



The Diggle-Kenward model for dropout [Link](#)  
**Peer-reviewed author version**

Made available by Hasselt University Library in [Document Server@UHasselT](#)

**Reference** (Published version):

Molenberghs, Geert & Verbeke, G(2006) The Diggle-Kenward model for dropout. In:  
Salkind, Neil J. (Ed.) Encyclopedia of Measurement and Statistics.

DOI: 10.1177/009286151004400406

Handle: <http://hdl.handle.net/1942/11171>

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/280228329>

# The Diggle–Kenward model for dropout

Article · September 2010

---

CITATION

1

---

READS

231

2 authors:



[Geert Molenberghs](#)

Universiteit Hasselt and University of Leuven

892 PUBLICATIONS 20,546 CITATIONS

[SEE PROFILE](#)



[Geert Verbeke](#)

University of Leuven

345 PUBLICATIONS 13,426 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Collaboration project [View project](#)



Implementation of pattern-mixture models [View project](#)

All content following this page was uploaded by [Geert Molenberghs](#) on 26 July 2015.

The user has requested enhancement of the downloaded file.

# The Diggle-Kenward Model for Dropout

## 1 Introduction

In medical research, studies are often designed in which specific parameters are measured repeatedly over time in the participating subjects. This allows modeling the process of change within each subject separately, based on both subject-specific factors (such as gender) and experiment-specific factors (such as treatment). The analysis of this type of longitudinal data requires statistical models which take into account the association between the measurements within subjects. During the last decade, a lot of effort has been put into the search for flexible longitudinal models (Diggle *et al* 2002, Verbeke and Molenberghs 2000, Molenberghs and Verbeke 2005).

In practice, longitudinal studies often suffer from attrition, i.e., subjects dropping out earlier than scheduled, for reasons outside the control of the investigator. The resulting data are then unbalanced with unequal numbers of measures for each participant. Nowadays, several statistical packages can handle unbalanced longitudinal data. However, they only yield valid inferences under specific assumptions for the dropout process.

Generally, valid inferences can only be obtained by modeling the response measurements and the dropout process simultaneously. Making various assumptions about the dropout mechanism, a large variety of models for continuous as well as categorical outcomes have been proposed in the statistical literature (Verbeke and Molenberghs 2000, Molenberghs and Ver-

beke 2005). With the volume of literature on models for incomplete data increasing, there has been growing concern about the critical dependence of many of these models on the validity of the underlying assumptions. To compound the issue, the data often have very little to say about the correctness of such assumptions (Little and Rubin 2002, Chap. 11, Verbeke and Molenberghs 2000).

When referring to the missing-value, or non-response, process we will use terminology of Little and Rubin (2002, Chapter 6). A non-response process is said to be *missing completely at random* (MCAR) if the missingness is independent of both unobserved and observed data and *missing at random* (MAR) if, conditional on the observed data, the missingness is independent of the unobserved measurements. A process that is neither MCAR nor MAR is termed *non-random* (MNAR). In the context of likelihood inference, and when the parameters describing the measurement process are functionally independent of the parameters describing the missingness process, MCAR and MAR are *ignorable*, while a non-random process is non-ignorable. Ignorability implies that valid inferences about the measurement model parameters can be obtained by analyzing the observed data alone, obviating the need for formulation of a dropout model.

We will present one modeling framework that has been developed for incomplete longitudinal data of a continuous nature, proposed by Diggle and Kenward (1994). The model has been subject to criticism since it is rather vulnerable to modeling assumptions made. These concerns will be discussed and a number of ways for dealing with it explored, a prominent place given to sensitivity analysis.

## 2 The Diggle-Kenward Model for Longitudinal Data

### With Dropout

We assume that for subject  $i$  in the study,  $i = 1, \dots, N$ , a sequence of measurements  $Y_{ij}$  is designed to be measured at time points  $t_{ij}$ ,  $j = 1, \dots, n_i$ , resulting in a vector  $\mathbf{Y}_i = (Y_{i1}, \dots, Y_{in_i})'$  of measurements for each participant. If dropout occurs,  $\mathbf{Y}_i$  is only partially observed. We denote the occasion at which dropout occurs by  $D_i > 1$ , and  $\mathbf{Y}_i$  is split into the  $(D_i - 1)$ -dimensional observed component  $\mathbf{Y}_i^{obs}$  and the  $(n_i - D_i + 1)$ -dimensional missing component  $\mathbf{Y}_i^{mis}$ . In case of no dropout, we let  $D_i = n_i + 1$ , and  $\mathbf{Y}_i$  equals  $\mathbf{Y}_i^{obs}$ . The likelihood contribution of the  $i$ th subject, based on the observed data  $(\mathbf{y}_i^{obs}, d_i)$ , is proportional to the marginal density function

