Handling of missing data in long-term clinical trials: a case study

Peer-reviewed author version

JANSSENS, Mark; MOLENBERGHS, Geert & Kerstens, René (2012) Handling of missing data in long-term clinical trials: a case study. In: PHARMACEUTICAL STATISTICS, 11 (6), p. 442-448.

DOI: 10.1002/pst.1532
Handle: http://hdl.handle.net/1942/14443

Handling of missing data in long-term clinical trials: a case study

Short title: A case study on missing single-arm trial data

Mark Janssens[1], Geert Molenberghs[2], René Kerstens[1]

[1]Shire-Movetis NV, Veedijk 58, B-2300, Belgium

[2]I-BioStat, Universiteit Hasselt & Kaholieke Universiteit Leuven, Agoralaan 1, B-3590 Diepenbeek, Belgium

Correspondence: René Kerstens, Veedijk 58, B-2300 Turnhout, Belgium, tel:+32 14404397, fax: +32 14404391, email: rkerstens@shire.com

ABSTRACT

Missing data in clinical trials is a well known problem and the classical statistical methods used can be overly simple. This case study shows how well-established missing data theory can be applied to efficacy data collected in a long-term, open-label trial with a discontinuation rate of almost 50%. Satisfaction with treatment in chronically constipated patients was the efficacy measure assessed at baseline and every 3 months post baseline. The improvement in treatment satisfaction from baseline was originally analyzed with a paired *t*-test, ignoring missing data and discarding the correlation structure of the longitudinal data. As the original analysis started from Missing Completely At Random (MCAR) assumptions regarding the missing data process, the satisfaction data were re-examined and several Missing At Random (MAR) and Missing Not At Random (MNAR) techniques resulted in adjusted estimate for the improvement in satisfaction over 12 months. Throughout the different sensitivity analyses, the effect sizes remained significant and clinically relevant. Thus, even for an open label trial design, sensitivity analysis, with different assumptions for the nature of dropouts (MAR or MNAR) and with different classes of models (selection, pattern-mixture, or multiple imputation models), has been found useful and provides evidence towards the robustness of the original analyses; additional sensitivity analyses could be undertaken to further qualify robustness.

KEYWORDS longitudinal data; missing data; sensitivity analysis; pattern-mixture models

1. INTRODUCTION

In clinical trials it is common to have incomplete data collection. Historically, much emphasis has been given to simple methods of data interpretation; for example, an analysis of completers, or a cross-sectional analysis at the last visit preceded by a single imputation technique, such as last observation carried forward. In a long-term, open-label trial of prucalopride, which had a high dropout rate, the improvement in treatment satisfaction was originally evaluated using a paired *t*-test. More adequate methods have since been proposed and have now become well established.

1.1. Background: pivotal double-blind, placebo-controlled prucalopride trials

Prucalopride was clinically tested in patients with chronic constipation in three confirmatory phase 3 trials with identical trial designs (12-week, randomized, double-blind, placebo-controlled, parallel-group studies). The primary efficacy endpoint was the proportion of subjects with an average of at least three bowel movements per week during the 12-week treatment period [1, 2, 3].

Health-related quality of life constituted a secondary efficacy endpoint in the prucalopride trials, and was measured using the Patient Assessment of Constipation – Quality of Life Questionnaire (PAC-QOL). The PAC-QOL is a self-administered questionnaire, developed and validated as a constipation-specific assessment tool for use in clinical studies [4]. The dissatisfaction subscale of the PAC-QOL questionnaire (five items) provides a metric score between 0 (low dissatisfaction) and 4 (high dissatisfaction). The dissatisfaction score, a key secondary endpoint in the trials that was associated with the primary efficacy endpoint, is the parameter of interest throughout this article. An improvement in satisfaction of ≥ 1 is considered to be a clinically meaningful outcome [5].

At the end of the 12-week, double-blind treatment period, 22% of patients who received placebo had an improvement of ≥ 1 on the satisfaction scale compared with 44% of those treated with prucalopride. At the same time, the median changes from baseline were 0.00 for patients taking placebo and 0.60 for patients treated with prucalopride. Furthermore, the PAC-QOL dissatisfaction subscale (key secondary endpoint) was associated with the frequency of bowel movements (primary endpoint): responders achieving the primary

endpoint had an increase on the satisfaction scale of 1.18 points compared with 0.46 points for non-responders. Because of proprietary reasons, the data cannot be made available.

