Statistical Modelling 2001; **1**: 125–142

A local influence approach to sensitivity analysis of incomplete longitudinal ordinal data

Kristel van Steen[1], Geert Molenberghs[1], Geert Verbeke[2] and Herbert Thijs[1]
[1]Biostatistics, Center for Statistics, Limburgs Universitair Centrum, Diepenbeek, Belgium
[2]Biostatistical Centre, Katholieke Universiteit Leuven, Leuven, Belgium

Abstract: One of the major concerns when analysing incomplete longitudinal data is the fact that models necessarily rest on strong assumptions, unverifiable from the data. In response to these concerns, there is growing awareness of the usefulness of sensitivity analysis. In this paper we will focus on repeated ordinal data. Specifically, we implement a formal approach to such a sensitivity assessment, based on local influence, in the presence of multivariate categorical data. We explore the influence of perturbing a MAR dropout model in the direction of non-random dropout, and apply the proposed method to data from a longitudinal multicentre psychiatric study.

Key words: influence analysis; incomplete data; multivariate Dale model; missing data; perturbation scheme; sensitivity analysis

Data and software available from: http://stat.uibk.ac.at/SMIJ

Received August 2000; **revised** January 2001; **accepted** April 2001

1 Introduction

In a longitudinal experiment or study, each unit is measured on several occasions. In this paper we are concerned with the analysis of longitudinal data in which the response is *ordinal* and some sequences terminate early, a feature also referred to as *dropout*. Hence, for the analysis of such data, we need to accommodate dropout in the modelling process, in addition to accommodating the statistical dependence among the repeated measurements.

Extensive research activity over the past 10–15 years has led to different schools of thought regarding the best approach to the analysis of correlated categorical responses. Unlike in the normal setting, marginal, conditional, and random-effects approaches tend to give dissimilar results, as do likelihood, quasi-likelihood, and generalized estimating equation (GEE)-based inferential methods. There are many excellent reviews, notably Prentice (1988), Fitzmaurice *et al*. (1993), Diggle *et al*. (1994), Fahrmeir and Tutz (1994) and Pendergast *et al*. (1996).

Several *likelihood-based* methods have been proposed: Fitzmaurice and Laird (1993) combine marginal parameters for the main effects with conditional odds ratios for the association. Fully marginal models are presented by Bahadur (1961) and Cox (1972), using marginal correlations, and by Ashford and Sowden (1970), using a dichotomized version of

Address for correspondence: K van Steen, Center for Statistics, Limburgs Universitair Centrum, Universitaire Campus, Building D, B-3590 Diependeek, Belgium. E-mail: kristel.vansteen@luc.ac.be

 0962-2802(01)ST007RA

a multivariate normal to analyse multivariate binary data. Alternatively, marginal odds ratios can be used, as shown by Dale (1986) and Molenberghs and Lesaffre (1994, 1999). Cox (1972) also describes a model whose parameters have interpretations in terms of conditional probabilities. Similar models were proposed by Rosner (1984) and Liang and Zeger (1986). A fully exponential family model was proposed by Molenberghs and Ryan (1999) and Ryan and Molenberghs (1999). Random-effects approaches have been studied by Stiratelli *et al.* (1984), Zeger *et al.* (1988), Breslow and Clayton (1993), and Wolfinger and O'Connell (1993). *Pseudo-likelihood* methods were developed by Geys *et al.* (1997, 1999). GEEs were developed in Liang and Zeger (1986). A thorough account of work up until the middle of the previous decade is given in Fahrmeir and Tutz (1994).

As for modelling dropout, Rubin (1976) and Little and Rubin (1987, Chap. 6) make important distinctions between different missing values processes. A dropout process is said to be completely random (MCAR) if dropout is independent of both unobserved and observed data and random (MAR) if, conditional on the observed data, dropout is independent of the unobserved measurements; otherwise the dropout process is termed non-random (MNAR). If a dropout process is random, then a valid analysis can be obtained through a likelihood-based analysis that ignores the dropout mechanism, provided that the parameters describing the measurement process (pertaining to a particular mean and covariance structure) are functionally independent of the parameters describing the dropout process, the so-called parameter distinctness condition. This situation is termed ignorable by Rubin (1976) and Little and Rubin (1987) and leads to considerable simplification in the analysis. In many examples, however, the reasons for dropout will be many and varied. It is therefore difficult to justify on a priori grounds the assumption of random dropout. Arguably, in the presence of *non-random dropout*, a wholly satisfactory analysis of the data is not feasible.

In fact, modelling in this context often rests on strong (untestable) assumptions and relatively little evidence from the data themselves. Glynn *et al.* (1986) indicated that this is typical for so-called selection models, where the joint distribution of the measurements and the missingness process is factorized into the marginal distribution of the measurement process and the conditional process of the missingness, given the measurements. It is somewhat less the case for pattern-mixture models (Little, 1993; 1994; Hogan and Laird, 1997), where the alternative factorization is used, although caution should be used (Thijs *et al.*, 2000). This awareness and the resulting skepticism about fitting non-random dropout models initiated the search for methods to investigate the results with respect to model assumptions and for methods allowing to assess influences in the parameters describing the measurement process, as well as the parameters describing the non-random part of the dropout mechanism. Several authors have suggested various types of sensitivity analyses to address this issue (Molenberghs *et al.* 2001; Scharfstein *et al.*, 1999; Verbeke *et al.*, 2001). In this paper, we will explore local influence by perturbing a MAR dropout model in the direction of non-random dropout. This methodology was applied by Verbeke *et al.* (2001) in the case of the linear mixed model, for continuous longitudinal outcomes. The basis of the procedure is the local influence approach suggested by Cook (1986).

The data arise from a multicentre study including 315 patients with psychiatric symptoms who were recruited for treatment with a drug for the treatment of psychiatric symptoms, such as obsessive–compulsive disorder. The subjects who were recruited made three visits to the clinic and at each visit both therapeutic effect and severity of side effects were recorded, each on a four-category ordinal scale. Side effect is coded as

1) none;
2) not interfering with functionality of patient;
3) interfering significantly with functionality of patient;
4) the side-effect surpasses the therapeutic effect.

Similarly, the effect of therapy is recorded on a four point ordinal scale

1) no improvement or worsening;
2) minimal improvement (not changing functionality);
3) moderate improvement (partial disappearance of symptoms);
4) important improvement (almost total disappearance of symptoms).

Thus, a side effect occurs if new symptoms occur while there is therapeutic effect if old symptoms disappear. There is also baseline covariate information on each subject: sex, age, initial severity (scale 1–7) and duration of actual mental illness. Previous analyses of these data assuming random dropout are described in Molenberghs and Lesaffre (1994), Kenward *et al*. (1994), Molenberghs *et al*. (1997), Molenberghs and Lesaffre (1999). For our purposes, only side effect at the two subsequent visits after the initial visit is considered (referred to as Side 2 and Side 3 in Table 1), borrowing covariate information from age, sex, initial severity and duration of mental illness (respectively Age, Sex, Severity and Duration in Table 1). The data are modelled using a simple logistic regression formulation for the dropout process and using a multivariate Dale model for the response (Molenberghs and Lesaffre, 1994; 1999;

Table 1 Extract of the psychiatric data set

Patient ID	Covariates				Responses	
	Age	Sex	Duration	Severity	Side 2	Side 3
1	44	1	6	4	1	NA
2	28	2	1	NA	0	0
3	37	2	NA	7	1	0
4	NA	2	6	5	1	1
5	30	1	96	6	1	1
6	39	2	3	4	1	1
7	36	2	48	5	1	1
8	56	1	4	6	0	0
9	53	2	3	6	0	1
10	37	2	6	5	1	0
105	57	1	48	4	0	NA
106	58	1	1	5	1	0
107	25	2	13	6	0	0
108	58	1	4	4	3	NA
109	41	2	2	5	2	NA
110	29	2	3	5	1	1
111	50	2	8	4	2	3
112	56	2	3	5	3	NA
113	37	2	6	5	1	1
114	34	2	36	5	0	0
115	59	1	3	6	0	NA

Molenberghs *et al*., 1997). Within this framework, a sensitivity analysis will be conducted. The importance of such an analysis is even more pertinent in the light of Molenberghs *et al*. (1997), who show that non-random dropout cannot be ruled out, jeopardizing the validity of an ignorable analysis.