$$f(\mathbf{y}_i^{obs}, d_i | \boldsymbol{\theta}, \boldsymbol{\psi}) = \int f(\mathbf{y}_i, d_i | \boldsymbol{\theta}, \boldsymbol{\psi}) d\mathbf{y}_i^{mis} = \int f(\mathbf{y}_i | \boldsymbol{\theta}) f(d_i | \mathbf{y}_i, \boldsymbol{\psi}) d\mathbf{y}_i^{mis}, \quad (1)$$

in which a marginal model for  $\mathbf{Y}_i$  is combined with a model for the dropout process, conditional on the response, and where  $\boldsymbol{\theta}$  and  $\boldsymbol{\psi}$  are vectors of unknown parameters in the measurement model and dropout model, respectively.

Let  $\mathbf{h}_{ij} = (y_{i1}, \dots, y_{i;j-1})$  denote the observed history of subject  $i$  up to time  $t_{i;j-1}$ . The Diggle-Kenward model for the dropout process allows the conditional probability for dropout at occasion  $j$ , given that the subject was still observed at the previous occasion, to depend on the history  $\mathbf{h}_{ij}$  and the possibly unobserved current outcome  $y_{ij}$ , but not on future outcomes  $y_{ik}$ ,  $k > j$ . These conditional probabilities  $P(D_i = j | D_i \geq j, \mathbf{h}_{ij}, y_{ij}, \boldsymbol{\psi})$  can now be used to calculate the probability of dropout at each occasion:

$$P(D_i = j | \mathbf{y}_i, \boldsymbol{\psi}) = P(D_i = j | \mathbf{h}_{ij}, y_{ij}, \boldsymbol{\psi})$$

$$= \begin{cases} P(D_i = j | D_i \geq j, \mathbf{h}_{ij}, y_{ij}, \boldsymbol{\psi}) & j = 2, \\ P(D_i = j | D_i \geq j, \mathbf{h}_{ij}, y_{ij}, \boldsymbol{\psi}) \\ \quad \times \prod_{k=2}^{j-1} [1 - P(D_i = k | D_i \geq k, \mathbf{h}_{ik}, y_{ik}, \boldsymbol{\psi})] & j = 3, \dots, n_i, \\ \prod_{k=2}^{n_i} [1 - P(D_i = k | D_i \geq k, \mathbf{h}_{ik}, y_{ik}, \boldsymbol{\psi})] & j = n_i + 1. \end{cases}$$

Diggle and Kenward (1994) combine a multivariate normal model for the measurement process with a logistic regression model for the dropout process. More specifically, the measurement model assumes that the vector  $\mathbf{Y}_i$  of repeated measurements for the  $i$ th subject satisfies the linear regression model  $\mathbf{Y}_i \sim N(X_i\boldsymbol{\beta}, V_i)$ , ( $i = 1, \dots, N$ ). The matrix  $V_i$  can be left unstructured or assumed to be of a specific form, e.g., resulting from a linear mixed model, a factor-analytic structure, or spatial covariance structure (Verbeke and Molenberghs 2000). The logistic dropout model can, for example, take the form

$$\text{logit} [P(D_i = j | D_i \geq j, \mathbf{h}_{ij}, y_{ij}, \boldsymbol{\psi})] = \psi_0 + \psi_1 y_{ij} + \psi_2 y_{i,j-1}. \quad (2)$$

More general models can easily be constructed by including the complete history  $\mathbf{h}_{ij} = (y_{i1}, \dots, y_{i,j-1})$ , as well as external covariates, in the above conditional dropout model. Note also that, strictly speaking, one could allow dropout at a specific occasion to be related to all future responses as well. However, this is rather counter-intuitive in many cases. Moreover, including future outcomes seriously complicates the calculations since computation of the likelihood (1) then requires evaluation of a possibly high-dimensional integral. Note also that special cases of model (2) are obtained from setting  $\psi_1 = 0$  or  $\psi_1 = \psi_2 = 0$ , respectively. In the first case, dropout is no longer allowed to depend on the current measurement, implying MAR. In the second case, dropout is independent of the outcome, which corresponds to MCAR.