1.2. Open-label extension and original analysis

Patients completing the double-blind pivotal trials could continue into a 2-year, open-label safety extension in which all participants received prucalopride. A total of 78% of patients did this, with similar numbers from each treatment arm. A daily diary to collect information on bowel movements was not used in the open-label extension. However, patient responses to the five-item PAC-QOL dissatisfaction subscale were assessed every 3 months and provided a good proxy for long-term prucalopride efficacy.

The average dissatisfaction scores from baseline to month 12 are shown in Figure 1. The improvement in satisfaction (i.e. the change in dissatisfaction since the first use of prucalopride) was evaluated using a paired *t*-test. An improvement of 1.59 ($p < 0.0001$) was found after 12 months of treatment with prucalopride [6].

The original analysis has two limitations. The first limitation is the one-arm, open-label design. In the more stringent setting of the double-blind pivotal trials, the change in dissatisfaction correlated well with primary efficacy, and prucalopride outperformed placebo. However, using a one-arm design, there is no empirical guarantee that this will remain true *beyond* the length of the pivotal trials, although it helps to build confidence. The second limitation relates to the assumptions about missing data when performing a paired *t*-test. The paired nature of the test implies that only completers are retained in the analysis, which would result in unbiased estimates under the strong assumption of Missing Completely At Random only (see Methods). The assumptions about missing data are especially important given the high dropout rate (46% of patients withdrew from the study before completing 12 months of treatment) and increased dissatisfaction in subjects withdrawing from the trial. The plots in Figure 2 show that completers had lower dissatisfaction scores than dropouts, i.e. there is no evidence that dropout was completely at random.

The aim of this article is to explore the robustness of the original analysis (paired *t*-test) and to encourage the use of sensitivity analysis, classically applied to the primary endpoint of a confirmatory trial, in a broader setting.

2. METHODS

2.1. Types of missing data

Three types of missing data are routinely distinguished. Missing data are said to be Missing Completely At Random (MCAR) if the missingness is independent of both unobserved and observed outcomes, although a dependence on model covariates is allowed. Missing At Random (MAR) means that, conditional on the observed outcomes and covariates, missingness is independent of the unobserved measurements. Otherwise data are Missing Not At Random (MNAR). Under the parameter distinctness condition, valid MAR results can be obtained through a likelihood-based analysis that *ignores* the missingness process. The likelihood approach does not involve explicit modeling of the missingness process, so such an analysis is therefore referred to as *ignorable* [7, 8].

2.2. Sensitivity

Depending on the context, the term ‘sensitivity’ relates to various concepts: the estimand (i.e. the parameter of interest), ways to analyze the data, assumptions about the data, or assumptions about the *missing* data. The choice of estimand involves both the outcome measure and the analysis population [9]. In this article the estimand is not variable, and is defined as the change in dissatisfaction since the first use of prucalopride for all randomized subjects from the first prucalopride dose up to 12 months of treatment.

Eight models are fitted in this article. Models 1–3 refer to ways in which the data are analyzed. Models 4 and 5, which explore an alternative specification of the mean structure, amend the assumptions about the data. Models 6–8 modify the assumptions about the *missing* data, in the sense that MAR and MNAR assumptions are applied in the framework of an under-identified pattern-mixture model.

Models 1–3: basic models

Different frameworks may be used to analyze data. For this case study, these are confined to fitting a selection model directly (model 1), a pattern-mixture model (model 2), and a multiple imputation model (model 3). Shared-parameter models, where the measure of interest and dropout are modeled jointly, usually in a Bayesian setting, will not be discussed. Proper multiple imputation, which was originally conceived as a Bayesian technique, can be used in conjunction with a variety of analysis tools (e.g. cross-sectional analysis at endpoint, a selection model, a pattern-mixture model, and a shared-parameter model) [10, 11]. Finally, multiple imputation was applied and combined with a cross-sectional analysis.

Models 4 and 5: alternative specification of the mean structure

The second set of sensitivity analyses examines the treatment effect under a more relaxed mean structure (i.e. a mean structure that allows the dissatisfaction score to increase again after 3 months of treatment). A piecewise linear model with breakpoint at month 3 was fitted to all data (model 4) as well as for completers and dropouts separately (model 5). The latter model is still an *identifiable* pattern-mixture model; no identifying restrictions were needed at this point. Other options could be explored here, for instance the use of surge functions or inverse polynomial functions.

