Published in final edited form as:

Comput Methods Programs Biomed. 2018 October ; 164: 207–220. doi:10.1016/j.cmpb.2018.06.014.

Measuring the Impact of Nonignorable Missingness Using the R Package *isni*

Hui Xie¹, Weihua Gao², Baodong Xing³, Daniel F. Heitjan⁴, Donald Hedeker⁵, and Chengbo Yuan²

¹Division of Epidemiology & Biostatistics (M/C 923), School of Public Health, University of Illinois at Chicago, 1603 West Taylor Street, Room 984, Chicago, 60612-4336, IL; Faculty of Health Sciences, Simon Fraser University, Burnaby, BC, V6S0G6, Canada; Arthritis Research Canada, Richmond, BC, V6X 2C7, Canada. Electronic address: huixie@uic.edu.

²Division of Epidemiology & Biostatistics (M/C 923), School of Public Health, University of Illinois at Chicago, 1603 West Taylor Street, Room 984, Chicago, 60612-4336, IL.

³AbbVie Inc., Chicago, 60612, IL.

⁴Department of Statistical Science, Southern Methodist University, Dallas, 75275-0332, TX; Department of Clinical Sciences, University of Texas Southwestern Medical Center, Dallas, 75390-9066, TX. Electronic address: dheitjan@smu.edu.

⁵Department of Public Health Sciences, The University of Chicago, Chicago, 60637, IL.

Abstract

Background and Objective: The popular assumption of *ignorability* simplifies analyses with incomplete data, but if it is not satisfied, results may be incorrect. Therefore it is necessary to assess the sensitivity of empirical findings to this assumption. We have created a user-friendly and freely available software program to conduct such analyses.

Method: One can evaluate the dependence of inferences on the assumption of ignorability by measuring their sensitivity to its violation. One tool for such an analysis is the *index of local sensitivity to nonignorability (ISNI)*, which evaluates the rate of change of parameter estimates to the assumed degree of nonignorability in the neighborhood of an ignorable model. Computation of ISNI avoids the need to estimate a nonignorable model or to posit a specific magnitude of nonignorability. Our new R package, named *isni*, implements ISNI analysis for some common data structures and corresponding statistical models.

Result: The *isni* package computes ISNI in the generalized linear model for independent data, and in the marginal multivariate Gaussian model and the linear mixed model for longitudinal/clustered data. It allows for arbitrary patterns of missingness caused by dropout and/or intermittent missingness. Examples illustrate its use and features.

Corresponding Author: Hui Xie, PhD, Division of Epidemiology & Biostatistics (M/C 923), School of Public Health, University of Illinois at Chicago, 1603 West Taylor Street, Chicago, IL, 60612-4336 and Faculty of Health Sciences, Simon Fraser University, Burnaby, BC, V6S0G6, Arthritis Research Canada, Richmond, BC V6X 2C7, huixie@uic.edu, Phone: 312-413-2734.

Conclusions: The R package **isni** enables a systematic and efficient sensitivity analysis that informs evaluations of reliability and validity of empirical findings from incomplete data.

Keywords

Analytical reliability; data quality; missing data; missing not at random; multivariate normal; selection model

1. Introduction

Ignorability

is a key working assumption in the analysis of data with missing observations. When the data are *missing at random (MAR)* and the parameters of the ideal-data model and missing-data model are *distinct (PD)*, the missingness mechanism is ignorable in the sense that one need not model it in order to generate valid likelihood/Bayesian inferences (Rubin 1976; Heitjan and Rubin 1991). This greatly simplifies analyses, because the missing data process can be difficult to model and is rarely of primary interest. Yet in practice we often suspect that MAR does not hold, and therefore that the underlying missing data mechanism is nonignorable. For example, if patients drop out of a study because of excessive toxicity, their missingness may be related to their unobserved data values, even after conditioning on all available information. This constitutes a violation of MAR, and therefore calls ignorability into question. When ignorability does not hold, standard likelihood/Bayesian analyses can yield inferences that incorrectly summarize the information in the data.

Because we cannot test MAR robustly using only the observed data, it is critical to have the means to conduct an analysis of sensitivity to departures from ignorability (Daniels and Hogan 2008; Little, D'Agostino, Cohen, Dickersin, Emerson, Farrar, Frangakis, Hogan, Molenberghs, Murphy, Neaton, Rotnitzky, Scharfstein, Shih, Siegel, and Stern 2012). Indeed, an expert panel convened to study the issue has declared that “[s]ensitivity analyses should be part of the primary reporting of findings from clinical trials. Examining sensitivity to the assumptions about the missing data mechanism should be a mandatory component of reporting” (The Panel on Handling Missing Data in Clinical Trials of the National Research Council 2010). Moreover, “[s]ensitivity analysis [to missing data assumptions] is a relatively new area, and further research on the best methods is needed.” (Little *et al.* 2012)

Estimating nonignorable models of missingness is conceptually and computationally challenging, which limits the types of sensitivity analyses that one can perform (Xie 2012; Xie and Qian 2012). One approach is to compute the *index of local sensitivity to nonignorability*, or *ISNI*, first proposed by Troxel, Ma, and Heitjan (2004) in the context of the univariate generalized linear model. The ISNI method implements the local sensitivity approach that examines the effect on inferences of minor departures from ignorability (Copas, JB, and Li, HG 1997; Copas, JB, and Eguchi, S 2001; Verbeke, Molenberghs, Thijs, Lesaffre, and Kenward 2001). The method has since been extended to a range of statistical models and coarsening types (Xie and Heitjan 2004; Ma, Troxel, and Heitjan 2005; Xie and Heitjan 2009; Xie 2008, 2009, 2010, 2012; Xie, Qian, and Qu 2011; Xie and Qian 2012; Gao, Hedeker, Mermelstein, and Xie 2016). ISNI overcomes the computational obstacles to

sensitivity analysis by requiring only readily-available MAR model computations, thereby avoiding the estimation of complicated nonignorable models.

The absence of general software for computing ISNI has hampered its widespread adoption. This article describes a new R package, denoted **isni**, that performs ISNI computations for common models — currently, generalized linear models for cross-sectional data, and marginal multivariate Gaussian models and linear mixed-effects models for longitudinal/clustered data. This article describes the software and its application.

We organize the article as follows: Section 2 reviews the ISNI method. Section 3 describes various aspects of using main functions in **isni**. Section 4 illustrates the application of **isni** and the interpretation of its outputs in real-data examples. Section 5 offers a summary and discussion. An Appendix presents further details.

2. Review of the ISNI Method

Let Y be a vector of outcomes and G be a vector of missingness indicators with the same length as Y , where each element of G takes the value of 0(1) when the corresponding element of Y is observed(missing). We further define $Y = (Y^{\text{obs}}, Y^{\text{mis}})$, where Y^{obs} , Y^{mis} denote the observed and missing elements in Y , respectively. We specify the joint distribution of Y and G with a selection model; that is, the marginal density of Y is $f_{\theta}(y)$, indexed by parameter θ , and the conditional probability mass function of G given $Y = y$ is $f_{\gamma_0, \gamma_1}(g|y)$, indexed by parameters γ_0 and γ_1 . We further assume that we can write the

conditional probability function of G given $Y = y$ as $f_{\gamma_0, \gamma_1}(g|y^{\text{obs}}, y^{\text{mis}})$, such that the

parameter γ_0 reflects the effect of fully observed data elements on the probability of missingness, and the parameter γ_1 reflects the effect of potentially unobserved outcomes on the probability of missingness. We denote the conditional distribution of the missingness indicator to be the *missing data mechanism (MDM)*. Under these assumptions, the observed data (y^{obs}, g) is *missing at random (MAR)* if, for every possible value of the parameters, $f_{\gamma_0, \gamma_1}(g|y^{\text{obs}}, y^{\text{mis}})$ takes the same value for all y^{mis} . For example, if there are no missing observations, the observed data are MAR by default, even if the MDM stipulates a strong correlation of Y and G . An MDM is MAR if every possible data set generated under it is MAR.

Following common practice, we formulate the MDM to reduce to MAR when $\gamma_1 = 0$. When $\gamma_1 = 0$, the missingness probability depends on unobserved data, y^{mis} , even after conditioning on the observed data. Because γ_1 captures the magnitude of nonignorability, we denote it the *nonignorability parameter*. Under the general selection model, the loglikelihood $L(\theta, \gamma_0, \gamma_1; y^{\text{obs}}, g)$ is

$$L(\theta, \gamma_0, \gamma_1; y^{\text{obs}}, g) = \ln \int_{Y^{\text{mis}}} f_{\theta}(y^{\text{obs}}, y^{\text{mis}}) f_{\gamma_0, \gamma_1}(g|y^{\text{obs}}, y^{\text{mis}}) dy^{\text{mis}}, \quad (1)$$

where $\Omega_{Y^{\text{mis}}}$ denotes the sample space of Y^{mis} , and $f_{\theta}(y^{\text{obs}}, y^{\text{mis}})$ is $f_{\theta}(Y^{\text{obs}}, Y^{\text{mis}})$ evaluated at the observed data y^{obs} and with the missing value Y^{mis} substituted by a hypothetical posited value $y^{\text{mis}} \in \Omega_{Y^{\text{mis}}}$; the completeness indicator g dictates the identities of the observed and missing y values. Note that if we assume an MAR MDM (i.e., $\gamma_1 = 0$), we have $f_{\gamma_0, \gamma_1}(g|y^{\text{obs}}, y^{\text{mis}}) = f_{\gamma_0, \gamma_1=0}(g|y^{\text{obs}})$, because this conditional probability does not depend on values of y^{mis} . Thus we can move this term out of the integration in Eqn (1), resulting in the simpler loglikelihood

$$\begin{aligned} L_I(\theta, \gamma_0, \gamma_1; y^{\text{obs}}, g) &= \ln \left\{ \int_{\Omega_{Y^{\text{mis}}}} f_{\theta}(y^{\text{obs}}, y^{\text{mis}}) dy^{\text{mis}} \right\} f_{\gamma_0, \gamma_1=0}(g|y^{\text{obs}}) \\ &= \ln f_{\theta}(y^{\text{obs}}) + \ln f_{\gamma_0, \gamma_1=0}(g|y^{\text{obs}}). \end{aligned} \quad (2)$$

Under MAR and the additional assumption of *parameter distinctness* — i.e., that θ and γ are independent (for Bayesian inference) or have parameter spaces that factor (for likelihood inference) — $\ln f_{\theta}(y^{\text{obs}})$ contains the totality of information on θ , and therefore it is unnecessary to estimate the MDM. We denote such an analysis the *ignorability analysis*. The more general analysis that does not assume ignorability bases inferences on Eqn (1), which requires positing the MDM in detail. Because observed data alone provide no robust information to assess the dependence of the missingness probability on y^{mis} , one cannot identify or estimate such a nonignorable model without either additional data or untestable assumptions.

Practical analyses therefore typically assume MAR and base inferences on Eqn (2). To evaluate the robustness of such an inference, one can conduct a sensitivity analysis. One way to do this is to posit a range of values of γ_1 and determine the extent to which an estimate of θ , such as the MLE given γ_1 , $\hat{\theta}(\gamma_1)$, depends on the values of γ_1 . The ISNI approach is to execute this analysis in a neighborhood of the MAR model by determining the rate of change of $\hat{\theta}(\gamma_1)$ as a function of γ_1 at $\gamma_1 = 0$. That is, ISNI calculates the derivative of $\hat{\theta}(\gamma_1)$ with respect to γ_1 , evaluated locally at the ignorable model (Troxel *et al.* 2004). As we show in Appendix A, a general formula for ISNI is

$$\text{ISNI} \equiv \left. \frac{\partial \hat{\theta}(\gamma_1)}{\partial \gamma_1^T} \right|_{\gamma_1=0} = -\nabla^2 L_{\theta, \theta}^{-1} \nabla^2 L_{\theta, \gamma_1}, \quad (3)$$

where

$$\begin{aligned}\nabla^2 L_{\theta, \theta} &= \frac{\partial^2 L(\theta, \gamma_0, \gamma_1)}{\partial \theta \partial \theta^T} \bigg|_{\hat{\theta}(0), \hat{\gamma}_0(0), \gamma_1 = 0} = \frac{\partial \ln f_{\theta}(y^{\text{obs}})}{\partial \theta \partial \theta^T} \bigg|_{\hat{\theta}(0), \hat{\gamma}_0(0), \gamma_1 = 0}, \\ \nabla^2 L_{\theta, \gamma_1} &= \frac{\partial^2 L(\theta, \gamma_0, \gamma_1)}{\partial \theta \partial \gamma_1^T} \bigg|_{\hat{\theta}(0), \hat{\gamma}_0(0), \gamma_1 = 0},\end{aligned}$$

and $L(\theta, \gamma_0, \gamma_1)$ denotes the loglikelihood of Eqn (1). The computation of ISNI thus involves two parts: First one computes $\nabla^2 L_{\theta, \theta}^{-1}$, which is the inverse of the observed information matrix of θ under the MAR outcome model; this is readily available from standard statistical software. Second, one computes $\nabla^2 L_{\theta, \gamma_1}$, which measures the nonorthogonality of θ and γ_1 .

Below we show ISNI formulas for some popular statistical models.

