STATISTICS IN MEDICINE
Statist. Med. 2005; **24**:2129–2150
Published online 20 May 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/sim.2107

An index of local sensitivity to nonignorable drop-out in longitudinal modelling

Guoguang Ma[1,*,†], Andrea B. Troxel[2] and Daniel F. Heitjan[2]

[1]*Clinical Biostatistics, Merck & Co. Inc., BL3-2, Blue Bell, PA 19422, U.S.A.*
[2]*Department of Biostatistics & Epidemiology, University of Pennsylvania School of Medicine, Philadelphia, PA 19104, U.S.A.*

SUMMARY

In longitudinal studies with potentially nonignorable drop-out, one can assess the likely effect of the nonignorability in a sensitivity analysis. Troxel *et al.* proposed a general *index of sensitivity to nonignorability*, or ISNI, to measure sensitivity of key inferences in a neighbourhood of the ignorable, missing at random (MAR) model. They derived detailed formulas for ISNI in the special case of the generalized linear model with a potentially missing univariate outcome. In this paper, we extend the method to longitudinal modelling. We use a multivariate normal model for the outcomes and a regression model for the drop-out process, allowing missingness probabilities to depend on an unobserved response. The computation is straightforward, and merely involves estimating a mixed-effects model and a selection model for the drop-out, together with some simple arithmetic calculations. We illustrate the method with three examples. Copyright © 2005 John Wiley & Sons, Ltd.

KEY WORDS: ignorability; longitudinal data; missing data; sensitivity analysis; covariance structure

1. INTRODUCTION

Drop-out in longitudinal studies is a pervasive and vexing problem. It commonly occurs in outpatient drug trials, where subjects may discontinue treatment and stop appearing for scheduled study visits after some time. Although we seldom know the precise reasons for drop-out, we often suspect that the mechanism is nonignorable, in the sense that the probability of drop-out is associated with the value of the variable under study. For example, if patients

*Correspondence to: Guoguang Ma, Clinical Biostatistics, Merck & Co. Inc., BL3-2, 10 Sentry Parkway, Blue Bell, PA 19422, U.S.A.
†E-mail: guoguang_ma@merck.com

Contract/grant sponsor: National Institute of Health; contract/grant numbers: HL65365, HL68074, CA79470

Received April 2004
Accepted September 2004

dropout because they believe that the treatment is not helping them, the mechanism is very likely nonignorable.

It is well known that inferences that fail to account for nonignorability can yield invalid results [1, 2]. Although recent years have witnessed the development of several methods for detecting and handling nonignorability, no clear favourite has yet emerged. One option is to investigate nonignorability directly by positing and estimating nonignorable models [3–5]. Unfortunately, these models can be difficult to identify and estimate from real data because one cannot directly assess the dependence of missingness probabilities on observations that are themselves missing. Thus these models require strong assumptions that are largely unverifiable from the data, and invite criticism on that basis. A second option is to undertake global sensitivity analysis, estimating models under a range of assumptions about the nonignorability parameters and assessing the impact of these parameters on key inferences. If there is little effect, the analysis is robust to specification of the missing-data mechanism, and the simpler, ignorable analysis is deemed valid. Rotnitzky *et al.* [6, 7] and Scharfstein *et al.* [8] present methods of this kind that make use of semiparametric models. Their approach avoids some of the uncertainties of direct non-ignorable modelling, but at the expense of deliberately ignoring any information that the outcome model might reveal about the incompleteness mechanism.

Desiring a more practical approach, Troxel *et al.* [9] developed an index of local sensitivity to nonignorability that analysts can calculate from the observed data without estimating a full nonignorable model. The index measures the extent to which small departures from ignorability affect the maximum likelihood estimate of the parameter of interest. The index effects a kind of first-order sensitivity analysis and thus functions as a tool for screening data, allowing the analyst to quickly assess the need for complex nonignorable models at an early stage of the analysis.

Others have introduced similar ideas from slightly different perspectives. Verbeke *et al.* [10] used notions of local influence to generate methods for assessing sensitivity to departures from ignorability in individual observations. Copas and Li [11] proposed an approach in which the error terms from the two components of the model (a normal linear model for the complete data and a probit selection model for the missingness probability) are correlated. The correlation parameter constitutes their nonignorability parameter, and they examined the local sensitivity of various inferences in the neighbourhood of zero correlation. Copas and Eguchi [12] generalized this approach by introducing a sensitivity parameter that measures the dependence between the outcome and the missing data indicator on the log-odds scale. The method of Troxel *et al.* is in the same spirit but is more general in that it is applicable to any combination of parametric complete-data and selection models; moreover, it is easy to compute using only standard modelling software. Troxel *et al.* [9] illustrated their method by applying it to a range of generalized linear models with univariate outcomes.

In this article, we apply local sensitivity analysis to longitudinal data with nonignorable drop-outs. In Section 2, we begin by reviewing the method of Troxel *et al.* [9]. We discuss an extension to longitudinal modelling in Section 3. In Sections 4–6, we illustrate the idea by applying it to three longitudinal data sets. Our analysis will address the questions of which, if any, model-based inferences from these data can be considered sensitive to non-ignorability and whether the sensitivity analysis itself depends in important ways on the specification of the missing-data selection model. Section 7 concludes with a discussion.

2. ISNI AND ITS PROPERTIES

2.1. Definition of ISNI

Troxel *et al.* [9] considered estimation of a parameter θ of the distribution of an outcome variable $Y=(Y_1,\ldots,Y_n)$, whose independent components Y_i have densities $f_\theta^{Y_i\,|\,Z_i}(y_i\,|\,z_i)$, given the fully observed predictors $Z_i=z_i$. The parameter θ is potentially a vector—for example, a vector of regression coefficients θ_1 together with a dispersion parameter θ_2. A completeness indicator G_i takes the value 1 for subjects who are observed and 0 for subjects who are missing. An incomplete-data selection model assumes that the probability of being observed depends on the outcome Y_i and a set of predictors X_i (possibly, but not necessarily, overlapping with Z_i) according to a specified link function, as follows:

$$\Pr_\gamma[G_i=1\,|\,Y_i=y_i,X_i=x_i]=h(x_i\gamma_0+y_i\gamma_1)$$

Common choices for $h(\cdot)$ are the logit and probit.

The *index of sensitivity to nonignorability (ISNI)* measures the extent to which the maximum likelihood estimate (MLE) of θ for a given value γ_1 of the nonignorability parameter (denoted as $\hat\theta(\gamma_1)$) depends on γ_1. Specifically, it measures the sensitivity of $\hat\theta(\gamma_1)$ to small departures of γ_1 from its MAR value of zero. Troxel *et al.* [9] defined ISNI as the derivative of $\hat\theta$ with respect to γ_1 at $\gamma_1=0$, i.e.

$$\text{ISNI}=\left.\frac{\partial\hat\theta(\gamma_1)}{\partial\gamma_1}\right|_{\gamma_1=0}$$

One obtains $\hat\theta(\gamma_1)$ from a Taylor-series expansion of the log-likelihood around $\theta=\hat\theta_0$ (the MLE of θ assuming ignorability), $\gamma_0=\hat\gamma_{00}$ (the MLE of γ_0 assuming ignorability) and $\gamma_1=0$. ISNI describes the approximate change in $\hat\theta$ associated with a unit change in the nonignorability parameter γ_1. A large ISNI implies substantial sensitivity.

Troxel *et al.* [9] presented formulas for ISNI in a range of specific generalized linear models. The estimated dispersion parameter requires no adjustment when $h(\cdot)$ is linear in Y. Sections 2.2–2.4 below summarize some properties of ISNI.

2.2. ISNI for contrast parameters

ISNI inherits a linearity property from the linearity of the differentiation operator:

$$\text{ISNI}_{a^\text{T}\theta_1}=a^\text{T}\text{ISNI}_{\theta_1}$$

for any vector a. For example, consider the comparison of two mean parameters μ_1 and μ_2. Letting $\theta_1=(\mu_1,\mu_2)^\text{T}$ and $a=(1,-1)^\text{T}$, then $a^\text{T}\theta_1=\mu_1-\mu_2$ and

$$\text{ISNI}_{\mu_1-\mu_2}=(1\ \ -1)\begin{pmatrix}\text{ISNI}_{\mu_1}\\ \\ \text{ISNI}_{\mu_2}\end{pmatrix}=\text{ISNI}_{\mu_1}-\text{ISNI}_{\mu_2}$$

Therefore, if ISNI_{μ_1} and ISNI_{μ_2} are equal, no matter how large they are, $\mathrm{ISNI}_{\mu_1-\mu_2}$ equals zero—i.e. the contrast is insensitive to nonignorability.

