q.d.	Once daily
RD	Random dropout
SD	Standard deviation
VPC	Visual predictive check

INTRODUCTION

Patient dropouts are common and inevitable in the course of any clinical trial. These can be intermittent dropouts where a patient has missed visits, or terminal dropouts where a patient discontinues the trial altogether. The reasons for dropout may vary, ranging from staff recording error to unacceptable

adverse events or perceived lack of efficacy. In any circumstance, dropouts lead to missing data, and in self-report measurements, such as pain scores, this problem can be more pronounced since the patient is relied upon to provide data to inform the endpoint calculations. Analysis of clinical trial data utilises the intention-to-treat principle, where all collected information from randomized patients are included. However in this process missing data can either be ignored or often imputed using simple methods such as last observation carried forward (LOCF). These single imputation methods are not ideal as they may introduce bias when dropout numbers are substantial and/or when dropouts are not random (1).

Imputation methods for missing data are usually data-driven, since they frequently make use of patterns or observations in the dataset to make inferences or estimate missing data. Common techniques include single or multiple imputation and pattern mixture methods. Single imputation methods include a variety of techniques such as last observation carried forward (LOCF), baseline observation carried forward (BOCF) or mean value substitution. Multiple imputation methods can generally be divided into explicit techniques, which use regression models or predictive mean matching, or implicit techniques, which use data-based methods such as approximate Bayesian bootstrap (2). Pattern mixture models, on the other hand, treat the whole population as a mixture of two or more patterns of dropouts and non-dropouts and estimates the overall population parameters averaged across patterns (3).

Whilst imputation methods for handling missing data are useful in sensitivity analyses for evaluating model outputs, they do not easily allow simulations from the model to predict dropout rates and drug efficacy or disease progression for future trials. For this purpose, selection models are well suited, since they incorporate both the drug efficacy and dropout using a joint distribution. Weibull or Gompertz models are commonly used parametric approaches for describing baseline hazard and survival for dropouts (4). The other common method to describe dropouts uses logistic models, where the probability of dropout at a certain time point is modelled directly using a logit function.

Despite the variety of available methods to handle missing data, there is currently no clear consensus or standard in specifying which methods should be used in the various situations of missing data (5). A report on the subject was commissioned by the Food and Drug Administration (FDA) and published by the National Academy of Sciences recently (6,7). In 2001 the European Medicines Agency (EMA) published a guidance on points to consider on missing data (8) but have since revised this to a guidance on missing data in confirmatory clinical trials, with substantial additions to the methods section including sensitivity analyses (9). The FDA guidance on patient reported outcomes suggests two or more sensitivity analyses with different methods for missing data imputation (10).

Population PK/PD models are often developed using only non-missing data, with the occasional exception of imputation or substitution for concentrations that are below the limit of quantification. Since dropouts lead to early censoring of data, the robustness of the analysis and the conclusions drawn may be affected by missing data, particularly since in clinical trials, these are typically non-random. Moreover, when models that do not include dropout are used for clinical trial simulations, inaccuracies for protocol design such as misspecification of sample size may occur (11). Recently, concerns were raised by the FDA about the amount of missing data in Rivaroxaban trials for acute coronary syndrome (12), with the drug subsequently failing to gain approval. Gomeni *et al.* (13) have also shown that non-inclusion of dropout may affect characterisation of placebo response. Therefore, if a specific model for dropouts is not included in the population PK/PD model, it is important to incorporate sensitivity analyses into the modelling process, in order to test the effect of missing data on the final estimated PK/PD parameters. To overcome these limitations, a selection model is ideal since it incorporates both dropout and drug efficacy, and therefore allows further use of the model for simulations in varying circumstances. We have previously reported a duloxetine PK/PD efficacy model with non-missing data in the treatment of diabetic peripheral neuropathic pain (DPNP) (14). This model (henceforth referred to as the original PK/PD model) described the pain scores following placebo, 20 or 60 mg once daily (*q.d.*) and 60 mg twice-daily (*b.i.d.*) doses of duloxetine. Pain scores were recorded over 12 weeks on a 11-point numerical rating scale (NRS) ranging from 0 to 10. This paper extends the original analyses, firstly, to evaluate the impact of missing data on the estimated PK/PD parameters of the model through the use of varying imputation methods. Secondly, a selection model is developed to investigate the randomness of dropout and also to illustrate its utility for handling further clinical trial simulations.

METHODS

Description of Duloxetine PK/PD Model and Associated Studies

The original base population PK/PD model was developed using data from a total of 1,106 patients contributing 12,549 pain scores (out of a theoretical maximum of 13,272 observations), from four different treatment groups in three different

studies. Patients received placebo (n=327), 20 mg duloxetine *q.d.* (n=110), 60 mg duloxetine *q.d.* (n=335) or 60 mg duloxetine *b.i.d.* (n=334) over 12 weeks. Daily 24 h average pain scores were collected on the NRS scale, ranging from 0 (no pain) to 10 (worst possible pain), and these scores were averaged over a week for use as a primary efficacy endpoint. All patients with baseline NRS pain scores lower than three, or randomized patients who failed to provide any efficacy measurements post-treatment were excluded from the analyses. The original base population PK/PD model can be found in Equation 1 and 2. Further details of the model are described elsewhere (14), as with the details around the conduct of each trial (15–17).

$$Painscore_{ij} = \mathrm{Base}_i \times \left(1 + PE_i\left(1-\mathrm{e}^{-\left(\mathrm{C}*\mathrm{time}_j\right)}\right) + \mathrm{f(d)} \times \left(1-\mathrm{e}^{-\left(\mathrm{K_i}*\mathrm{time}_j\right)}\right) + \varepsilon_{ij} \quad (1)$$

$$and\ f(d) = \frac{E_{max,i} \times C_{ss,i}}{EC_{50,\mathrm{i}} + C_{ss,i}} \quad (2)$$

where $Painscore_{ij}$ describes the jth pain score of the ith individual, $Base_i$=pain score at baseline of the ith individual, PE_i=magnitude of placebo effect of the ith individual, C=first order rate constant describing the onset of placebo effect, $E_{max,i}$=maximum effect of the ith individual, $EC_{50,i}$=concentration required to achieve 50% of maximal effect of the ith individual, K_i=first order rate constant describing the onset of drug effect of the ith individual, ε_{ij}=residual error with a mean of 0 and variance of σ^2 and $C_{ss,i}$=duloxetine concentration at steady state for the ith individual. This was predicted based on individual demographic parameters using the population PK model described by Lobo *et al.* (18), where significant covariates affecting the PK model were found to be smoking and sex on bioavailability, age and dose on apparent clearance and ethnicity on apparent volume of distribution.

Analysis of Dropouts Using Imputation Methods

In the evaluation of dropout using imputation methods, missing data for each patient included in the dataset used to develop the original PK/PD model were imputed using three methods to obtain new complete datasets – simple imputation, multiple imputation and pattern mixtures. Both intermittent and terminal dropouts were included in the analyses involving simple and multiple imputations, whilst the analysis involving pattern mixtures was carried out using only terminal dropouts. Analyses and simulations were carried out using NONMEM version 6 (19) using the first order conditional estimation (FOCE) method.

In the simple imputation method, the LOCF technique was chosen, where the last observation for the individual patient was carried forward for each missing pain score. The PK/PD model described in Equation 1 and 2 were then fitted to the new complete dataset. For the multiple imputation method, five different pain scores were simulated for each missing data point using the PK/PD model described in Equation 1 and 2. This resulted in the creation of five new complete datasets, each containing a mixture of observed and simulated data (14,378 observations from 1,106 patients in each dataset). The PK/PD model in Equation 1 and 2 were then fitted to each dataset producing five unique set of PK/PD parameter estimates. These were subsequently averaged to obtain one set of pooled model parameter estimates. The variance of each parameter estimate (V_β) was computed according to Equation 3 and 4 (2):

$$V_\beta = \overline{U_\beta} + \left(1 + \frac{1}{M}\right) * B_\beta \quad (3)$$

$$and\ \ \overline{U_\beta} = \frac{1}{M}\sum\nolimits_{m=1}^{M} U_\beta{}^m, B_\beta = \frac{1}{M-1}\sum\nolimits_{m=1}^{M}\left(\widehat{\beta}^m - \bar{\beta}\right)^2, \quad (4)$$

$$\bar{\beta} = \frac{1}{M}\sum\nolimits_{m=1}^{M} \widehat{\beta}^{\,m},\ U_\beta{}^m = Var\left(\widehat{\beta}^m\right)$$

where $\overline{U_\beta}$ = pooled within imputation variance, M=no. of imputations (5 in this case), B_β = variance of estimates (between imputation), $\widehat{\beta}^m$ = parameter estimates (within imputation), $\overline{\beta}$= pooled estimates and U_β^m = variance of estimates (within imputation).