Section 2 introduces the multivariate Dale model. In Section 3, we sketch the missing data model. In Section 4, the local influence methodology as introduced by Cook (1986) is reviewed and applied to the Dale model. Finally, we exemplify the method on the psychiatric data set in Section 5.

2 A marginal model for multivariate ordinal data: the multivariate Dale model

The multivariate Dale model extends the bivariate global cross-ratio model described by Dale (1986) and McCullagh and Nelder (1989). It accounts for the dependence between multiple responses, as well as their dependence on covariate vector(s), which may be time-varying, continuous and/or discrete. The model arises from a decomposition of the joint probabilities into main effects (described by marginal probabilities) and interactions (described by cross-ratios of second and higher orders). The model will be introduced using a generalized linear model formulation (Molenberghs and Lesaffre, 1999).

Let $i = 1, \ldots, N$ indicate the covariate (design) level, containing n_i subjects. Every subject r in the ith level (group) is evaluated at T_i distinct time points and at each visit the subject is scored using a categorical outcome variable. Hence, the outcome for subject r in the ith level (group) is a series of measurements Y_{irt} $(t = 1, \ldots, T_i)$, where Y_{irt} can take on c_t distinct (possibly ordered) values k_t. Without loss of generality, we denote the category levels by $1, \ldots, c_t$. Along with the outcomes, a vector x_{it} of covariates is recorded, possibly time-dependent as the subscript t indicates.

For modelling purposes, it is convenient to summarize the categorical outcomes measured for subjects with covariate vector x_{it} in a cross-classification of the outcomes Y_{irt} into a $c_1 \times \cdots \times c_{T_i}$ dimensional contingency table with cell counts

$$Z_i^*(\boldsymbol{k}) \equiv Z_i^*(k_1, \ldots, k_{T_i}) \tag{2.1}$$

Obviously, $\sum_{\boldsymbol{k}} Z_i^*(\boldsymbol{k}) = n_i$. At every n_i-dimensional cutpoint, the data table is collapsed into a $2 \times 2 \times \cdots \times 2$ table, each of which is assumed to arise as a discretization of a multivariate Plackett distribution (Plackett, 1965). In line with the desire to use cumulative measures, given the outcomes are ordinal, a data table of cumulative counts can be constructed

$$Z_i(\boldsymbol{k}) = \sum_{\boldsymbol{\ell} \leq \boldsymbol{k}} Z_i^*(\boldsymbol{\ell})$$

in which $\boldsymbol{\ell} \leq \boldsymbol{k}$ can be written out as $(l_1, \ldots, l_{T_i}) \leq (k_1, \ldots, k_{T_i})$, and means that $l_j \leq k_j, j = 1, \ldots, T_i$. Thus, $Z_i(\boldsymbol{k})$, where $\boldsymbol{k} = (k_1, \ldots, k_{T_i})$, is just the number of individuals in group i whose observed response vector is $\boldsymbol{\ell}$ with $\boldsymbol{\ell} \leq \boldsymbol{k}$. Note that $Z_i^*(\boldsymbol{\ell})$ represents the number of individuals in group i whose observed response vector is $\boldsymbol{\ell}$. The corresponding probabilities are

$$\mu_i(\boldsymbol{k}) = \text{pr}(\boldsymbol{Y}_{ir} \leq \boldsymbol{k} \mid x_i, \boldsymbol{\theta}) \tag{2.2}$$

and $\mu_i^*(\boldsymbol{k}) = \text{pr}(\boldsymbol{Y}_{ir} = \boldsymbol{k} \mid x_i, \boldsymbol{\theta})$. Note that $Z_i(c_1, \ldots, c_{T_i}) = n_i$ and $\mu_i(c_1, \ldots, c_{T_i}) = 1$.

The preceding description very naturally combines univariate, bivariate, and multivariate information. Indeed, the marginal counts are given by all counts for which all but one indexes are equal to their maximal value: $Z_{itk} \equiv Z_i(c_1, \ldots, c_{t-1}, k, c_{t+1}, \ldots, c_{T_i})$. Bivariate cell counts, i.e., cell counts of a cross-classification of a pair of outcomes, follow from setting all but two indexes k_s equal to c_s, etc. Similarly, e.g., bivariate probabilities pertaining to the tth and sth outcomes, are denoted by $\mu_{i,ts,k\ell} = \mu_i(c_1, \ldots, c_{t-1}, k, c_{t+1}, \ldots, c_{s-1}, \ell, c_{s+1}, \ldots, c_{T_i})$. Generalizations to other orders are straightforward. The order of the components is immaterial but should be carried through the computations in a consistent fashion.

To complete the model description, we need to incorporate link functions and linear predictors for both marginal and association parameters. For the vector of links $\boldsymbol{\eta}_i$ we consider a function mapping the C_i-vector $\boldsymbol{\mu}_i(C_i = c_1 \times c_2 \times \cdots \times c_{T_i})$ to $\boldsymbol{\eta}_i = \boldsymbol{\eta}_i(\boldsymbol{\mu}_i)$, a C_i'-vector. Often, $C_i = C_i'$, and $\boldsymbol{\eta}_i$ and $\boldsymbol{\mu}_i$ have the same ordering. A counterexample is provided by the probit model, where the number of link functions is smaller than the number of mean components, as soon as $T_i > 2$. For the univariate marginal links, a convenient choice is the logistic

$$\eta_{itk} = \text{logit}(\mu_{itk} \mid x_{it}) = \beta_{0\,itk} + \beta_{itk} x_{it} \quad (1 \leq t \leq n, \;\; 1 \leq k < c_t) \tag{2.3}$$

If evidence is found that the regression parameters are consistent across the cutpoints k, β_{itk} in equation (2.3) may be replaced by β_{it}, implying a proportional odds model for the response.

Full specification of the association is done in terms of marginal global odds ratios

$$\psi_{i,ts,k\ell} = \frac{(\mu_{i,ts,k\ell})(1 - \mu_{itk} - \mu_{is\ell} + \mu_{i,ts,k\ell})}{(\mu_{is\ell} - \mu_{i,ts,k\ell})(\mu_{itk} - \mu_{i,ts,k\ell})} \tag{2.4}$$

They are usefully modelled on the log scale as

$$\eta_{i,ts,k\ell} = \ln(\mu_{i,ts,k\ell}) - \ln(\mu_{itk} - \mu_{i,ts,k\ell}) - \ln(\mu_{is\ell} - \mu_{i,ts,k\ell}) + \ln(1 - \mu_{itk} - \mu_{is\ell} + \mu_{i,ts,k\ell})$$

Higher order global odds ratios are easily introduced using ratios of conditional odds (ratios). For

$$\mu_{it\,|\,s}(z_s) = \text{pr}(Z_{irtk_t} = 1 \mid Z_{irsk_s} = z_s, x_i, \boldsymbol{\theta}) \tag{2.5}$$

the conditional probability of observing a success at occassion t, given the value z_s is observed at occasion s, and writing the corresponding conditional odds as

$$\psi_{it\,|\,s}(z_s) = \mu_{it\,|\,s}(z_s)/(1 - \mu_{it\,|\,s}(z_s))$$

the pairwise marginal odds ratio, for occasions t and s, is defined as

$$\psi_{its} = \frac{\{\text{pr}(Z_{irtk_t} = 1, Z_{irsk_s} = 1)\}\{\text{pr}(Z_{irtk_t} = 0, Z_{irsk_s} = 0)\}}{\{\text{pr}(Z_{irtk_t} = 0, Z_{irsk_s} = 1)\}\{\text{pr}(Z_{irtk_t} = 1, Z_{irsk_s} = 0)\}} = \frac{\psi_{it\,|\,s}(1)}{\psi_{it\,|\,s}(0)}$$

in accordance with equation (2.4). This formulation can be exploited to define the higher order marginal odds ratios in a recursive fashion

$$\psi_{it_1 \cdots t_m\, t_{m+1}} = \frac{\psi_{it_1 \cdots t_m\,|\,t_{m+1}}(1)}{\psi_{it_1 \cdots t_m\,|\,t_{m+1}}(0)} \tag{2.6}$$

where $\psi_{it_1 \cdots t_m\,|\,t_{m+1}}(z_{m+1})$ is defined by conditioning all probabilities occurring in the expression for $\psi_{it_1 \cdots t_m}$ on $Z_{irt_{m+1}} = z_{t_{m+1}}$. The choice of the variable to condition on is immaterial. Observe that multi-way marginal global odds ratios are defined solely in terms of conditional probabilities. Marginal local odds ratios can be defined similarly.