Diggle and Kenward (1994) obtained parameter and precision estimates by maximum likelihood. The likelihood involves marginalization over the unobserved outcomes  $\mathbf{Y}_i^{mis}$ . Practically, this involves relatively tedious and computationally demanding forms of numerical integration. This, combined with likelihood surfaces tending to be rather flat, makes the model difficult to use. These issues are related to the problems to be discussed next.

### 3 Remarks on Sensitivity Analysis and Other Models

Apart from the technical difficulties encountered during parameter estimation, there are further important issues surrounding MNAR based models. Even when the measurement model (e.g., the multivariate normal model) would beyond any doubt be the choice of preference for describing the measurement process *should the data be complete*, then the analysis of the actually observed, incomplete version is, in addition, subject to further untestable modeling assumptions.

When missingness is MAR, the problems are less complex, since it has been shown that, in a likelihood or Bayesian framework, it is sufficient to analyze the observed data, without explicitly modeling the dropout process (Rubin 1976, Molenberghs and Verbeke 2000). However, the very assumption of MAR is itself untestable. Therefore, ignoring MNAR models is as little an option as blindly shifting to one particular MNAR model. A sensible compromise between considering a single MNAR model on the one hand or excluding such models from consideration on the other hand, is to study the nature of such sensitivities and, building on this knowledge, formulate ways for conducting sensitivity analyses. Indeed, a strong conclusion, arising from

most sensitivity analysis work, is that MNAR models have to be approached cautiously. This was made clear by several discussants to the original paper by Diggle and Kenward (1994), in particular by Laird, Little, and Rubin. An implication is that, for example, formal tests for the null hypothesis of MAR versus the alternative of MNAR, should be approached with the utmost caution, a topic studied in detail by Jansen *et al* (2005).

Verbeke, Lesaffre, and Spiessens (2001) have shown, in the context of an onychomycosis study, that excluding a small amount of measurement error, drastically changes the likelihood ratio test statistics for the MAR null hypothesis. Kenward (1998) revisited the analysis of the mastitis data performed by Diggle and Kenward (1994). In this study, the milk yields of 107 cows were to be recorded during two consecutive years. While data were complete in the first year, 27 animals were missing in the second year because they developed mastitis and their milk yield was no longer of use. While in Diggle and Kenward (1994) there was strong evidence for MNAR, Kenward (1998) showed that removing two out of 107 anomalous profiles, completely removed this evidence. In addition, he showed that changing the conditional distribution of the year 2 yield, given the year 1 yield, from a normal distribution to a heavy-tailed  $t$ , also led to the same result of no residual evidence for MNAR. This particular conditional distribution is of great importance, because a subject with missing data does not contribute to it, and hence is a source of sensitivity issues. Once more, the conclusion is that fitting a MNAR model should be subject to careful scrutiny.

In addition to the instances described above, sensitivity to model assumptions has been reported for about two decades (Verbeke and Molenberghs 2000, Molenberghs and Verbeke 2005). In an attempt to formulate an answer to these concerns, a number of authors have pro-



posed strategies to study sensitivity. We broadly distinguish between two types. A first family of approaches can be termed *substantive-driven* in the sense that they start from particularities of the problem at hand. Kenward's (1998) approach falls within this category. Arguably, such approaches are extremely useful, both in their own right and as a preamble to using the second family, where what could be termed *general purpose* tools are used.

Broadly, we could define a sensitivity analysis as one in which several statistical models are considered simultaneously and/or where a statistical model is further scrutinized using specialized tools (such as diagnostic measures). This rather loose and very general definition encompasses a wide variety of useful approaches. The simplest procedure is to fit a selected number of (MNAR) models which are all deemed plausible or one in which a preferred (primary) analysis is supplemented with a number of variations. The extent to which conclusions (inferences) are stable across such ranges provides an indication about the belief that can be put into them. Variations to a basic model can be constructed in different ways. The most obvious strategy is to consider various dependencies of the missing data process on the outcomes and/or on covariates. Alternatively, the distributional assumptions of the models can be changed.

Several authors have proposed the use of global and local influence tools (Verbeke *et al* 2001, Verbeke and Molenberghs 2000, Molenberghs and Verbeke 2005). Molenberghs *et al* (2001) revisited the mastitis example. They were able to identify the same two cows found by Kenward (1998), in addition to another one. Thus, an important question is to what exactly are the sources causing an MNAR model to provide evidence for MNAR against MAR. There is evidence to believe that a multitude of outlying aspects, but not necessarily the (outlying)

nature of the missingness mechanism in one or a few subjects, is responsible for an apparent MNAR mechanism (Jansen *et al* 2005). The consequence of this is that local influence should be applied and interpreted with due caution.