For the pattern mixtures models method, three cases were investigated – complete case missing value (CCMV), available case missing value (ACMV) and neighbouring case missing value (NCMV). In order to limit the number of possible patterns, all intermittent dropouts were excluded from these analyses, leaving a total of 12 possible patterns depending on when the patient dropped out from the trial (Table I, top half). Since some patterns contained too few patients to inform the PK/PD models (for example, there were no patients in pattern 2 in the placebo and 20 mg *q.d.* groups), several patterns were combined and collapsed into groups (3). For these reasons and the purposes of further analyses in all pattern mixture cases, patterns A to E (Table I, bottom half) are used rather than the individual patterns 1 to 12. For the CCMV case, the PK/PD model described in Equation 1 and 2 were fitted to the data from pattern A. The resulting model parameters were then used to simulate five different pain scores for each missing data point in patterns B to E, resulting in five new complete datasets. For the ACMV case, the PK/PD model described in Equation 1 and 2 were fitted in turn to data from

Table I Possible and combined patterns for pattern mixture methods (intermittent dropouts excluded)

Pattern	Week where pain score is available													No. of patients with this pattern				
	0	1	2	3	4	5	6	7	8	9	10	11	12	Placebo	20 mg *q.d.*	60 mg *q.d.*	60 mg *b.i.d.*	Total
1	X	X	X	X	X	X	X	X	X	X	X	X	X	244	87	257	237	825
2	X	X	X	X	X	X	X	X	X	X	X	X		0	0	3	4	7
3	X	X	X	X	X	X	X	X	X	X	X			5	2	2	6	15
4	X	X	X	X	X	X	X	X	X	X				1	0	1	0	2
5	X	X	X	X	X	X	X	X	X					3	5	7	2	17
6	X	X	X	X	X	X	X	X						2	0	1	0	3
7	X	X	X	X	X	X	X							9	2	3	5	19
8	X	X	X	X	X	X								1	2	1	1	5
9	X	X	X	X	X									11	0	7	13	31
10	X	X	X	X										5	3	4	4	16
11	X	X	X											9	1	12	11	33
12	X	X												10	4	19	33	66
Combined/Grouped Patterns																		
A	X	X	X	X	X	X	X	X	X	X	X	X	X	244	87	257	237	825
B	X	X	X	X	X	X	X	X	X	X	X	**X**		5	2	5	10	22
C	X	X	X	X	X	X	X	X	X	**X**				4	5	8	2	19
D	X	X	X	X	X	X	**X**	**X**						12	4	5	6	27
E	X	X	**X**	**X**	**X**									35	8	42	61	146

Bold underlined crosses denote observations that were not available in all subjects at that week since subjects were combined into patterns where appropriate

pattern A, patterns A+B, patterns A+B+C and patterns A+B+C+D. This resulted in four different sets of PK/PD parameter estimates (arbitrarily named models 1 to 4 in the above order). Parameters from model 1 were used to simulate five values for each missing data point in pattern B, whilst parameters from model 2 were used to simulate five values for each missing data point in pattern C and so on. At the end of the process, five new complete datasets were again obtained. For the NCMV case, the PK/PD model described in Equation 1 and 2 were now fitted in turn to data from each of the patterns A, B, C and D. Imputation for missing data was generally carried out using models fitted to data from neighbouring patterns, resulting in five new complete datasets. The details of the patterns used for model fitting and simulations in each of the above cases are described in Table II. All simulations were carried out using 5 iterations with the same seed number. For each of the five new complete datasets from each pattern mixture subtype (each comprising 13,507 observations from 1,039 patients), the PK/PD model described in Equation 1 and 2 were fitted, with resulting parameter estimates averaged and variances calculated as previously described for the multiple imputation methods in Equation 3 and 4.

Overall comparison between the various imputation methods was carried out by means of graphical measures. The percentage change in PD parameter estimates obtained from the new models (fitted to datasets with imputed data) *versus* the original PK/PD model (fitted to non-missing data) were plotted. To provide an additional comparison, the original PK/PD model was also fitted to a dataset comprised only of patients who completed the study (827 patients with 10,751 observations).

Table II Groupings for pattern mixture methods

Method	Model derived from non-missing data in… (number of patients in pattern)	Used to simulate missing data in…
CCMV	Pattern A (825)	Patterns B+C+D+E
ACMV	Pattern A (825)	Pattern B
	Patterns A+B (847)	Pattern C
	Patterns A+B+C (866)	Pattern D
	Patterns A+B+C+D (893)	Pattern E
NCMV	Pattern A (825)	Weeks 11–12 in Pattern B
		Week 12 in Patterns C+D+E
	Pattern B (22)	Weeks 9–11 in Patterns C+D+E
	Pattern C (19)	Weeks 6–8 in Patterns D+E
	Pattern D (27)	Weeks 2–5 in Pattern E

Analysis of Dropouts Using Selection Models

In order to identify possible predictors of dropout, the observed survival was first plotted as Kaplan-Meier graphs. Pain scores of dropouts *versus* non-dropouts were also plotted by time to identify the possible dependence of dropout on pain scores (either previous or predicted at time of dropout). The dataset used in the selection models included only terminal dropouts, consisting of 1,039 patients with 13,507 observations.

Logistic regression models were used to model the probability of dropout using NONMEM version 6 (19). This was conducted independently of the pain scores model. Model building started with a baseline dropout probability:

$$logit(Pr) = \theta_1 \tag{5}$$

$$and\ Pr = \frac{\exp(logit(Pr))}{1 + \exp(logit(Pr))} \tag{6}$$

where Pr=probability of dropout and θ_1=intercept. Potential predictors, including age, time, dose, duloxetine concentration at steady state (C_{ss}), change from baseline of last observed pain score prior to dropout (Y_{PBSL}) and change from baseline of model-predicted score at time of dropout (Y_{UBSL}), were then added one at a time. If the objective function decreased by more than 6.63 points ($p \leq 0.01$, $df=1$) with the addition of each parameter, then the covariate was retained. In the backward elimination step, each covariate was removed separately from the model, if its removal caused an increase in objective function of at least 10.83 points ($p \leq 0.001$, $df=1$), then it was retained for the final model. Visual predictive checks (VPCs) for dropout were carried out using 100 iterations for each subject to assess model fits.

In order to determine the influence of dropout on the drug efficacy model and to support the approach of separate development of each model, a joint model was evaluated in WinBUGS Version 1.4.3 (20). The PK/PD model describing the pain scores (14) was transferred into WinBUGS with non-informative priors (with the exception of EC_{50} which required an informative prior using mean parameter estimates from the original model), and had the following structure:

$$Pain\ score_{ij} \sim normal\left(\mu_{ij}, \sigma^2\right) \tag{7}$$

$$\text{and}\ \mu_{ij} = Base_i \times \left[1 + PE_i\left(1 - e^{-(C \times \text{time}_j)}\right) + \left(\frac{E_{max,i} * C_{ss_i}}{EC_{50,i} + C_{ss_i}} * \left(1 - e^{(-K_i \times \text{time}_j)}\right)\right)\right] \tag{8}$$

where μ_{ij}=jth pain score for the ith individual and the other parameters defined as before in Equation 1. Maximum likelihood expressions for Bayesian methods can be further found in Lunn *et al.* (20) and Best *et al.* (21).

The dropout model was first developed separately, before being combined with the pain scores model, with the following structure:

$$logit(Pr) = \theta_1 + \theta_2 * \exp(-\theta_3 * time) + \theta_4 * \left(Painscore_{ij} - Base\right) \tag{9}$$

$$\text{and}\ Pr = \frac{\exp(\text{logit}(Pr))}{1 + \exp(logit(Pr))} \tag{10}$$

where Pr=probability of dropout, θ_1=intercept, θ_2 and θ_3=parameters associated with time and θ_4=parameter associated with the change from baseline model-predicted score at time of dropout.