When contrasts of log probabilities are used as link functions (e.g., cumulative logit links for the marginal probabilities and log cross-ratios for the associations), the model can be written using composite links (Thompson and Baker, 1981; McCullagh and Nelder, 1989)

$$\boldsymbol{\eta}_i(\boldsymbol{\mu}_i) = C_i \ln(A_i \boldsymbol{\mu}_i)$$

where the probabilities involved are linear combinations $A_i\boldsymbol{\mu}$. The multidimensional vector $\boldsymbol{\mu}_i$ of probabilities is expanded to a vector containing all probabilities of dimensions 1 to $T = \max(T_i)$ by multiplying $\boldsymbol{\mu}_i$ with an appropriate matrix of constants A_i. Contrasts of the log probabilities, by means of C_i, are linked to linear predictors.

3 A model for dropout

When dropout occurs, the hypothetical full data consist of complete data and a dropout indicator D, which can take on values $2, \ldots, T+1$, with $D = T+1$ indicating no dropout at all and $D = j$ indicating that measurements are available up to and including time $j-1$. It is convenient to denote the full data again by Z_i^*, $i = 1, \ldots, N$ referring to a particular covariate level, now containing components $Z_i^*(d, \boldsymbol{k})$ (note that the cell counts $Z_i^*(d, \boldsymbol{k})$ are obtained after cross-classifying both the outcome for the dropout indicator D and the categorical outcomes in a $T \times c_1 \times \cdots \times c_{T_i}$ dimensional contingency table) and having joint probabilities

$$\begin{aligned} \nu_i^*(d, \boldsymbol{k}) &= \text{pr}(D = d, \boldsymbol{Y}_{ir} = \boldsymbol{k} \,|\, x_i, \boldsymbol{\theta}, \boldsymbol{\psi}) \\ &= \text{pr}(\boldsymbol{Y}_{ir} = \boldsymbol{k} \,|\, x_i, \boldsymbol{\theta})\, \text{pr}(D = d \,|\, \boldsymbol{Y}_{ir} = \boldsymbol{k} \,|\, x_i, \boldsymbol{\psi}) \\ &= \mu_i^*(\boldsymbol{k}, \boldsymbol{\theta})\ \phi_i(d \,|\, \boldsymbol{k}, \boldsymbol{\psi}) \end{aligned} \tag{3.1}$$

where $\boldsymbol{\psi}$ parameterizes the dropout probabilities $\phi_i(d \,|\, \boldsymbol{k}, \boldsymbol{\psi})$. Parameters $\boldsymbol{\theta}$ and $\boldsymbol{\psi}$ are assumed to be distinct. We assume that the distribution of D may depend both on the past history of the process, denoted by $H_d = (k_1, \ldots, k_{d-1})$ for $D = d$, and the current outcome category k_d, but not on the process after that time. Consequently

$$\phi_i(d \mid \boldsymbol{k}, \boldsymbol{\psi}) = \begin{cases} \prod_{t=2}^{d-1} \{1 - p_{it}(H_t, k_t; \boldsymbol{\psi})\} p_{id}(H_d, k_d; \boldsymbol{\psi}) & \text{if } D \leq T \\ \prod_{t=2}^{T} \{1 - p_{it}(H_t, k_t; \boldsymbol{\psi})\} & \text{if } D = T+1 \end{cases} \tag{3.2}$$

where

$$p_{id}(H_d, k_d; \boldsymbol{\psi}) = \text{pr}(D = d \mid D \geq d, \boldsymbol{Y}_{ir}^{\text{obs}} = H_d, Y_{ird} = k_d; x_i; \boldsymbol{\psi})$$

and $\boldsymbol{Y}_{ir}^{\text{obs}}$ refers to the subvector of $\boldsymbol{Y}_{ir}$ that is observed.

We specify the model for the dropout probabilities by logit links of the form

$$\xi_{idk_1 \cdots k_d} = \text{logit}\{p_{id}(H_d, k_d; \boldsymbol{\psi})\}$$

Choosing linear predictors completes the model specification

$$\xi_{idk_1 \cdots k_d} = \text{logit}\{p_{id}(H_d, k_d; \boldsymbol{\psi})\} = \boldsymbol{Y}_{ir}^{\text{obs}} \boldsymbol{\psi} + \omega Y_{ird} \tag{3.3}$$

with $\boldsymbol{Y}_{ir}^{\text{obs}}$ referring to the observed subvector of $\boldsymbol{Y}_{ir}$, and $\boldsymbol{Y}_{ir}$ in turn referring to the vector of measurements Y_{irt}, $t = 1, \ldots, T_i$, for subject r in the ith group. Note that, *under the posited model*, $\omega \neq 0$ is equivalent to a non-random dropout process. The heart of the problem resides in the fact that one can never fully verify the adequacy of the posited model from the observed data alone, once data are incomplete.

4 Local influence

Cook (1986) suggests that more confidence can be put in a model which is relatively stable under small modifications. The best known perturbation schemes are based on case-deletion (Cook and Weisberg, 1982) in which the effect is studied of completely removing cases from the analysis. A quite different paradigm is the local influence approach where one investigates how the results of an analysis are changed under small perturbations of the model. In the framework of the linear mixed model Beckman *et al.* (1987) used local influence to assess the effect of perturbing the error variances, the random-effects variances and the response vector. In the same context, Lesaffre and Verbeke (1998) have shown that the local influence approach is also useful for the detection of influential subjects in a longitudinal data analysis. Verbeke *et al.* (2001) and Verbeke and Molenberghs (2000) use the same idea to explore the sensitivity of a selection model for repeated continuous outcomes. The principal idea is to explore how small perturbations around MAR, in the direction of MNAR, can have a large impact. These authors have shown that various types of influential subjects can cause a model to appear of the MNAR type. This implies that caution should be used before concluding that the model really is MNAR, since many types of influential subjects, different from an MNAR mechanism, can force such a conclusion.

Specifically, we consider the following perturbed dropout model

$$\xi_{idk_1\cdots k_d} = \text{logit}\{\, p_{id}(H_d, k_d; \boldsymbol{\psi})\} = \boldsymbol{Y}_{ir}^{\text{obs}}\boldsymbol{\psi} + \omega_i Y_{ird} \tag{4.1}$$

Indeed, $\omega_i = 0$ for all i corresponds to an MAR process, which cannot influence the measurement model parameters. When small perturbations in a specific ω_i lead to relatively large differences in the model parameters, then this suggests that these subjects may have a large impact on the final analysis. Note that the ω_i are not to be seen as fixed or random subject-specific parameters, but rather as (infinitesimal) perturbations, to which differential geometry will be applied, rather than ordinary parameter estimation.

We first give a general description of the local influence methodology as introduced by Cook (1986). In Section 5, we will apply it to the dropout model presented in Section 3.

We denote the log-likelihood function corresponding to model equation (4.1) by $\ell(\boldsymbol{\gamma} \,|\, \boldsymbol{\omega}) = \sum_{i=1}^{N} \ell_i(\boldsymbol{\gamma} \,|\, \omega_i)$, in which $\ell_i(\boldsymbol{\gamma} \,|\, \omega_i)$ is the contribution of the ith individual to the log-likelihood, and where $\boldsymbol{\gamma} = (\boldsymbol{\theta}, \boldsymbol{\psi})$ is the s-dimensional vector, grouping the parameters of the measurement model and the dropout model, not including the $N \times 1$ vector $\boldsymbol{\omega} = (\omega_1, \omega_2, \ldots, \omega_N)'$ of weights defining the perturbation of the MAR model. It is assumed that $\boldsymbol{\omega}$ belongs to an open subset Ω of $\mathbb{R}^N$. For $\boldsymbol{\omega}$ equal to $\boldsymbol{\omega_0} = (0, 0, \ldots, 0)'$, $\ell(\boldsymbol{\gamma} \,|\, \boldsymbol{\omega_0})$ is the log-likelihood function which corresponds to a MAR dropout model.