Models 6–8: departures from MAR (pattern-mixture models)

The third set of sensitivity analyses explores departures from MAR in the framework of an *under-identified* pattern-mixture model. The starting point for these models is the basic pattern-mixture model for completers and dropouts (model 2, two patterns). The number of patterns is expanded from two to five (each dropout time defines a pattern), resulting in an over-specified model. Next, classical identifying restrictions are applied: Available Case Missing Value (ACMV) restrictions (model 6); Nearest Case Missing Value (NCMV) restrictions (model 7); and Complete Case Missing Value (CCMV) restrictions (model 8). The terms Available Case, Nearest Case, and Complete Case refer to the patterns from which information is derived in order to identify the model. When a model effect for an early dropout pattern cannot be estimated, CCMV restrictions will reach out to the estimated model effect of the fully observed pattern (i.e. the completers); NCMV will

borrow effect estimates from the nearest pattern; and ACMV will, in the absence of an estimate, average over all patterns where the model effect could be estimated. The average is weighted in the sense that patterns weigh on the ACMV estimate proportional to the number of subjects that they represent. It has been shown that, within the realm of identifying restrictions, ACMV restrictions come down to MAR dropout assumptions, while NCMV and CCMV restrictions lead to strict MNAR models [12].

Pattern-mixture models allow both the identifying restrictions and the mean/covariance structure to be specified. Both elements can be drawn in a Bayesian framework (see, for example, Demirtas 2005 [13]). Bayesian constraints for inestimable parameters allow for a natural amount of random variability across patterns. Demirtas also notes that sophisticated polynomial mean structures tend to cause instability whereas simple ones may be unable to capture genuine data trends [13]. We agree with the latter observation and have noticed that (inverse) polynomials, both for the identifiable pattern-mixture model (model 5) and the under-identified models (models 6–8), tend to generate poor fit and instability. For reasons of fit and stability, a piecewise linear approach was selected for model 5 and a sparse specification of the mean structure for models 6–8 (see Analysis and Results).

All models, apart from the multiple imputation model (model 3), are fitted using direct likelihood techniques. Within the pattern-mixture models, the standard errors for pattern-averaged estimates are adjusted using the delta method. The delta adjustment reflects the fact that pattern proportions are unknown quantities, estimated from the data. As a final methodological note, there is no empirical evidence that the true underlying model is MAR or MNAR. The aim of sensitivity analysis is to explore plausible departures from MAR (and their effect on the conclusion) rather than to confirm that a particular MNAR model is correct [11].

We developed a macro using Statistical Analysis Software (SAS version 9.2, SAS Institute Inc., Cary, USA) to generate these pattern-mixture estimates automatically from the analysis data set, which decreased the work required and reduced coding error. In addition, the macro provided adjusted standard errors for pattern-averaged estimates

(using the delta method). The code, including an example, is available in the Supplementary Material.

3. ANALYSIS AND RESULTS

3.1. Models 1–3: basic models

The first model is the random intercept model:

$$Y_{ij} = \beta_0 + b_{0i} + \beta_1 Y\_BAS_i + \beta_2 SBM\_BAS_i + \beta_3 (t_{ij})^{-1} + \varepsilon_{ij}$$

where i is the subject indicator and j is the within-subject indicator for the repeated measures. Y_{ij} is the longitudinal dissatisfaction score, Y_BAS_i is the dissatisfaction at baseline, and SBM_BAS_i is the number of spontaneous bowel movements (SBM) at baseline. t_{ij} is the time point indicator; the values of t_{ij} are fixed by the value of j, which means that $t_{ij} = 1$ at baseline ($j = 1$), $t_{ij} = 2$ at the second measurement ($j = 2$), $t_{ij} = 3$ at the third measurement ($j = 3$), and so forth. β_0, β_1, and β_2 are the fixed effects; b_0 is the random intercept and $b_0 \sim N(0, \sigma_b^2)$ with σ_b^2 the between-subject variance; ε_{ij} is the residual error and $\varepsilon_{ij} \sim N(0, \sigma^2)$ with σ^2 the residual variance.