Two sets of priors for θ_1 to θ_4 were investigated to test the sensitivity of the model to priors. The first set consisted of informative priors (using parameter estimates and standard errors from the original model): $\theta_1 \sim$ normal (−4.38, 0.063); $\theta_2 \sim$ normal (3.77, 0.797); $\theta_3 \sim$ normal (0.38, 0.02) and $\theta_4 \sim$ normal (0.0787, 0.002). The second set comprised of non-informative priors where θ_1 to $\theta_4 \sim$ uniform (0, 10,000). Two iteration chains with different initial starting values were used and sufficient numbers of iterations performed until convergence was achieved. This was inspected by a number of visual and diagnostic checks, including visual examination of history plots (which should show a snake-like appearance) and Gelman-Rubin convergence (*bgr*) diagnostic plots. In the latter, the green line and blue lines show the width, and average width, respectively, of 80% intervals of pooled chains and should be stable, whilst the red line shows the ratio of pooled/within and should be close to 1. In addition, the resulting Monte

Carlo (MC) error should be 5% or less of the posterior standard deviation (SD).

In order to illustrate a potential application of selection models in clinical trial simulations, predictions were performed for various doses of a hypothetical drug Z. The drug is assumed to be at the end of Phase 1 development and has a similar indication for pain in patients with DPNP. Table III shows the model parameters and its associated source/assumptions for this hypothetical drug. The aim was to evaluate the pain score profiles of various dose levels in a dose-ranging proof of concept study. PK and PD parameters, together with their pain score profiles were generated for 200 patients according to the mean and variance values described in Table III using Matlab version 2007b (The MathWorks Inc., Natick, MA), at dose levels 0, 2.5, 10, 20, 40 and 60 mg *q.d.* (40 subjects per dose level). The dropouts for each subject were then simulated in NONMEM, with the simulated pain scores changed to missing following a subject dropout. The resultant median pain scores with 80% prediction intervals for each dose level were plotted.

RESULTS

Summary of Dropouts

The percentage and reasons for dropouts across the three individual trials are shown in Table IV. The overall completion rate across the 3 studies was approximately 75%, with rates of at least 60% for every treatment arm across all studies. In each trial, 12% or more of patients discontinued from the 60 mg *q.d.* and *b.i.d.* arms due to adverse events, whilst in the placebo and 20 mg *q.d.* treatment arms, this figure was 7% or less. For lack of efficacy, placebo arms in each study had the highest discontinuation rates, ranging from 1 to 6%.

Table III Model parameter assumptions for hypothetical drug Z

Parameter	Assumption	Mean value (sd)
Pharmacokinetics		
CL/F (L/h)	From healthy volunteer data	30 (9)
Pharmacodynamics		
Base	Similar to duloxetine	5.74 (1.72)
PE	Similar to duloxetine	−0.0994 (0.03)
C ($week^{-1}$)	Similar to duloxetine	0.289 (0.09)
E_{max}	Similar to duloxetine	−0.576 (0.17)
EC_{50} (ng/mL)	50% more potent than duloxetine from animal models	11.8 (5.9)
K ($week^{-1}$)	20% faster onset than duloxetine from animal models	0.53 (0.159)
Dropout model		
θ_1	Similar to duloxetine	−4.25
θ_2	Similar to duloxetine	3.50
θ_3	Similar to duloxetine	0.377
θ_4	Similar to duloxetine	0.152

sd standard deviation, *CL/F* apparent plasma clearance of drug after extravascular administration, *Base* pain score at baseline, *PE* magnitude of placebo effect, *C* first order rate constant describing onset of placebo effect, E_{max} maximum effect, EC_{50} concentration for one-half the maximum effect, *K* first order rate constant describing onset of drug effect, θ_1 intercept, θ_2 and θ_3 parameters associated with time and θ_4 parameter associated with the change from baseline score prior to dropout

Mean pain scores of dropouts *versus* non-dropouts by week are shown in Fig. 1 and Kaplan-Meier graphs of dropouts grouped by various potential predictors are shown in Fig. 2. These figures show that the possible predictors of dropout may include dose (or C_{ss}) and/or age, and that dropouts appeared to have higher pain scores than non-dropouts, thus highlighting a possible dependency of dropout on previous scores and/or unobserved missing scores.

Comparison of Imputation Methods for the PK/PD model

Pharmacodynamic parameter estimates and their associated standard deviations/errors obtained from fitting the original PK/PD model to the imputed datasets from each described method are shown in Table V. The corresponding percentage change in parameter values from the original PK/PD model are shown in Fig. 3, and simulations of typical pain score profiles using parameters obtained from the different methods of imputation are shown in Fig. 4. Across the different imputation methods, the model parameters *Base*, E_{max}, *PE* and *K* were largely similar. The parameter which varied most between the different imputation methods was EC_{50}, with the completers, multiple imputation and pattern mixtures models showing approximately 30–60% lower EC_{50} estimates than the original model. In general, between subject variabilities and residual errors were largely similar across the various models. Overall, the various methods used to impute missing data produced similar PD parameter estimates, largely within 40% of the original PK/PD model. All new model estimates were within 2-fold of the original model, and the simulations of typical pain score profiles using the various new model estimates did not show any visible difference to pain score profiles.

Analysis of dropouts Using A Selection Model

In agreement with the graphical assessment, the final logistic regression model developed for dropout showed that the probability of dropout was only dependent on time and the previous observed pain score (Y_{PBSL}) prior to dropout. Final model estimates obtained from NONMEM are shown in Table VI with the model fits and 90% prediction intervals for a few Y_{PBSL} values presented in Fig. 5.

For the joint-efficacy and dropout modelling in WinBUGS, convergence was deemed satisfactory after

Table IV Percentage and Reasons for Dropout Across the 3 Trials

Study	Treatment	Completion rate (%)	% of patients who completed 6 weeks of trial	Reason for discontinuation (%)		
				Adverse event	Lack of efficacy	Others
Wernicke *et al.* (16)	Placebo	69	86	7	6	18
	60 mg *q.d.*	69	85	13	1	17
	60 mg *b.i.d.*	62	81	18	3	17
Raskin *et al.* (15)	Placebo	74	89	3	1	22
	60 mg *q.d*	82	92	4	0	14
	60 mg *b.i.d.*	77	84	12	0	11
Goldstein *et al.* (17)	Placebo	72	84	4	3	21
	20 mg *q.d.*	75	88	4	2	19
	60 mg *q.d.*	73	85	13	1	13
	60 mg *b.i.d.*	70	77	19	2	9

35,000 iterations with the first 5,000 iterations discarded for the PK/PD model describing the pain scores only. For the dropout model, this took 40,000 iterations with the first 20,000 iterations discarded (non-joint models). For the joint