Let $\widehat{\boldsymbol{\gamma}}$ be the maximum likelihood estimator for $\boldsymbol{\gamma}$, obtained by maximizing $\ell(\boldsymbol{\gamma} \,|\, \boldsymbol{\omega_0})$, and let $\widehat{\boldsymbol{\gamma}}_{\boldsymbol{\omega}}$ denote the maximum likelihood estimator for $\boldsymbol{\gamma}$ under $\ell(\boldsymbol{\gamma} \,|\, \boldsymbol{\omega})$. The local influence approach now compares $\widehat{\boldsymbol{\gamma}}_{\boldsymbol{\omega}}$ with $\widehat{\boldsymbol{\gamma}}$. Similar estimates indicate that the parameter estimates are robust with respect to perturbations of the MAR model in the direction of MNAR. Very different estimates suggest that the estimation procedure is highly sensitive to such perturbations, which suggests that the choice between a random and a non-random dropout model greatly affects the results of the analysis. Cook (1986) proposed to measure the distance between $\widehat{\boldsymbol{\gamma}}_{\boldsymbol{\omega}}$ and $\widehat{\boldsymbol{\gamma}}$ by the so-called likelihood displacement, defined by $LD(\boldsymbol{\omega}) = 2[\ell(\widehat{\boldsymbol{\gamma}} \,|\, \boldsymbol{\omega_0}) - \ell(\widehat{\boldsymbol{\gamma}}_{\boldsymbol{\omega}} \,|\, \boldsymbol{\omega})]$. This takes into account the variability of $\widehat{\boldsymbol{\gamma}}$. Indeed, $LD(\boldsymbol{\omega})$ will be large if $\ell(\boldsymbol{\gamma} \,|\, \boldsymbol{\omega_0})$ is strongly curved at $\widehat{\boldsymbol{\gamma}}$, which means that $\boldsymbol{\gamma}$ is estimated with high precision, and small otherwise. Therefore, a graph of $LD(\boldsymbol{\omega})$ versus $\boldsymbol{\omega}$ contains essential information on the influence of perturbations. It is useful to view this graph as the geometric surface formed by the values of the $N + 1$ dimensional vector $\boldsymbol{\xi}(\boldsymbol{\omega}) = (\boldsymbol{\omega}', LD(\boldsymbol{\omega}))'$ as $\boldsymbol{\omega}$ varies throughout Ω. Since this so-called influence graph can only be depicted when $N = 2$, Cook (1986) proposed to consider local influence, i.e., at the normal curvatures $C_{\boldsymbol{h}}$ of $\boldsymbol{\xi}(\boldsymbol{\omega})$ in $\boldsymbol{\omega_0}$, in the direction of some N dimensional vector $\boldsymbol{h}$ of unit length. Let $\boldsymbol{\Delta}_i$ be the s dimensional vector defined by

$$\boldsymbol{\Delta}_i = \left. \frac{\partial^2 \ell_i(\boldsymbol{\gamma} \,|\, \omega_i)}{\partial \omega_i \partial \boldsymbol{\gamma}} \right|_{\boldsymbol{\gamma} = \hat{\boldsymbol{\gamma}},\, \omega_i = 0}$$

and define Δ as the $(s \times N)$ matrix with $\boldsymbol{\Delta}_i$ as its ith column. Further, let $\ddot{L}$ denote the $(s \times s)$ matrix of second order derivatives of $\ell(\boldsymbol{\gamma} \,|\, \boldsymbol{\omega_0})$ with respect to $\boldsymbol{\gamma}$, also evaluated at $\boldsymbol{\gamma} = \widehat{\boldsymbol{\gamma}}$. Cook (1986) has then shown that $C_{\boldsymbol{h}}$ can be easily calculated by $C_{\boldsymbol{h}} = 2\,|\, \boldsymbol{h}'\Delta'\ddot{L}^{-1}\Delta\boldsymbol{h}\,|$.

Obviously, $C_{\boldsymbol{h}}$ can be calculated for any direction $\boldsymbol{h}$. One evident choice is the vector $\boldsymbol{h_i}$ containing one in the ith position and zero elsewhere, corresponding to the perturbation of the ith weight only. This reflects the influence of allowing the ith subject to drop out non-randomly, while the others can only drop out at random. The corresponding local influence

measure, denoted by C_i, then becomes $C_i = 2 \mid \boldsymbol{\Delta}_i' \ddot{L}^{-1} \boldsymbol{\Delta}_i \mid$. Another important direction is the direction $\boldsymbol{h}_{\max}$ of maximal normal curvature $C_{\max}$. It shows how to perturb the MAR model to obtain the largest local changes in the likelihood displacement. It is readily seen that $C_{\max}$ is the largest eigenvalue of $-2\Delta' \ddot{L}^{-1} \Delta$, and that $\boldsymbol{h}_{\max}$ is the corresponding eigenvector.

When a subset $\boldsymbol{\gamma}_1$ of $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1', \boldsymbol{\gamma}_2')'$ is of special interest, a similar approach can be used, replacing the log-likelihood by the profile log-likelihood for $\boldsymbol{\gamma}_1$, and the methods discussed above for the full parameter vector directly carry over. Details can be found in Lesaffre and Verbeke (1998) and in Verbeke *et al.* (2001).

It will be clear from the previous derivations that calculation of local influence measures merely reduces to evaluation of Δ and $\ddot{L}$. In the linear mixed model case, Verbeke *et al.* (2000) and Verbeke and Molenberghs (2000) have proposed closed form expressions, with some emphasis on the case of compound symmetry. For the multivariate Dale model, as will be the case for many other non-normal models, this is algebraically very involved and may not yield the same type of insightful expressions. However, when a program is available to fit the full non-random model (3.3), a particularly convenient computational scheme can be used. Indeed, in this case there are usually tools available to obtain a Hessian matrix evaluated in a point of interest (e.g., through EM-aided differentiation). Note that in our situation, it suffices to compute the second derivatives of the likelihood, for each observation separately, after which the subvector $\boldsymbol{\Delta}_i$ pertaining to the $(\boldsymbol{\gamma}, \boldsymbol{\omega})$-block can be selected.

In practice, the parameter $\boldsymbol{\theta}$ in the measurement model is often of primary interest. Since $\ddot{L}$ is block-diagonal with blocks $\ddot{L}(\boldsymbol{\theta})$ and $\ddot{L}(\boldsymbol{\psi})$, we have that for any unit vector $\boldsymbol{h}$, $C_{\boldsymbol{h}}$ equals $C_{\boldsymbol{h}}(\boldsymbol{\theta}) + C_{\boldsymbol{h}}(\boldsymbol{\psi})$, with

$$C_{\boldsymbol{h}}(\boldsymbol{\theta}) = -2\boldsymbol{h}' \left[\left. \frac{\partial^2 \ell_{i\omega}}{\partial \boldsymbol{\theta} \partial \omega_i} \right|_{\omega_i = 0} \right]' \ddot{L}^{-1}(\boldsymbol{\theta}) \left[\left. \frac{\partial^2 \ell_{i\omega}}{\partial \boldsymbol{\theta} \partial \omega_i} \right|_{\omega_i = 0} \right] \boldsymbol{h}$$

$$C_{\boldsymbol{h}}(\boldsymbol{\psi}) = -2\boldsymbol{h}' \left[\left. \frac{\partial^2 \ell_{i\omega}}{\partial \boldsymbol{\psi} \partial \omega_i} \right|_{\omega_i = 0} \right]' \ddot{L}^{-1}(\boldsymbol{\psi}) \left[\left. \frac{\partial^2 \ell_{i\omega}}{\partial \boldsymbol{\psi} \partial \omega_i} \right|_{\omega_i = 0} \right] \boldsymbol{h}$$

evaluated at $\boldsymbol{\gamma} = \widehat{\boldsymbol{\gamma}}$.