2.1. ISNI for independent data

We first consider the case where Y follows a generalized linear model (GLM) (McCullagh and Nelder 1989) that assumes that scalar Y_i , given predictors x_i , $i = 1, \dots, N$, are independent draws from the exponential family

$$f_{\theta}(y_i | x_i) = \exp \left\{ \frac{y_i \Psi_i(\beta, x_i) - b(\Psi_i(\beta, x_i))}{a_i(\tau)} + c(y_i, \tau) \right\}, \quad (4)$$

where Ψ_i is the canonical parameter as a function of the regression coefficient parameter β ; functions $b(\cdot)$ and $c(\cdot, \cdot)$ identify a distribution in the exponential family; and $a_i(\tau) = \tau/w_i$ with the dispersion parameter τ and a known weight w_i . We further assume that the MDM is a logistic regression

$$P(G_i = 1 | s_i, y_i) = h(\gamma_0^T s_i + \gamma_1 y_i) = \frac{1}{1 + \exp[-(\gamma_0^T s_i + \gamma_1 y_i)]}, \quad (5)$$

where $G_i = 0(1)$ if the i th observation is observed(missing), and s_i includes a set of observed predictors. The parameters γ_0 and γ_1 associate the probability of missingness with the set of fully observed missingness predictors in s_i and the partially observed outcome variable y_i , respectively. Throughout this paper and in the **isni** package, we assume the inverse logit form for $h(\cdot)$, as it is popular and robust, and it simplifies interpretation (Xie and Heitjan 2009). Following Troxel *et al.* (2004), under independence over units i we have

$$\begin{aligned}\nabla^2 L_{\theta, \theta} &= \frac{\partial^2 L(\theta, \gamma_0, \gamma_1)}{\partial \theta \partial \theta^T} \bigg|_{\hat{\theta}(0), \hat{\gamma}_0(0), \gamma_1 = 0} = \sum_{i: g_i = 0} \frac{\partial \ln f_{\theta_y}(y_i^{\text{obs}} | x_i)}{\partial \theta \partial \theta^T} \bigg|_{\gamma_1 = 0}, \\ \nabla^2 L_{\theta, \gamma_1} &= \frac{\partial^2 L(\theta, \gamma_0, \gamma_1)}{\partial \theta \partial \gamma_1^T} \bigg|_{\hat{\theta}(0), \hat{\gamma}_0(0), \gamma_1 = 0} = \sum_{i: g_i = 1} (1 - h_i) \frac{\partial E(Y_i^{\text{mis}} | x_i)}{\partial \theta} \bigg|_{\gamma_1 = 0}.\end{aligned}\quad (6)$$

ISNI for $\theta = (\beta, \tau)$ simplifies to

$$\text{ISNI} = \left(- \left[\sum_i (1 - g_i) \left(y_i \frac{\partial^2 \Psi_i}{\partial \beta^2} - \frac{\partial^2 b}{\partial \beta^2} \right) a_i^{-1}(\tau) \right]^{-1} \sum_i g_i (1 - h_i) \frac{\partial^2 b}{\partial \Psi_i \partial \beta} \right)_{\hat{\beta}(0), \hat{\tau}(0)},$$

where $h_i = h(\hat{\gamma}_0^T s_i)$ is the predicted probability of being missing under MAR, and $\hat{\beta}(0)$ and $\hat{\tau}(0)$ are MLEs under MAR. This formula is the same as that presented in Troxel *et al.* (2004) except that we reverse the signs to reflect the reversed role of the indicator G .

2.2. ISNIs for longitudinal/clustering data

Several authors have described generalizations of ISNI to longitudinal/clustering data (Ma *et al.* 2005; Xie and Heitjan 2009; Xie 2008, 2009, 2012; Xie and Qian 2012). We describe below the ISNI method for the setting of longitudinal data with non-monotone missingness; adaptation to other types of clustering data is straightforward.

2.2.1. Models for the notional complete data—Let $Y_i = (Y_{i1}, \dots, Y_{in_i})$ denote the notional complete data for subject i , where Y_{ij} is the outcome at measurement occasion j , $i = 1, \dots, N$, $j = 1, \dots, n_i$. We assume that the density function of Y_i is $f_{\theta}(y_i | x_i)$, where θ is the vector of the parameters of interest with length p_{θ} , and x_i is a matrix of fully observed predictors. We describe below the ISNI analysis of two popular classes of models for data of this form.

Marginal multivariate Gaussian model (MMGM). The MMGM for a continuous outcome is

$$Y_i | x_i \stackrel{\text{ind}}{\sim} \text{MVN}(x_i \theta_1, \Sigma_i(\theta_2)), \quad (7)$$

where θ_1 and θ_2 are parameters of the population mean and the variance-covariance, respectively. The matrix $\Sigma_i(\theta_2)$ must be positive definite. Under ignorable missingness, one typically estimates this model by generalized least-squares, for example with the R function `gls()`.

Linear mixed model (LMM).: The LMM for a continuous outcome is (Laird and Ware 1982)

$$Y_i | b_i, x_i, z_i \stackrel{\text{ind}}{\sim} \text{MVN}(x_i \beta + z_i b_i, \Lambda_i), \quad b_i \sim N(0, V_b). \quad (8)$$

Here β is a vector of p fixed population parameters; b_i is a vector of q random effects associated with individual i ; x_i and z_i are predictor matrices for the fixed and random effects, respectively, where z_i is a subset of x_i , and V_b and Λ_i are variance-covariance matrices for the random effects and residuals, respectively. Λ_i depends on i only in that its size is $n_i \times n_i$. Marginally,

$$Y_i | x_i, z_i \stackrel{\text{ind}}{\sim} \text{MVN}(x_i \beta, \Lambda_i + z_i V_b z_i^T).$$

We set $\theta = (\beta, D)$ where D denotes the parameters in the variance-covariance matrices Λ_i and V_b . One can estimate the LMM using R function lme().

2.2.2. An MDM for non-monotone missing data—Longitudinal studies typically exhibit two types of missingness: Intermittently missing observations, for example from missed visits; and dropout, from subjects who leave the study permanently before completing follow-up. We therefore define an MDM that allows for both types of missingness by means of a general transitional model (Xie 2012). Let $G_i = (G_{i1}, \dots, G_{in_i})$ denote the vector of missingness status variables for subject i , where $G_{ij}, j = 1, \dots, n_i$ denotes the missingness status of subject i at occasion j , and

$$G_{ij} = \begin{cases} O & \text{if subject } i \text{ is observed at occasion } j, \\ I & \text{if subject } i \text{ is intermittently missing at occasion } j, \\ D & \text{if subject } i \text{ has dropped out at occasion } j. \end{cases}$$

Xie (2012) described an approach that writes the MDM as a product of transition probabilities:

$$f_Y(g_{i1}, \dots, g_{in_i} | y_i, x_i) = f(g_{i1} | y_i, x_i) \prod_{j=2}^{n_i} f_Y(g_{ij} | g_{i1}, \dots, g_{i,j-1}, y_i, x_i).$$

We typically assume that all units are observed at baseline; i.e., $f(O | y_i, x_i) = 1$. One can then model the remaining univariate conditional distributions separately. We note first that each conditional probability can include all past missingness status variables, thereby naturally incorporating information on how past missingness affects current missingness. Second, one needs to decide what variables in Y_i enter each conditional distribution. Let $s_{ij} = (y_{i(j)}^{\text{obs}}, x_i)$ be a matrix containing fully observed predictors for missingness up to visit j for subject i , where $y_{i(j)}^{\text{obs}}$ includes all observed outcome measurements prior to visit j for subject i . Because

we condition on all the past missingness status variables within each past missingness pattern, $y_{i(j)}^{\text{obs}}$ can include all past observed outcomes. If any future outcome, y_{iJ} , where $J > j$, is observed for all the subjects with the same past missingness pattern, y_{iJ} can also be included in $y_{i(j)}^{\text{obs}}$ for those observations with that same past missingness pattern.

We further let a numerical variable u index missingness status with $u = 0, 1, 2$ representing O, I, D , respectively. Our MDM then assumes the following conditional distribution of G_{ij} :

$$[G_{ij} | S_{ij} = s_{ij}, Y_{ij} = y_{ij}, G_{i(j)} = g_{i(j)}] \sim \text{Multinomial} \left(1, \left[P_{ij}^{0g_{i(j)}}, P_{ij}^{1g_{i(j)}}, P_{ij}^{2g_{i(j)}} \right] \right), \quad (9)$$

where $g_{i(j)} = (g_{i,j-1}, \dots, g_{i1})$ denotes the past missingness pattern prior to visit j , and the cell probabilities $[P_{ij}^{0g_{i(j)}}, P_{ij}^{1g_{i(j)}}, P_{ij}^{2g_{i(j)}}]$ are specified as

$$P_{ij}^{ug_{i(j)}} = \frac{\phi_{ij}^{ug_{i(j)}}}{\sum_{U=0}^2 \phi_{ij}^{Ug_{i(j)}}}, \quad u = 0, 1, 2; \quad (10)$$

$$\text{with } \phi_{ij}^{ug_{i(j)}}(s_{ij}, y_{ij}) = \exp(\gamma_0^{ug_{i(j)}} s_{ij} + \gamma_1^{ug_{i(j)}} y_{ij}).$$

When $\gamma_1^{ug_{i(j)}} = 0$ for all values of u and $g_{i(j)}$, the MDM does not depend on the potentially

unobserved outcome and thus is MAR. When $\gamma_1^{ug_{i(j)}} \neq 0$ for some u and $g_{i(j)}$, the model allows that, given the missingness pattern prior to time j and the other fully observed predictor s_{ij} , the missingness status depends on the potentially unobserved outcomes through the contemporaneous outcome Y_{ij} . An alternative MDM specification would let the missingness depend on both past and future unobserved outcomes. We chose the former model for two reasons: First, it reduces the number of parameters for nonignorable missingness, allowing us to more easily interpret the sensitivity analysis. As others (e.g., Vansteelandt, Rotnitzky, and Robins (2007)) have noted, parsimony is desirable in sensitivity analysis. Second, one can always take the integration of the latter model with respect to past and future unobserved outcomes so that the resulting selection model depends only on the outcome at the current visit. In this sense, our model can be viewed as an approximation to the model in which the probability of missingness depends also on past and future unobserved outcomes; see Xie (2012).

Xie (2012) applied the transitional MDM to a longitudinal dataset with non-monotone missingness arising from a design with $n_j = n$; this enables conditioning on the entire past missingness vector $g_{i(j)}$. In the setting with varying n_j , conditioning on $g_{i(j)}$ may be impossible. In this case, a convenient alternative is a first-order transition model, where one assumes that, conditional on $(s_{ij}, y_{ij}, g_{i,j-1})$, the missingness status G_{ij} at the current visit is independent of missingness status at all other prior visits; thus Eqn (9) reduces to

$$G_{ij}|S_{ij} = s_{ij}, Y_{ij} = y_{ij}, G_{i,j-1} = v \sim \text{Multinomial} \left(1, [P_{ij}^{0v}, P_{ij}^{1v}, P_{ij}^{2v}] \right);$$

$$P_{ij}^{uv} = \frac{\phi_{ij}^{uv}}{\sum_{U=0}^2 \phi_{ij}^{Uv}}, \text{ where } u = 0, 1, 2; v = 0, 1, \quad (11)$$

$$\text{and } \phi_{ij}^{uv}(s_{ij}, y_{ij}) = \exp(\gamma_0^{uv} s_{ij} + \gamma_1^{uv} y_{ij}).$$

By the definition of dropout, $G_{ij}=2$ deterministically when $G_{i,j-1}=2$ (the prior visit is a dropout); by the definition of intermittent missingness, $\phi_{ij}^{21}=0$; and because the response probabilities must sum to unity, for identification purposes $\phi_{ij}^{00}=\phi_{ij}^{01}=1$. Here s_{ij} is a vector of fully observed predictors for missingness at time j for subject i , which commonly includes the history of the predictors x in the ideal-data model up to and including time j as well as the history of the observed prior outcomes (*i.e.*, the observed elements in $(Y_{i1}, \dots, Y_{i,j-1})$).

The conditional probability $f_{\gamma}(g_{ij} | s_{ij}, y_{ij}, g_{i,j-1})$ then takes the form

$$f_{\gamma}(g_{ij}|s_{ij}, y_{ij}, g_{i,j-1}) = \quad (12)$$

$$\left\{ \begin{array}{ll} \frac{1}{1 + \exp(\gamma_0^{10}s_{ij} + \gamma_1^{10}y_{ij}) + \exp(\gamma_0^{20}s_{ij} + \gamma_1^{20}y_{ij})} & \text{if } g_{i,j-1} = 0, g_{ij} = 0, \\ \frac{\exp(\gamma_0^{g_{ij}^0}s_{ij} + \gamma_1^{g_{ij}^0}y_{ij})}{(1 + \exp(\gamma_0^{10}s_{ij} + \gamma_1^{10}y_{ij}) + \exp(\gamma_0^{20}s_{ij} + \gamma_1^{20}y_{ij}))} & \text{if } g_{i,j-1} = 0, g_{ij} \neq 0, \\ \frac{1}{1 + \exp(\gamma_0^{11}s_{ij} + \gamma_1^{11}y_{ij})} & \text{if } g_{i,j-1} = 1, g_{ij} = 0, \\ \frac{\exp(\gamma_0^{11}s_{ij} + \gamma_1^{11}y_{ij})}{1 + \exp(\gamma_0^{11}s_{ij} + \gamma_1^{11}y_{ij})} & \text{if } g_{i,j-1} = 1, g_{ij} = 1, \\ 0 & \text{if } g_{i,j-1} = 1, g_{ij} = 2, \\ 0 & \text{if } g_{i,j-1} = 2, g_{ij} \neq 2, \\ 1 & \text{if } g_{i,j-1} = 2, g_{ij} = 2. \end{array} \right.$$

The package **isni** allows computation of ISNI statistics under both the general transitional MDM and the simpler first-order transitional MDM.