Of course, the above discussion is not limited to the Diggle-Kenward model. A variety of other models have been proposed for incomplete longitudinal data. First, the model has been formulated within the selection model framework, in which the joint distribution of the outcome and dropout processes is factorized as the marginal distribution of the outcomes  $f(\mathbf{y}_i|\boldsymbol{\theta})$  and the conditional distribution of the dropout process, given the outcomes  $f(d_i|\mathbf{y}_i, \boldsymbol{\psi})$ . Molenberghs, Kenward, and Lesaffre (1997) proposed a version for categorical data. Within the selection model framework, models have been proposed for non-monotone missingness as well (Jansen and Molenberghs 2005), and further a number of proposals have been made for non-Gaussian outcomes (Molenberghs and Verbeke 2005). Apart from the selection model framework, so-called pattern-mixture models have gained popularity (Verbeke and Molenberghs 2000, Molenberghs and Verbeke 2005), where the reverse factorization is applied, with factors  $f(\mathbf{y}_i|d_i, \boldsymbol{\theta})$  and  $f(d_i|\boldsymbol{\psi})$ . Also within this framework, both models and sensitivity analysis tools for them have been formulated. A third framework consists of so-called *shared parameter models*, where random effects are employed to describe the relationship between the measurement and dropout processes (Wu and Carroll 1988, DeGruttola and Tu 1994).

Geert Molenberghs and Geert Verbeke

See Longitudinal/Repeated Measures Data, Missing Values, Mixed Models, Repeated Measures Analysis of Variance,

## Further Readings

- De Gruttola, V. and Tu, X.M. (1994) Modelling progression of CD4-lymphocyte count and its relationship to survival time. *Biometrics*, **50**, 1003–1014.
- Diggle, P. and Kenward, M.G. (1994) Informative drop-out in longitudinal data analysis (with discussion). *Applied Statistics*, **43**, 49–93.
- Diggle, P.J., Heagerty, P.J., Liang, K.-Y., and Zeger, S.L. (2002) *Analysis of Longitudinal Data (2nd ed.)*. Oxford Science Publications. Oxford: Clarendon Press.
- Jansen, I., Hens, N., Molenberghs, G., Aerts, M., Verbeke, G., and Kenward, M.G. (2005) The nature of sensitivity in missing not at random models. *Computational Statistics and Data Analysis*, **00**, 000–000.
- Jansen, I. and Molenberghs, G. (2005) A flexible marginal modeling strategy for non-monotone missing data. *Submitted for publication*.
- Kenward, M.G. (1998) Selection models for repeated measurements with nonrandom dropout: an illustration of sensitivity. *Statistics in Medicine*, **17**, 2723–2732.
- Little, R.J.A. and Rubin, D.B. (2002) *Statistical analysis with missing data (2nd ed.)*. New York: John Wiley.
- Molenberghs, G., Kenward, M. G., and Lesaffre, E. (1997) The analysis of longitudinal ordinal data with non-random dropout. *Biometrika*, **84**, 33–44.
- Molenberghs, G. and Verbeke, G. (2005) *Discrete Longitudinal Data*. New York: Springer.

- Molenberghs, G., Verbeke, G., Thijs, H., Lesaffre, E., and Kenward, M.G. (2001) Mastitis in dairy cattle: Local influence to assess sensitivity of the dropout process. *Computational Statistics and Data Analysis*, **37**, 93–113.
- Rubin, D.B. (1976) Inference and missing data. *Biometrika*, **63**, 581–592.
- Verbeke, G., Lesaffre, E., and Spiessens, B. (2001) The practical use of different strategies to handle dropout in longitudinal studies. *Drug Information Journal*, **35**, 419–434.
- Verbeke and Molenberghs (2000) *Linear Mixed Models for Longitudinal Data*. New York: Springer.
- Verbeke, G., Molenberghs, G., Thijs, H., Lesaffre, E., and Kenward, M.G. (2001) Sensitivity analysis for non-random dropout: A local influence approach. *Biometrics*, **57**, 7–14.
- Wu, M.C. and Carroll, R.J. (1988) Estimation and comparison of changes in the presence of informative right censoring by modeling the censoring process. *Biometrics*, **44**, 175–188.