2.3. The sensitivity transformation

Because ISNI depends on the units of measurement of Y, Troxel *et al.* [9] proposed a scale-free measure called the sensitivity transformation c, defined as

$$c = |\sigma_Y \mathrm{SE}_Y / \mathrm{ISNI}_Y|$$

where σ_Y is the conditional standard deviation (SD.) of Y given Z, SE_Y is the standard error (SE) of a scalar parameter of interest, and ISNI_Y is the ISNI for that parameter. Under a logit link, the model implies that if a change of $1/c$ SD. in Y is associated with a change in the odds of being observed of $e^1 = 2.7$, then the ISNI is equal to 1 SE. Thus, large values of c suggest that sensitivity occurs only in cases of extreme nonignorability, whereas small values suggest that sensitivity may be a problem even when the nonignorability is modest. Troxel *et al.* [9] have suggested using $c < 1$ as a cutoff value for important sensitivity.

2.4. Sensitivity of ISNI to model specifications

Although previous applications of ISNI have used a logistic link, in principle it would be straightforward to use other links, such as the probit. It is not immediately obvious that the well-known near-equivalence of probit and logit models [13] will translate into equivalence in sensitivity analyses as well. To consider how one would execute a sensitivity analysis using a probit model, consider first this result approximating the standard normal cumulative distribution function (cdf) $\Phi(\cdot)$ by the logistic cdf [14]:

$$\Phi(y) \approx \frac{\exp(qy)}{1 + \exp(qy)} \tag{1}$$

where $q = \pi/\sqrt{3} = 1.81$. Therefore, under a probit link, a one-unit increase in Y corresponds to a change in the odds of missingness of $\exp(q) = 6.1$ if $\gamma_1 = 1$. We thus propose a sensitivity transformation c_Φ under a probit link,

$$c_\Phi = |q\sigma_Y \mathrm{SE}_Y / \mathrm{ISNI}_Y|$$

which is comparable to c under a logit link. In addition, from the approximation (1), we have

$$\frac{\phi(y)}{1 - \Phi(y)} \approx q\Phi(y)$$

Thus the ISNI under a probit link is approximately q times that under a logit link.

The choice of predictors in the missing-data selection model is another area of potential sensitivity. As we have indicated, it is not necessary to use the same set of predictors in the complete-data and selection models. Although we commonly think of nonignorability as a dependence of the missing-data mechanism on the values of the missing outcome data, it is often more fruitful to view it as a dependence on one or more predictors that are associated both with the outcome and the missingness, but are absent from the prediction models. Thus one strategy for avoiding nonignorability bias is to employ ignorable models that have large numbers of predictors, as suggested in Reference [15]. In the application to be discussed below,

we will demonstrate the effect on the ISNI statistic of varying the number of predictors in the selection model.

3. ISNI ANALYSIS IN LONGITUDINAL DATA

The idea is straightforward to carry through for longitudinal modelling. Suppose that a study proposes to measure n subjects at m_i times, $i = 1, \ldots, n$. Let $Y_i = (Y_{i1}, \ldots, Y_{im_i})^{\mathrm{T}}$ be a vector of intended measurements for the ith subject. Let $t_i = (t_{i1}, \ldots, t_{im_i})^{\mathrm{T}}$ be the corresponding set of times at which measurements are taken. We assume that Y_i follows a multivariate Gaussian distribution,

$$Y_i \sim \mathrm{MVN}(Z_i\theta_1, \Sigma_i(\theta_2))$$

where Z_i is an $m_i \times p$ predictor matrix for the ith subject, θ_1 is the parameter of primary interest, θ_2 is a nuisance parameter governing the variance of the complete-data model, and Σ_i is a function describing the dependence on θ_2 of the variance–covariance matrix for subject i.

Assume that subject i has been observed at consecutive times $t_1 < \cdots < t_{k_i}$, where $1 \leqslant k_i \leqslant m_i$ identifies the time of the last observation and $k_i = m_i$ identifies no drop-out. Let $z_i^{(k_i)}$ be the first k_i rows of z_i, $y_i^{(k_i)}$ be the first k_i elements of y_i, and $\Sigma_i^{(k_i)}$ be the submatrix of the first k_i rows and columns of Σ_i. Let $f_i(y_i^{(k_i)})$ be the joint probability density function of $y_i^{(k_i)}$,

$$\ln f_i(y_i^{(k_i)}) = -\frac{k_i}{2}\ln(2\pi) - \frac{1}{2}\ln\left|\Sigma_i^{(k_i)}(\theta_2)\right| - \frac{1}{2}(y_i^{(k_i)} - z_i^{(k_i)}\theta_1)^{\mathrm{T}}\Sigma_i^{(k_i)}(\theta_2)^{-1}(y_i^{(k_i)} - z_i^{(k_i)}\theta_1)$$

Under ignorability, the observation Y_{i,k_i+1} at time t_{k_i+1} given the data up to that point $Y_i^{(k_i)}$ follows a conditional normal distribution.

We assume that a subject can dropout with probability depending on the current values of predictors, the previous outcome observation, and the current outcome observation, and the dependence is governed by a set of parameters γ that is distinct from $\theta = (\theta_1,\ \theta_2)^{\mathrm{T}}$. More specifically, let G_{ij} be an indicator for whether the ith subject completes observation j: $G_{ij} = 1$ if the subject completes the observation; $G_{ij} = 0$ if the subject drops out before the observation is made. Let X_i be a set of predictors that possibly, but not necessarily, overlaps with Z_i. The drop-out model is defined as

$$\Pr_\gamma[G_{ij} = 1 \mid X_{ij} = x_{ij}, Y_i = y_i, G_{i,j-1} = 1] = h(x_{ij}\gamma_{01} + y_{i,j-1}\gamma_{02} + y_{ij}\gamma_1)$$

where $j = 2, \ldots, m_i$, and h is a link function, possibly the logit or probit. This model, which is essentially that of Diggle and Kenward [4], is nonignorable if $\gamma_1 \neq 0$. We also assume that any intermittent missing observations are governed by a separate, MAR selection process.

Let L be the log-likelihood for $\theta = (\theta_1, \theta_2), \gamma_0 = (\gamma_{01}, \gamma_{02})$ and γ_1, which can be partitioned as

$$L(\theta, \gamma_0, \gamma_1) = L_1(\theta) + L_2(\gamma_0, \gamma_1) + L_3(\theta, \gamma_0, \gamma_1)$$

where

$$L_1(\theta) = \sum_{i=1}^{n} \ln f_i(y_i^{(k_i)})$$

$$L_2(\gamma_0, \gamma_1) = \sum_{i=1}^{n} \sum_{j=2}^{k_i} \ln h(x_{ij}\gamma_{01} + y_{i,j-1}\gamma_{02} + y_{ij}\gamma_1)$$

$$L_3(\theta, \gamma_0, \gamma_1) = \sum_{i:k_i<m_i} \ln \int \phi_\theta^{Y_{i,k_i+1} \mid Y_i^{(k_i)}}(u \mid y_i^{(k_i)})[1 - h(x_{i,k_i+1}\gamma_{01} + y_{i,k_i}\gamma_{02} + u\gamma_1)]\,\mathrm{d}u$$

and $\phi_\theta^{Y_{i,k_i+1} \mid Y_i^{(k_i)}}$ is the conditional normal density of the observation at time t_{k_i+1} given the data up to that point for subject i. If there is no drop-out, then L_3 evaluates to zero and the data are automatically MAR.

Let $\hat{\theta}_0, \hat{\gamma}_{01}$ and $\hat{\gamma}_{02}$ be the MLEs of θ, γ_{01}, and γ_{02}, respectively, under MAR. Let $h_{ij} = h(x_{ij}\hat{\gamma}_{01} + y_{i,j-1}\hat{\gamma}_{02})$ and $h'_{ij} = h'(x_{ij}\hat{\gamma}_{01} + y_{i,j-1}\hat{\gamma}_{02})$, for $j = 2, \ldots, k_i + 1$. Then,

$$\mathrm{ISNI} = \left.\frac{\partial \hat{\theta}(\gamma_1)}{\partial \gamma_1}\right|_{\gamma_1=0} = -\nabla^2 L_{11}^{-1} \nabla^2 L_{13}$$

where

$$\nabla^2 L_{11} = \left.\frac{\partial^2 L}{\partial\theta\partial\theta^{\mathrm{T}}}\right|_{\hat{\theta}_0, \hat{\gamma}_0, 0} = \sum_{i=1}^{n} \left.\frac{\partial^2 \ln f_i(y_i^{(k_i)})}{\partial\theta\partial\theta^{\mathrm{T}}}\right|_{\hat{\theta}_0}$$

and

$$\nabla^2 L_{13} = \left.\frac{\partial^2 L}{\partial\theta\partial\gamma_1}\right|_{\hat{\theta}_0, \hat{\gamma}_0, 0} = -\sum_{i:k_i<m} \frac{h'_{i,k_i+1}}{1 - h_{i,k_i+1}} \left.\frac{\partial \mathrm{E}_\theta^{Y_{i,k_i+1} \mid Y_i^{(k_i)}}}{\partial\theta}\right|_{\hat{\theta}_0}$$

ISNI can include both fixed effects (θ_1) and variance–covariance components (θ_2), and the properties described in Sections 2.2–2.4 hold equally well in the longitudinal setting. The computation is only slightly more difficult in the mixed model, requiring some additional simple arithmetic calculations. For details see the appendix.