2.2.3. ISNI with the transitional MDM—The nonorthogonality term $\nabla^2 L_{\theta, \gamma_1}$ in Eqn

(3) has been derived for longitudinal data with monotone (Ma *et al.* 2005) and non-monotone missingness patterns (Xie 2012; Xie and Qian 2012); see Appendix B. For longitudinal models with the first-order transitional MDM and $\gamma_1 = (\gamma_1^{10}, \gamma_1^{20}, \gamma_1^{11})$,

$$\text{ISNI} = \left(\text{ISNI}_{\gamma_1^{10}}, \text{ISNI}_{\gamma_1^{20}}, \text{ISNI}_{\gamma_1^{11}} \right) \text{ and}$$

$$\begin{aligned}
\nabla^2 L_{\theta, \gamma_1} &= \frac{\partial^2 L}{\partial \theta \partial \gamma_1^T} \bigg|_{\hat{\theta}(0), \hat{\gamma}_0(0), \gamma_1 = 0} = \left(\frac{\partial^2 L}{\partial \theta \partial \gamma_1^{10}}, \frac{\partial^2 L}{\partial \theta \partial \gamma_1^{20}}, \frac{\partial^2 L}{\partial \theta \partial \gamma_1^{11}} \right) \bigg|_{\hat{\theta}(0), \hat{\gamma}_0(0), \gamma_1 = 0} \\
&= \sum_{i: K_i < n_i} \frac{\partial E \left((Y_i^{\text{mis}})^T | y_i^{\text{obs}} \right)}{\partial \theta} \bigg|_{\gamma_1 = 0} \cdot [A_i^{10}, A_i^{20}, A_i^{11}],
\end{aligned} \tag{13}$$

where K_i is the length of y_i^{obs} and $\hat{\theta}(0)$ and $\hat{\gamma}_0(0)$ are MLEs of θ and γ_0 under the MAR assumption.

The term $\partial E \left((Y_i^{\text{mis}})^T | y_i^{\text{obs}} \right) / \partial \theta$ is a $p_\theta \times d_i$ matrix where d_i is the number of missing outcomes for subject i . Under MAR in the complete-outcome models of Section 2.2.1, $(Y_i^{\text{mis}})^T | y_i^{\text{obs}}$ is a Gaussian distribution involving θ parameters only and $\partial E \left((Y_i^{\text{mis}})^T | y_i^{\text{obs}} \right) / \partial \theta$ can be derived in a closed form involving matrix multiplication and differentiation as shown in Appendix C. The term $A_i = [A_i^{10}, A_i^{20}, A_i^{11}]$ is a $d_i \times 3$ matrix. For the l th ($l = 1, \dots, d_i$) component of y_i^{mis} that corresponds to the j th element of y_i , one can show that (Xie 2012; Xie and Qian 2012)

$$\begin{aligned}
A_{il}^{10} &= I(g_{i,j-1} = 0) [I(g_{i,j} = 1) - P_{ij}^{10}]_{\hat{\gamma}_0(0), \gamma_1 = 0}, \\
A_{il}^{20} &= I(g_{i,j-1} = 0) [I(g_{i,j} = 2) - P_{ij}^{20}]_{\hat{\gamma}_0(0), \gamma_1 = 0}, \\
A_{il}^{11} &= I(g_{i,j-1} = 1) [P_{ij}^{01}]_{\hat{\gamma}_0(0), \gamma_1 = 0}.
\end{aligned}$$

Here, P_{ij}^{10} , P_{ij}^{20} and P_{ij}^{01} are the missingness status transitional probabilities from Eqn (11), calculated under MAR (i.e., with $\gamma_1 = 0$). We note that because dropout and intermittent missingness are competing causes of missingness, A_{il}^{10} and A_{il}^{20} are of opposite sign when $g_{i,j-1} = 0$.

When γ_1 is a scalar, ISNI approximates the changes of the estimates when γ_1 is perturbed from 0 to 1. In our model, the nonignorability parameter γ_1 is the vector $(\gamma_1^{10}, \gamma_1^{20}, \gamma_1^{11})$. Thus, ISNI is a vector of three elements where each approximates the changes in estimates when only the corresponding element in γ_1 is perturbed from 0 to 1. We suggest the following strategies to produce a parsimonious sensitivity analysis with multiple nonignorability parameters. To summarize the joint effects of all three nonignorable parameters, one can approximate the change in the MAR estimates, $\hat{\theta}(\gamma_1) - \hat{\theta}(0)$ by $\frac{\partial \hat{\theta}(\gamma_1)}{\partial \gamma_1^T} \bigg|_{\gamma_1 = 0} \gamma_1$. When one is willing to assume that intermittent missingness and dropout have roughly the same

nonignorability mechanism, we can fix $\gamma_1^{10}, \gamma_1^{20}$ and γ_1^{11} at the same largest perturbation value. In this scenario, γ_1 becomes a scalar and A_i a vector where $A_i = A_i^{10} + A_i^{20} + A_i^{11}$ and the components of A_i become $P_{ij}^{0, g_i^{j-1}}$, the estimated probability of being observed for the missing observations under ignorability. For its simplicity and ease of interpretation, this is the default method in our package.

Alternatively, one could consider all perturbations of the elements of γ_1 that are within a hypercube of size 1 from the origin. These perturbations include scenarios under which dropout and intermittent missingness can have different or even opposite nonignorable missingness mechanisms. For this strategy, we suggest the following index of sensitivity:

$$\text{MISNI}(\hat{\theta}) = \sum_{i=1}^q \left| \frac{\partial \hat{\theta}(\gamma_1)}{\partial \gamma_{1i}} \right|_{\hat{\theta}(0), \hat{\gamma}_0(0), \gamma_1 = 0}, \quad (14)$$

where q is the length of γ_1 ; here, $q = 3$. MISNI has the interpretation of maximum sensitivity, $\max_{|\gamma_{1i}|=1, i=1, \dots, q} (\hat{\theta}(\gamma_1) - \hat{\theta}(0))$, when each element of γ_1 is perturbed between -1 and 1 . Our package also produces MISNI. Xie and Heitjan (2004) consider a perturbation scheme when the nonignorability parameter lies on a hyperball of radius $\sqrt{q}$ around the origin, that is, $\|\gamma_1\| = \sqrt{q}$. A potential issue with this scheme is that elements in the nonignorability parameter vector can have extreme and implausible values, especially when q , the dimension of the nonignorability parameter, is large. An alternative is to consider a hyperball of size that is independent of q (Gao *et al.* 2016). We note that a hypercube of size r as used here contains a hyperball of radius r .

One can extend these strategies to the general transitional MDM. Recall that under Eqn (10), the nonignorability parameter $\gamma_1 = \cup_{j, g_{i(j-1)}} \left(\gamma_1^{10g_{i(j-1)}}, \gamma_1^{20g_{i(j-1)}}, \gamma_1^{11g_{i(j-1)}} \right)$, and the dimensionality of γ_1 depends on the number of unique past missing data patterns and can be large. For a more parsimonious sensitivity analysis, one option is to reduce the number of nonignorability parameters such that $\gamma_1^{10g_{i(j-1)}} = \gamma_1^{10}, \gamma_1^{20g_{i(j-1)}} = \gamma_1^{20}$ and $\gamma_1^{11g_{i(j-1)}} = \gamma_1^{11}, \forall j$ and $g_{i(j-1)}$. This reduces γ_1 to $(\gamma_1^{10}, \gamma_1^{20}, \gamma_1^{11})$, and we can then use the strategies described for the first-order transitional model.

We note moreover that in the special case of no intermittent missingness, the equations further simplify: $A_i^{10} = A_i^{11} = 0$ and $\partial E\left((Y_i^{\text{mis}})^T | Y_i^{\text{obs}}\right) / \partial \theta$ reduces to a $p_\theta \times 1$ vector. The calculation of $\nabla^2 L_{\theta, \gamma_1}$ then reduces to that of Ma *et al.* (2005) and Xie (2008).

2.3. Calibrating ISNI

With a logit link in the MDM, a scalar γ_1 is the log odds ratio in the probability of being missing associated with a one-unit change in y ; when $\gamma_1 = 1$, a one-unit change in y is associated with a 2.7-fold increase in the odds of being missing. For outcomes with a single natural scale, such as the Poisson and binomial, one can interpret ISNI directly in this manner. For continuous Y , this interpretation is inadequate because the value of ISNI depends on the scale of measurement. We describe below two calibration approaches that facilitate interpretation by creating a scale-free index.

The first approach evaluates changes in estimates of θ for a magnitude of nonignorability where a one-SD change in Y is associated with an odds ratio of $e^1 = 2.7$ in the probability of being missing, i.e., when $\gamma_1 = \pm 1/\sigma_Y$, or one standardized unit of nonignorability.

The second approach is to approximate the minimum standardized magnitude of nonignorability that is needed for the change in $\hat{\theta}$ to equal one standard error (SE). One can then assess sensitivity by evaluating whether this level of nonignorability is plausible.

Specifically, note that for θ^j (the j th element of θ) we have $\hat{\theta}^j(\gamma_1) - \hat{\theta}^j(0) \approx \gamma_1 \text{ISNI}_Y(\hat{\theta}^j)$, where $\text{ISNI}_Y(\cdot)$ refers to the ISNI computed with data on the Y scale. We compute the smallest absolute value of γ_1 that gives a 1-SE change as $\tilde{\gamma}_1 = \frac{\text{SE}}{\text{ISNI}}$. To put the magnitude of nonignorability in the standardized scale, we define the sensitivity transformation c statistic as

$$c = |\tilde{\gamma}_1 \times \sigma_Y| = \left| \frac{\sigma_Y \text{SE}}{\text{ISNI}} \right|. \quad (15)$$

The c statistic (which is not to be confused with the concordance statistic of the same name) informs us that in order for the effect of nonignorability to be as large as the sampling error, the magnitude of nonignorability needs to be at least as large as that where a $\frac{1}{c}$ SD change in

Y is associated with an odds ratio of 2.7 in the probability of being missing. For MISNI

from Eqn (14), $c = \left| \frac{\sigma_Y \text{SE}}{\text{MISNI}} \right|$. The MISNI c statistic informs us that for effect of

nonignorability to be as large as the sampling error, the vector nonignorability parameter γ_1 needs to lie on a hypercube with its size from the origin at least as large as c standardized units of nonignorability as defined above; the hypercube of this size implies that a $\frac{1}{c}$ SD change in Y is associated with an odds ratio of 2.7 for dropout (or intermittent missingness) versus being observed.

For large c , only extreme nonignorability can make the estimate change substantially, and thus sensitivity to nonignorability is of little concern. For example, $c = 10$ implies that in order for the error in an MAR estimate to be the same size as its sampling error, the nonignorability needs to be strong enough that an SD/10 change in Y causes a significant change in the odds of being missing. For small c , modest departure from MAR can cause the

estimate to change substantially. For example, $c = \frac{1}{10}$ implies that when even when we need a 10-SD change in Y to cause a significant change in the odds of being missing, the estimate may change substantially. As this degree of nonignorability is plausible in many applications, $c = \frac{1}{10}$ signals sensitivity. Troxel *et al.* (2004) suggest using $c < 1$ as a rule of thumb to identify sensitivity.

Because ISNI is a derivative, we have in general that $\text{ISNI}(A \cdot \theta) = A \cdot \text{ISNI}(\theta)$ for any matrix A that has the same number of columns as there are elements in θ (where $\cdot$ here represents matrix multiplication). Thus, we can readily conduct a separate ISNI analysis (including computation of c) for any element of θ .

3. Program Description and Usage

We describe here **isni**'s three main functions: `isniglmm()`, `isnimgm()`, and `isnilmm()`.

3.1. Function interface

The **isni** functions share a common interface that enables specification of the complete-data model for y , the MDM, and the data.

Function `isniglmm()` computes ISNI for the univariate GLM:

```
isniglmm = function(formula, family=gaussian, data, weights, subset,
start=NULL, offset).
```

formula is an object of model formulas — at a minimum a two-sided formula that specifies the complete-data model using the variable names for the outcome y_i and the predictors x_i in Eqn (4). Alternatively one can specify a two-equation model formula that additionally specifies the MDM using the variable name for the missingness indicator g_i and the missingness predictor s_i in Eqn (5). Details appear in Section 3.2.

family describes the error distribution to be used in the GLM for the outcome y in Eqn (4). The options are `gaussian` (the default), `poisson`, `binomial`, `gamma` and `inverse.gaussian`. The current version implements ISNI computation with the canonical link.

data names the data frame containing the model variables; the `y` element must include both the non-missing and missing observations, the latter indicated by NA.

weights is an optional vector of “prior weights” to be used in the fitting process for the complete-data model and the MDM.

subset is an optional vector specifying a subset of observations to be used in the fitting process for the outcome model and the MDM.

start is a starting value for the parameters in the linear predictor of the outcome model.

offset is an optional vector specifying an *a priori* known component in the linear predictor of the outcome GLM. It can be NULL or a numeric vector of length equal to the number of observations.