The number of SBM at baseline is an indicator of the severity of constipation. More severely affected patients are likely to have higher dissatisfaction scores, and this is confirmed by the trial data. A bowel movement is termed spontaneous when it is not induced by laxative medication. The average number of SBM at baseline was observed for every patient, hence SBM_BAS_i is included in the model as a time-independent covariate without missing values. Other covariates such as sex or age were not statistically significant.

The time point indicator t was used to mimic empirically the initial drop in dissatisfaction followed by more stable levels of dissatisfaction (see Figure 1). A different view of dissatisfaction development arises, however, when the data are plotted according to dropout pattern (see Figure 2). The sensitivity of the results to the specification of the mean structure is considered in models 4 and 5.

The first model is conditional on the random effect; it is a selection model. The second model belongs to the class of pattern-mixture models and is convenient to use because it is fully identified. Two patterns are considered, completers and dropouts, and no identifying restrictions are needed to fit this model. At first sight, model 2 is similar to the first model but considers completers and dropouts separately. The fixed effects are indeed fitted *by pattern*; however, the random intercept and residual error terms remain shared parameters in the pattern-mixture model. Model 2 takes all discontinued subjects together and assumes a common average evolution *irrespective of* the time of dropout. This may be an over simplification (see Figure 2). However, model 2 has some appealing features. First, it belongs to a different class of models to model 1. Secondly, it is fully identified, so no assumptions about identifying restrictions need to be made. Thirdly, the two patterns are quite balanced (54% completers versus 46% dropouts), so no parameter is estimated based on a small subset of subjects.

The third model is a multiple imputation model based on an assumed multivariate normal distribution of Y_{ij}. While small numbers of imputation are often recommended, it is prudent to use moderate to large numbers (e.g. 50–200), which also helps to stabilize the estimates of variability in the sense that numerically trustworthy quantities are obtained [11]. The parameter variance is higher compared with models 1 and 2, even with as many as 200 imputations (Table 1).

Models 1 and 2 (i.e. the basic selection and pattern-mixture models) do not treat time as a factor. A saturated model was therefore fitted for comparison, through direct likelihood, and the resulting estimand was in the range found for the basic models (estimand saturated model: –1.48; estimand models 1–3: –1.61 to –1.42). Note that, in contrast to models 1 and 2, model 3 *does* rely on a multivariate normal distribution through the use of multiple imputation; in fact, model 3 is a generic saturated model. The model 3 estimand *was* more conservative compared with models 1 and 2 (estimand model 3: –1.42; estimand models 1 and 2: –1.61 and –1.47, respectively). Hence, the time constraints in models 1 and 2 seem to induce a certain amount of bias. The remaining models (models 4–8) help to qualify this bias. The advantage of model 1 is that it is very sparse, which

reduces the complexity (and volatility) of the more advanced pattern-mixture models (models 6–8).

Because we consider change versus baseline, one might wonder about the inclusion of baseline into the model. Of course, it is possible that the change versus baseline itself is a function of baseline. This is allowed for. Comparing models with and without baseline shows a significant difference (given the large sample size), but parameter estimates change in the third decimal place only. Because baseline exhibits low variability in this study, it is not a surprise to see that, while significant, baseline explains little variability.

3.2. Models 4 and 5: alternative specification of the mean structure

Figure 2 shows a decrease (drop) in dissatisfaction followed by a relapse in patients who discontinued participation in the trial. Models 1–3 start from the *overall* evolution of dissatisfaction, and do not take this observation into account. Model 4 is a piecewise linear model allowing the dissatisfaction score to increase again after 3 months of treatment:

$$Y_{ij} = \beta_0 + b_{0i} + \beta_1 Y\_BAS_i + \beta_2 SBM\_BAS_i + \beta_3 t_{1ij} + \beta_4 t_{2ij} + \varepsilon_{ij}$$

where $t_{1ij} = 1$ at baseline and 0 otherwise; and where $t_{2ij} = 1$ at month 3, $t_{2ij} = 2$ at month 6, $t_{2ij} = 3$ at month 9, $t_{2ij} = 4$ at month 12, and 0 otherwise.

Model 4 is a useful way of showing the piecewise linear model *in concept*; however, model 5 is more interesting because the breakpoint at month 3 is fitted for completers and dropouts separately (in line with Figure 2). The assumptions about the treatment effect over time are relaxed and result in a lower estimate of the treatment effect (Table 2).