5 Analysis of the psychiatric study

We apply the local influence approach to the psychiatric data, introduced in Section 1. In the analysis we restrict attention to patients with monotone dropout, leaving a total of 287 subjects for study. The response of interest is side effect, recorded at two subsequent visits (say visit 1 and visit 2) after the initial visit. A multivariate Dale model is used for the measurement model. Regression parameters (intercept parameters excluded) are assumed to be constant across the cutpoints k (based on Kenward *et al.*, 1994), implying a proportional odds model for the response. Effects for age and sex are held constant over the two visits (previous analyses indicated that the effects of age and sex differ little from visit to visit), whereas the effects for duration and severity are allowed to change over time. No other covariates are included in the measurement model, nor in the model for dropout. Assessing the effects of covariates in the dropout model on influence measures is subject to further research. Specifically, the measurement model can be written in terms of univariate marginal

links, as in equation (2.3), and the single marginal odds ratio ψ_{i12}

$$\eta_{i11} = \text{logit}(\mu_{i11} \mid x_{i1}) = \text{intercept 1} + (\text{age sex duration 1 severity 1})x_{i1}$$
$$\eta_{i12} = \text{logit}(\mu_{i12} \mid x_{i1}) = \text{intercept 2} + (\text{age sex duration 1 severity 1})x_{i1}$$
$$\eta_{i13} = \text{logit}(\mu_{i13} \mid x_{i1}) = \text{intercept 3} + (\text{age sex duration 1 severity 1})x_{i1}$$

$$\eta_{i21} = \text{logit}(\mu_{i21} \mid x_{i2}) = \text{intercept 1} + (\text{age sex duration 2 severity 2})x_{i2}$$
$$\eta_{i22} = \text{logit}(\mu_{i22} \mid x_{i2}) = \text{intercept 2} + (\text{age sex duration 2 severity 2})x_{i2}$$
$$\eta_{i23} = \text{logit}(\mu_{i23} \mid x_{i2}) = \text{intercept 3} + (\text{age sex duration 2 severity 2})x_{i2}$$

In practice, equation (3.3) reduces to

$$\xi_{idk_1 \cdots k_d} = \text{intercept} + k_{d-1}\ \text{prev. measurement} + k_d\ \text{curr. measurement}$$

since the conditional dependence of dropout on observations preceding k_{d-1} is considered to be negligible given k_{d-1} and k_d. The case 'curr. measurement = 0' corresponds to a random dropout process (MAR), the case 'prev. measurement = curr. measurement = 0' to a completely random dropout process (MCAR). The estimates under different assumptions of the dropout mechanism are listed in Table 2. Provided MAR is the correct alternative hypothesis and provided the parametric form for the MAR process is correct, there seems to be little evidence for MAR and one could adopt the simpler MCAR assumption. The likelihood ratio test statistic to compare MCAR with MAR is $G^2 = 2.92$ on 1 degree of freedom ($p = 0.087$). A comparison between the non-random and random dropout model produces a likelihood ratio test statistic of $G^2 = 0.15$ with 1 degree of freedom ($p = 0.694$). Hence, there hardly seems to be any evidence for MNAR. Note that the first two models in

Table 2 Maximum likelihood estimates and standard errors (SEs) of random and non-random dropout models

Effect	MCAR		MAR		MNAR	
	Estimate	SE	Estimate	SE	Estimate	SE
Measurement model						
Intercept 1	−0.434	0.853	−0.434	0.853	−0.572	0.807
Intercept 2	1.713	0.859	1.713	0.859	1.520	0.797
Intercept 3	2.953	0.876	2.953	0.876	2.705	0.797
Age	−0.020	0.008	−0.020	0.008	−0.020	0.008
Duration 1	−0.014	0.005	−0.014	0.005	−0.014	0.005
Duration 2	−0.022	0.006	−0.022	0.006	−0.023	0.006
Severity 1	0.263	0.135	0.263	0.135	0.284	0.130
Severity 2	0.329	0.135	0.329	0.135	0.341	0.132
Sex	−0.108	0.226	−0.108	0.226	−0.090	0.222
Association	3.151	0.297	3.151	0.297	3.242	0.289
Dropout model						
Intercept	−4.494	0.212	−4.494	0.567	−5.093	0.683
Prev. measurement	–	–	1.091	0.222	0.431	0.548
Curr. measurement	–	–	–	–	0.972	0.610
−2 observed log likelihood	1129.114		1126.194		1126.039	

Table 2 yield exactly the same values for the mean parameter estimates, since MCAR and MAR are ignorable. Even if the dropout model parameters are not explicitly estimated, the same parameter estimates would be obtained.

To investigate the sensitivity of inferences with respect to modelling assumptions for the dropout process, the overall C_i, influences $C_i(\boldsymbol{\theta})$ and $C_i(\boldsymbol{\psi})$ for the measurement parameters and dropout parameters, as well as $\boldsymbol{h}_{\max}$ of maximal curvature are displayed in Figure 1. Note that the largest C_i are observed for patients #34 and #252 (both having side effects surpassing the therapeutic effect at visit 1 and visit 2), followed by patients #182, #64, #122, #28, #108, #287, #232, #112 and #245, all of whom yield the worst score on side effects at visit 1 and drop out at visit 2. We pay special attention to patient #239, showing side effects interfering significantly with functionality at visit 1, after which dropout occurs.

In addition, Figure 1 shows some evidence of the fact that influence on measurement model parameters can theoretically only arise from those measurement occasions at which dropout occurs, a fact already observed by Verbeke *et al.* (2001). Nevertheless, it should be

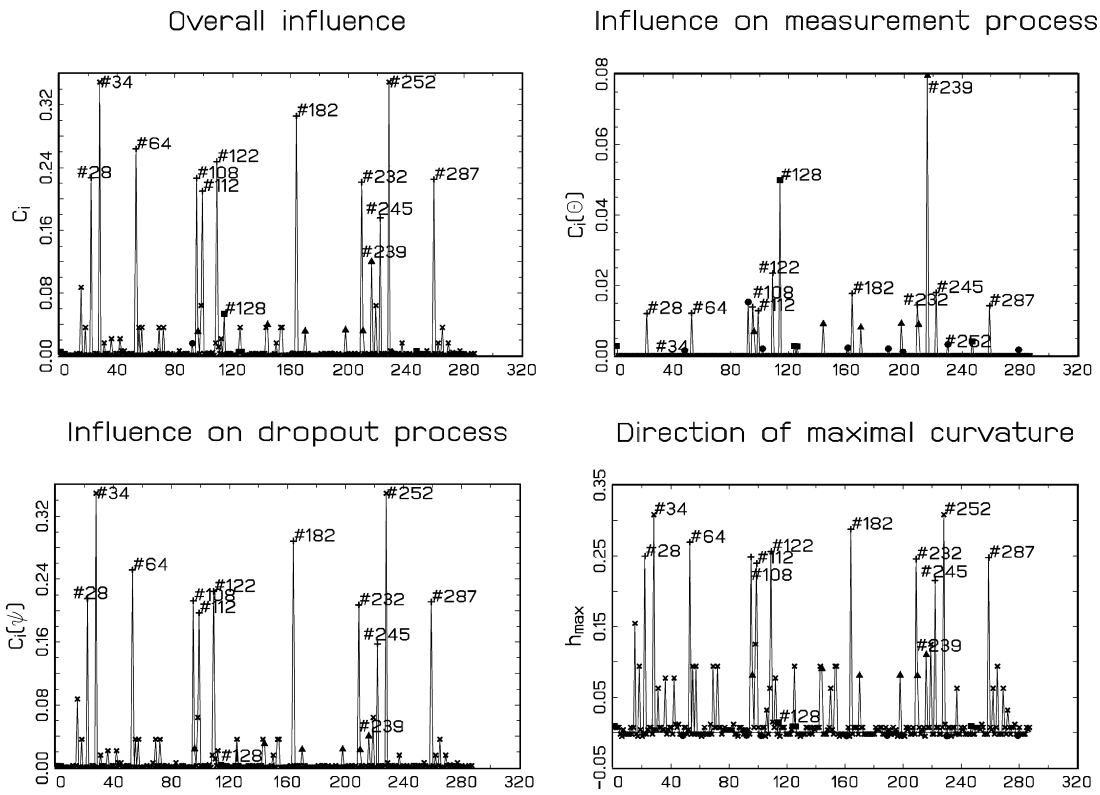

Figure 1 Index plots of C_i, $C_i(\boldsymbol{\theta})$, $C_i(\boldsymbol{\psi})$, and of the components of the direction $\boldsymbol{h}_{\max}$ of maximal curvature. The *x*-axis merely contains sequential indicators. Relevant patient IDs have been added to the plot. Completers (patients with observed responses at visit 1 and visit 2) are indicated with a solid star. A solid circle, a solid square, a solid triangle or a solid plus is used for subjects whose score on side effects at visit 1 respectively ranges from (1) to (4). Patients with a non-monotone dropout pattern are discarded

noted that influence on the measurement model parameters can also arise from complete observations. Indeed, when small perturbations in a specific ω_i lead to relatively large differences in the model parameters, the subject's impact on dropout parameters indirectly influences all functions that include these dropout parameters. An example of such a function is the conditional mean of an unobserved measurement, given the observed measurements and given the fact that the patient belongs to a certain dropout pattern. As a consequence, the corresponding measurement model parameters will *indirectly* be affected as well (Verbeke *et al*., 2001).