Function `isnimgm()` computes ISNI for the MMGM:

```
isnimgm=function(formula, data, cortype="CS", id, subset, weights,
predprobobs, misni=FALSE)
```

formula is an object of model formulas — at a minimum a two-sided formula that specifies the complete-data model using the variable names for the outcome y_i and the predictors x_i in Eqn (7). Alternatively, one can specify a two-equation model formula that additionally specifies the MDM using the variable names for the missing status variable g_{ij} and the missingness predictor s_{ij} in Eqn (11). Details appear in Section 3.2.

data is as above.

cortype describes the within-subject correlation structure of Σ_i in Eqn (7). The currently-available options are “CS” (the default), “AR1”, and “UN” for compound symmetry, autoregressive order 1, and unstructured, respectively.

id designates the level-2 clustering variable.

weights is an optional vector of frequency weights to be assigned to each id. When absent, each id is weighted equally.

subset is as above.

predprobobs is NULL if using the built-in multinomial first-order transitional logistic model to obtain predicted probabilities of being observed; otherwise, the user can supply the name of the variable in data that gives these probabilities.

misni is FALSE if using the default approach to computing ISNI with a scalar nonignorability parameter, and TRUE when computing ISNI with multiple nonignorability parameters.

Function `isnilmm()` computes ISNI for the LMM:

```
isnilmm=function(formula, data, random, id, weights, subset, predprobobs,
misni=FALSE)
```

formula is an object of model formulas — at a minimum a two-sided formula that specifies the complete-data model using the variable names for the outcome y_i and the predictors x_i in Eqn (8). Alternatively, one can specify a two-equation model formula that additionally specifies the MDM using the variable names for the missing status variable g_{ij} and the missingness predictor s_{ij} in Eqn (11). Details appear in Section 3.2.

data is as above.

random is a one-sided formula that specifies the random-effects part of the complete-data model LMM using the variables names for z_i in Eqn (8). The program uses a general positive-definite variance-covariance matrix V_b for the random effects.

id designates the level-2 clustering variable.

subset is as above.

weights is an optional vector of frequency weights to be assigned to each id. When absent, each id is weighted equally.

predprobabs NULL if using the built-in multinomial first-order transitional logistic model to obtain predicted probabilities of being observed; otherwise, the user can supply the name of the variable in data that gives these probabilities.

misni is FALSE if using the default approach to computing ISNI with a scalar nonignorability parameter, and TRUE when computing ISNI with multiple nonignorability parameters.

3.2. Model speci_cation

One specifies the variables in the complete-data model and MDM in the formula argument. The user must supply at least a single equation for the complete-data model in the form $\text{response} \sim \text{Xterms}$ where response is the (numeric or factor) vector for the outcome, and Xterms is a series of terms, separated by + operators, that specify the linear predictor. For the longitudinal setting, both `isnimgm` and `isnilmm` by default use the utility function `definemissingstatus` provided in the package to generate the missingness status variables g_{ij} , and then use Xterms as the predictors s_j for fitting a first-order transitional MDM in Eqn (11). With longitudinal data, it is necessary to sort observations by time within id.

The single-equation model specification uses the x_i defined for the complete-data model as s_i in the MDM. To use different sets of predictors, one specifies a two-equation formula. For this purpose, the formula argument in the **isni** functions uses the R package **Formula** (Zeileis and Croissant 2010). For `isniglm`, one can specify the model with response $| \text{is.na}(\text{response}) \sim \text{Xterms} | \text{Sterms}$, which designates the complete-data model as $\text{response} \sim \text{Xterms}$ and the MDM under MAR as $\text{is.na}(\text{response}) \sim \text{Sterms}$. For both `isnimgm` and `isnilmm`, one can write $\text{response} | \text{miss} + \text{missprior} \sim \text{Xterms} | \text{Sterms}$. This specifies the outcome model as $\text{response} \sim \text{Xterms}$, and gives the MDM of Eqn (11) as $\text{miss} \sim \text{missprior} + \text{Sterms}$, where miss and missprior are the variable names in data denoting the missingness status at the current visit and prior visits, respectively, and Sterms is the MDM predictor s_i .

For `isnilmm`, $\text{response} \sim \text{Xterms}$ identifies the fixed-effect part of the linear mixed-effects model for the outcome. One specifies the random-effect part of the model as a one-sided formula via the argument **random**.

3.3. Formatting data for an ISNI analysis

The argument **data** names the input data frame. The user must observe two rules when preparing it for input: First, except for the missingness status variables, the columns in the master dataset should include all the variables (i.e., response and explanatory variables) in

both the complete-data model and the MDM. For independent data, the missingness status variable is simply an indicator for missingness and is generated as `is.na(response)` inside `isniglm` if not provided in the formula argument. For longitudinal data, the missingness status variable G_{ij} can have three categories: “0” (being observed), “I” (intermittently missingness) and “D” (dropout). Users can define and supply the missingness status variables at the current and the prior visit as two separate columns in data and pass them in via the formula argument to estimate a first-order transitional MDM. **isni** provides a utility function `definemissingstatus` to generate the missingness status variables for users. `isnimgm` and `isnilmm` call `definemissingstatus` to generate missingness status variables if the user does not supply them. Again, it is necessary to sort beforehand by time within `id`.

Independent data.—We array the data in a rectangular matrix, with rows representing units. Fields include the main outcome of interest with missing values denoted as NA, the predictors x for the outcome model, the predictors s in the MDM, and optionally the missingness indicator variable.

Longitudinal/clustered data.—Here we have multiple level-1 observations (e.g., repeated measures) within a level-2 unit. The name of the level-2 variable appears in the argument `id`. The level-1 observations for the unit occupy a set of rows identified by a common value of `id`. Missing values are denoted NA.

Observations with missing outcomes contribute to the sensitivity analysis through $\nabla^2 L_{\theta, \gamma_1}$.

Thus, the master dataset must include places for all planned observations, present or missing. This differs from a standard ignorable analysis, where often one can simply omit the missing observations. An exception occurs in longitudinal data with dropout, because when a subject leaves the study, the probability of observing outcomes is 0 for all subsequent times. Thus one can omit all visits after the dropout visit from the database.

The fields in the master dataset for longitudinal/clustered data consist of the level-2 variable, the dependent variable y , the predictors x for the outcome model and S for the MDM, and optionally the variables for missingness status at the current and prior visit. The order is irrelevant. Both models include an intercept by default.

If users supply the missingness status variable in the master dataset, it should be a character or factor with permitted values 0 (observed), I (intermittently missing), and D (dropout). Because the built-in MDM employs a first-order Markov model that depends on the previous missingness status, users should also supply a variable that denotes missingness status at the prior visit; this is a character/factor variable taking values of 0, I, D, or U (for the baseline observation, which has no prior visit).

3.4. Handling missing values in covariates

Currently, **isni** assumes that predictors are fully observed and handles missing values in the supplied x and s as follows: Observations with missing values in x are excluded, and the program notes in the R console that it has dropped observations due to missing values in the outcome covariates. If in the remaining data there are missing values in s , to avoid further

reducing the sample size the program detects and removes any variables that contain missing values. In the future we plan to implement an alternative ISNI method that can incorporate observations with missing values in covariates; see the Discussion below.

4. Examples

Next we describe some applications of the **isni** package.

4.1. ISNI analysis of a GLM for a cross-sectional survey

Raab and Donnelly (1999) analyzed a cross-sectional survey of sexual practices among students at the University of Edinburgh. The response variable is the students' answer to the question "Have you ever had sexual intercourse?". Many declined to answer, leading to substantial missing data. We consider a simplified data set consisting of the response variable, with the student's sex and faculty as predictors.

```
# load the library and data set

> library(isni)

> data(sos)

> sos[sample(nrow(sos),10),]
```

	sexact	gender	faculty
2845	<NA>	male	other
1249	yes	male	other
5406	<NA>	female	other
3132	<NA>	male	mdv
225	yes	male	other
5812	<NA>	female	other
5115	<NA>	female	other
769	yes	male	other
2366	<NA>	male	other
2622	<NA>	male	other

This code loads **isni** and the data frame `sos`, and displays a random subsample of 10 records. `sos` includes the following factor variables: `sexact` is the response to the question "Have you ever had sexual intercourse?" (no [reference], yes); `gender` is the student's sex (male [reference], female); `faculty` is the student's faculty (medical/dental/veterinary, other [reference]).

We estimated an ignorable model predicting response from sex, faculty and their interaction:

```
> ymodel= sexact ~ gender*faculty

> summary(glm(ymodel,family=binomial, data=sos))
```

Call:

```
glm(formula=ymodel, family=binomial, data=sos)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.6713	-1.3282	0.7540	0.7642	1.0338

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	1.08153	0.05561	19.448	< 2e-16 ***
genderfemale	0.03081	0.07958	0.387	0.699
facultymdv	-0.73389	0.14921	-4.918	8.73e-07 ***
genderfemale:facultymdv	0.10213	0.20670	0.494	0.621

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 4450.3 on 3827 degrees of freedom

Residual deviance: 4408.2 on 3824 degrees of freedom

AIC: 4416.2

Number of Fisher Scoring iterations: 4

The estimates show that students in a medical faculty were less likely to report having had sexual intercourse; the other factors are not statistically significant.

With only 62.4% responding to the sexual practice question, there is concern that this analysis is sensitive to the assumption of ignorability. Troxel *et al.* (2004) conducted an ISNI analysis, which we replicate here with `isniglm()`. We posit a nonignorable selection model in the form of Eqn (5) with the missingness predictor s_i including gender, faculty and their interaction:

```
> sos.isni<-isniglm(ymodel, family=binomial, data=sos)
```

```
> sos.isni
```

Call:

```
isniglm(formula=ymodel, family=binomial, data=sos)
```

ISNIs:			
(Intercept)	genderfemale	facultymdv	genderfemale:facultymdv
0.410141	-0.038983	-0.169859	0.027542
c statistics:			
(Intercept)	genderfemale	facultymdv	genderfemale:facultymdv
0.13559	2.04146	0.87846	7.50482

Residual Deviance of the MAR model: 4408.2

AIC of the MAR model: 4416.2

The **isni** summary function expresses the isniglm() object:

```
> summary(sos.isni)
```

Call:

```
isniglm(formula=ymodel, family=binomial, data=sos)
```

	MAR Est.	Std. Err	ISNI	c
(Intercept)	1.081531	0.055611	0.410141	0.1356
genderfemale	0.030808	0.079583	-0.038983	2.0415
facultymdv	-0.733886	0.149215	-0.169859	0.8785
genderfemale:facultymdv	0.102133	0.206696	0.027542	7.5048

The columns “MAR Est.” and “Std. Err” denote the logistic model estimates and their standard errors under MAR; “ISNI” and “c” denote the ISNI and c statistics. The ISNIs are equal in absolute values to those reported in Troxel *et al.* (2004), but with opposite signs because our package models the probability that an observation is missing rather than the probability that it is observed. Recall that ISNI denotes the approximate change in the MLEs when γ_1 in the selection model is changed from 0 to 1. Under our nonignorable selection model, the assumption $\gamma_1 = 1$ means that a student whose answer is “yes” has an increase of 2.7-fold in the odds of nonresponse. Thus, subjects whose true value is “yes” are more likely to have a missing value, and the naïve MAR estimate for (Intercept) should be less than the (Intercept) estimate under the correct nonignorable model. The positive sign of the ISNI value for (Intercept) is consistent with this prediction. The ISNI for the faculty predictor is -0.17, indicating that if, as is more plausible here, $\gamma_1 = 1$, the MLE for the estimate should change from -0.73 to -0.90. If $\gamma_1 = -1$, the estimate would change from -0.73 to -0.56. The c statistics for (Intercept) and faculty are both less than 1, suggesting that these coefficients are sensitive to nonignorability. Raab and Donnelly also found that neither the gender nor the gender:faculty interaction term should be sensitive, as our findings confirm.

In the above we do not explicitly specify an MDM. One can generate an identical analysis with the two-equation model formula `sexact | is.na(sexact) ~ gender*faculty | gender*faculty`, which uses the operator `|` to separately specify variables used in the complete-data model and the MDM:

```
> ygmodel= sexact | is.na(sexact) ~ gender*faculty | gender*faculty
> summary(isniglm(ygmodel, family=binomial, data=sos))
```

Call:

```
isniglm(formula=ygmodel, family=binomial, data=sos)
```

	MAR Est.	Std. Err	ISNI	c
(Intercept)	1.081531	0.055611	0.410141	0.1356
genderfemale	0.030808	0.079583	-0.038983	2.0415
facultymdv	-0.733886	0.149215	-0.169859	0.8785
genderfemale:facultymdv	0.102133	0.206696	0.027542	7.5048

Because all the covariates are categorical, one can also analyze the data using a grouped binomial regression with the `weights` argument in `isniglm`:

```
> gender <- c(0,0,1,1,0,0,1,1)
> faculty <- c(0,0,0,0,1,1,1,1)
> gender <- factor(gender, levels=c(0, 1), labels =c("male", "female"))
> faculty <- factor(faculty, levels=c(0, 1), labels =c("other", "mdv"))
> SAccount <- c(NA, 1277, NA, 1247, NA, 126, NA, 152)
> total <- c(1189,1710,978,1657,68,215,73,246)
> sosgrp <- data.frame(gender=gender, faculty=faculty, SAccount=SAccount,
total=total)
> ymodel <- SAccount/total ~gender*faculty
> sosgrp.isni<-isniglm(ymodel, family=binomial, data=sosgrp, weights=total)
> summary(sosgrp.isni)
```

Call:

```
isniglm(formula=ymodel, family=binomial, data=sosgrp, weights=total)
```

	MAR Est.	Std. Err	ISNI	c
(Intercept)	1.081531	0.055611	0.410141	0.1356
genderfemale	0.030808	0.079583	-0.038983	2.0415
facultymdv	-0.733886	0.149215	-0.169859	0.8785
genderfemale:facultymdv	0.102133	0.206696	0.027542	7.5048

4.2. Illustration: An ISNI analysis of simulated data

Here we illustrate ISNI analysis using simulated data with known MDMs. Specifically, we generated one complete data set with a sample size of 10,000 under the logistic model $\text{logit}(\Pr[Y = 1|x]) = \beta_0 + \beta_1 x$ with $\beta_0 = 1$, $\beta_1 = -0.7$. The covariate x was generated as Bernoulli with probability 0.5. The MDM was the logistic regression $\text{logit}(\Pr[G = 1|y]) = \gamma_{00} + \gamma_{01}x + \gamma_{1y}$. We simulated one missingness pattern under each of the four parameter settings $\gamma = (\gamma_{00}, \gamma_{01}, \gamma_1) = (-2, -0.5, -1)$, $(-2, -0.5, 0)$, $(1.2, -0.5, -1)$, and $(1.2, -0.5, 0)$. The datasets had amounts of missingness ranging from 5% to 70%. We applied the `isniglm()` to each simulated dataset; results appear in Table 1.

The columns “Est” and “SE” denote the outcome logistic model estimate for β_0 or β_1 (as listed under “ θ ”) and its standard error under MAR; “%[$G = 1$]”, “ISNI”, and “ c ” denote the missing proportion, ISNI, and c statistic, respectively. One can use $\text{ISNI} \times \gamma_1$ to approximate the change in the estimate under the posited value of γ_1 . For example, using the true value $\gamma_1 = -1$ from Dataset C, the estimate of β_0 changes to $\text{Est} + \text{ISNI} \times \gamma_1 = 1.651 + 0.607 \times (-1) = 1.044$, which is close to the true value of 1.0; the estimate of β_1 changes to $-0.782 + (-0.081) \times (-1) = -0.701$, which is close to the true value of -0.7 . Within each dataset, the estimates of β_0 and β_1 can have quite different ISNI and c values, suggesting that the missing proportion alone is insufficient to capture the individual sensitivity of different model parameters. The four datasets are sorted by missing proportion from low to high. We observe that for both β_0 and β_1 , the absolute ISNI value increases and c value decreases as the fraction of data missing increases.

4.3. ISNI analysis of an MMGM for longitudinal data

Ma *et al.* (2005) analyzed data from a randomized trial comparing flutamide with a placebo in the treatment of prostate cancer. A sub-study collected QoL outcomes at baseline and 1, 3, and 6 months after randomization. We focus here on the emotional functioning (EF) scale outcome. Below is a sample of the data set `qolef`.

```
> qolef[1:12,]
```

	id	y	time	group	perf	sever	yp	g	gp	basey
1	117938	8.485281	0	0	0	1	0.000000	0	U	8.485281
2	117938	8.485281	1	0	0	1	8.485281	0	0	8.485281
3	117938	NA	3	0	0	1	8.485281	1	0	8.485281

4	117938	8.717798	6	0	0	1	8.485281	0	I	8.485281
5	124149	10.000000	0	1	0	1	0.000000	0	U	10.000000
6	124149	10.000000	1	1	0	1	10.000000	0	0	10.000000
7	124149	8.000000	3	1	0	1	10.000000	0	0	10.000000
8	124149	9.797959	6	1	0	1	8.000000	0	0	10.000000
9	124674	9.591663	0	0	0	1	0.000000	0	U	9.591663
10	124674	9.380832	1	0	0	1	9.591663	0	0	9.591663
11	124674	NA	3	0	0	1	9.380832	I	0	9.591663
12	124674	9.165151	6	0	0	1	9.380832	0	I	9.591663

The variables in `qo le f` are as follows:

id — patient id

y — EF score

time — time in months since randomization

group — placebo (0) or flutamide (1)

perf — baseline performance score

sever — baseline disease severity

yp — most recently observed prior outcome

g — missingness status (O=observed, D=dropout, I=intermittently missing)

gp — missingness status in the prior visit (as above, plus U=undefined)

basey — EF at baseline

We seek to evaluate the drug effect on EF over time. The original EF is on a scale of 0 (worst) to 100 (best); as in Ma *et al.* (2005), we transform to the square-root scale. Table 2 presents the mean and SD of transformed EF and the fraction observed, by treatment arm and visit. We see that missingness percentages are comparable between arms but increase over time, with almost a quarter of the subjects missing by the final visit. Our analysis excludes a small number of patients ($\approx 3\%$) whose EF data were missing at baseline. Table 3 presents missing data patterns for the longitudinal EF outcomes (omitting subjects with missing baseline) and shows that there are both dropouts and intermittently missing items.

Xie (2012) presents an ISNI analysis of the impact of nonignorable nonmonotone missingness on the MAR estimates for this dataset. Here we conduct an ISNI analysis for the marginal multivariate Gaussian model, using function `isningm()`. We take the predictor vector to be

$$x_i = (\text{Intercept, perf, sever,} \\ \text{T1.0(pb), T3.0(pb), T6.0(pb),} \\ \text{T0(fl) - T0(pb), T1.0(fl) - T1.0(pb), T3.0(fl) - T3.0(pb), T6.0(fl) - T6.0(pb)})_i.$$

The two predictors “perf” and “sever” are baseline covariates. The predictor $T0(a)$ is an indicator for the baseline observation in arm a . The predictors $Tt0(a)$ are contrasts of time t vs. baseline in arm a , $t = 1, 3, 6$. Thus the third line of predictors gives contrasts of these contrasts between arms. One can specify the complete-data model as $y \sim \text{perf} + \text{sever} + \text{as.factor}(\text{time}) + \text{group} + \text{as.factor}(\text{time}): \text{group}$. To evaluate the robustness of the MAR analysis to nonignorable missingness, we assume the first-order transitional model of Eqn (11), where the missingness status variables at the current visit and at the prior visit are g and gp in `qolef`, respectively; the missingness predictor s_i is $\text{as.factor}(\text{time}) * \text{group} + \text{yp} + \text{perf} + \text{sever}$; and yp is the most recently observed outcome prior to the current visit. We apply `isnimgm()` to perform the ISNI analysis:

```
> models= y | g + gp ~ p erf + sever + as.factor(time) + group
+as.factor(time):group | group*as.factor(time) + yp + perf + sever

> qolef.isni=isnimgm (models, data=qolef, id=id)

# weights: 36 (22 variable)

initial value 2144.491187

iter 10 value 797.136438

iter 20 value 780.985037

iter 30 value 779.568898

final value 779.561688

converged

# weights: 12 (11 variable)

initial value 52.679186

iter 10 value 19.792933

iter 20 value 19.572566

iter 30 value 19.570615

iter 40 value 19.570445

final value 19.570432

converged

> qolef.isni
```


Call:

```
isningm(formula = models, data = qolef, id = id)
```

ISNIs:

[1]	-0.0267878	0.0873426	0.0306558	0.1413056	0.1363404	0.1790565
[7]	-0.0011816	-0.0144459	-0.0210095	-0.0308886	-0.0072580	-0.0076511

c statistics:

[1]	5.12102	3.35252	4.35833	0.66656	0.70197	0.55226	112.65390
[8]	9.13534	6.36963	4.45768	4.64884	3.43062		

logLik of the MAR model: -3859.8

AIC of the MAR model: 7743.6

BIC of the MAR model: 7813.6

We summarize the results as follows:

```
> summary(qolef.isni)
```

Call:

```
isningm(formula=models, data=qolef, id=id)
```

	MAR Est.	Std. Err	ISNI	c
(Intercept)	8.430611	0.102762	-0.0267878	5.1210
perf	-0.270986	0.219348	0.0873426	3.3525
sever	-0.146629	0.100085	0.0306558	4.3583
as.factor(time)1	0.452168	0.070556	0.1413056	0.6666
as.factor(time)3	0.493949	0.071693	0.1363404	0.7020
as.factor(time)6	0.362189	0.074074	0.1790565	0.5523
group	-0.020181	0.099713	-0.0011816	112.6539
as.factor(time)1:group	-0.220661	0.098857	-0.0144459	9.1353
as.factor(time)3:group	-0.220821	0.100246	-0.0210095	6.3696
as.factor(time)6:group	-0.226147	0.103144	-0.0308886	4.4577
sigma	1.328627	0.025275	-0.0072580	4.6488
rho	0.548157	0.019662	-0.0076511	3.4306

isningm() estimates the MAR model using gls () in R package **nlme**. The ignorable analysis suggests that, after adjustment for baseline performance and severity, placebo gives statistically significantly better EF at all three follow-up visits (estimates for

as.factor(time)1:group, as.factor(time)3:group, and as.factor(time)6:group are significant and negative).

“ISNI” values quantify the potential change from the MAR estimates with $\gamma_1 = 1$. The scale-independent c statistics suggest that the time effect estimates as.factor(time) 1, as.factor(time)3, and as.factor(time)6 are sensitive to nonignorability. The treatment contrasts at follow-up — as.factor(time)1:group, as.factor(time)3:group, and as.factor(time)6:group — are insensitive, with $c > 1$.

One can designate the model formula without identifying the missingness variable. In this case, users should first sort data by time within id:

```
> qolef <- qolef[order(qolef$id, qolef$time),]

> models1=y ~ perf + sever + as.factor(time) + group + as.factor(time):group
| as.factor(time)*group + yp + perf + sever

> qolef.isni <- isnimgm(models1, data=qolef, id=id)
```

To posit an MDM other than that of Eqn (11), one can supply predicted probabilities of being observed under the desired model through the optional argument predprobobs. The R code below uses function tmdm() to obtain the predicted probabilities of being observed for all observations and passes them in as the vector predprobobs. isnimgm() then refrains from fitting the default first-order MDM and instead uses predprobobs to evaluate ISNI:

```
> predprobobs=tmdm(g+gp ~ as.factor(time)*group + yp + perf + sever,
data=qolef)$obsprob

> summary(isnimgm(models, data=qolef, id=id, predprobobs=predprobobs))
```

Call:

```
isnimgm(formula=models, data=qolef, id=id, predprobobs=predprobobs)
```

	MAR Est.	Std. Err	ISNI	c
(Intercept)	8.430611	0.102762	-0.0267878	5.1210
perf	-0.270986	0.219348	0.0873426	3.3525
sever	-0.146629	0.100085	0.0306558	4.3583
as.factor(time)1	0.452168	0.070556	0.1413056	0.6666
as.factor(time)3	0.493949	0.071693	0.1363404	0.7020
as.factor(time)6	0.362189	0.074074	0.1790565	0.5523
group	-0.020181	0.099713	-0.0011816	112.6539
as.factor(time)1:group	-0.220661	0.098857	-0.0144459	9.1353
as.factor(time)3:group	-0.220821	0.100246	-0.0210095	6.3696
as.factor(time)6:group	-0.226147	0.103144	-0.0308886	4.4577

sigma	1.328627	0.025275	-0.0072580	4.6488
rho	0.548157	0.019662	-0.0076511	3.4306

By default, `isningm()` uses the compound symmetry correlation structure in the outcome model of Eqn (7). One can specify alternative correlation structures via the argument `cortype`. The current implementation permits two other correlation models: AR1 and unstructured. The code below illustrates the use of this optional argument.