In the following sections, we perform ISNI analyses on three real longitudinal datasets. Although we only present the results of the fixed effects part, which is normally of primary interest, we note here that analyses of variance parameters do not show strong sensitivity ($c > 1$). The S-Plus code and data are available at

`http://www.cceb.upenn.edu/heitjan/isni`.

4. EXAMPLE 1: MILK PROTEIN TRIAL

The milk protein trial example is prominent in the literature of longitudinal modelling [4, 16] (see Figure 1). The data are from a randomized study of the effects of three different diets

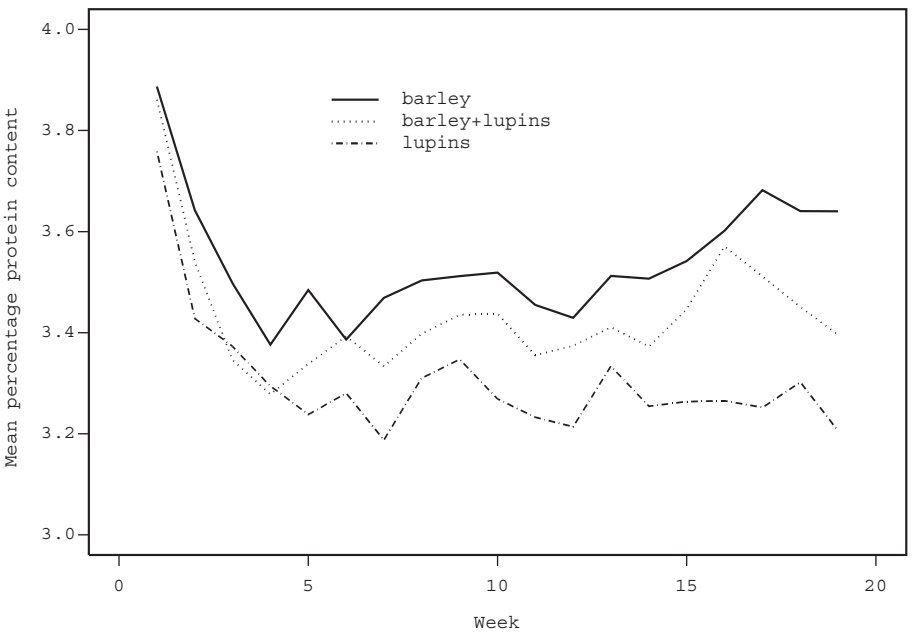

Figure 1. Observed mean response profiles for the milk protein data.

(barley, lupins, and mixed barley–lupins) on milk protein content in dairy cattle. Measurements were taken for up to 19 weeks on 79 cows, with 25, 27, and 27 in the three groups, but 38 (48 per cent) cows stopped producing milk from week 15. Figure 1 shows the mean response profiles.

After fitting several models, Diggle and Kenward [4] concluded that the drop-out is non-ignorable: When the protein content is low, the cows are more likely to stop producing milk and be removed from the study. The question we address here through ISNI analyses is to what extent the estimates of the longitudinal models are affected by nonignorability.

In their paper, Diggle and Kenward assumed that the mean response profiles take the following form:

$$\mu_{g,t} = \begin{cases} \beta_{0g} - t\beta_1, & t < 3 \\ \beta_{0g} - 3\beta_1 + \beta_2(t-3) + \beta_3(t-3)^2, & t \geqslant 3 \end{cases}$$

where $g = 1, 2$ and 3 denotes the treatment groups barley, mixed barley–lupins and lupins, respectively. In this case, $\theta_1 = (\beta_{01}, \beta_{02}, \beta_{03}, \beta_1, \beta_2, \beta_3)$. The covariance structure of the complete measurement sequence Y was defined as

$$\mathrm{Cov}(Y_{ij}, Y_{ik}) = \begin{cases} \sigma^2 + \tau^2, & j = k \\ \sigma^2 \exp[-(j-k)^2/\rho^2], & j \neq k \end{cases}$$

Table I. MLEs under ignorable and nonignorable models and ISNI analysis for the milk protein data.

	Estimates of regression coefficients					
	β_{01}	β_{02}	β_{03}	β_1	$\beta_2 \times 100^*$	$\beta_3 \times 100^*$
Ignorable model	4.16	4.05	3.94	−0.23	0.72	−0.06
SE of coefficients	0.052	0.051	0.051	0.015	0.009	0.0005
ISNI∗1000	0.14	0.40	0.44	−0.11	0.58	−0.08
Sensitivity						
Transformation (c)	111.91	39.25	36.14	44.30	4.76	2.06
Nonignorable model	4.15	4.05	3.93	−0.23	0.51	−0.02
ISNI estimates	4.16	4.05	3.94	−0.23	0.40	−0.01
(Ign + ISNI × $\hat{\gamma}_1$)†						

*The mean estimates of β_2 and β_3 are multiplied by 100.
†Ign stands for estimates under ignorable model; $\hat{\gamma}_1 = -5.65$.

for the ith subject. Finally, the missing-data model assumed that the drop-out probability depends on the current and immediately previous protein content through a logit link h, i.e.

$$\Pr_\gamma[G_l = 1 \mid Y = y^{(l)}, G_{l-1} = 1] = h(\gamma_{0,l-14} + y_{l-1}\gamma_{0,6} + y_l\gamma_1), \quad l = 15, 16, 17, 19$$

They fitted a nonignorable model by estimating γ_1 directly. Our approach is to estimate the models under MAR, or $\gamma_1 = 0$, and calculate the ISNI statistics.

Table I presents the MLEs from Diggle and Kenward's nonignorable model, and results from the ISNI analysis. The sensitivity transformations for the estimates of $\beta_{01}, \beta_{02}, \beta_{03}$, and β_1 are large, which suggests that the estimated treatment effect and the time slope before week 3 are insensitive to nonignorable drop-out. This is not surprising, because the drop-outs occur during the last five weeks of the study, and most of the information about these parameters would come from times where there is no missing data. The transformations for the estimates of β_2 and β_3 are relatively smaller, suggesting that the estimates of the linear and quadratic rise in the mean response toward the end can be potentially affected. Yet because all values exceed one, the influence may be unimportant. In every case, the ISNI analysis agrees with the magnitudes of the changes in the estimates as presented in Table I.

ISNI also correctly predicts the direction of changes in the estimates of the regression coefficients. Diggle and Kenward [4] computed the MLE of the nonignorability parameter to be $\hat{\gamma}_1 = -5.65$. Therefore, we would expect a smaller estimate under the nonignorable model if ISNI is positive, and a larger estimate if ISNI is negative. This is exactly what we see in Table I.

We can approximate the estimates under the nonignorable model by simply adding the MAR estimates to the product of $\hat{\gamma}_1$ and ISNI. For the first four parameters, the ISNI estimates and the estimates from the full model are close, whereas the estimates of β_2 or β_3 are appreciably different, suggesting that a linear approximation to $\hat{\theta}(\gamma_1)$ is inadequate globally.

5. EXAMPLE 2: DRUG TREATMENT IN PSYCHIATRY

The data are from a 12-week randomized trial of the tricyclic antidepressant desipramine in outpatients with current diagnosis of both cocaine dependence and a depressive disorder. The study hypothesizes that desipramine is superior to placebo in reducing the symptoms of depression, and consequently reduces cocaine use. One hundred and eleven participants were randomized to either desipramine (DMI) or placebo (PBO). Five subjects dropped out at week one. Therefore, we consider the 106 patients (DMI = 52; PBO = 54) who had at least one observation: 85 (DMI = 43, 83 per cent; PBO = 42, 78 per cent) were observed for four or more weeks; whereas only 47 (DMI = 25, 48 per cent; PBO = 22, 41 per cent) completed 12 weeks. We are interested in looking at the drug effect on one of the outcome measures, average dollars per day spent on cocaine use. Every patient in the analysis has a baseline value for this variable at week 0. Figure 2 plots the raw response profiles, roughly showing that the treatment group begins with a higher baseline on average and reduces its cocaine use more, compared to the placebo group.