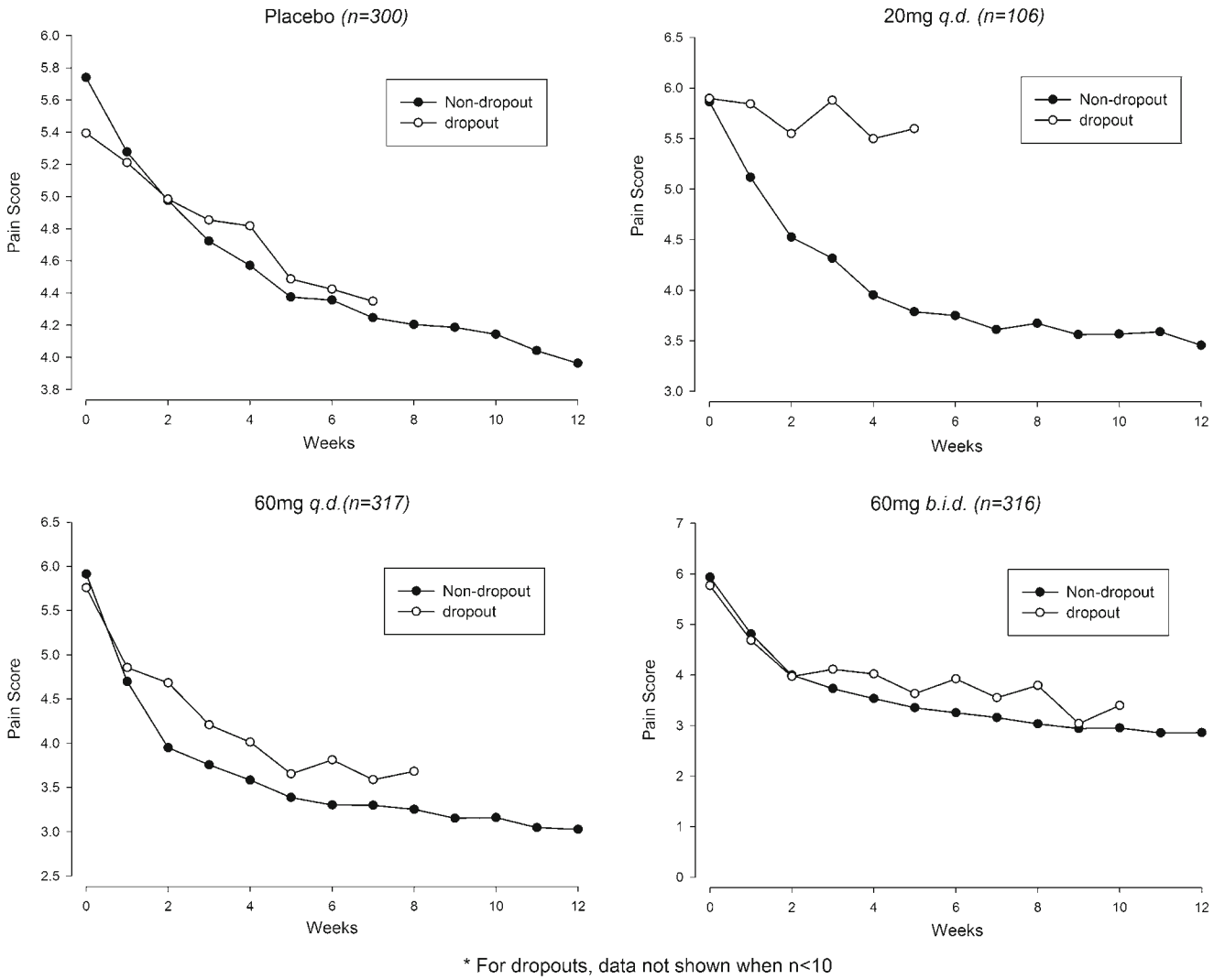

Fig. 1 Mean pain scores for dropouts *versus* non-dropouts by treatment group.

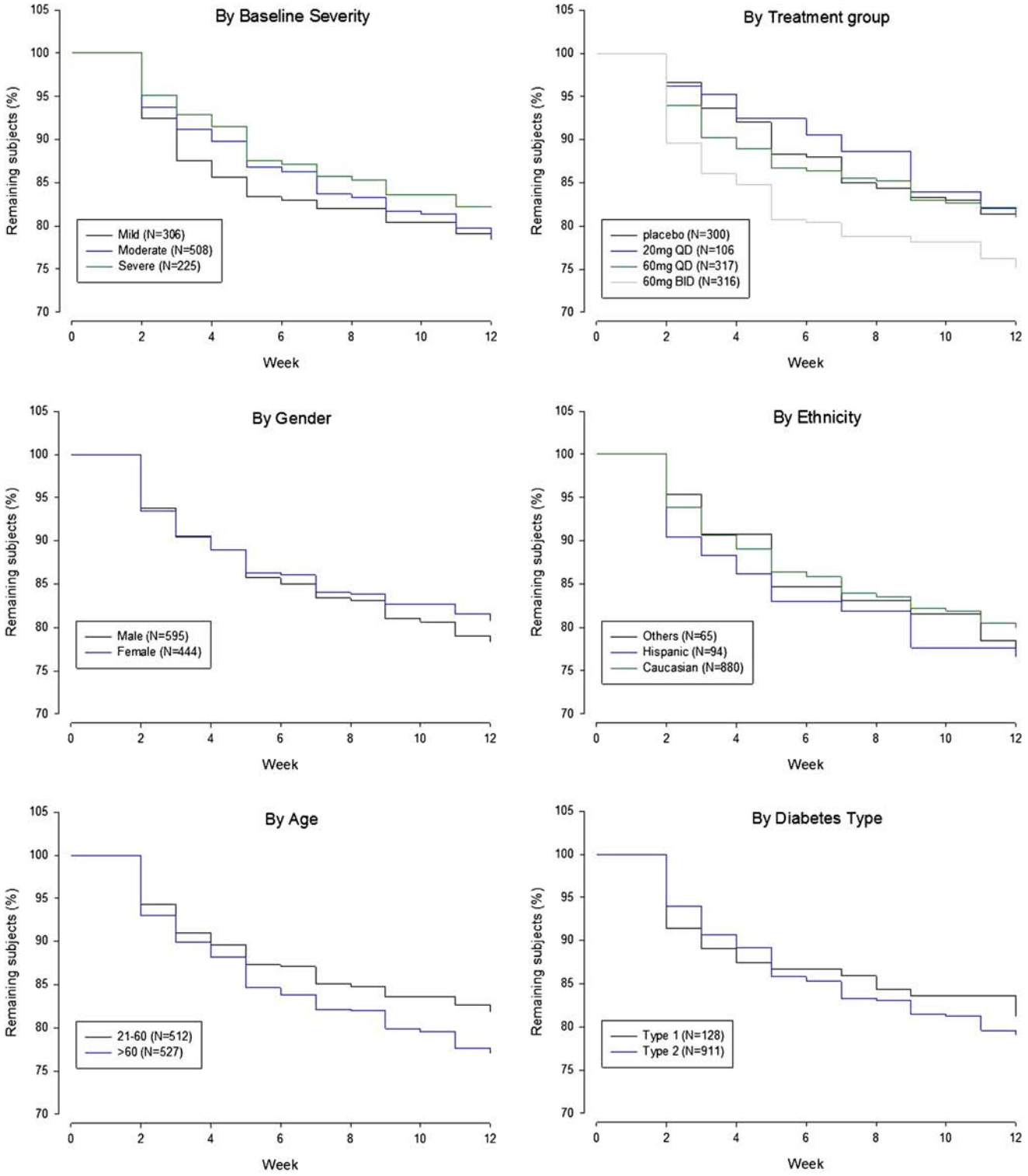

Fig. 2 Kaplan-Meier plots of possible predictors against dropout.

efficacy – dropout models, convergence was deemed satisfactory after 40,000 iterations with the first 12,000, and 10,000 iterations discarded for the uninformative and informative priors, respectively. Table VII shows the priors and posterior distributions for the various models investigated. MC error/SD calculations, along with example history and *bgr* diagnostic plots to assess convergence can be found in Fig. 6. Regardless of the set of priors used, all the joint models produced pain scores model parameter estimates similar to those of the separate model, with differences of up to 3%. This was higher for the dropout models where differences of up

Table V Pharmacodynamic parameter estimates from the various methods of imputing missing data

Parameter	*Original* PK/PD Model [14]		Completers Only		LOCF		Multiple Imputation		Pattern Mixture (CCMV)		Pattern Mixture (ACMV)		Pattern Mixture (NCMV)	
	Estimate	%SE	Estimate	%SE	Estimate	%SE	Estimate	%SD	Estimate	%SD	Estimate	%SD	Estimate	%SD
Base	5.74	0.8	5.75	0.9	5.75	0.8	5.73	0.8	5.73	0.9	5.7	0.8	5.73	0.8
PE	−0.0994	11.2	−0.102	11.8	−0.0752	28.5	−0.121	12.8	−0.11	14.8	−0.125	12	−0.153	14.2
C (week^{-1})	0.289	9.1	0.288	11.2	0.31	11.9	0.238	12.2	0.216	20	0.177	18	0.227	12.3
E_{max}	−0.576	4.6	−0.565	5	−0.558	9.8	−0.503	7.6	−0.53	7	−0.493	8.2	−0.487	7.8
EC_{50} (ng/mL)	23.6	23.0	15.5	26.4	23.7	43.5	11.59	29.8	11.6	25.9	9.84	37	14.7	29.2
K (week^{-1})	0.438	6.2	0.448	6.8	0.495	6.0	0.418	6.7	0.418	6.4	0.422	6.1	0.364	7.9
Between-subject variabilities (%)														
Base	23	4.4	23	5.2	24	4	23	5	24	5.3	24	4.9	23	5.6
PE	145	10.8	141	11.7	164	19.7	135	12.3	147	12	144	10.3	111	16.4
E_{max}	34	15.2	35	16.5	37	24.1	47	30.3	46	27.5	51	30.3	41	28.4
EC_{50} (ng/mL)	404	16.9	399	18.4	456	29.5	444	22.2	424	18.1	473	24.3	458	22.8
K (week^{-1})	108	8.3	115	8.2	11	9.8	108	8.6	108	8.4	108	9	11	8.2
Residual error														
Additive	0.705	4.9	0.701	5.3	0.661	4.9	0.91	11.3	0.845	11.2	0.851	8.5	1.15	19.7