Influential completers occur in the index plots of C_i, $C_i(\boldsymbol{\psi})$, and of the components of the direction $\boldsymbol{h}_{\max}$ of maximal curvature, but are absent in the index plot for $C_i(\boldsymbol{\theta})$. Focusing on $C_i(\boldsymbol{\theta})$, Figure 1 reveals the highest peaks for patients #239 and #128. It appears that the influence of allowing subject #239 to drop out non-randomly, is best visible on the measurement model parameters. Patient #128 has an incomplete sequence, with a relatively mild score for side effects (side effects not interfering with functionality). Hence, the relatively large value for $C_i(\boldsymbol{\theta})$ is somewhat unusual, especially since other index plots do not show evidence of any influential effect, not even globally. One could ask the question whether other, unmeasured factors could have caused this phenomenon.

Before addressing this question, we turn attention to $C_i(\boldsymbol{\psi})$ and $\boldsymbol{h}_{\max}$. To avoid confusion, observe that the scale is different from that of $C_i(\boldsymbol{\theta})$. The most influential patients appear to be the same as for the overall C_i (#34, #252 and #182, #64, #122, #28, #108, #287, #232, #112, #245). The same patients are also shown in the index plot for $\boldsymbol{h}_{\max}$.

Observe that in all plots, 'layers' of influential cases may be distinguished. The higher the layers, the more they seem to be associated with particular response levels. For instance, in Figure 1, patients #34 and #252 give rise to components of $\boldsymbol{h}_{\max}$ that are larger than 0.3. Patients #182, #64, #122, #28, #108, #287, #232, #112 and #245 (corresponding to the influential patients in the previous paragraph) refer to $\boldsymbol{h}_{\max}$ components that are all smaller than 0.3 but larger than 0.2. The layer formation is not clear though, and recalling the particular behaviour of patient #128, one is led to believe that another distorting factor is involved, blurring the picture. Therefore, we investigate the effect of covariates on the ability to interpret influence plots.

To this end, we consider two additional models. The first one includes sex as the only covariate in the measurement model, the second one uses age as the only covariate. These models perform worse than the model including both age and sex, augmented with duration and severity, but they are merely intended for illustrative purposes. The resulting influence plots are enlightening. Figure 2 shows the index plots when age is included as the only covariate, Figure 3 displays the corresponding pictures in case sex is the only source of covariate information. In both cases, much smaller values are obtained for $C_i(\boldsymbol{\theta})$. The high peaks for patients #239 and #128 have disappeared. Patients #122, #245 and #182 also show up in Figure 2 with the highest peaks for $C_i(\boldsymbol{\theta})$, although hard to distinguish from the peaks for patients #287, #232, #28, #108, #64 and #112. The variability observed in $C_i(\boldsymbol{\theta})$ values also appears in Figure 3. However, in this case, it seems to be caused by the fact that patients #108, #182, #287 and #232 have $C_i(\boldsymbol{\theta})$ equal to about 0.0116 compared to approximately 0.0097 for patients #28, #245, #64, #122 and #112. This layer effect may be explained by the binary character of sex as opposed to age, the latter of which entered the model as a continuous variable. Also note that patients #108, #182, #287 and #232 are all male, whereas patients #28, #245, #64, #122 and #112 are all female. All these patients drop out at

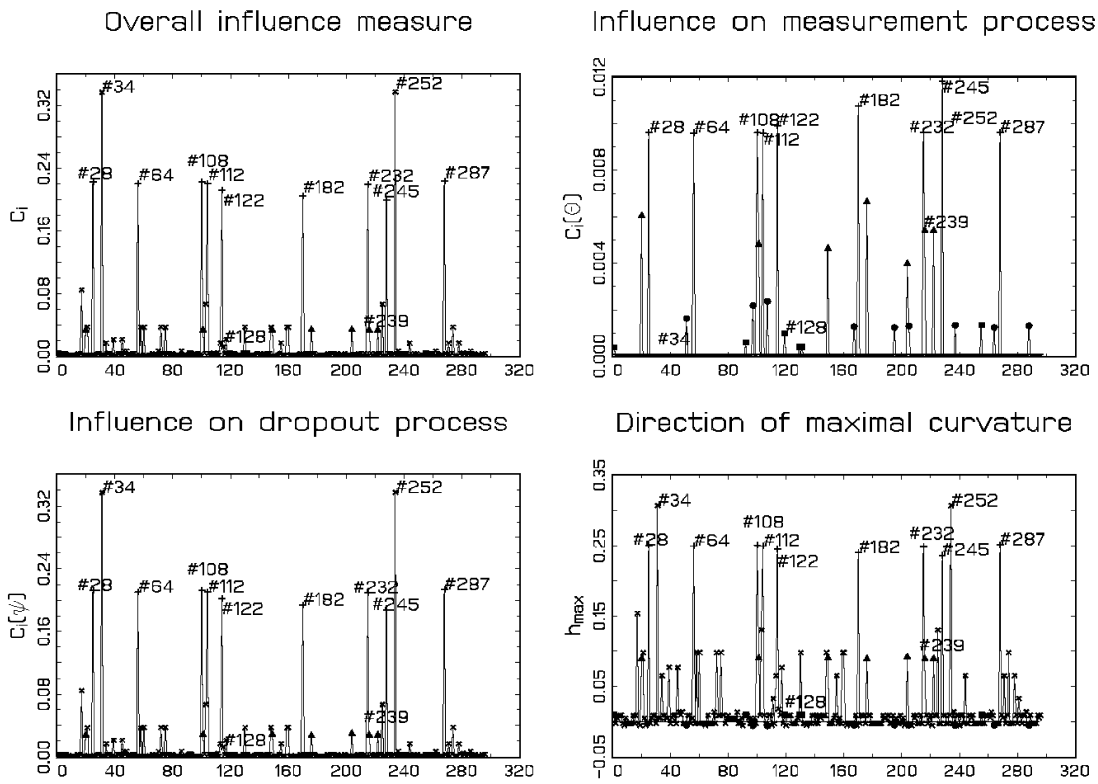

Figure 2 Index plots of C_i, $C_i(\boldsymbol{\theta})$, $C_i(\boldsymbol{\psi})$, and of the components of the direction $\boldsymbol{h}_{\max}$ of maximal curvature, where age is considered as the only covariate in the Dale model. The *x*-axis contains sequential indicators. Completers are indicated with a solid star. A solid circle, a solid square, a solid triangle or a solid plus is used for subjects whose score at visit 1 on side effects respectively ranges from (1) to (4)

visit 2 and showed side effects surpassing therapeutic effect at visit 1. In Figures 2 and 3, the same patient group (i.e., patients #34, #252, #287, #108, #28, #112, #64, #232, #122, #182 and #245) is distinguished as globally influential, with highest C_i values for #34 and #252. The layering effect is again the most explicit when sex is considered as the only covariate (Figure 3). Influential patients for $C_i(\boldsymbol{\psi})$ and $\boldsymbol{h}_{\max}$ appear to be the same as before, where sex and age were both considered in the pool of covariates, with the exception of subject #239 whose corresponding component in $\boldsymbol{h}_{\max}$ is now less than 0.1000. The distribution over potential values becomes more discrete when age is considered to be the only covariate in the multivariate Dale model. Changing age for sex causes the distribution to be even more discrete and therefore the layer effect more explicit.