To illustrate the formula, $\mathrm{ISNI}_{a^{\mathrm{T}}\theta_1} = a^{\mathrm{T}}\mathrm{ISNI}_{\theta_1}$ in Section 2, we fit a model with two parameterizations. We first estimated the treatment contrasts adjusted by baseline values, and set $z = (\mathrm{Intercept}, \mathrm{Group}, \mathrm{Time}, \mathrm{Group} \times \mathrm{Time}, \mathrm{Baseline})$, where Group equals 1 for DMI and 0 for PBO and Time is the corresponding vector of time points at which the outcome values are measured. Equivalently, we use a parameterization with $z = (\mathrm{Int(PBO)}, \mathrm{Int(DMI)}, \mathrm{Time(PBO)},$

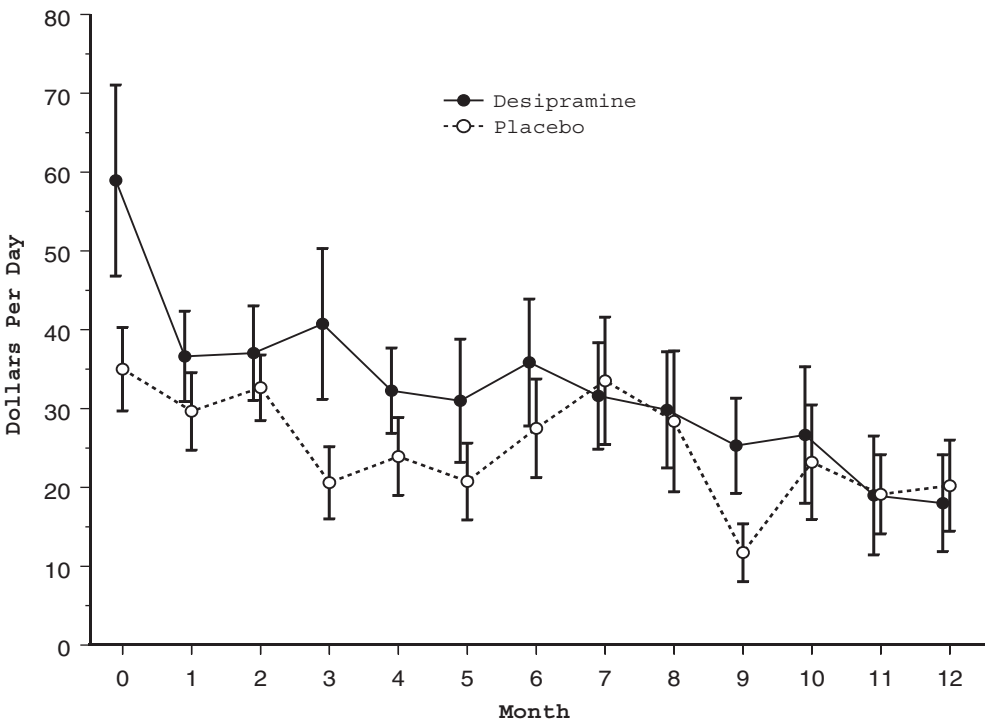

Figure 2. Observed mean(SE) profiles for the cocaine data.

Table II. MLEs under an ignorable model and ISNI analysis for the cocaine data.

Predictor	Ignorable model		ISNI	c
	Estimate	SE		
Parameterization I				
Intercept	24.26	4.59	−46.75	3.85
Group	8.96	6.26	24.02	10.21
Time	−0.17	0.49	−17.20	1.12
Group × Time	−1.36	0.68	2.19	12.19
Baseline	0.13	0.04	0.01	191.49
Parameterization II				
Int(PBO)	24.26	4.59	−46.75	3.85
Int(DMI)	33.21	4.95	−22.73	8.52
Time(PBO)	−0.17	0.49	−17.20	1.12
Time(DMI)	−1.52	0.47	−15.00	1.23
Baseline	0.13	0.04	0.01	191.49

Time(DMI), Baseline), to estimate the intercept and slope of each group separately. Our particular interest is to see how much the estimates of the coefficients under different parameterizations will be influenced by the drop-outs, and how the sensitivities are related to each other.

We assume a compound symmetry structure for the variance–covariance matrix, i.e.

$$\mathrm{Cov}(Y_{ij}, Y_{ik}) = \begin{cases} \sigma^2, & j = k \\ \sigma^2\rho, & j \neq k \end{cases}$$

for subject i. We assume a logistic model h for the drop-out process which depends on z_{ij}, the current outcome y_{ij}, and the immediately previous outcome $y_{i,j-1}$,

$$\Pr_\gamma[G_{ij} = 1 \mid Z_{ij} = z_{ij}, Y_i = y_i, G_{ij-1} = 1] = h(z_{ij}\gamma_{01} + y_{i,j-1}\gamma_{02} + y_{ij}\gamma_1)$$

for $i = 1, \ldots, 106$ and $j = 2, \ldots, 12$. Estimates under an MAR model and ISNI statistics are given in Table II.

In parameterization I, the sensitivity transformations for Group and Group × Time are 10.21 and 12.19, respectively, suggesting that the treatment contrasts are insensitive to nonignorability. In parameterization II, a small c for the intercept of the placebo group suggests modest sensitivity; however, the intercept of the desipramine group is again hardly affected. The slope estimates can be potentially sensitive, with c of 1.12 and 1.23, respectively. The estimate of the baseline coefficient is least likely to be affected, with a large c of 191.49.

Furthermore, we observe that

$$\mathrm{ISNI}_{\mathrm{Group}} = \mathrm{ISNI}_{\mathrm{Int(DMI)}} - \mathrm{ISNI}_{\mathrm{Int(PBO)}}$$

and

$$\mathrm{ISNI}_{\mathrm{Group}\times\mathrm{Time}} = \mathrm{ISNI}_{\mathrm{Time(DMI)}} - \mathrm{ISNI}_{\mathrm{Time(PBO)}}$$

which verifies the equation $\mathrm{ISNI}_{a^{\mathrm{T}}\theta_1} = a^{\mathrm{T}}\mathrm{ISNI}_{\theta_1}$. Therefore, the contrasts are less influenced when the linear parameters are affected in the same directions and to comparable degrees.

For this example, the ISNI analysis gives $c = 12.19$ for the slope comparison between groups, which means that the contrast would be insensitive unless a change of $1/12.19$ SDs is associated with an odds ratio of 2.7 in the observation probability. Thus unless one expects such strong nonignorability, the MAR analysis is credible for this application. If we want to look at the group parameters—for example, the slope of each group—we may have to undertake more elaborate modelling, however, because these parameters are potentially sensitive to nonignorable drop-out with c values slightly over 1. Note that the sensitivity transformation does not directly assess the impact of nonignorability on the significance of treatment effects, but rather the magnitude of the potential change in the MLE relative to its standard error under MAR.

6. EXAMPLE 3: THE SWOG 9039 QoL DATA

Quality-of-life (QoL) data in clinical trials are often subject to various kinds of missingness. QoL is typically a secondary outcome, especially in chronic diseases like cancer and AIDS where the focus of data collection is on survival and objective indicators of disease status. Moreover, even patients who attend all scheduled clinic visits may refuse to complete QoL questionnaires, particularly if they are feeling poorly. If this happens, missing observations are more likely to occur in subjects who would have given lower scores on the QoL assessment instruments, leading to nonignorably missing data.

Southwest Oncology Group (SWOG) study S9039 was a companion study to INT-0105, an intergroup, randomized, double-blind, placebo-controlled trial in patients with stage M_1 carcinoma of the prostate. Patients were treated with bilateral orchiectomy and randomized to receive either flutamide or a placebo. The trial accrued 1387 patients, and the survival comparison was not statistically significant [17].

Of 737 eligible patients who also enrolled in the QoL component of the study, 370 were randomized to flutamide (fl) and 367 to placebo (pb). The SWOG QoL questionnaire was to be administered at randomization and at 1, 3, and 6 months after randomization. Our analysis will concentrate on two primary QoL domains: emotional functioning (EF) and physical functioning (PF). The scales range from 0 to 100, with higher scores reflecting better functioning. Cross-sectional and longitudinal analyses identified statistically significant mean differences favouring the placebo arm for EF at 3 and 6 months [18].

Ninety-eight per cent of the subjects submitted baseline questionnaires; submission rates at 1, 3, and 6 months were 88, 86, and 81 per cent, respectively. Missing items within a questionnaire were handled as follows: when subjects failed to answer more than 20 per cent of the items, the total score was considered missing; when patients failed to answer fewer than 20 per cent of the items, the total score was calculated as the mean of the remaining answered items. Fewer than 11 per cent of the subjects had one or two intermittent missing observations. The drop-out rates were similar in the treatment arms (Table III). For each outcome, roughly 6 per cent of remaining subjects dropped out at months 1 and 3, and about double that fraction dropped out at month 6.

Table III. Drop-out rates by time and treatment group.