SE standard error of the estimate, *SD* standard deviation, *Base* pain score at baseline, *PE* magnitude of placebo effect, *C* 1st order rate constant describing onset of placebo effect, E_{max} maximum effect, EC_{50} concentration for one half the maximum effect, *K* 1st order rate constant describing onset of drug effect

to 54% in model parameters were seen in the joint models compared to those modelled separately. Some dropout parameters estimated using uninformative priors had larger standard deviations compared to the estimates using vague or informative priors.

In clinical trial simulations, median and 80% prediction intervals for each dose level for hypothetical drug Z are shown in Fig. 7a. A line representing 2 points decrease from baseline is included in the graph as this has been shown to be associated with clinically meaningful pain relief (22). The no-effect dose level was 2.5 mg, and there appeared to be large overlaps in the 40 and 60 mg doses, suggesting little benefit in increasing doses beyond 40 mg. The effect of dropout on clinical trials simulations is further explored at the 20 mg dose level, with the median and 80% prediction intervals of pain scores at this dose level shown in Fig. 7b. In this particular example, median pain scores were up to 0.3 points lower and coefficient of variation up to 3% lower over the 12-week simulations when dropout was included.

DISCUSSION

A population PK/PD model describing the efficacy of duloxetine in the treatment of DPNP was previously developed and reported (14). However, not all patients completed the study, and the

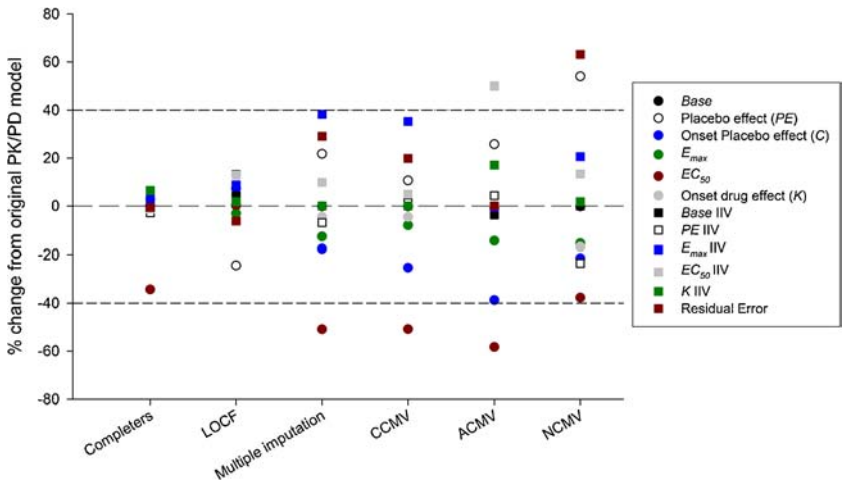

Fig. 3 Percentage change from original PK/PD model for PD parameter estimates across different methods for imputing missing data.

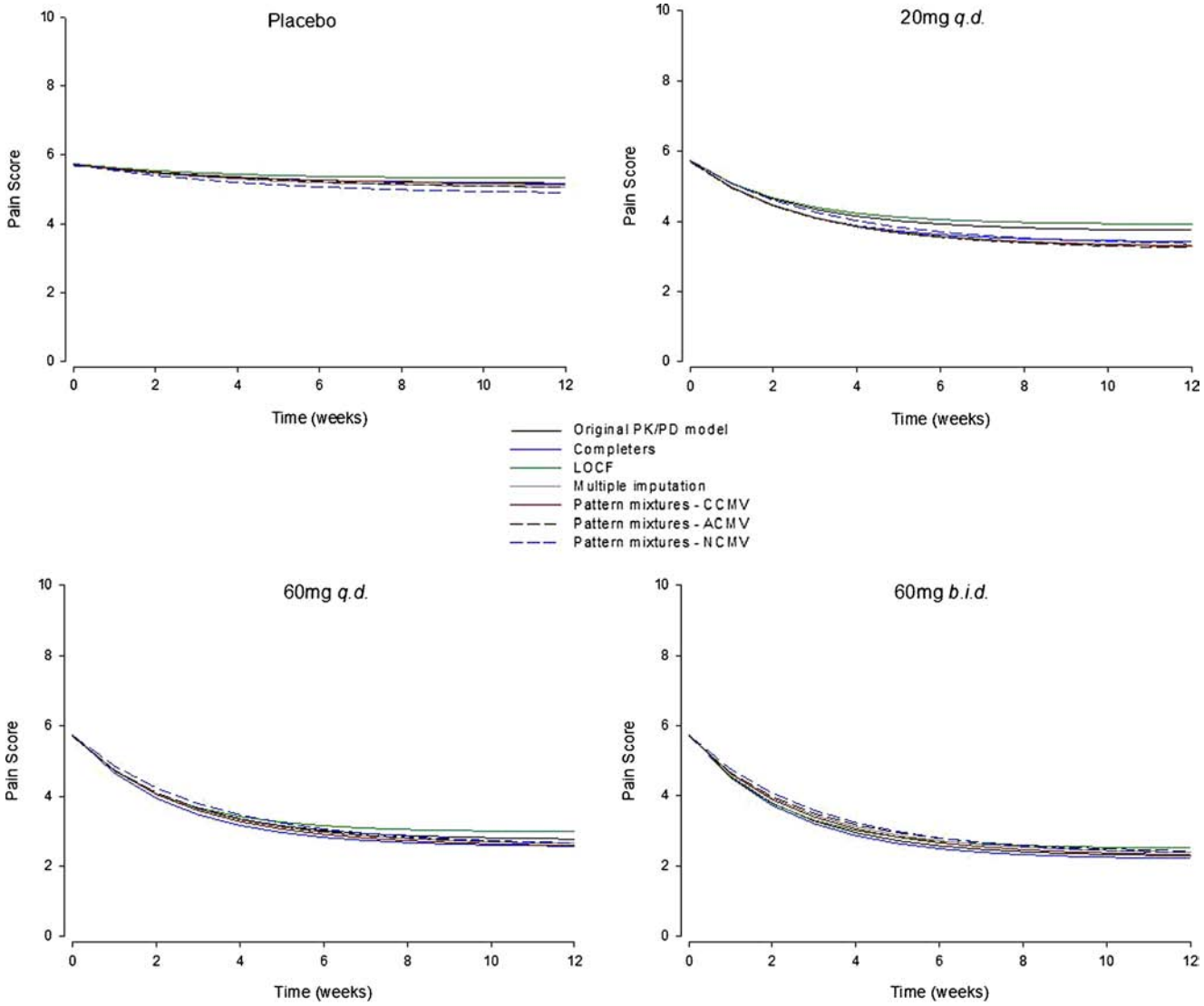

Fig. 4 Simulations of typical pain score profiles from models using different methods of handling missing data.

analyses included only non-missing observations from patients. The aims and objectives of the current analyses were to examine the influence of different methods for handling missing data, and how this might affect the results of the original population PK/PD analyses.

Table VI Parameter estimates from logistic dropout model

logit (Pr) = $\theta_1 + \theta_2$*exp ($-\theta_3$*time) + θ_4* Y_{PBSL}		
Parameter	Estimate	SE (%)
θ_1 (intercept)	−4.25	5.69
θ_2 (time)	3.50	25.2
θ_3 (time)	0.377	40.3
θ_4 (Y_{PBSL})	0.152	30.6

SE standard error of the estimate, θ_1 intercept, θ_2 and θ_3 parameters associated with time and θ_4 parameter associated with the change from baseline score prior to dropout (Y_{PBSL}).