3.3. Models 6–8: departures from MAR (pattern-mixture models)

Finally, the initial model was fitted for each dropout pattern separately (Table 3). Model 6 is a sparse model; nevertheless no curve can be fitted for the first pattern (where only baseline is observed). The information on the model parameters for the first pattern is

deduced through identifying restrictions, which draw the sensitivity analysis away from MAR assumptions (with the exception of ACMV restrictions).

The pattern-mixture models were fitted using straightforward direct likelihood tools. Therein, the specification of the *pattern x time point* estimates and the specification of the *pattern-averaged* estimates are not straightforward with three or more patterns.

4. CONCLUSIONS

Sensitivity analysis has been found to be useful and feasible in settings broader than those of confirmatory randomized, controlled trials. In this case study, the estimate from a simple paired *t*-test was examined in several ways. The change in dissatisfaction since the first use of prucalopride was originally estimated to be –1.59. The effect size, across the eight sensitivity models, varied between –1.31 and –1.61. These results remained statistically significant, and clinically relevant.

The basic models in this article included a selection model, an identified pattern-mixture model, and multiple imputation combined with a cross-sectional analysis at endpoint. The latter two models provided a simple way of checking the selection model. The multiple imputation model (assuming multivariate normality without parametric time structure) is especially useful when the selection model has a parametric time structure or a specific covariance structure (in other words, when the selection model is more restrictive than the multiple imputation model). The pattern-mixture model was extended to contain one pattern per time of dropout, resulting in an under-identified pattern-mixture model. Estimation involved identifying restrictions (ACMV, NCMV and CCMV), and worked well given the sparse specification of the mean structure. Pattern-mixture models with sophisticated mean structures tend to cause instability [13], which was also observed in this case study. Pattern-mixture models with patterns by reason of dropout, although not explored in this article, provide another viable option for sensitivity analysis. Finally, sensitivity techniques are commonly used in trials with a parallel-group design, where information on missing outcomes can, for example, be distilled from one treatment arm (usually the control arm). This is not possible in a long-term, open-label trial. Nevertheless, a sensitivity focusing on the model specification itself, on the class of

models (selection, pattern-mixture, or multiple imputation model), and on the assumed nature of missing values (MAR or MNAR), remains a worthwhile exercise.

REFERENCES

1. Camilleri M, Kerstens R, Rykx A, Vandeplassche L. A placebo-controlled trial of prucalopride for severe chronic constipation. *New England Journal of Medicine* 2008; **358**:2344–2354.

2. Quigley EM, Vandeplassche L, Kerstens R, Ausma J. Clinical trial: the efficacy, impact on quality of life, and safety and tolerability of prucalopride in severe chronic constipation – a 12-week, randomized, double-blind, placebo-controlled study. *Alimentary Pharmacology and Therapeutics* 2009; **29**:315–328.

3. Tack J, Van Outryve M, Beyens G, Kerstens R, Vandeplassche L. Prucalopride (Resolor) in the treatment of severe chronic constipation in patients dissatisfied with laxatives. *Gut* 2009; **58**:357–365.

4. Marquis P, De La Loge C, Dubois D, McDermott A, Chassany O. Development and validation of the Patient Assessment of Constipation Quality of Life questionnaire. *Scandinavian Journal of Gastroenterology* 2005; **40**:540–551.

5. Dubois D, Gilet H, Viala-Danten M, Tack J. Psychometric performance and clinical meaningfulness of the Patient Assessment of Constipation-Quality of Life questionnaire in prucalopride (RESOLOR) trials for chronic constipation. *Neurogastroenterology and Motility* 2010; **22**:e54–e63.

6. Camilleri M, Van Outryve MJ, Beyens G, Kerstens R, Robinson P, Vandeplassche L. Clinical trial: the efficacy of open-label prucalopride treatment in patients with chronic constipation; follow-up of patients from the pivotal studies. *Alimentary Pharmacology and Therapeutics* 2010; **32**:1113–1123.

7. Rubin DB. Inference and missing data. *Biometrika* 1976; **63**:581–592.

8. Little RJA, Rubin DB. *Statistical Analysis with Missing Data* (2nd edn). Wiley: New York, NY, 2002.

9. National Research Council/Committee on National Statistics/Panel on Handling Missing Data in Clinical Trials. *The Prevention and Treatment of Missing Data in Clinical Trials*. The National Academies Press: Washington, DC, 2010.