Group	Month 1	3	6
Emotional functioning			
Placebo($N = 359$)	20 (5.6 per cent)	20 (5.6 per cent)	45 (12.5 per cent)
Flutamide($N = 366$)	22 (6.0 per cent)	15 (4.1 per cent)	38 (10.4 per cent)
Physical functioning			
Placebo($N = 360$)	21 (5.8 per cent)	19 (5.3 per cent)	45 (12.5 per cent)
Flutamide($N = 366$)	23 (6.3 per cent)	14 (3.8 per cent)	42 (11.5 per cent)

Note: Drop-outs at Month i are defined as patients who had monotonely missing scores at Month $\geqslant i$, $i = 1, 3$, and 6; (per cent) = number of drop-outs at Month i/total number of patients in each group.

6.1. The complete-data model

Our complete-data models describe the dependence of EF and PF on treatment arm and time, with adjustment for the baseline predictors performance status (`perf`) and disease severity (`sever`). Preliminary analyses suggested that for both outcomes a multivariate Gaussian model would be appropriate following a square root transformation. Figure 3 suggests that the response profiles are neither linear nor quadratic in time, so our models incorporate time as a categorical factor.

To illustrate the dependence of the sensitivity on the parameterization, we estimated the same model under three parameterizations. In parameterization PI,

$$z = (\mathrm{T0(pb)}, \mathrm{T1(pb)}, \mathrm{T3(pb)}, \mathrm{T6(pb)}, \mathrm{T0(fl)}, \mathrm{T1(fl)}, \mathrm{T3(fl)}, \mathrm{T6(fl)}, \texttt{perf}, \texttt{sever})$$

that is, the first eight parameters represent the means of the outcome variable at the eight combinations of treatment and time, with the final two parameters representing adjustments for the baseline predictors. In parameterization PII,

$$\begin{aligned} z = (&\mathrm{T0(pb)}, \mathrm{T1.0(pb)}, \mathrm{T3.0(pb)}, \mathrm{T6.0(pb)}, \\ &\mathrm{T0(fl)}, \mathrm{T1.0(fl)}, \mathrm{T3.0(fl)}, \mathrm{T6.0(fl)}, \\ &\texttt{perf}, \texttt{sever}) \end{aligned}$$

by which we mean that the first parameter is the baseline mean, the second is the difference between month 1 and baseline, and so on. Thus parameters 2–4 are the time effects for placebo and parameters 6–8 are the time effects for flutamide. In parameterization PIII,

$$\begin{aligned} z = (&\mathrm{T0(pb)}, \mathrm{T1.0(pb)}, \mathrm{T3.0(pb)}, \mathrm{T6.0(pb)}, \\ &\mathrm{T0(fl)\_T0(pb)}, \mathrm{T1.0(fl)\_T1.0(pb)}, \mathrm{T3.0(fl)\_T3.0(pb)}, \mathrm{T6.0(fl)\_T6.0(pb)}, \\ &\texttt{perf}, \texttt{sever}) \end{aligned}$$

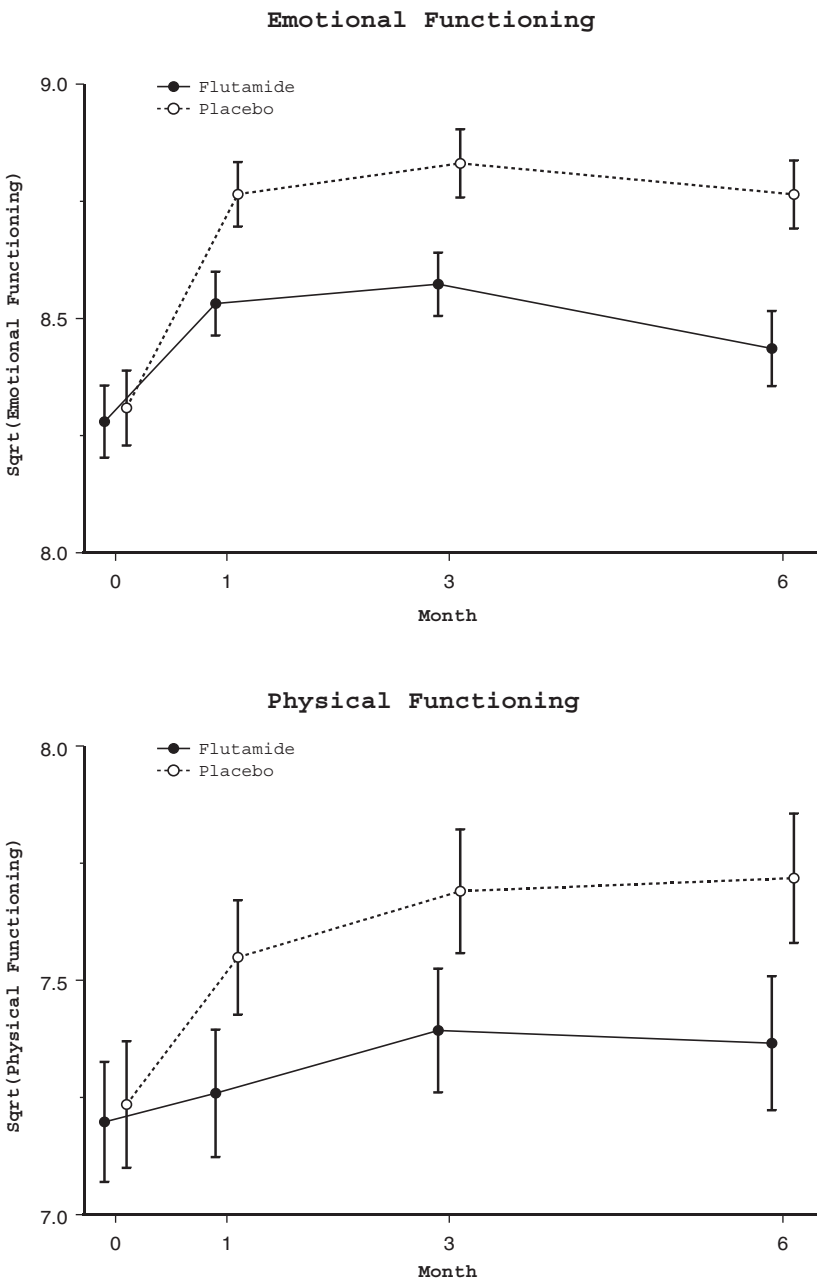

Figure 3. Observed mean(SE) profiles for SWOG 9039 data.

by which we mean that the first four coefficients are placebo time effects as in PII, and 5–8 are the differences between the baseline and time effect parameters for flutamide as compared to placebo.

Our basic complete-data longitudinal model assumes a general symmetry structure with no random effects:

$$\mathrm{Cov}(Y_{ij}, Y_{ik}) = \begin{cases} \sigma^2, & j = k \\ \sigma^2 \rho_{jk}, & j \neq k \end{cases}$$

for all subjects and $\rho_{jk} = \rho_{kj}$ for $j \neq k$, where $j, k = 0, 1, 2$, and 3 correspond to baseline, 1, 3, and 6 months after randomization.

6.2. The drop-out selection model

No subject missed the baseline value, so $G_{i0} = 1$ for all i. We assume the following missing-data model, which is essentially that of Diggle and Kenward [4],

$$\Pr_\gamma[G_{ij} = 1 \mid Y_i = y_i, G_{i,j-1} = 1] = h(\gamma_{0j} + y_{i,j-1}\gamma_{04} + y_{ij}\gamma_1) \tag{2}$$

for $j = 1, 2, 3$. That is, we assume a separate drop-out probability at each time, potentially modified by the last previous observation and the next potential observation. The drop-out is ignorable if $\gamma_1 = 0$.

6.3. Sensitivity analysis of the MAR parameter estimates

We first estimated the ignorable model and calculated ISNI statistics under each of the three parameterizations assuming a logistic link in the selection model. Table IV presents the ML estimates of the longitudinal model for EF together with their ISNI values and sensitivity transformations for the three parameterizations.

Recall that parameterization PI estimates a separate mean for each combination of time and treatment group, after adjusting for baseline performance status and disease severity. Because the fraction of drop-outs increases monotonically from baseline to month 6, the ISNI values increase—and the sensitivity transformations decline—as one proceeds from the mean at baseline to the mean at 6 months, for both placebo and flutamide. By month 6 in the placebo group, the mean root EF has $c = 0.93$, which we interpret as meaning that if a $1/0.93 = 1.08$ SD change in root EF is associated with a change in the odds of being observed of $e^1 = 2.7$, the estimated mean under MAR will be off by one standard error. That is, a strong but not outlandish degree of nonignorability would change this MLE by one SE or more, suggesting that there is cause for concern about nonignorability in the estimation of this parameter.