From the observed completion rates in each trial (Table IV), more patients in the 60 mg treatment groups (*q.d.* and/or *b.i.d.*) discontinued the trials due to adverse events compared to the placebo and 20 mg *q.d.* treatment groups. For lack of efficacy, more patients in the placebo group discontinued the trials compared to the duloxetine-treated groups. Together, these suggest that dropouts may not be entirely random. The observed completion rates of 60 to 80% are consistent with similar DPNP trials (23,24). Plots of individual parameter estimates of the original PK/PD model *versus* reason for dropout (graphs not shown) also did not reveal any visible trends. This could be because the numbers of dropouts were relatively small compared to non-dropouts, and also we have shown that the dropout model can be developed separately from the pain scores model, therefore dropouts are not expected to have an effect on parameter estimates of the pain scores model.

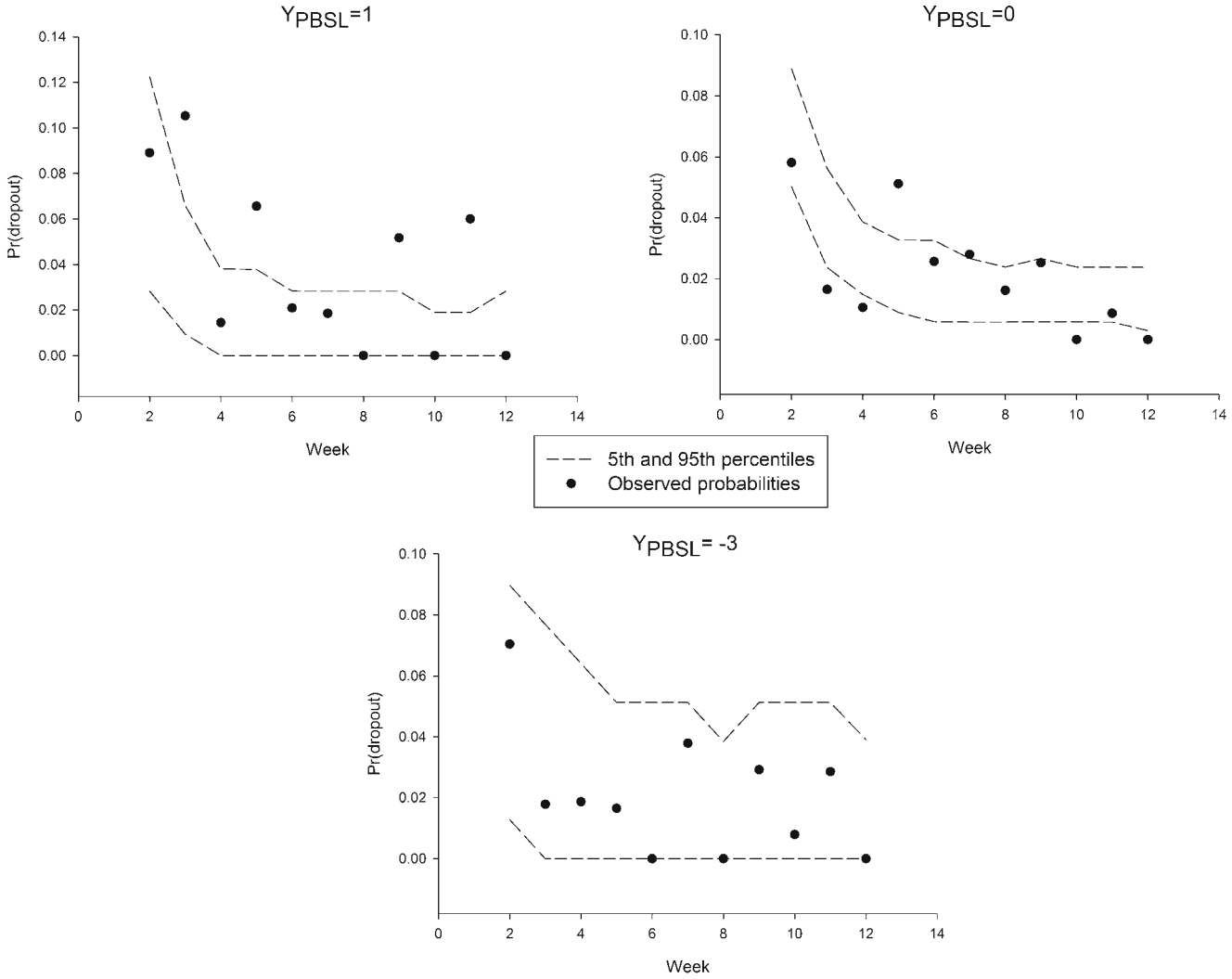

Fig. 5 Observed probabilities with 90% prediction intervals for a few Y_{PBSL} values.

The various methods used to impute missing data produced similar PD parameter estimates, mostly within 40% of the original PK/PD model. The completers model had a 34% lower EC_{50} estimate since the dataset did not include patients who had dropped out due to lack of efficacy, thus leading to over-estimation of drug potency. Indeed, complete-case analysis should not be used routinely, since it ignores all patients with missing data, but this does not satisfy the intent-to-treat principle as it does not include all randomized patients (25).

In the LOCF model, EC_{50} estimates were very similar to the original PK/PD model, in contrast to common opinion that LOCF provides a more conservative approach - an issue which had been previously discussed by Mallinckrodt *et al.* (26). This situation is possible since in the duloxetine trials, observed dropout rate due to lack of efficacy is minimal compared to those of other reasons or adverse events. It is intuitively expected that LOCF would be a worst case scenario analysis only if a substantial number of patients dropped out of the trial due to lack of efficacy. In fact, in the current analyses, LOCF was the only imputation method where all the parameters were within the 40% boundaries from the original PK/PD model (Fig. 3). This brings into question the appropriateness of LOCF as a conservative method, which is an approach that has been regularly expected by regulatory authorities for submissions. Molnar *et al.* (27) have suggested that regulatory agencies should re-consider accepting analyses based on the LOCF approach since it has been shown to be biased. The main drawback of simple imputation methods such as LOCF lies in the underestimation of variance and thus overstating precision, plus the assumption that data are not missing at random (28,29).

The multiple imputation and pattern mixtures methods used five imputed datasets since the use of more than 5 to 10 imputations provide little extra benefit unless rates of missing information are unusually high (30). Across the pattern mixture methods, although the estimates and variabilities were largely comparable, those from NCMV had a larger residual error. This could be due to the fact that imputation using the NCMV method was less robust, since a fairly large proportion of missing data were imputed from models fitted

Table VII Parameter estimates from Joint Models (WinBUGS).

Model Type	Separate modelling of Pain scores and dropout		Joint model Uninformative priors		Joint model Informative priors	
Parameter	Priors	Posterior mean (SD)	Priors	Posterior mean (SD)	Priors	Posterior mean (SD)
Pain scores model						
Base	Unif (0,100)	6.02 (0.05)	Same as separately-developed Pain model	6.02 (0.05)	Same as separately-developed Pain model	6.02 (0.05)
PE	Unif (0,100)	−0.301 (0.01)		−0.301 (0.01)		−0.301 (0.01)
C	Unif (0,100)	0.246 (0.01)		0.246 (0.01)		0.247 (0.01)
E_{max}	Unif (0,100)	−0.393 (0.02)		−0.392 (0.02)		−0.392 (0.02)
EC_{50}	Norm (23.6, 29.4)	34.5 (4. 9)		34.0 (4.9)		33.8 (4.9)
K	Unif (0,100)	1.07 (0.07)		1.08 (0.07)		1.07 (0.07)
IIV Base	Unif (0,100)	148%		148%		148%
IIV PE	Unif (0,100)	24%		24%		24%
IIV E_{max}	Unif (0,100)	25%		25%		25%
IIV EC_{50}	Gamma (0.1,0.1)	1,280%		1,250%		1,240%
IIV K	Unif (0,100)	53%		53%		53%
Dropout model						
θ_1	Norm (0,10,000)	−4.34 (0.3)	Norm (0,10,000)	−4.34 (0.5)	Norm (−4.38,15.9)	−4.27 (0.3)
θ_2	Norm (0,10,000)	4.71 (2.0)	Norm (0,10,000)	5.91 (8.0)	Norm (3.77,1.25)	4.24 (1.5)
θ_3	Norm (0,10,000)	0.471 (0.2)	Norm (0,10,000)	0.544 (0.3)	Norm (0.380,48.9)	0.479 (0.2)
θ_4	Norm (0,10,000)	0.0848 (0.05)	Norm (0,10,000)	0.131 (0.05)	Norm (0.0787,531)	0.129 (0.05)
Residual error	Gamma (0.1, 0.1)	0.745 (0.009)	Gamma (0.1, 0.1)	0.735 (0.008)	Gamma (0.1, 0.1)	0.735 (0.008)