In an attempt to improve insight into the driving forces present in the set of data, which may explain possible deviations from a random dropout process, we exclude patients #34 and #252 from the data set and apply the same measurement model as in the beginning of Section 5 (thus including the covariates age, sex, duration and severity). The results after computing the corresponding parameter estimates under different assumptions of the dropout mechanism (MCAR, MAR or MNAR) are listed in Table 3. Provided MAR is the

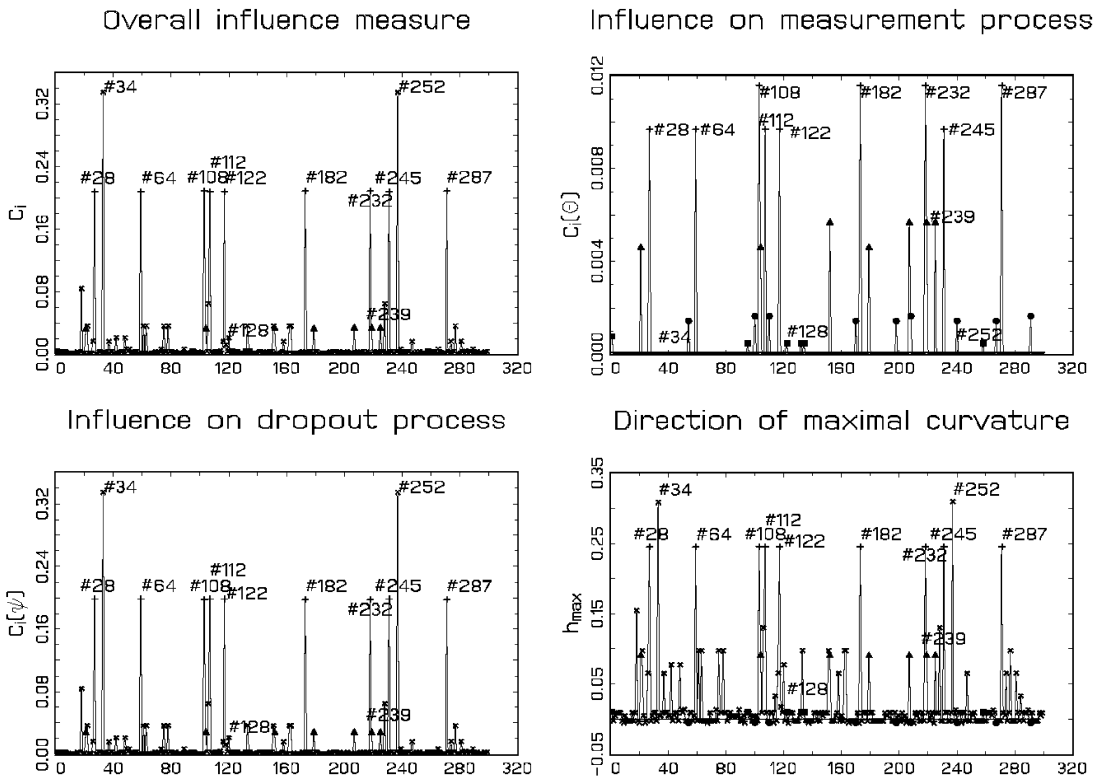

Figure 3 Index plots of C_i, $C_i(\boldsymbol{\theta})$, $C_i(\boldsymbol{\psi})$, and of the components of the direction $\boldsymbol{h}_{\max}$ of maximal curvature, where sex is considered as the only covariate in the Dale model. The *x*-axis contains sequential indicators. Completers are indicated with a solid star. A solid circle, a solid square, a solid triangle or a solid plus is used for non-completers whose score at visit 1 on side effects respectively ranges from (1) to (4)

correct alternative hypothesis and provided the parametric form for the MAR process is correct (again, no covariates were included), there seems to be even less evidence for MAR; the likelihood ratio test statistic comparing MCAR with MAR equals $G^2 = 0.94$, based on 1 degree of freedom ($p = 0.333$). Note that now borderline evidence for MNAR is observed, since a comparison between the non-random and random dropout models generates a likelihood ratio test statistic of $G^2 = 3.74$ with 1 degree of freedom ($p = 0.053$). Hence, the suggested local influence approach bridges the gap between the random and the non-random model: some of the mechanisms that cannot be explained by the random model and are captured by the non-random model, the latter resting on untestable assumptions, can be attributed to the observations for patients #34 and #252.

Repeating the previous analysis on a reduced data set, where patient #239 is excluded instead of patients #34 and #252, we find no evidence for MAR against MCAR (Table 4; $G^2 = 0.01, p = 0.913$). After investigating the likelihood ratio test statistic for comparing the non-random with the random dropout model ($G^2 = 2.13, p = 0.145$), we may conclude (as in Table 2) that the MCAR assumption is fairly plausible. It is not surprising that similar conclusions can be drawn as for Table 2. Indeed, although patient #239 appeared to be the most influential patient with respect to the measurement model parameters, it should

Table 3 Maximum likelihood estimates and standard errors (SEs) of random and non-random dropout models, after exclusion of patients #34 and #252, cidr = 0.349

Effect	MCAR		MAR		MNAR	
	Estimate	SE	Estimate	SE	Estimate	SE
Measurement model						
Intercept 1	−0.254	0.846	−0.254	0.846	−0.412	0.822
Intercept 2	1.946	0.838	1.946	0.838	1.729	0.817
Intercept 3	3.308	0.843	3.308	0.843	3.024	0.829
Age	−0.020	0.008	−0.020	0.008	−0.020	0.008
Duration 1	−0.013	0.005	−0.013	0.005	−0.013	0.005
Duration 2	−0.021	0.006	−0.021	0.006	−0.022	0.006
Severity 1	0.235	0.133	0.235	0.133	0.259	0.131
Severity 2	0.302	0.134	0.302	0.134	0.317	0.133
Sex	−0.125	0.226	−0.125	0.226	−0.104	0.224
Association	3.040	0.306	3.040	0.306	3.141	0.292
Dropout model						
Intercept	−4.451	0.211	−4.644	0.585	−5.390	0.731
Prev. measurement	–	–	1.174	0.231	0.576	0.473
Curr. Measurement	–	–	–	–	0.988	0.541
−2 observed log likelihood	1112.092		1111.964		1108.227	

Table 4 Maximum likelihood estimates and standard errors (SEs) of random and non−random dropout models, after exclusion of patient #239

Effect	MCAR		MAR		MNAR	
	Estimate	SE	Estimate	SE	Estimate	SE
Measurement model						
Intercept 1	−0.351	0.851	−0.351	0.851	−0.461	0.818
Intercept 2	1.799	0.846	1.799	0.846	1.642	0.809
Intercept 3	3.016	0.860	3.016	0.860	2.820	0.814
Age	−0.020	0.008	−0.020	0.008	−0.020	0.008
Duration 1	−0.013	0.005	−0.013	0.005	−0.013	0.005
Duration 2	−0.021	0.006	−0.021	0.006	−0.021	0.006
Severity 1	0.250	0.135	0.250	0.135	0.268	0.131
Severity 2	0.317	0.135	0.317	0.135	0.325	0.133
Sex	−0.117	0.225	−0.117	0.225	−0.102	0.223
Association	3.154	0.302	3.154	0.302	3.240	0.292
Dropout model						
Intercept	−4.529	0.215	−4.472	0.569	−4.961	0.665
Prev. measurement	–	–	1.067	0.223	0.542	0.513
Curr. Measurement	–	–	–	–	0.795	0.584
−2 observed log likelihood	1123.067		1123.055		1120.926	

be noted that (i) the value for $C_i(\boldsymbol{\theta})$ is 'only' 0.079 (further investigation is required to define some critical value above which $C_i(\boldsymbol{\theta})$ can be said to be statistically significantly large) and that (ii) patient #239 did not appear to be influential overall.

6 Concluding remarks

In this paper, we have proposed a possible method to assess influence, via local influence methods, in the case of selection models for incomplete longitudinal ordinal measurements. Such an approach has been followed before by Verbeke *et al.* (2001) and Verbeke and Molenberghs (2000). This approach, while only one way of studying sensitivity, is very broad and can potentially take many forms. Our method is based on the concept of individual-specific infinitesimal perturbations around the MAR model. Technically, our method assigns a perturbation, within the linear predictor of the dropout model, to the so-called current, potentially unobserved measurement. The advantage of the approach is its computational simplicity, given the availability of code to maximize the likelihood of the MNAR version of the model. In such a case, only small modifications suffice.