In parameterization PII, we examine the differences in mean score from baseline, again adjusted for the predictors. Here the sensitivities of the time effects are generally greater than they were in PI, because the ISNI values for baseline and subsequent times are of opposite sign. The c values are less than 1 for the month 6 versus baseline comparisons in both arms. These suspect comparisons are statistically significant, or nearly so. Together these facts suggest that there are big differences from baseline but that their estimates may be spoiled if the nonignorability is strong. The negative ISNI implies that the MLEs are too large if $\gamma_1 > 0$—i.e. if the propensity to give complete data increases with the value of EF. Because this is the type of nonignorability that we expect to encounter in this study, we are

Table IV. Parameter estimates and ISNI analysis for EF—logit link.

Predictor	Estimate	SE	ISNI	c
Parameterization PI				
T0(pb)	8.44	0.10	0.03	4.81
T1(pb)	8.89	0.11	−0.05	2.88
T3(pb)	8.90	0.11	−0.07	2.05
T6(pb)	8.79	0.11	−0.16	0.93
T0(fl)	8.42	0.11	0.03	4.93
T1(fl)	8.65	0.11	−0.05	2.82
T3(fl)	8.69	0.11	−0.06	2.58
T6(fl)	8.56	0.11	−0.13	1.21
`perf`	−0.29	0.22	−0.12	2.45
`sever`	−0.16	0.10	−0.03	4.28
Parameterization PII				
T0(pb)	8.44	0.10	0.03	4.81
T1.0(pb)	0.45	0.07	−0.08	1.28
T3.0(pb)	0.47	0.08	−0.10	1.06
T6.0(pb)	0.36	0.08	−0.19	0.59
T0(fl)	8.42	0.11	0.03	4.93
T1.0(fl)	0.23	0.07	−0.08	1.19
T3.0(fl)	0.27	0.08	−0.09	1.17
T6.0(fl)	0.14	0.08	−0.16	0.69
`perf`	−0.29	0.22	−0.12	2.45
`sever`	−0.16	0.10	−0.03	4.28
Parameterization PIII				
T0(pb)	8.44	0.10	0.03	4.81
T1.0(pb)	0.45	0.07	−0.08	1.28
T3.0(pb)	0.47	0.08	−0.10	1.06
T6.0(pb)	0.36	0.08	−0.19	0.59
T0(fl)_T0(pb)	−0.02	0.10	0.001	217.70
T1.0(fl)_T1.0(pb)	−0.23	0.10	−0.004	32.00
T3.0(fl)_T3.0(pb)	−0.19	0.11	0.01	13.37
T6.0(fl)_T6.0(pb)	−0.22	0.11	0.03	5.10
`perf`	−0.29	0.22	−0.12	2.45
`sever`	−0.16	0.10	−0.03	4.28

concerned that the within-group MLEs overestimate the effects of both placebo and flutamide treatment over time. Moreover the size of ISNI is such that a value of $\gamma_1 = 1$ (i.e. a one-unit change in root EF multiplies the odds of staying on study by 2.7) would be sufficient to nullify the treatment effect at 6 months.

The sensitivities for the placebo parameters in PIII are the same as in PII because these parameters are the same. The sensitivities for the flutamide–placebo contrasts are smaller, however, because the ISNI values are much the same in the two groups and consequently the ISNIs for the contrasts of these time effects are nearly zero. Thus although the month 6 contrasts from baseline are suspect in both arms ($c = 0.59$ for placebo and $c = 0.69$ for flutamide), the contrast between them is resistant to nonignorability ($c = 5.10$).

Table V. Parameter estimates and ISNI analysis for EF.

			Probit link		Intercept-only selection model	
Predictor	Estimate	SE	ISNI	c_Φ	ISNI	c
T0(pb)	8.44	0.10	0.06	4.22	0.03	4.76
T1(pb)	8.89	0.11	−0.11	2.42	−0.05	3.00
T3(pb)	8.90	0.11	−0.15	1.74	−0.07	2.11
T6(pb)	8.79	0.11	−0.30	0.87	−0.16	0.89
T0(fl)	8.42	0.11	0.06	4.32	0.03	4.88
T1(fl)	8.65	0.11	−0.11	2.41	−0.05	2.93
T3(fl)	8.69	0.11	−0.12	2.17	−0.06	2.70
T6(fl)	8.56	0.11	−0.24	1.13	−0.13	1.17
`perf`	−0.29	0.22	−0.25	2.17	−0.12	2.45
`sever`	−0.16	0.10	−0.07	3.75	−0.03	4.22

6.4. *Sensitivity of ISNI to model specification*

To assess the effect of the selection-model link function on the sensitivity analysis, we repeated the ISNI analysis assuming a probit form for $h(\cdot)$ (Table V). Compared to Table IV, the signs of the ISNI values remain the same but the magnitudes are about twice as large as before. The adjusted sensitivity transformations c_Φ are consistently 10–20 per cent smaller than the c values generated under the logit model. The difference is not enough to affect our conclusions about the analysis, and supports the hypothesis that probit and logit links will give similar results.

To assess the sensitivity of the analysis to different sets of linear predictors in the selection model, we calculated ISNI statistics under a logistic model that included only a single common intercept term (Table V). The logistic regression gives $h_i = n_o/n$ for all subjects, where n is the total number of potential observations and n_o is the number of observed values. Although this missing-data model fits significantly worse than model (2), with a $1098 - 1059 = 39$ increase in deviance on 3 df, nevertheless the differences in ISNI and c from Table IV are negligible.

6.5. *Analysis of physical functioning*

We present analyses for the PF outcome in Table VI. Because the designs and models for the two outcomes are identical, and the drop-out patterns are nearly so, any differences in the sensitivity analyses mainly reflect differences in the MAR parameter estimates. One obvious difference is that the $\hat{\sigma}$ value is larger in the PF analysis, as suggested by comparing the SE columns of Tables IV and VI (the actual values are $\hat{\sigma} = 1.35$ for EF and $\hat{\sigma} = 2.35$ for PF). Recall that ISNI equals the matrix of second partials of the log-likelihood with respect to θ and γ_1 times the observed-information-based variance–covariance matrix of θ. Because the latter term is roughly proportional to σ^2, we expect ISNI to also be larger when the variability is greater. This is in fact what we observe in the data. Because c is $\hat{\sigma}$ times the SE divided by ISNI, this effect mainly vanishes in the sensitivity transformation; thus we see that the c values for PF are generally larger but in most cases only slightly so.

A second prominent difference is the negligible effect of flutamide on PF across time; unlike EF, the flutamide time curve is nearly flat and there are no significant differences from

Table VI. Parameter estimates and ISNI analysis for PF—logit link.

Predictor	Estimate	SE	ISNI	c
Parameterization PI				
T0(pb)	7.85	0.18	0.07	6.21
T1(pb)	8.10	0.19	−0.12	3.56
T3(pb)	8.22	0.19	−0.18	2.46
T6(pb)	8.09	0.19	−0.42	1.07
T0(fl)	7.85	0.19	0.07	6.27
T1(fl)	7.87	0.20	−0.13	3.64
T3(fl)	8.05	0.20	−0.15	3.02
T6(fl)	7.90	0.20	−0.37	1.27
`perf`	−2.35	0.40	−0.28	3.28
`sever`	−0.70	0.18	−0.08	5.47
Parameterization PII				
T0(pb)	7.85	0.18	0.07	6.20
T1.0(pb)	0.25	0.11	−0.19	1.31
T3.0(pb)	0.36	0.12	−0.25	1.14
T6.0(pb)	0.23	0.13	−0.49	0.63
T0(fl)	7.85	0.19	0.07	6.27
T1.0(fl)	0.02	0.11	−0.20	1.25
T3.0(fl)	0.20	0.12	−0.23	1.24
T6.0(fl)	0.06	0.13	−0.44	0.68
`perf`	−2.35	0.40	−0.28	3.28
`sever`	−0.70	0.18	−0.08	5.47
Parameterization PIII				
T0(pb)	7.85	0.18	0.07	6.21
T1.0(pb)	0.25	0.11	−0.19	1.31
T3.0(pb)	0.36	0.12	−0.25	1.14
T6.0(pb)	0.23	0.13	−0.49	0.63
T0(fl)_T0(pb)	−0.01	0.18	0.003	140.24
T1.0(fl)_T1.0(pb)	−0.22	0.15	−0.01	49.15
T3.0(fl)_T3.0(pb)	−0.16	0.17	0.02	18.07
T6.0(fl)_T6.0(pb)	−0.18	0.18	0.04	9.55
`perf`	−2.35	0.40	−0.28	3.28
`sever`	−0.70	0.18	−0.08	5.47

baseline in this arm. As in the EF analysis, means at month 6 are potentially sensitive, as are differences from baseline at the later times, but between-group differences in time effects are insensitive. Because ISNI values are larger than they were for EF, the general pattern for PF is one of increased potential sensitivity; that is, the numerical values of the estimates are subject to considerable changes if nonignorability is present.