Unif uniform distribution, *norm* normal distribution, *IIV* interindividual variability, *Base* pain score at baseline, *PE* magnitude of placebo effect, *C* 1st order rate constant describing onset of placebo effect, E_{max} maximum effect, EC_{50} concentration for one half the maximum effect, *K* 1st order rate constant describing onset of drug effect, θ_1 intercept, θ_2 and θ_3 parameters associated with time and θ_4 parameter associated with the unobserved change from baseline score at time of dropout

using data from fewer than 150 patients. In contrast, the ACMV and CCMV methods involved missing data imputed from model estimates derived from analyses of at least 825 patients.

Selection models have an important advantage over imputation methods for missing data, since they allow characterisation of the randomness of dropout and the models to be used for clinical trial simulations. It has been noted that models that do not include dropout in clinical trial simulations may not fully capture the observations (11). Thus, in order to accurately reproduce the clinical trial results, or simulate with better predictive capabilities for future trials, it is therefore necessary to describe the effect of drug efficacy/disease progression on the probability of dropout and vice versa.

Logistic models were chosen to model dropout in these analyses since they offer greater flexibility for inclusion of many predictors at a time and have shorter run times (31,32). Besides, they also provide an output that may be more intuitive and understandable for non-pharmacometricians or statisticians. Hazard models have the advantage of taking into account interval censoring and are more parsimonious, but also require longer run-times when modelling. The results for the logistic model showed that the probability of dropout was only dependent on time and previous observed pain score prior to dropout. Addition of the latter as a predictor caused a larger drop in objective function value compared to the use of unobserved missing pain score at time of dropout as a predictor. This is consistent with the plots of mean pain score between dropouts and non-dropouts (Fig. 1), where dropouts tended to have a higher pain score compared to trial completers. Dropouts in clinical trials are rarely completely

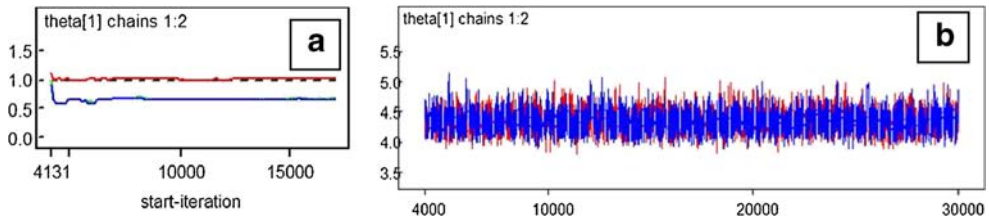

Fig. 6 Example *bgr* (**a**) and history (**b**) plots to assess drug efficacy/dropout model convergence in WinBUGS.

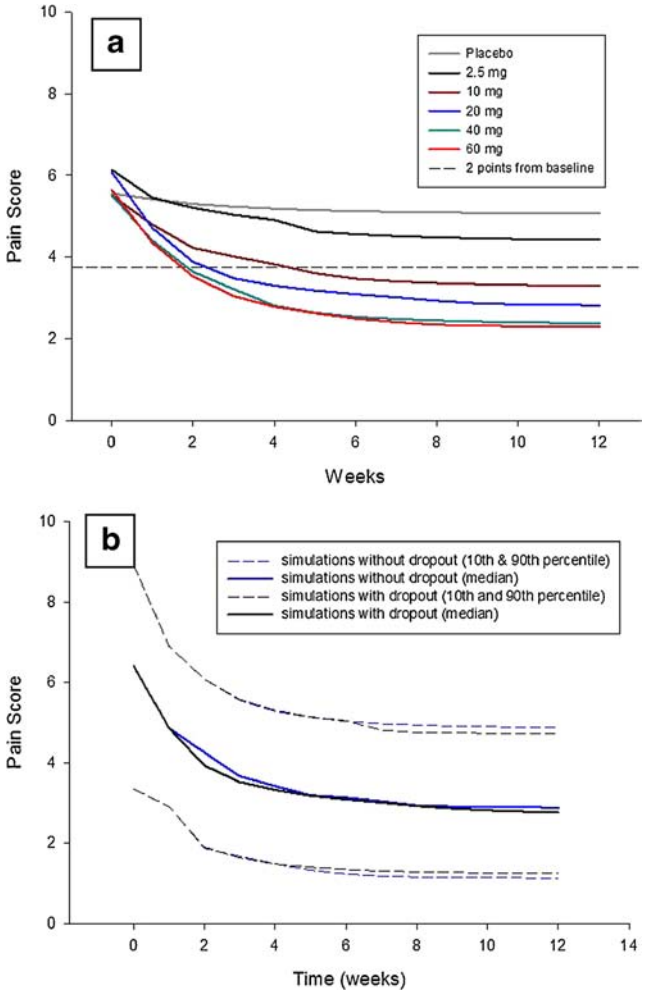

Fig. 7 Simulations for various dose levels for hypothetical drug Z (**a**) and simulations for 20 mg dose with and without dropout (**b**) (N = 40 per dose level).

random, since in reality, adverse events or lack of efficacy often lead patients to discontinue clinical trials.

In the investigation of joint models for non-linear longitudinal data with informative dropout, Hu and Sale (11) described missing data as completely random (CRD), random (RD) or informative (ID). Dropout is CRD when it is independent of observed and missing data, RD when it is dependent on observed data, and ID when it is dependent on missing data. Under CRD and RD, the dropout likelihood and the disease progression likelihood can be optimised separately, whereas under ID, the dropout and disease progression should be modelled together since they contain information about each other. However, it has also been shown that a sufficiently high observation density can justify a sequential rather than a joint modelling approach of PKPD and dropout data (33). Others have also shown it is often adequate to assume that the conditional hazard of dropout between the ith and $(i+1)$ th time points, given Y at times prior to and including the ith, depends only on the score Y_i (34). Therefore, in order to determine the influence of dropout on the drug efficacy model and to support the modelling approach of separate dropout/drug efficacy model development, a joint model was built and compared to the previous results from separate modelling. This joint model was developed under a Bayesian framework since it is a natural way to handle shared-parameter models (29). The pain score and dropout model parameters used mainly uninformative priors since this would mirror the nonlinear mixed effects estimation methods. However, for EC_{50}, an informative prior was required in order for the model to be stable, possibly due to the large inter-individual variability originally observed for this parameter (14). The results showed that the use of informative or uninformative priors for dropout models produced similar parameter estimates for both the pain score and dropout models. Higher uncertainties in parameter estimation with the uninformative priors were observed, which is consistent with current knowledge that dropout data provide little information for estimating dropout model parameters, thus informative prior distributions may be necessary (21). The results of the joint modelling exercise support the use of separate models for dropout and drug efficacy in this particular analysis of dropouts and duloxetine drug efficacy in DPNP, in concordance with reported literature for other drugs (33,34).

A simple case study was presented to illustrate the utility of a selection model in clinical trial simulations. This was to satisfy the main aim of illustrating how a joint efficacy-dropout model can be used for simulations as opposed to imputation methods for describing dropout which cannot perform a similar task. In reality, these simulations can be made simpler or more complicated depending on data availability and current stage of drug development. Lockwood *et al.* (35) have used a similar approach for clinical trial simulations using data from gabapentin for pregabalin clinical trial simulations in chronic neuropathic pain. However dropouts were not included in those clinical trial simulations. In our example, failure to include the dropouts would have caused a slight over-estimation of variability and underestimation of drug efficacy. Although minor, these would have potential implications in calculations of sample sizes required to detect significant changes in pain scores when designing future clinical trials. The sensitivity of the trial outcome to dropouts can also be investigated with these models by changing the dropout rate.