```
> qolef.isni.ar=isningm(models, id=id, cortype='AR1', data=qolef)

> qolef.isni.un=isningm(models, id=id, cortype='UN', data=qolef)

> summary(qolef.isni.ar)
```

Call:

```
isningm(formula=models, data=qolef, cortype="AR1", id=id)
```

	MAR Est.	Std. Err	ISNI	c
(Intercept)	8.428061	0.102070	-0.0302202	4.5088
perf	-0.233914	0.216802	0.0976118	2.9650
sever	-0.154242	0.097230	0.0346554	3.7453
as.factor(time)1	0.455742	0.064515	0.1157535	0.7440
as.factor(time)3	0.488756	0.082559	0.1395210	0.7899
as.factor(time)6	0.385020	0.094246	0.2064043	0.6095
group	-0.012408	0.101323	-0.0014306	94.5480
as.factor(time)1:group	-0.227693	0.090418	-0.0075279	16.0341
as.factor(time)3:group	-0.211977	0.115574	-0.0195549	7.8899
as.factor(time)6:group	-0.243204	0.131493	-0.0327322	5.3628
sigma	1.350494	0.025311	-0.0118938	2.8408
rho	0.634149	0.016649	-0.0105955	2.0976

```
> summary(qolef.isni.un)
```

Call:

```
isningm(formula=models, data=qolef, cortype="UN", id=id)
```

	MAR Est.	Std. Err	ISNI	c
(Intercept)	8.434487	0.105008	-0.0260590	5.3800
perf	-0.310585	0.224305	0.09139557	3.2763
sever	-0.150384	0.102388	0.02918823	4.6828
as.factor(time)1	0.456740	0.074576	0.13510182	0.7369

as.factor(time)3	0.474532	0.077807	0.12967221	0.8010
as.factor(time)6	0.350548	0.081839	0.17288868	0.6319
group	-0.019578	0.101797	-0.00099203	136.9847
as.factor(time)1:group	-0.226479	0.104531	-0.00830169	16.8090
as.factor(time)3:group	-0.203433	0.108956	-0.01523747	9.5456
as.factor(time)6:group	-0.213838	0.114211	-0.02430005	6.2743
sigma	1.355958	0.025921	-0.00669802	5.1661
cor(1,2)	0.508482	0.024711	-0.00656804	5.0225
cor(1,3)	0.472153	0.026098	0.00164597	21.1662
cor(1,4)	0.445383	0.028275	-0.00068605	55.0178
cor(2,3)	0.707501	0.018507	-0.00792585	3.1171
cor(2,4)	0.599332	0.027904	-0.00927305	4.0171
cor(3,4)	0.734721	0.016464	-0.00439757	4.9979

Our results suggest that the choice of correlation model has little impact on either the MAR parameter estimates or the sensitivity to nonignorability.

The QoL dataset had both intermittent missingness ($\approx 10\%$) and dropout ($\approx 20\%$). In some applications, the missingness can be of only one type or the other, and thus the MDM reduces to a special case of the more general multinomial transitional model used in our package. To demonstrate the use of `isnimgm()` in these situations, we conduct ISNI analysis on two subsamples of QoL datasets; the first contains only complete observations and dropouts, and the second contains only complete observations and those with intermittent missingness.

```
> ## Run ISNI analysis on the subset that excludes intermittent missingness.

> qolefD.isni=isnimgm(models, id=id, data=qolef, subset= g!="I")

> summary(qolefD.isni)

Call:
isnimgm(formula=models, data=qolef, id=id, subset= g!="I")
```

	MAR Est.	Std. Err	ISNI	c
(Intercept)	8.430611	0.102762	-0.02718564	5.0461
perf	-0.270986	0.219348	0.10189567	2.8737
sever	-0.146629	0.100085	0.03030158	4.4093
as.factor(time)1	0.452168	0.070556	0.07554213	1.2468
as.factor(time)3	0.493949	0.071693	0.08496271	1.1265
as.factor(time)6	0.362189	0.074074	0.18210075	0.5430
group	-0.020181	0.099713	-0.00097502	136.5219
as.factor(time)1:group	-0.220661	0.098857	0.00082456	160.0474

as.factor(time)3:group	-0.220821	0.100246	-0.01431582	9.3479
as.factor(time)6:group	-0.226147	0.103144	-0.02851807	4.8282
sigma	1.328627	0.025275	-0.00821263	4.1085
rho	0.548157	0.019662	-0.00866653	3.0287

```
> ## Run ISNI analysis on the subset that excludes dropouts.

> qolefI.isni=isnimgm(models, id=id, data=qolef, subset= g != "D")

> summary(qolefI.isni)
```

Call:

```
isnimgm(formula=models, data=qolef, id=id, subset= g!="D")
```

	MAR Est.	Std. Err	ISNI	c
(Intercept)	8.430611	0.102762	-0.00138914	98.7522
perf	-0.270986	0.219348	-0.00974692	30.0420
sever	-0.146629	0.100085	0.00246156	54.2778
as.factor(time)1	0.452168	0.070556	0.07476190	1.2598
as.factor(time)3	0.493949	0.071693	0.05836248	1.6399
as.factor(time)6	0.362189	0.074074	-0.00056459	175.1457
group	-0.020181	0.099713	-0.00030265	439.8188
as.factor(time)1:group	-0.220661	0.098857	-0.01647702	8.0092
as.factor(time)3:group	-0.220821	0.100246	-0.00869283	15.3946
as.factor(time)6:group	-0.226147	0.103144	-0.00321366	42.8457
sigma	1.328627	0.025275	0.00029717	113.5440
rho	0.548157	0.019662	0.00025485	102.9953

ISNIs from both `summary(qolefD.isni)` and `sum MARY(qolefI.isni)` are generally smaller than those from the original dataset as shown in `sum MARY(qolef.isni)`. This is expected, because sensitivity increases with the proportion of missingness in a dataset.

Finally, we illustrate the computation of MISNI. As compared with `sum mary(qolef.isni)`, which considers a scalar γ_1 , MISNI shows a slight increase in sensitivity:

```
> qolef.misni=isnimgm(models, id=id, data=qolef, misni=T)
```

```
> summary(qolef.misni)
```

Call:

```
isnimgm(formula=models, data=qolef, id=id, misni=T)
```

	MAR Est.	Std. Err	MISNI	c
(Intercept)	8.430611	0.102762	0.0269852	5.0836
perf	-0.270986	0.219348	0.1154852	2.5355
sever	-0.146629	0.100085	0.0306558	4.3583
as.factor(time)1	0.452168	0.070556	0.1413056	0.6666
as.factor(time)3	0.493949	0.071693	0.1363404	0.7020
as.factor(time)6	0.362189	0.074074	0.1853873	0.5334
group	-0.020181	0.099713	0.0011816	112.6539
as.factor(time)1:group	-0.220661	0.098857	0.0172323	7.6582
as.factor(time)3:group	-0.220821	0.100246	0.0210095	6.3696
as.factor(time)6:group	-0.226147	0.103144	0.0318087	4.3287
sigma	1.328627	0.025275	0.0092060	3.6652
rho	0.548157	0.019662	0.0097592	2.6896

In the first-order transitional MDM, $ISNI = (ISNI_{\gamma_1^{10}}, ISNI_{\gamma_1^{20}}, ISNI_{\gamma_1^{11}})$ (Eqn (13)) is a

sensitivity vector whose components approximate the change in the estimate when the corresponding element in γ_1 is perturbed from 0 to 1. The above ISNI analysis provides parsimonious measures that combine these three elements to quantify sensitivity to both nonignorable dropout and intermittent missingness. The program outputs the sensitivity vector as a component called “isnivec” for the object “isnimgm” (and also the object “isnilmm” described in the next subsection), which allows one to assess sensitivity by missingness types as follows:

```
> qolef.isni$isnivec
```

	ISNI_IO	ISNI_DO	ISNI_II
(Intercept)	-0.0003187537	-0.0265677652	9.870601e-05
perf	-0.0109753859	0.1014139023	-3.095925e-03
sever	0.0010985018	0.0295008814	5.645147e-05
as.factor(time)1	0.0666834484	0.0713312390	3.290869e-03
as.factor(time)3	0.0394565261	0.0828477934	1.403606e-02
as.factor(time)6	-0.0023044904	0.1822219019	-8.609470e-04
group	-0.0002171978	-0.0009218452	-4.254895e-05
as.factor(time)1:group	-0.0145586659	0.0013932094	-1.280454e-03
as.factor(time)3:group	-0.0015131409	-0.0137899008	-5.706474e-03
as.factor(time)6:group	-0.0030790572	-0.0282696341	4.600442e-04
sigma	0.0009739717	-0.0081477743	-8.421882e-05
rho	0.0010540298	-0.0085868571	-1.182775e-04

The column names “ISNI_IO”, “ISNI_DO” and “ISNI_II” represent sensitivity associated with nonignorability parameter $\gamma_1^{10}, \gamma_1^{20}, \gamma_1^{11}$, respectively. Thus they indicate sensitivity

caused by missingness configuration “IO”, “DO” and “II”, respectively with the first and second letter inside each quote denoting the missingness types (O (observed), I (intermittently missing), and D (dropout)) in the current and prior visits, respectively. One can also assess the total sensitivity to intermittent missingness by combining “ISNI_IO” and “ISNI_II”.

4.4. ISNI analysis of an LMM for longitudinal data

Finally, we illustrate the LMM analysis using the QoL data. We first consider a random intercept model with the output below. This model is equivalent to the marginal multivariate model with compound symmetry correlation structure illustrated above. Consequently they produce similar MAR and ISNI results.

```
> data(qolef)

> models=y | g + gp ~ perf + sever + as.factor(time) + group +
as.factor(time):group | group*as.factor(time) + yp + perf + sever

## Random intercept model

> result=isnilmm(models, random=~1, id=id, data=qolef)

> summary(result)

Call:
isnilmm(formula=models, data=qolef, random=~1, id=id)
```

	MAR Est.	Std. Err	ISNI	c
(Intercept)	8.430611	0.102761	-0.0267777	5.1229
perf	-0.270986	0.219351	0.0874180	3.3497
sever	-0.146629	0.100085	0.0306377	4.3609
as.factor(time)1	0.452168	0.070555	0.1413020	0.6666
as.factor(time)3	0.493949	0.071709	0.1363900	0.7019
as.factor(time)6	0.362189	0.074101	0.1791559	0.5522
group	-0.020181	0.099713	-0.0011793	112.8717
as.factor(time)1:group	-0.220661	0.098839	-0.0144065	9.1587
as.factor(time)3:group	-0.220821	0.100255	-0.0210332	6.3630
as.factor(time)6:group	-0.226147	0.103170	-0.0309658	4.4477
sigmav	0.983684	0.033229	-0.0135647	3.2702
sigmae	0.893093	0.014957	0.0036156	5.5222

We next consider a model with random effects for both intercept and time slope. The ISNI analysis shows that the time estimate has $c < 1$, suggesting sensitivity to nonignorable missingness. It also shows that the time:group estimate for measuring the treatment group

differences over time has $c > 1$, suggesting that sensitivity to nonignorability is not of concern for this estimate. These results echo our findings under the MMGM.

```
## Random intercept and slope model

> models1=y |g + gp ~ time*group + perf + sever

> result1=isnilmm(models1,random= ~ 1+time, id=id, data=qolef)

> summary(result1)

Call:

isnilmm(formula=models1, data=qolef, random= ~1 + time, id=id)
```

	MAR Est.	Std. Err	ISNI	c
(Intercept)	8.633868	0.099125	0.0134140	9.8648
time	0.047974	0.013094	0.0294745	0.5930
group	-0.103960	0.092752	-0.0033128	37.3758
perf	-0.263945	0.221257	0.1001395	2.9495
sever	-0.152619	0.099815	0.0326422	4.0820
time:group	-0.032053	0.018209	-0.0064265	3.7824
sigmav1	1.043033	0.039745	-0.0124504	4.2614
sigmav2	0.118714	0.014123	0.0025849	7.2935
rho12	-0.327000	0.080491	-0.0344655	3.1176
sigmae	0.855890	0.017576	0.0045477	5.1592

4.5. Hardware and software specifications

isni is written in R, and can be used under Windows, Linux and Macintosh. As a requirement to pass the R CRAN submission tests, for no individual example in the package does the ISNI computation time exceed 5 seconds.

4.6. Mode of availability

isni is freely available at the R CRAN repository (<https://CRAN.R-project.org/package=isni>) and can be installed as an add-on within the R console.

5. Discussion

There is a pressing need for statistical software for analysis of sensitivity to departures from the assumption of ignorable missingness. Our new R package **isni** implements such computations in three common settings: The GLM for univariate data and the MMGM and LMM for longitudinal/clustering data. Our next planned version will cover the generalized linear mixed model (Xie 2008). Other potential extensions are to ISNI analysis under the general coarse-data model (Heitjan and Rubin 1991; Zhang and Heitjan 2005, 2006, 2007;

Liu and Heitjan 2012) and to Bayesian versions of ISNI (Zhang and Heitjan 2007; Xie 2009).

We emphasize that ISNI is a measure of *local sensitivity*, in that it quantifies the extent to which MLEs should change to reflect small departures from MAR. It is not intended to detect or estimate nonignorability; rather, it alerts the user to situations where, should nonignorability exist, estimates under an MAR model are unreliable.

ISNI is an increasing function of the fraction of missing observations, the predicted probabilities of being observed for the missing observations, and the influence of the missing observations on estimates. The most readily apparent component of sensitivity is the the fraction of missingness; indeed, in the estimation of a univariate normal mean, ISNI is proportional to the fraction of missing observations (Troxel *et al.* 2004). The positioning of missing information in the data set also affects the value of ISNI. For example, holding the number of missing observations fixed, ISNI for a slope is large if the missing data are concentrated among high-influence points, but small if they are concentrated among low-influence points; see Troxel *et al.* (2004). Note also that $\text{ISNI} = 0$, and thus $c = \infty$, whenever there are no missing observations, regardless of ignorability of the underlying true MDM.

In practice, the most effective measure one can take to reduce sensitivity is to avoid missing observations. Failing that, the next best thing is to prospectively gather robust information on the causes of missingness, which can then inform judgments on the plausible degree of nonignorability in the data.

As with regression influence diagnostics (Cook 1986), ISNI is concerned only with the data at hand — defined here as (y^{obs}, g) — and not with other data sets that might have been gathered but were not. From this perspective, the sampling distribution of ISNI is irrelevant. Again, consider the case of a univariate normal mean, where ISNI is proportional to the fraction of missing observations. Clearly, it matters not that this proportion may have been higher or lower; for the analysis at hand, it only matters what the proportion *is*.

When ISNI identifies an estimate as sensitive to nonignorability, the investigator typically has only two choices: Flag the MAR analysis as unreliable, or conduct a speculative nonignorable analysis. When further data collection is possible, one may have the additional option of taking a refreshment sample (Hirano, Ridder, and Rubin 2001), which can shed light on the MDM and inform a more robust nonignorable modeling exercise. In such a situation, ISNI can be useful in identifying targets for additional data collection. For example, an ISNI analysis of US young adult smoking data (Xie and Qian 2012) revealed that the impact of nonignorable missingness was largest in black males, suggesting the need for more persistent efforts to acquire data in that group.

One cannot always distinguish intermittent missing data from true dropout on the basis of the missingness pattern alone. For example, a missing observation at only the final visit could represent dropout (abandoning the study) or intermittent missing data (skipping a single visit). Accurate analysis thus depends not just on specifying a suitably flexible MDM, but on correctly gleaned the reasons for missed visits. Ideally, one would obtain such data

directly from the subjects; if this is not possible, one can compute the range of ISNIs for possible combinations of missingness types (Xie 2012).

Our use of a logistic model for $f_{\gamma}(g|y)$ raises the concern that ISNI itself may be sensitive to untestable assumptions about the MDM. Although investigations to date have suggested that results are, as a practical matter, insensitive to the link function and the form of the linear predictor (Xie and Heitjan 2009; Xie *et al.* 2011; Xie and Qian 2012), robustness is not infinite. Therefore it may be desirable to extend the package to incorporate generalizations of the basic MDM (Xie and Heitjan 2009; Xie *et al.* 2011). We caution, however, that this is not simply a matter of model fit; the interpretation of the logistic γ_1 as a log odds ratio at every level of y greatly simplifies the conceptualization of a sensitivity analysis.

Another limitation of our package is that it allows missing data in the outcome only, and therefore it uses an *ad hoc* approach to repairing the data for analysis when there is missingness in the predictors. The method of Gao *et al.* (2016), which addresses this problem to some extent, has not yet been implemented in a distribution package. With many observational data sets, a variable may be an “outcome” in one study and a “covariate” in another, depending on the context and objectives. Thus it is desirable to have methods that can assess sensitivity in a principled way regardless of the configuration of missing items. A general solution is to construct a comprehensive model for all the random variables, and to conduct an ISNI analysis for the parameters of the conditional distribution of the “outcomes” given the “predictors”.

There are various other approaches to the analysis of sensitivity to nonignorability. Verbeke *et al.* (2001) proposed a model that generates individual-level influence diagnostics. Recently, Mallinckrodt, Roger, Chuang-Stein, Molenberghs, O’Kelly, Ratitch, Janssens, and Bunouf (2014) proposed an approach that uses imputations from pattern-mixture models to assess how strong departures from MAR must be in order to affect primary conclusions. Although these approaches differ from ours in form and concept, they could be valuable components of a comprehensive sensitivity analysis package.

The ISNI approach is consistent with recent recommendations for handling missing data in that it involves reporting the MAR inference together with analyses that assess the impact of alternative MDM assumptions. In this spirit, we intend that the availability of our software will make such analyses a routine part of statistical practice.

Acknowledgments

The US National Cancer Institute supported our research under awards R01CA178061 and P01CA098262.

Appendix A:: Derivation of ISNI

For fixed γ_1 , the conditional maximum likelihood estimates $\hat{\theta}(\gamma_1)$ and $\hat{\gamma}_0(\gamma_1)$ satisfy

$$\frac{\partial L(\hat{\theta}(\gamma_1), \hat{\gamma}_0(\gamma_1), \gamma_1)}{\partial (\theta^T, \gamma_0^T)^T} = 0,$$

where $L(\theta, \gamma_0, \gamma_1)$ is the loglikelihood for the selection model (Eqn 1). Differentiating both sides with respect to γ_1 and noting that $\hat{\theta}(\gamma_1)$, and $\hat{\gamma}_0(\gamma_1)$ are implicit functions of γ_1 , we have

$$\frac{\partial^2 L(\hat{\theta}(\gamma_1), \hat{\gamma}_0(\gamma_1), \gamma_1)}{\partial(\theta^T, \gamma_0^T)^T \partial \gamma_1^T} + \frac{\partial^2 L(\hat{\theta}(\gamma_1), \hat{\gamma}_0(\gamma_1), \gamma_1)}{\partial(\theta^T, \gamma_0^T)^T \partial(\theta^T, \gamma_0^T)} \frac{\partial(\hat{\theta}^T(\gamma_1), \hat{\gamma}_0^T(\gamma_1))^T}{\partial \gamma_1^T} = 0.$$

Thus for any γ_1 , we have

$$\frac{\partial(\hat{\theta}^T(\gamma_1), \hat{\gamma}_0^T(\gamma_1))^T}{\partial \gamma_1^T} = - \left[\frac{\partial^2 L(\hat{\theta}(\gamma_1), \hat{\gamma}_0(\gamma_1), \gamma_1)}{\partial(\theta^T, \gamma_0^T)^T \partial(\theta^T, \gamma_0^T)} \right]^{-1} \frac{\partial^2 L(\hat{\theta}(\gamma_1), \hat{\gamma}_0(\gamma_1), \gamma_1)}{\partial(\theta^T, \gamma_0^T)^T \partial \gamma_1^T}. \quad (16)$$

In our local sensitivity analysis, the primary interest is to investigate sensitivity around the MAR model, i.e., $\gamma_1 = 0$. This local sensitivity can be captured by the derivatives at this point. In particular, we define the first derivative evaluated at $\gamma_1 = 0$ as ISNI and

$$\begin{aligned} \text{ISNI} &= \frac{\partial(\hat{\theta}^T(\gamma_1), \hat{\gamma}_0^T(\gamma_1))^T}{\partial \gamma_1^T} \Big|_{\gamma_1=0} \\ &= - \begin{bmatrix} \nabla^2 L_{\theta, \theta} & \nabla^2 L_{\theta, \gamma_0} \\ \nabla^2 L_{\gamma_0, \theta} & \nabla^2 L_{\gamma_0, \gamma_0} \end{bmatrix}^{-1} \begin{bmatrix} \nabla^2 L_{\theta, \gamma_1} \\ \nabla^2 L_{\gamma_0, \gamma_1} \end{bmatrix}, \end{aligned}$$

where $\nabla^2 L_{a,b} = \frac{\partial^2 L(\theta(\gamma_1), \gamma_0(\gamma_1), \gamma_1)}{\partial a \partial b} \Big|_{\hat{\theta}(0), \hat{\gamma}_0(0), 0}$. Under MAR, we have $\nabla^2 L_{\theta, \gamma_0} = 0$, and thus

the ISNI for $\hat{\theta}$, the parameter estimates of primary interest, is

$$\frac{\partial \hat{\theta}(\gamma_1)}{\partial \gamma_1^T} \Big|_{\gamma_1=0} = - \nabla^2 L_{\theta, \theta}^{-1} \nabla^2 L_{\theta, \gamma_1},$$

where

$$\nabla^2 L_{\theta, \theta} = \frac{\partial^2 L(\theta, \gamma_0, \gamma_1)}{\partial \theta \partial \theta^T} \Big|_{\hat{\theta}(0), \hat{\gamma}_0(0), \gamma_1=0} = \frac{\partial \ln f_{\theta}(y^{\text{obs}})}{\partial \theta \partial \theta^T} \Big|_{\hat{\theta}(0), \hat{\gamma}_0(0), \gamma_1=0}$$

and

$$\nabla^2 L_{\theta, \gamma_1} = \frac{\partial^2 L(\theta, \gamma_0, \gamma_1)}{\partial \theta \partial \gamma_1^T} \bigg|_{\hat{\theta}(0), \hat{\gamma}_0(0), \gamma_1 = 0}.$$

Appendix B:: Derivation of $\nabla^2 L_{\theta, \gamma_1}$

To derive the likelihood of the nonignorable selection model for longitudinal data, adopt the notation of $Y_i = (Y_i^{\text{obs}}, Y_i^{\text{mis}})$, where Y_i^{obs} refers to the components of Y_i that are observed, and Y_i^{mis} refers to the components of Y_i that are missing. Let K_i be the length of Y_i^{obs} . If subject i completed all the intended visits of the study, then $K_i = n_i$ and Y_i^{mis} vanishes; otherwise, $K_i < n_i$. Let L be the correct loglikelihood for $(\theta, \gamma_0, \gamma_1)$ under the nonignorable selection model specified in Sections 2.2.1 and 2.2.2. Then

$$\begin{aligned} L(\theta, \gamma_0, \gamma_1) &= \sum_{i=1}^N L_i(\theta, \gamma_0, \gamma_1; y_i^{\text{obs}}, g_i) \\ &= \sum_{i=1}^N \ln \left(\int \prod_{j=1}^{n_i} f_{\gamma}(g_{ij} | s_{ij}, y_{ij}, g_{i,j-1}) f_{\theta}(y_i^{\text{obs}}, y_i^{\text{mis}} | x_i) dy_i^{\text{mis}} \right) \\ &= \sum_{i=1}^N \ln f_{\theta}(y_i^{\text{obs}} | x_i) + \\ &\quad \sum_{i=1}^N \sum_{j: g_{ij}=0} \ln f_{\gamma}(g_{ij} | s_{ij}, y_{ij}, g_{i,j-1}) + \\ &\quad \sum_{i: K_i < n_i} \ln \left(\int \prod_{j: g_{ij} \neq 0} f_{\gamma}(g_{ij} | s_{ij}, y_{ij}, g_{i,j-1}) f_{\theta}(y_i^{\text{mis}} | y_i^{\text{obs}}, x_i) dy_i^{\text{mis}} \right), \end{aligned} \quad (17)$$