6.6. *Consequences for the QoL analysis*

Our ISNI analysis suggests that sensitivity to nonignorability in these models follows a common pattern, in that the month 6 means are relatively sensitive, as are contrasts between

month 6 and baseline. Between-group contrasts are insensitive, even when the components being compared are themselves sensitive.

Although we never actually estimated a nonignorable model, implicitly there are two such models in our analysis: one where the drop-out depends on EF, and another where drop-out depends on PF. If the EF model is correct, then the observed flutamide effect on EF could be explainable as an artifact of nonignorability; that is, the subjects with lower EF scores avoided the study questionnaires, thereby biasing mean EF estimates. The direction of the potential bias is the same in the PF scores, but the fact that the flutamide effect is minimal suggests that either PF is unrelated to drop-out or that PF affects drop-out but that average PF scores are declining over time; we do not see this in the observed data because the nonignorability artificially lifts the means.

7. DISCUSSION

We have applied local sensitivity analysis using ISNI statistics to longitudinal modelling with potentially nonignorable drop-outs. Our model assumes a joint Gaussian distribution for the response and a regression model for the drop-out process, and allows drop-out to depend on fully observed predictors and current and immediately previous responses. In our applications, the means and time effects within groups were potentially quite sensitive, especially at later times when more subjects had dropped out. Contrasts between groups, on the other hand, were generally insensitive to nonignorability. We anticipate that these features will hold in other studies having similar designs and outcome variables.

The choice of factors in the selection model had little effect on the sensitivity measures, even when these factors contributed significantly to the fit of the model. Experience suggests that selection-model factors can have a substantial impact on ISNI, but only if their predictive power is very strong. For example, we have observed that the presence of a perfect baseline predictor of drop-out sets all ISNI coefficients to zero [9], essentially because there is no further scope for the missing variable to improve predictions. As our QoL example shows, statistical significance alone is not sufficient to influence ISNI substantially. Nevertheless, in applied work it seems prudent to include as many potential predictors of drop-out as is practical.

Assumptions about the missingness model can affect the sensitivity analysis. If the 'true' missingness mechanism is not monotone in y, as is the case with self-reported income data, our analysis assuming monotonicity may fail to reveal the true degree of sensitivity. However, with QoL and most other kinds of data one expects that nonignorability, if it exists, is in the form of a monotone relationship between y and the probability that y is missing [19]. In our examples the probit and logit link functions gave similar results. This is not surprising, because both functions correspond to symmetric, unimodal distributions, and they are known to give similar predictions in samples of practical size. We have not yet explored the use of asymmetric link functions such as the complementary log–log. Yet our results so far confirm those of Kenward [20], who found that the form of the complete-data model is a more potent source of sensitivity than the form of the selection model.

We emphasize again that our approach differs from both full parametric modelling and semiparametric modelling in that we adopt a provisional nonignorable model but attempt neither to estimate it nor to execute a global sensitivity analysis. Instead, we view ISNI as a

practical tool to screen data sets for sensitivity to nonignorability. A useful analogy is with the fraction of missing data in a survey. When the fraction is large, the potential bias is substantial and the results have poor credibility, whereas when the fraction is small the selection, even if biased, is likely to have only a small effect on estimates. Similarly with ISNI.

Unlike methods that estimate the full nonignorable model or that effect a global sensitivity analysis, our approach measures sensitivity only in the neighborhood of the MAR model. We acknowledge that it is possible for the $\hat{\theta}(\gamma_1)$ curve to be flat near $\gamma_1 = 0$ but steep elsewhere, a circumstance that would in principle defeat our method. Nevertheless we believe that our emphasis on local sensitivity has more to recommend it than just computational simplicity. Unlike most of the common statistical models, with nonignorable models there is essentially never sufficient information in the data to robustly assess fit. Thus if one found that a nonignorable model fits better than an ignorable model (e.g. by a likelihood-ratio test), it is not clear that the nonignorable model should be considered superior in any practical sense. Moreover if a global sensitivity analysis identified strong sensitivity in an area remote from the MAR model, it would be difficult to know what to do about it, as the finding itself may be sensitive to aspects of the model that are not subject to test. Thus we continue to view nonignorable modelling as a speculative enterprise, and hold a strong predilection for MAR modelling even when the proportion of missing data is large. This view motivates our strategy, which is to estimate the MAR model and check locally to see whether admitting the possibility of nonignorability would affect our estimate. If there is sensitivity locally, then global sensitivity is moot. If there is no sensitivity locally, then global sensitivity is unlikely to be of interest. In our opinion, modest sensitivity at the MAR model (i.e. a small ISNI) trumps strong sensitivity at some distance from the MAR model, unless there is compelling prior information to suggest that nonignorability is extreme.

ISNI analysis addresses sensitivity of parameter estimates compared to their standard errors under the MAR model. To assess sensitivity of statistical significance would require joint evaluation of the sensitivities of the estimates and their SEs. Further research will be needed to address this aim. Ma and Heitjan [21] have developed ISNI-type measures to evaluate the effects of non-MCAR sampling on test power and interval coverage probability. These approaches are primarily valuable in study design.

APPENDIX A: CALCULATING ISNI IN THE MIXED MODEL

In Section 3, we derived the ISNI formula for longitudinal modelling, assuming a drop-out model,

$$\Pr_\gamma[G_{ij} = 1 \mid X = x_i, Y = y_i, G_{i,j-1} = 1] = h(x_{ij}\gamma_{01} + y_{i,j-1}\gamma_{02} + y_{ij}\gamma_1)$$

where $j = 2, \ldots, m_i$, and a normally distributed complete-data model, with the density $f_i(y_i^{(k_i)})$, where

$$\ln f_i(y_i^{(k_i)}) = -\frac{k_i}{2}\ln(2\pi) - \frac{1}{2}\ln|\Sigma_i^{(k_i)}(\theta_2)| - \frac{1}{2}(y_i^{(k_i)} - z_i^{(k_i)}\theta_1)^{\mathrm{T}}\Sigma_i^{(k_i)}(\theta_2)^{-1}(y_i^{(k_i)} - z_i^{(k_i)}\theta_1)$$

Under these assumptions, we have

$$\mathrm{ISNI} = \left.\frac{\partial\hat{\theta}(\gamma_1)}{\partial\gamma_1}\right|_{\gamma_1=0} = -\nabla^2 L_{11}^{-1}\nabla^2 L_{13}$$

where

$$\nabla^2 L_{11} = \left.\frac{\partial^2 L}{\partial\theta\partial\theta^{\mathrm{T}}}\right|_{\hat{\theta}_0,\hat{\gamma}_0,0} = \sum_{i=1}^{n}\left.\frac{\partial^2 \ln f_i(y_i^{(k_i)})}{\partial\theta\partial\theta^{\mathrm{T}}}\right|_{\hat{\theta}_0}$$

$$\nabla^2 L_{13} = \left.\frac{\partial^2 L}{\partial\theta\partial\gamma_1}\right|_{\hat{\theta}_0,\hat{\gamma}_0,0} = -\sum_{i:k_i<m_i}\frac{h'_{i,k_i+1}}{1-h_{i,k_i+1}}\left.\frac{\partial \mathrm{E}_\theta^{Y_{i,k_i+1}\mid Y_i^{(k_i)}}}{\partial\theta}\right|_{\hat{\theta}_0}$$

In more detail,

$$\frac{\partial^2 \ln f_i(y_i^{(k_i)})}{\partial\theta\partial\theta^{\mathrm{T}}} = \begin{pmatrix} \dfrac{\partial^2 \ln f_i(y_i^{(k_i)})}{\partial\theta_1\partial\theta_1^{\mathrm{T}}} & \dfrac{\partial^2 \ln f_i(y_i^{(k_i)})}{\partial\theta_1\partial\theta_2^{\mathrm{T}}} \\ \dfrac{\partial^2 \ln f_i(y_i^{(k_i)})}{\partial\theta_2\partial\theta_1^{\mathrm{T}}} & \dfrac{\partial^2 \ln f_i(y_i^{(k_i)})}{\partial\theta_2\partial\theta_2^{\mathrm{T}}} \end{pmatrix}$$

and

$$\frac{\partial^2 \ln f_i(y_i^{(k_i)})}{\partial\theta_1\partial\theta_1^{\mathrm{T}}} = 2z_i^{(k_i)\mathrm{T}}\Sigma_i^{(k_i)-1}z_i^{(k_i)} \tag{A1}$$

$$\frac{\partial^2 \ln f_i(y_i^{(k_i)})}{\partial\theta_1\partial\theta_2^{\mathrm{T}}} = \frac{\partial^2 \ln f_i(y_i^{(k_i)})}{\partial\theta_2\partial\theta_1^{\mathrm{T}}} = 2z_i^{(k_i)\mathrm{T}}\Sigma_i^{(k_i)-1}\frac{\partial\Sigma_i^{(k_i)}}{\partial\theta_2}\Sigma_i^{(k_i)-1}(y_i^{(k_i)} - z_i^{(k_i)}\theta_1) \tag{A2}$$

$$\frac{\partial^2 \ln f_i(y_i^{(k_i)})}{\partial\theta_2\partial\theta_2^{\mathrm{T}}} = -\frac{1}{2}\frac{\partial^2 \log|\Sigma_i^{(k_i)}|}{\partial\theta_2\partial\theta_2^{\mathrm{T}}} + \frac{1}{2}(y_i^{(k_i)} - z_i^{(k_i)}\theta_1)^{\mathrm{T}}\frac{\partial^2\Sigma_i^{(k_i)}}{\partial\theta_2\partial\theta_2^{\mathrm{T}}}(y_i^{(k_i)} - z_i^{(k_i)}\theta_1) \tag{A3}$$