10. Kenward MG, Molenberghs G, Thijs H. Pattern-mixture models with proper time dependence. *Biometrika*; **90**:53–71.

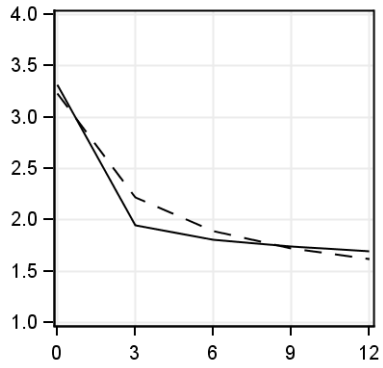
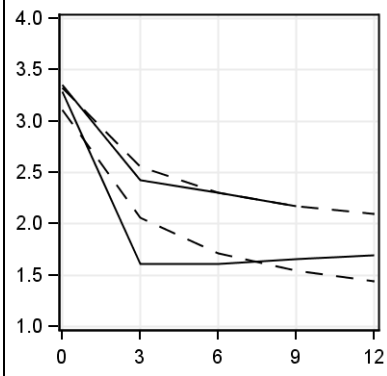
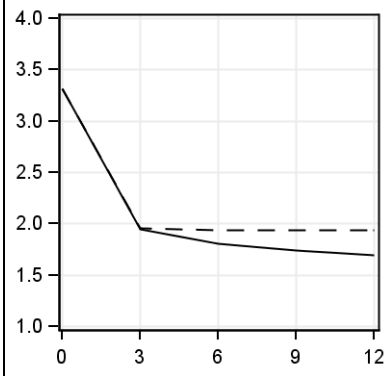
11. Carpenter JR, Kenward MG. *Missing Data in Randomized Controlled Trials: a Practical Guide*. Medical Statistics Unit/London School of Hygiene & Tropical Medicine: London, 2007.

12. Molenberghs G, Michiels B, Kenward MG, Diggle PJ. Missing data mechanisms and pattern-mixture models. *Statistica Neerlandica* 1998; **52**:153–161.

13. Demirtas H. Multiple imputation under Bayesianly smoothed pattern-mixture models for non-ignorable drop-out. *Statistics in Medicine* 2005; **24**:2345–2363.

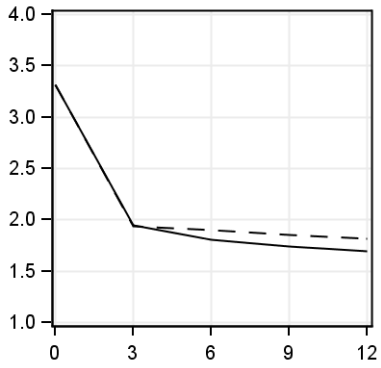
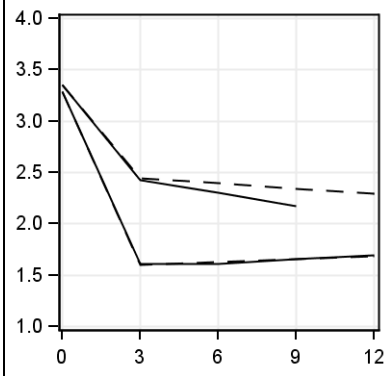
TABLES

Table 1. Details of models 1–3, and the results seen with these models

	Model 1	**Model 2**	**Model 3**
	—— Observed - - - Estimated	*—— Observed - - - Estimated*	*—— Observed - - -Estimated*
	Dissatisfaction (mean)	**Dissatisfaction (mean)**	**Dissatisfaction (mean)**
			
	Months on prucalopride	**Months on prucalopride**	**Months on prucalopride**
Method	Selection model	Pattern-mixture model	Multiple imputation + ANOVA model
MAR/MNAR	MAR	MNAR	MAR
Time structure	1/*t* *with* t = *1, 2, 3,...* *the time point indicator*	1/*t* *i.e. model 1 applied for* *completers and dropouts separately*	Multivariate normal
Month 12 *vs.* baseline			
Mean (SE)	–1.61 (0.030)	–1.47 (0.033)	–1.42 (0.055)
[LCL; UCL]	[–1.67; –1.55]	[–1.53; –1.40]	[–1.53; –1.31]