In all cases, influence decomposes into a measurement and dropout part, the first of which is zero in the case of a complete observation. The latter comment needs careful qualification because it may seem counterintuitive at first sight. However, influence on the dropout parameters translates, through the conditional expectation of unobserved measurements, given dropout, into influence on the measurement model parameters and functions thereof. Therefore, the study of local influence, together with its indirect implications, can provide valuable insight into which observations may lead to a seemingly non-random dropout model, as illustrated in Section 5.

There are important qualitative differences between an influence plot for continuous outcomes (Verbeke *et al.*, 2001) and the ones presented here for ordinal outcomes. In some cases, a layering effect can be seen. This is likely to be the case when the influence of the outcomes and/or of categorical covariates dominates, since then there only a limited number of profiles. When a continuous covariate is a dominating factor, a more scattered picture will be obtained. Further research is ongoing with respect to the stochastic behaviour of the influence measures. This would give more insight into developing a statistically sound cut-off between individuals that are influential and those that are not.

The authors have developed GAUSS code that is available on the Journal website.

Acknowledgements

We gratefully acknowledge support from FWO-Vlaanderen Research Project 'Sensitivity Analysis for Incomplete and Coarse Data'. The first author wishes to thank the Vlaamse Interuniversitaire Raad for granting support. The fourth author gratefully acknowledges support from a research grant of Vlaams Instituut voor de Bevordering van het Wetenschappelijk-Technologisch Onderzoek in de Industrie.

References

Ashford JR, Sowden RR (1970) Multivariate probit analysis. *Biometrics* **26**, 535–46.

Bahadur RR (1961) A representation of the joint distribution of responses to *n* dichotomous items. In: Solomon H, ed. *Studies in item analysis and prediction*. Stanford Mathematical Studies in the Social Sciences VI, Stanford, CA: Stanford University Press.

Beckman RJ, Nachtsheim, Cook RD (1987) Diagnostics for mixed-model analysis of variance. *Technometrics* **29**, 413–26.

Breslow NE, Clayton DG (1993) Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association* **88**, 9–25.

Cook RD (1986) Assessment of local influence. *Journal of the Royal Statistical Society, Series B* **48**, 133–69.

Cook RD, Weisberg S (1982) *Residuals and influence in regression*. London: Chapman and Hall.

Cox DR (1972) The analysis of multivariate binary data. *Applied Statistics* **21**, 113–20.

Dale JR (1986) Global cross-ratio models for bivariate, discrete, ordered responses. *Biometrics* **42**, 909–17.

Diggle PJ, Liang K-Y, Zeger SL (1994) *Analysis of longitudinal data*. Oxford: Oxford Science Publications.

Fahrmeir L, Tutz G (1994) *Multivariate statistical modelling based on generalized linear models*. Heidelberg: Springer-Verlag.

Fitzmaurice GM, Laird NM (1993) A Likelihood-based method for analysing longitudinal binary responses. *Biometrika* **80**, 141–51.

Fitzmaurice GM, Laird NM, Rotnitzky A (1993) Regression models for discrete longitudinal responses. *Statistical Science* **8**, 284–309.

Geys H, Molenberghs G, Ryan LM (1997) Pseudo-likelihood inference for clustered binary data. *Communications in Statistics: Theory and Methods* **26**, 2743–67.

Geys H, Molenberghs G, Ryan L (1999) Pseudolikelihood modelling of multivariate outcomes in developmental toxicology. *Journal of the American Statistical Association* **94**, 734–45.

Glynn RJ, Laird NM, Rubin DB (1986) Selection modelling versus mixture modelling with non-ignorable nonresponse. In: Wainer H, ed. *Drawing inferences from self selected samples*. New York: Springer-Verlag, 115–42.

Hogan JW, Laird NM (1997) Mixture models for the joint distribution of repeated measures and event times. *Statistics in Medicine* **16**, 239–58.

Kenward MG, Molenberghs G, Lesaffre E (1994) An application of maximum likelihood and estimating equations to the analysis of ordinal data from a longitudinal study with cases missing at random. *Biometrics* **50**, 945–53.

Lesaffre E, Verbeke G (1998) Local influence in linear mixed models. *Biometrics* **54**, 570–82.

Liang K-Y, Zeger SL (1986) Longitudinal data analysis using generalized linear models. *Biometrika* **73**, 13–22.

Little RJA, Rubin DB (1987) *Statistical analysis with missing data*. New York: Wiley.

Little RJA (1993) Pattern-mixture models for multivariate incomplete data. *Journal of the American Statistical Association* **88**, 125–34.

Little RJA (1994) A class of pattern-mixture models for normal incomplete data. *Biometrika* **81**, 471–83.

McCullagh P, Nelder JA (1989) *Generalized linear models*. London: Chapman and Hall.

Molenberghs G, Goetghebeur E, Kenward MG (2001) Sensitivity analysis for incomplete contingency tables. The Slovenian plebiscite case. *Journal of the Royal Statistical Society – Series C Applied Statistics* **50**, 15–30.

Molenberghs G, Goetghebeur E, Lipsitz SR, Kenward MG, Lesaffre E, Michiels B (1999) Missing data perspectives of the fluvoxamine data set: a review. *Statistics in Medicine* **18**, 2449–64.

Molenberghs G, Kenward MG, Lesaffre E (1997) The analysis of longitudinal ordinal data with nonrandom dropout. *Biometrika* **84**, 33–44.

Molenberghs G, Lesaffre E (1994) Marginal modelling of correlated ordinal data using a multivariate Plackett distribution. *Journal of the American Statistical Association* **89**, 633–44.

Molenberghs G, Lesaffre E (1999) Marginal modelling of multivariate categorical data. *Statistics in Medicine* **18**, 2237–55.

Molenberghs G, Ryan LM (1999) Likelihood inference for clustered multivariate binary data. *Environmetrics* **10**, 279–300.

Pendergast JF, Gange SJ, Newton MA, Lindstrom MJ, Palta M, Fisher MR (1996) A survey of methods for analyzing clustered binary response data. *International Statistical Review* **64**, 89–118.

Plackett RL (1965) A class of bivariate distributions. *Journal of the American Statistical Association* **60**, 516–22.

Prentice RL (1988) Correlated binary regression with covariates specific to each binary observation. *Biometrics* **44**, 1033–48.

Rosner B (1984) Multivariate methods in ophtalmology with applications to other paired-data situations. *Biometrics* **40**, 1025–35.

Rubin DB (1976) Inference and missing data. *Biometrika* **63**, 581–92.

Ryan LM, Molenberghs G (1999) Statistical methods for developmental toxicity: analysis of clustered multivariate binary data. *Annals of the New York Academy of Sciences* **895**, 196–211.

Scharfstein DO, Rotnitzky A, Robins JM (1999) Adjusting for non-ignorable drop-out using semiparametric nonresponde models (with discussion). *Journal of the American Statistical Association* **94**, 1096–146.

Stiratelli R, Laird N, Ware J (1984) Random effects models for serial observations with dichotomous response. *Biometrics* **40**, 961–72.

Thijs H, Molenberghs G, Michiels B, Verbeke G, Curran D (2000) Strategies to fit pattern-mixture models. (*Submitted*).

Thompson R, Baker, RJ (1981) Composite link functions in generalized linear models. *Applied Statistics* **30**, 125–31.

Verbeke G, Molenberghs G (2000) *Linear mixed models for longitudinal data*. New York: Springer-Verlag.

Verbeke G, Molenberghs G, Thijs H, Lesaffre E, Kenward MG (2001) Sensitivity analysis for non-random dropout: a local influence approach. *Biometrics* **57**, 7–14.

Wolfinger R, O'Connell M (1993) Generalized linear mixed models: a pseudo-likelihood approach. *Journal of Statistical Computation and Simulation* **48**, 233–43.

Zeger SC, Liang K-Y, Albert PS (1988) Models for longitudinal data: a generalized estimating equation approach. *Biometrics* **44**, 1049–60.