[22]. The quantities in equations (A1) and (A3) can be obtained directly from the Splus functions `gls` and `lme`. These functions use unconstrained parameterizations to ensure that the estimated variance–covariance matrix is positive semi-definite [23]. The parameterizations vary according to the structure being modelled. For example, in the case of the Gaussian correlation structure (and all spatial correlation structures), the following unconstrained parameters are used: the natural log of the range (ρ); and the logit of the ratio between the nugget (τ^2) and the error variance ($\sigma^2 + \tau^2$). Therefore, we rewrite the variance–covariance matrix, $\Sigma_i^{(k_i)}$ in terms of the unconstrained parameters, and take the derivatives of $\Sigma_i^{(k_i)}$ with respect to these parameters. Equation (A2) is not provided by the software. However, the calculation is straightforward when the variance–covariance matrix is prespecified [24].

The term $\nabla^2 L_{13}$ involves the MAR estimates of the probability of being observed for the drop-out subjects, together with partial derivatives of the conditional mean of the first missing

observation given observed data. Under normality, we know that

$$\mathrm{E}_{\theta}^{Y_{i,k_i+1} \mid Y_i^{(k_i)}} = z_{i,k_i+1}\theta_1 + c_i^{(k_i)\mathrm{T}} \Sigma_i^{(k_i)-1} (y_i^{(k_i)} - z_i^{(k_i)}\theta_1)$$

where $c_i^{(k_i)}$ is the k_i-element vector of $\Sigma_i^{(k_i)}$. Hence,

$$\frac{\partial \mathrm{E}_{\theta}^{Y_{i,k_i+1} \mid Y_i^{(k_i)}}}{\partial \theta_1} = z_{i,k_i+1} - c_i^{(k_i)\mathrm{T}} \Sigma_i^{(k_i)-1} z_i^{(k_i)}$$

$$\frac{\partial \mathrm{E}_{\theta}^{Y_{i,k_i+1} \mid Y_i^{(k_i)}}}{\partial \theta_2} = \frac{\partial c_i^{(k_i)\mathrm{T}}}{\partial \theta_2} \Sigma_i^{(k_i)-1} (y_i^{(k_i)} - z_i^{(k_i)}\theta_1)$$

$$- c_i^{(k_i)\mathrm{T}} \Sigma_i^{(k_i)-1} \frac{\partial \Sigma_i^{(k_i)}}{\partial \theta_2} \Sigma_i^{(k_i)-1} (y_i^{(k_i)} - z_i^{(k_i)}\theta_1)$$

where $\partial c_i^{(k_i)}/\partial \theta_2$ is the k_i-element vector of $\partial \Sigma_i^{(k_i)}/\partial \theta_2$. The derivatives are easy to calculate for the elements of θ_1, which concern only the mean function, but are more complicated for the elements of θ_2, which involve the variance–covariance matrix and parameterizations. When the variance–covariance matrix is given, however, the calculation still is direct.

ACKNOWLEDGEMENTS

We are grateful to Dr Edward Nunes for allowing us to use the cocaine data. We also thank SWOG and its director, Dr Charles Coltman, for providing the data. The National Institutes of Health supported this research under USPHS grants HL65365, HL68074 and CA79470. We thank the referees and the editor for valuable suggestions. We also thank SWOG and its director, Dr. Charles Coltman, for providing this QoL data.

REFERENCES

1. Rubin DB. Inference and missing data. *Biometrika* 1976; **63**:581–592.
2. Little RJA, Rubin DB. *Statistical Analysis with Missing Data* (2nd edn). Wiley: New York, 2002.
3. Schluchter MD. Methods for the analysis of informatively censored longitudinal data. *Statistics in Medicine* 1992; **11**:1861–1870.
4. Diggle PJ, Kenward MG. Informative dropout in longitudinal data analysis (with discussion). *Applied Statistics* 1994; **43**:49–93.
5. Troxel AB, Harrington DP, Lipsitz SR. Analysis of longitudinal measurements with non-ignorable non-monotone missing values. *Applied Statistics* 1998; **47**:425–438.
6. Rotnitzky A, Robins JM. Analysis of semiparametric regression models with nonignorable nonresponse. *Statistics in Medicine* 1997; **16**:81–102.
7. Rotnitzky A, Robins JM, Scharfstein DO. Semiparametric regression for repeated outcomes with nonignorable nonresponse. *Journal of the American Statistical Association* 1998; **93**:1321–1339.
8. Scharfstein DO, Rotnitzky A, Robins JM. Adjusting for nonignorable drop-outs using semiparametric nonresponse models. *Journal of the American Statistical Association* 1999; **94**:1096–1120.
9. Troxel AB, Ma G, Heitjan DF. An index of sensitivity to nonignorability. *Statistica Sinica* 2004; **14**:1221-1237.
10. Verbeke G, Molenberghs G, Thijs H, Lesaffre E, Kenward MG. Sensitivity analysis for nonrandom dropout: a local influence approach. *Biometrics* 2001; **57**:7–14.
11. Copas JB, Li HG. Inference for non-random samples (with discussion). *Journal of the Royal Statistical Society*, *Series B* 1997; **59**:55–95.
12. Copas JB, Eguchi S. Local sensitivity approximations for selectivity bias. *Journal of the Royal Statistical Society*, *Series B* 2001; **63**:871–895.

13. Chambers EA, Cox DR. Discrimination between alternative binary response models. *Biometrika* 1967; **54**: 573–578.
14. Malik HJ. Logistic distribution. *Encyclopedia of Statistical Sciences*, vol. 5. Wiley: New York, 1995.
15. David M, Little RJA, Samuhel ME, Triest RK. Alternative methods for CPS income imputation. *Journal of the American Statistical Association* 1986; **81**:29–41.
16. Verbyla AP, Cullis BR. Modeling in repeated measures experiments. *Applied Statistics* 1990; **39**:341–356.
17. Eisenberger MA, Blumenstein BA, Crawford ED, Miller G, McLeod DG, Loehrer PJ, Wilding G, Sears K, Culkin DJ, Thompson Jr IM, Bueschen AJ, Lowe BA. Bilateral orchiectomy with or without flutamide for metastatic prostate cancer. *New England Journal of Medicine* 1998; **339**:1036–1042.
18. Moinpour CM, Savage MJ, Troxel A, Lovato LC, Eisenberger M, Veith RW, Higgins B, Skeel R, Yee M, Blumenstein B, Crawford ED, Meyskens Jr FL. Quality of life in advanced prostate cancer: results of a randomized therapeutic trial. *Journal of the National Cancer Institute* 1998; **90**:1537–1544.
19. Troxel AB. A comparative analysis of quality of life data from a Southwest Oncology Group randomized trial of advanced colorectal cancer. *Statistics in Medicine* 1998; **17**:767–779.
20. Kenward MG. Selection models for repeated measurements with non-random dropout: an illustration of sensitivity. *Statistics in Medicine* 1998; **17**:2723–2732.
21. Ma G, Heitjan DF. Sensitivity to nonignorability in frequentist inference. *Applied Bayesian Modeling and Causal Inference From Incomplete-Data Perspectives*, Chapter 16, Andrew Gelman, Xiao-Li Meng (eds). Wiley: New York, 2004.
22. Pinheiro JC, Bates DM. *Mixed-effects Models in S and S-PLUS*. Springer: New York, 2000.
23. Pinheiro JC, Bates DM. Unconstrained parameterizations for variance–covariance matrices. *Statistics and Computing* 1996; **6**:289–296.
24. Lindstrom MJ, Bates DM. Newton-Raphson and EM algorithms for linear mixed-effects models for repeated-measures data. *Journal of the American Statistical Association* 1988; **83**:1014–1022.