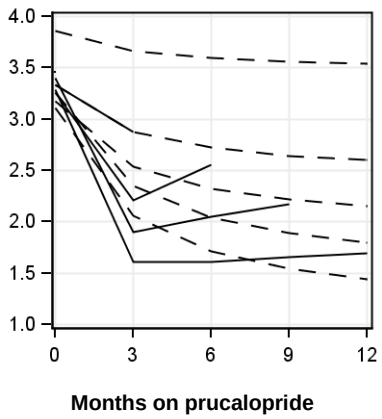
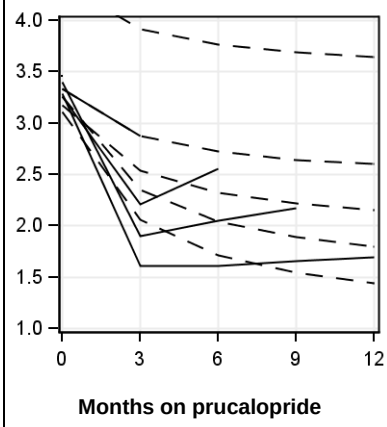
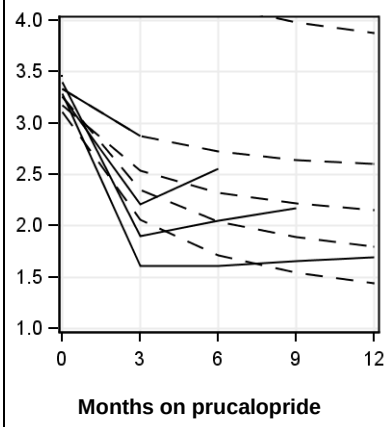
ANOVA, analysis of covariance; LCL, 95% lower confidence limit; MAR, Missing At Random; MNAR, Missing Not At Random; SE, standard error; UCL, 95% upper confidence limit.

Table 2. Details of models 4 and 5, and the results seen with these models

	Model 4 *—— Observed - - - Estimated* **Dissatisfaction (mean)**  **Months on prucalopride**	**Model 5** *—— Observed - - - Estimated* **Dissatisfaction (mean)**  **Months on prucalopride**
Method	Selection model	Pattern-mixture model
MAR/MNAR	MAR	MNAR
Time structure	Piecewise linear *with breakpoint* *at month 3*	Piecewise linear *i.e. model 4 applied for* *completers and dropouts separately*
Month 12 ***vs.* baseline** Mean (SE) [LCL; UCL]	 –1.50 (0.033) [–1.57; –1.44]	 –1.35 (0.050) [–1.44; –1.25]

LCL, 95% lower confidence limit; MAR, Missing At Random; MNAR, Missing Not At Random; SE, standard error; UCL, 95% upper confidence limit.

Table 3. Details of models 6–8, and the results seen with these models

	Model 6	Model 7	Model 8
	Observed - - - Estimated **Dissatisfaction (mean)**  **Months on prucalopride**	*Observed - - - Estimated* **Dissatisfaction (mean)**  **Months on prucalopride**	*Observed - - - Estimated* **Dissatisfaction (mean)**  **Months on prucalopride**
Method	Pattern-mixture model	Pattern-mixture model	Pattern-mixture model
MAR/MNAR	MAR (ACMV restrictions)	MNAR (NCMV restrictions)	MNAR (CCMV restrictions)
Time structure	1/*t* *for each dropout pattern separately*	1/*t* *for each dropout pattern separately*	1/*t* *for each dropout pattern separately*
Month 12 *vs.* baseline Mean (SE) [LCL; UCL]	 –1.31 (0.035) [–1.38; –1.24]	 –1.34 (0.035) [–1.41; –1.28]	 –1.42 (0.034) [–1.48; –1.35]

ACMV, Available Case Missing Value; CCMV, Complete Case Missing Value; LCL, 95% lower confidence limit; MAR, Missing At Random; MNAR, Missing Not At Random; NCMV, Nearest Case Missing Value; SE, standard error; UCL: 95% upper confidence limit.

FIGURE LEGENDS

Figure 1. Observed dissatisfaction at (a) baseline and (b) for up to 12 months of open-label treatment with prucalopride. SD, standard deviation.

Figure 2. Observed evolution of dissatisfaction for (a) all subjects, (b) completers versus dropouts, and (c) completers versus dropout patterns.