CONCLUSION

In summary, for the analyses on impact of missing data using imputation methods, the relative change in PK/PD parameter estimates from models fitted to different imputed datasets showed that most were within 40% of the original PK/PD model. These support the fact that at the current observed

overall trial completion rate of approximately 75%, missing data had relatively little impact on the original PK/PD analyses. It is uncertain, however, if these conclusions would remain should there be an increase in the percentage of missing data. It is therefore important, that sensitivity analyses be conducted to test the effect of dropout on the estimated PK/PD parameters in modelling exercises, since these approaches frequently only include non-missing data. Whilst these imputation methods can be useful in sensitivity analyses for determining appropriate choices for handling missing data, they however do not easily facilitate simulations and prediction of dropout in future trials. Selection models, on the other hand, are best poised to handle these situations, including investigation of the randomness of dropout. The selection model developed for duloxetine in DPNP has shown that the dropout mechanism for these trials was random, ie. dependent on previous observed pain score prior to dropout. Also, the logistic model for dropout included separate components for the different reasons for dropout, and finally, the advantage of selection models in clinical trial simulations was highlighted via a hypothetical case study.

ACKNOWLEDGMENTS AND DISCLOSURES

This work was financially supported by Eli Lilly and Co. Eunice Yuen and Ivelina Gueorguieva are employees of Eli Lilly & Co. and Leon Aarons is a paid consultant to Eli Lilly & Co.

REFERENCES

1. Hedden SL, Woolson RF, Malcolm RJ. A comparison of missing data methods for hypothesis tests of the treatment effect in substance abuse clinical trials: a Monte-Carlo simulation study. Subst Abuse Treat Prev Policy. 2008;3:13.
2. Fairclough DL. Design and analysis of quality of life studies in clinical trials. Boca Raton: Chapman and Hall/CRC; 2002.
3. Siddiqui O, Ali MW. A comparison of the random-effects pattern mixture model with last-observation carried forward (LOCF) analysis in longitudinal clinical trials with dropouts. J Biopharm Stat. 1998;8:545–63.
4. Lindsey JK. Dropouts in longitudinal studies: definitions and models. J Biopharm Stat. 2000;10:503–25.
5. Food and Drug Administration. Guidance for Industry: E9 Statistical Principles for Clinical Trials; 1998. Available from http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm073137.pdf.
6. O'Neill RT, Temple R. The prevention and treatment of missing data in clinical trials: an FDA perspective on the importance of dealing with it. Clin Pharm Ther. 2012;91:550–4.
7. Panel on Handling Missing Data in Clinical Trials; National Research Council. The Prevention and Treatment of Missing Data in Clinical Trials; 2010. Available from http://www.nap.edu/catalog.php?record_id=12955#toc.
8. European Agency for the Evaulation of Medicinal Products. Points to consider on missing data; 2001. Available from http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003641.pdf.
9. European Medicines Agency. Guideline on missing data in confirmatory clinical trials; 2009. Available from http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003642.pdf.
10. Food and Drug Administration. Guidance for Industry: Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims; 2009. Available from http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf.
11. Hu C, Sale ME. A joint model for nonlinear longitudinal data with informative dropout. J Pharmacokinet Pharmacodyn. 2003;30:83–103.
12. Food and Drug Administration. FDA Briefing Document for the Cardiovascular and Renal Drugs Advisory Committee (CRDAC); 2012. Available from http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/CardiovascularandRenalDrugsAdvisoryCommittee/UCM304755.pdf.
13. Gomeni R, Lavergne A, Merlo-Pich E. Modelling placebo response in depression trials using a longitudinal model with informative dropout. Eur J Pharm Sci. 2009;36:4–10.
14. Yuen E, Gueorguieva I, Bueno-Burgos L, Iyengar S, Aarons L. Population pharmacokinetic/pharmacodynamic models for duloxetine in the treatment of diabetic peripheral neuropathic pain. Eur J Pain. 2013;17:382–93.
15. Raskin J, Pritchett YL, Wang F, D'Souza DN, Waniger AL, Iyengar S, *et al.* A double-blind, randomized multicenter trial comparing duloxetine with placebo in the management of diabetic peripheral neuropathic pain. Pain Med. 2005;6:346–56.
16. Wernicke JF, Pritchett YL, D'Souza DN, Waniger AL, Tran P, Iyengar S, *et al.* A randomized controlled trial of duloxetine in diabetic peripheral neuropathic pain. Neurology. 2006;67:1411–3.
17. Goldstein D, Lu Y, Detke M, Lee TC. Iyengar S. Duloxetine vs placebo in patients with painful diabetic neuropathy Pain. 2005;116:109–18.
18. Lobo E, Quinlan T, O'Brien L, Knadler M, Heathman M. Population pharmacokinetics of orally administered duloxetine in patients. Clin Pharmacokinet. 2009;48:189–97.
19. Beal S, Sheiner LB, Boeckmann A. NONMEM Users Guides. Ellicott City: Icon Development Solutions; 1989–2006.
20. Lunn D, Thomas A, Best N, Spiegelhalter D. WinBUGS – a Bayesian modelling framework: concepts, structure and extensibility. Stat Comput. 2000;10:325–37.
21. Best NG, Speigelhalter D, Thomas A, Brayne CEG. Bayesian analysis of realistically complex models. J R Statist Soc A. 1996;159:323–42.
22. Farrar JT, Young JP, LaMoreaux L, Werth JL, Poole RM. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. Pain. 2001;94:149–58.
23. Wymer JP, Simpson J, Sen D, Bongardt S. Efficacy and Safety of Lacosamide in Diabetic Neuropathic Pain. Clin J Pain. 2009;25:376–85.
24. Freynhagen R, Strojek K, Griesing T, Whalen E, Balkenohl M. Efficacy of pregabalin in neuropathic pain evaluated in a 12-week, randomised, double-blind, multicentre, placebo-controlled trial of flexible- and fixed-dose regimens. Pain. 2005;115:254–63.
25. Liu G, Gould AL. Comparison of alternative strategies for analysis of longitudinal trials with dropouts. J Biopharm Stat. 2002;12:207–26.
26. Mallinckrodt CH, Sanger TM, Dube S, DeBrota DJ, Molenberghs G, Carrol RJ, *et al.* Assessing and Interpreting Treatment Effects in Longitudinal Clinical trials with missing data. Biol Psychiatry. 2003;53:754–60.
27. Molnar FJ, Man-Son-Hing M, Hutton B, Fergusson DA. Have last observation-carried-forward analyses caused us to favour more toxic

dementia therapies over less toxic alternatives? A systematic review. Open Med. 2009;3:e31–50.
28. Lane P. Handling drop-out in longitudinal clinical trials: a comparison of the LOCF and MMRM approaches. Pharmaceut Statist. 2008;7:93–106.
29. Molenberghs G, Kenward MG. Missing Data in Clinical Studies. West Sussex: John Wiley and Sons; 2007.
30. Schafer JL. Multiple imputation: A primer. Stat Methods Med Res. 1999;8:3–15.
31. Ridout MS, Diggle P. Testing for random dropouts in repeated measurement data. Biometrics. 1991;47:1617–21.
32. Reddy VP, Kozielsak M, Johnson M, Vermuelen A, de Greef R, Liu J, *et al.* Structural models describing placebo treatment effects in schizophrenia and other neuropsychiatric disorders. Clin Pharmacokinet. 2011;80:429–50.
33. Diderichsen P, Dutta S. Sufficiently high observation density justifies a sequential modelling approach of PKPD and dropout data; 2010. Available from http://www.page-meeting.org/default.asp?abstract=1836.
34. Sheiner LB, Beal SL, Dunne A. Analysis of nonrandomly censored ordered categorical longitudinal data from analgesic trials. J Am Stat Assoc. 1997;92:1235–43.
35. Lockwood PA, Cook JA, Ewy WE, Mandema JW. The use of clinical trial simulation to support dose selection: application to development of a new treatment for chronic neuropathic pain. Pharm Res. 2003;20:1752–9.