where $g_i = (g_{i1}, \dots, g_{in_i})$ is a vector of discrete variables for the missingness status of subject i ; $f_{\theta}(y_i^{\text{obs}}, y_i^{\text{mis}} | x_i)$ is the density function of the outcome model defined above; $f_{\gamma}(g_{ij} | s_{ij}, y_{ij}, g_{i,j-1})$ is the probability mass function of the missing-data model defined in Eqn (12), and if the general transitional model as specified in Eqn (9) is used for modeling G_{ij} , $f_{\gamma}(g_{ij} | s_{ij}, y_{ij}, g_{i,j-1})$ is then replaced with $f_{\gamma}(g_{ij} | s_{ij}, y_{ij}, g_{i(j)})$. We intend the integral sign to refer to summation with discrete outcomes.

The components of y_i^{mis} is after dropout do not enter the integral in Eqn (17), because these outcomes are deterministically missing. Thus, the dimensionality of the integration for the i th unit is $d_i = \sum_j I(g_{ij} = 1) + I(\text{any of } g_{ij} \text{ is } 2)$. Henceforth, the notation y_i^{mis} is includes only the intermittent missing outcomes and the outcome at the time of dropout. With nonignorable missingness, the integral with respect to y_i^{mis} is does not have a closed form,

and we require a numerical method for its evaluation. The computational workload for such integration increases exponentially with the number of intermittent missing outcomes, rendering the evaluation of L difficult with even moderate intermittent missingness.

To derive $\nabla^2 L_{\theta, \gamma_1}$, we note that $\nabla^2 L_{\theta, \gamma_1} = \left(\nabla^2 L_{\theta, \gamma_{10}}, \nabla^2 L_{\theta, \gamma_{20}}, \nabla^2 L_{\theta, \gamma_{11}} \right)$, where

$$\begin{aligned}
 \nabla^2 L_{\theta, \gamma_1^{10}} &= \sum_{i: K_i < n_i} \frac{\partial^2}{\partial \theta \partial \gamma_1^{10}} \ln \left(\int \prod_{j: g_{ij} \neq 0} f_{\gamma}(g_{ij} | s_{ij}, y_{ij}, g_{i, j-1}) f_{\theta}(y_i^{\text{mis}} | y_i^{\text{obs}}, x_i) dy_i^{\text{mis}} \right) \Big|_{\gamma_1 = 0} \\
 &= \sum_{i: K_i < n_i} \frac{\partial}{\partial \theta} \left(\frac{\int \frac{\partial \prod_{j: g_{ij} \neq 0} f_{\gamma}(g_{ij} | s_{ij}, y_{ij}, g_{i, j-1})}{\partial \gamma_1^{10}} f_{\theta}(y_i^{\text{mis}} | y_i^{\text{obs}}, x_i) dy_i^{\text{mis}}}{\int \prod_{j: g_{ij} \neq 0} f_{\gamma}(g_{ij} | s_{ij}, y_{ij}, g_{i, j-1}) f_{\theta}(y_i^{\text{mis}} | y_i^{\text{obs}}, x_i) dy_i^{\text{mis}}} \right) \Big|_{\gamma_1 = 0} \\
 &= \sum_{i: K_i < n_i} \frac{\partial}{\partial \theta} \left(\int \sum_{j: g_{ij} \neq 0} \frac{I(g_{i, j-1} = 0) \frac{\partial P_{ij}^{g_{ij}, g_{i, j-1}}}{\partial \phi_{ij}} \phi_{ij}^{1, g_{i, j-1}}}{P_{ij}^{g_{ij}, g_{i, j-1}}} y_{ij} f_{\theta}(y_i^{\text{mis}} | y_i^{\text{obs}}, x_i) dy_i^{\text{mis}} \right) \Big|_{\gamma_1 = 0} \\
 &= \sum_{i: K_i < n_i} \frac{\partial}{\partial \theta} \mathbb{E} \left((Y_i^{\text{mis}})^T A_i^{10} \mid y_i^{\text{obs}}, x_i \right) \Big|_{\gamma_1 = 0} \\
 &= \sum_{i: K_i < n_i} \frac{\partial \mathbb{E} \left((Y_i^{\text{mis}})^T \mid y_i^{\text{obs}}, x_i \right)}{\partial \theta} \Big|_{\gamma_1 = 0} \cdot A_i^{10}.
 \end{aligned}
 \tag{18}$$

If the k th component of Y_i^{mis} corresponds to the j th element of y_i , the k th element of A_i^{10} is

$$\begin{aligned}
 A_{il}^{10} &= \frac{I(g_{i, j-1} = 0) \frac{\partial P_{ij}^{g_{ij}, g_{i, j-1}}}{\partial \phi_{ij}} \phi_{ij}^{1, g_{i, j-1}}}{P_{ij}^{g_{ij}, g_{i, j-1}}} \Big|_{\hat{\gamma}_0(0), \gamma_1 = 0} \\
 &= I(g_{i, j-1} = 0) \left[I(g_{i, j} = 1) - P_{ij}^{10} \right] \Big|_{\hat{\gamma}_0(0), \gamma_1 = 0}.
 \end{aligned}$$

$\nabla^2_{\theta, \gamma_1^{20}} L$ and $\nabla^2_{\theta, \gamma_1^{11}} L$ are derived similarly to Eqn (18) with A_i^{10} replaced by A_i^{20} and A_i^{11} , respectively, where

$$A_{il}^{20} = I(g_{i,j-1} = 0) \left[I(g_{i,j} = 2) - P_{ij}^{20} \right]_{\hat{\gamma}_0(0), \gamma_1 = 0},$$

$$A_{il}^{11} = I(g_{i,j-1} = 1) \left[P_{ij}^{01} \right]_{\hat{\gamma}_0(0), \gamma_1 = 0}.$$

Appendix C:: Derivation of $E((Y_{\text{imis}})^T | y_{\text{iobs}}, x_i) \mid \gamma_1 = 0$

We note that Y_i^{mis} is a vector of length $d_i = \sum_j I(g_{ij} = 1) + I(\text{any of } g_{ij} \text{ is } 2)$.

C.1 MMGM.

Because

$$E\left((Y_i^{\text{mis}})^T \mid y_i^{\text{obs}}, x_i\right)_{\gamma_1 = 0} = \theta_1^T x_{i,M}^T + \left((y_i^{\text{obs}})^T - \theta_1^T x_{i,O}^T\right) \Sigma_{i,OO}^{-1} \Sigma_{i,OM},$$

by vector differentiation, we have

$$\begin{aligned} \frac{\partial E\left((Y_i^{\text{mis}})^T \mid y_i^{\text{obs}}, x_i\right)}{\partial \theta_1} \bigg|_{\gamma_1 = 0} &= x_{i,M}^T - x_{i,O}^T \Sigma_{i,OO}^{-1} \Sigma_{i,OM} \\ \frac{\partial E\left((Y_i^{\text{mis}})^T \mid y_i^{\text{obs}}, x_i\right)}{\partial \theta_2} \bigg|_{\gamma_1 = 0} &= -\left((y_i^{\text{obs}})^T - \theta_1^T x_{i,O}^T\right) \Sigma_{i,OO}^{-1} \frac{\partial \Sigma_{i,OO}}{\partial \theta_2} \Sigma_{i,OO}^{-1} \Sigma_{i,OM} \\ &\quad + \left((y_i^{\text{obs}})^T - \theta_1^T x_{i,O}^T\right) \Sigma_{i,OO}^{-1} \frac{\partial \Sigma_{i,OM}}{\partial \theta_2}, \end{aligned}$$

where $x_{i,O}$ and $x_{i,M}$ are the predictor matrices for Y_i^{obs} and Y_i^{mis} is, respectively, and

$$\text{Var}\left(Y_i^{\text{obs}}, Y_i^{\text{mis}} \mid X_i\right) = \begin{pmatrix} \Sigma_{i,OO} & \Sigma_{i,OM} \\ \Sigma_{i,MO} & \Sigma_{i,MM} \end{pmatrix}.$$

C.2 LMM.

One can rewrite the model of Eqn (8) as a marginal multivariate normal model with $\Sigma_i = \Lambda_i + Z_i V_b Z_i^T$. Then we apply the above result to obtain

$$\frac{\partial \Sigma_i}{\partial \theta_2} = \frac{\partial \Lambda_i}{\partial \theta_2} + Z_i \frac{\partial V_b}{\partial \theta_2} Z_i^T$$

References

- Cook R (1986). "Assessment of local influence." *Journal of the Royal Statistical Society B*, 48, 133–169.
- Copas JB, and Eguchi S (2001). "Local sensitivity approximations for selectivity bias." *Journal of the Royal Statistical Society B*, 63, 871–895.
- Copas JB, and Li HG (1997). "Inference for non-random samples (with discussion)." *Journal of the Royal Statistical Society B*, 59, 55–95.
- Daniels M, Hogan J (2008). *Missing Data in Longitudinal Studies: Strategies for Bayesian Modeling and Sensitivity Analysis*. Chapman and Hall, Boca Raton.
- Gao W, Hedeker D, Mermelstein R, Xie H (2016). "A scalable approach to measuring the impact of nonignorable nonresponse with an EM A application." *Statistics in Medicine*, 35, 5579–5602. [PubMed: 27538504]
- Heitjan D, Rubin D (1991). "Ignorability and coarse data." *Annals of Statistics*, 19, 2244–2253.
- Hirano K, Ridder G, Rubin D (2001). "Combining panels with attrition and refreshment samples." *Econometrica*, 69, 1645–1659.
- Laird N, Ware J (1982). "Random effects models for longitudinal data." *Biometrics*, 38, 963–974. [PubMed: 7168798]
- Little R, D Agostino R, Cohen M, Dickersin K, Emerson S, Farrar J, Frangakis C, Hogan J, Molenberghs G, Murphy S, Neaton J, Rotnitzky A, Scharfstein D, Shih W, Siegel J, Stern H (2012). "The prevention and treatment of missing data in clinical trials." *New England Journal of Medicine*, 367, 1355–1360. [PubMed: 23034025]
- Liu T, Heitjan D (2012). "Sensitivity of the discrete-time Kaplan-Meier estimate to nonignorable censoring: Application in a clinical trial." *Statistics in Medicine*, 31, 2998–3010. [PubMed: 22806989]
- Ma G, Troxel AB, Heitjan D (2005). "An index of local sensitivity to nonignorable dropout in longitudinal modeling." *Statistics in Medicine*, 24, 2129–2150. [PubMed: 15909292]
- Mallinckrodt C, Roger J, Chuang-Stein C, Molenberghs G, O'Kelly M, Ratitch B, Janssens M, Bunouf P (2014). "Recent Developments in the prevention and treatment of missing data." *Therapeutic Innovation & Regulatory Science*, 48, 68–80. [PubMed: 30231419]
- McCullagh P, Nelder J (1989). *Generalized Linear Models*. Chapman and Hall, New York.
- Raab G, Donnelly C (1999). "Information on sexual behaviour when some data are missing." *Applied Statistics*, 48, 117–133.
- Rubin DB (1976). "Inference and missing data (with discussion)." *Biometrika*, 63, 581–592.
- The Panel on Handling Missing Data in Clinical Trials of the National Research Council (2010). *The Prevention & Treatment of Missing Data in Clinical Trials*. National Academies Press, Washington, DC.
- Troxel A, Ma G, Heitjan D (2004). "An index of local sensitivity to nonignorability." *Statistica Sinica*, 14, 1221–1237.
- Vansteelandt S, Rotnitzky A, Robins J (2007). "Estimation of regression models for the mean of repeated outcomes under nonignorable nonmonotone nonresponse." *Biometrika*, 94, 841–860. [PubMed: 27453583]
- Verbeke G, Molenberghs G, Thijs H, Lesaffre E, Kenward MG (2001). "Sensitivity analysis for nonrandom dropout: A local influence approach." *Biometrics*, 57, 7–14. [PubMed: 11252620]
- Xie H (2008). "A local sensitivity analysis approach to longitudinal non-Gaussian data with non-ignorable dropout." *Statistics in Medicine*, 27, 3155–3177. [PubMed: 17948917]
- Xie H (2009). "Bayesian inference from incomplete longitudinal data: A simple method to quantify sensitivity to nonignorable dropout." *Statistics in Medicine*, 28, 2725–2747. [PubMed: 19572257]

- Xie H (2010). "Adjusting for nonignorable missingness when estimating generalized additive models." *Biometrical Journal*, 52, 186–200. [PubMed: 20422637]
- Xie H (2012). "Analyzing longitudinal clinical trial data with nonignorable missingness and unknown missingness reasons." *Computational Statistics & Data Analysis*, 56, 1287–1300.
- Xie H, Heitjan DF (2004). "Sensitivity analysis of causal inference in a clinical trial subject to crossover." *Clinical Trials*, 1, 21–30. [PubMed: 16281459]
- Xie H, Heitjan DF (2009). "Local sensitivity to nonignorability: Dependence on the assumed dropout mechanism." *Statistics in Biopharmaceutical Research*, 29, 243–257.
- Xie H, Qian Y (2012). "Measuring the impact of nonignorability in panel data with non-monotone nonresponse." *Journal of Applied Econometrics*, 27, 129–159.
- Xie H, Qian Y, Qu L (2011). "A semiparametric approach for analyzing nonignorable missing data." *Statistica Sinica*, 21, 1881–1899.
- Zeileis A, Croissant Y (2010). "Extended model formulas in R: Multiple parts and multiple responses." *Journal of Statistical Software*, 34(1), 1–13. doi:10.18637/jss.v034.i01.
- Zhang J, Heitjan D (2005). "Nonignorable censoring in randomized clinical trials." *Clinical Trials*, 2, 488–496. [PubMed: 16422309]
- Zhang J, Heitjan D (2006). "A simple local sensitivity analysis tool for nonignorable coarsening: Application to dependent censoring." *Biometrics*, 62, 1260–1268. [PubMed: 17156301]
- Zhang J, Heitjan D (2007). "Impact of nonignorable coarsening on Bayesian inference." *Biostatistics*, 8, 722–743. [PubMed: 17215331]

Table 1:

ISNI Analysis Using Simulated Data From a Logistic Regression: $\text{logit}(\Pr(Y = 1)) = \beta_0 + \beta_1 x$ with $\beta_0 = 1$ and $\beta_1 = -0.7$.

Data	γ_{00}	γ_{01}	γ_1	%[G = 1]	θ	Est	SE	ISNI	c
A	-2.0	-0.5	-1	6.0	β_0	1.089	0.074	0.068	1.090
					β_1	-0.754	0.092	-0.017	5.375
B	-2.0	-0.5	0	9.6	β_0	0.986	0.034	0.117	0.287
					β_1	-0.740	0.045	-0.044	1.011
C	1.2	-0.5	-1	56.7	β_0	1.651	0.140	0.607	0.230
					β_1	-0.782	0.162	-0.081	2.007
D	1.2	-0.5	0	72.2	β_0	1.028	0.071	0.773	0.092
					β_1	-0.756	0.088	-0.102	0.865

Table 2:

Summary Statistics in the Prostate Cancer QoL Dataset.

Month	Placebo (<i>n</i> = 367)		Flutamide (<i>n</i> = 370)	
	Mean (SD)	<i>n</i> observed (%)	Mean (SD)	<i>n</i> observed (%)
0	8.31 (1.50)	352 (96)	8.28 (1.46)	363 (98)
1	8.76 (1.22)	315 (86)	8.54 (1.23)	325 (88)
3	8.83 (1.26)	301 (82)	8.57 (1.20)	313 (85)
6	8.76 (1.20)	274 (75)	8.43 (1.37)	291 (79)

Table 3:

Missing Patterns in the Prostate Cancer QoL Dataset.

EF0	EF1	EF2	EF3	Placebo (<i>n</i> = 352)		Flutamide (<i>n</i> = 363)	
				Freq	Percent	Freq	Percent
P	P	P	P	239	67.8	258	71.1
P	P	P	A	38	10.8	37	10.2
P	P	A	A	19	5.4	14	3.8
P	A	A	A	20	5.6	22	6.1
P	P	A	P	13	3.7	13	3.6
P	A	P	P	11	3.1	15	4.1
P	A	P	A	7	2.0	1	0.3
P	A	A	P	5	1.4	3	0.8

Note: "P" indicates presence in the visit and "A" indicates absence in the visit.