29
Selection Models

29.1 Introduction

Chapters 27 and 28 have shown that, if MAR can be guaranteed to hold, a standard analysis would follow. This is certainly true for likelihood methods, while others, in particular GEE, can be adjusted for the MAR case (Section 27.5).

However, only rarely is such an assumption known to hold (Murray and Findlay 1988). Nevertheless, ignorable analyses may provide reasonably stable results, even when the assumption of MAR is violated, in the sense that such analyses constrain the behavior of the unseen data to be similar to that of the observed data (Mallinckrodt *et al* 2001ab). A discussion of this phenomenon in the survey context has been given in Rubin, Stern, and Vehovar (1995). These authors argue that, in rigidly controlled experiments (some surveys and many clinical trials), the assumption of MAR is often reasonable. Second, and very importantly for such studies as confirmatory trials, an MAR analysis can be specified *a priori* without additional work relative to a situation with complete data. Third, though MNAR models are more general and explicitly incorporate the dropout mechanism, the inferences they produce are typically highly dependent on untestable and often implicit assumptions regarding the distribution of the unobserved measurements given the observed measurements. The quality of the fit to the observed data need not reflect at all the appropriateness of the implied structure governing the unobserved data. This point is irrespective of the MNAR route taken, whether a parametric model of the type of

Diggle and Kenward (1994) or Molenberghs, Kenward, and Lesaffre (1997) is chosen, or a semiparametric approach such as in Robins, Rotnitzky, and Scharfstein (1998). Hence, in incomplete-data settings, a definitive MNAR analysis does not exist. To explore the impact of deviations from the MAR assumption on the conclusions, one should ideally conduct a sensitivity analysis (Chapter 31), within which MNAR models of the selection type as described in this chapter and pattern-mixture models (Chapter 30) can play a major role. See also Verbeke and Molenberghs (2000, Chapter 17–20), for a discussion in the context of continuous longitudinal data.

Diggle and Kenward (1994) describe a modeling procedure for continuous longitudinal data, also discussed in Diggle *et al* (2002, Chapter 11) and Verbeke and Molenberghs (2000, Chapter 17). Based on the multivariate Dale model (Section 7.7), Molenberghs, Kenward, and Lesaffre (1997) proposed a model for repeated ordinal outcomes with MNAR dropout. This model will be described in Section 29.2. The work on incomplete categorical data is vast. Baker and Laird (1988) develop the original work of Fay (1986) and give a thorough account of the modelling of contingency tables in which there is one response dimension and an additional dimension indicating whether the response is absent. Baker and Laird use loglinear models and the EM algorithm for the analysis. They pay particular attention to the circumstances in which no solution exists for the non-random dropout models. Such non-estimability is also a feature of the models we use below, but the more complicated setting makes a systematic account more difficult. Stasny (1986) and Conaway (1992, 1993) consider non-random missingness models for categorical longitudinal data. Baker (1995) allows for intermittent missingness in repeated categorical outcomes. Baker, Rosenberger, and DerSimonian (1992) present a method for incomplete bivariate binary outcomes with general patterns of missingness. The model was adapted for the use of covariates by Jansen *et al* (2003) and is presented in Section 29.3. In both cases, the method is illustrated using the fluvoxamine study, introduced in Section 2.4 and analyzed before in Sections 6.5, 7.2.4, and 7.11. These methods will be employed in Chapter 32 to develop sensitivity analysis tools.

29.2 An MNAR Dale Model

Molenberghs, Kenward, and Lesaffre (1997) proposed a model for longitudinal ordinal data with non-random dropout, i.e., the missingness mechanism was assumed to be MNAR, which combines the multivariate Dale model for longitudinal ordinal data with a logistic regression model for dropout. The resulting likelihood can be maximized relatively simply, using the fact that all stochastic outcomes are of a categorical type, using the EM algorithm. It

means that the integration over the missing data, needed to maximize the likelihood of Diggle and Kenward (1994), is replaced by finite summation.

29.2.1 Likelihood Function

We will derive a general form for the likelihood for longitudinal categorical data with non-random dropout and introduce particular functional forms for the response, using the multivariate Dale model developed by Molenberghs and Lesaffre (1994), see also Section 7.7, and for the dropout process, using a simple logistic regression formulation.

We adopt the contingency table notation, outlined in Section 7.1. Assume we have $r = 1, \ldots, N$ design levels in the study, characterized by covariate information X_r. Let there be N_r subject at design level r. Let the outcome for subject i at level r be a c level ordinal categorical outcome is designed to be measured at occasions $j = 1, \ldots, n$, denoted by Y_{rij}. In principle, we could allow the number of measurement occasions to be different across subjects, but in an incomplete data setting, it is often sensible to assume that the number of measurements at the design stage is constant. Extension to the more general case is straightforward.

As in (7.1), the outcomes at level r are grouped into a contingency table $Z_r^{c*}(k_1 \ldots k_n)$. The cumulative version is $Z_r^c(k_1 \ldots k_n)$ as in (7.2). We have added the superscript c to refer to the (possibly hypothetical) complete data. Shorthand notation is $Z_r^{c*}(\boldsymbol{k})$ and $Z_r^c(\boldsymbol{k})$, and the corresponding cell probabilities are $\mu_r^{c*}(\boldsymbol{k})$ and $\mu_r^c(\boldsymbol{k})$. The corresponding vectors are $\boldsymbol{Z}^{c*}$, $\boldsymbol{Z}^c$, $\boldsymbol{\mu}^{c*}$, and $\boldsymbol{\mu}^c$, respectively.

Any model of the general family described in Section 7.3 can be used, with in particular the multivariate Dale model. The essence is a set of link functions:

$$\boldsymbol{\eta}_r^c(\boldsymbol{\mu}_r^c) = X_r^c \boldsymbol{\beta}. \tag{29.1}$$

Specific choices are discussed in Section 7.3, with in particular the multivariate probit model (Section 7.6) and the multivariate Dale model (Section 7.7). Also the Bahadur model (Section 7.2) can be employed.

We now also need to model the missingness or, in this particular case, the dropout process. Assume the random variable D can take values $2, \ldots, n+1$, the time at which a subject drops out, where $D = n + 1$ indicates no dropout. The value $D = 1$ is not included since we assume at least one follow-up measurement is available. The hypothetical full data consist of complete data and the dropout indicator. The full data, $\boldsymbol{Z}_r^{c*}$, contain components $Z_{rdk_1 \ldots k_n}^{c*}$ with joint probabilities:

$$\nu_{rdk_1 \ldots k_n}^{c*} = \mu_{rk_1 \ldots k_n}^{c*}(\boldsymbol{\beta})\ \phi_{rd|k_1 \ldots k_n}(\boldsymbol{\psi}), \tag{29.2}$$

where the $\boldsymbol{\psi}$ parameterizes the dropout probabilities $\phi_{rd|k_1 \ldots k_n}$. We typically assume both parameters are distinct but this is, strictly speaking, not necessary.

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. The advantage in modeling terms is that the set of unobserved outcomes, relevant to the modeling taks, is a singleton. Also, it is usually deemed plausible in time-ordered longitudinal data, that there is no additional information on the dropout process in the future measurements, given the history and the current, possibly unobserved, measurement.

Factorization (29.2) was made in terms of cell probabilities, superscripted with $*$. The factorization in terms of cumulative probabilities is identical and obtained upon dropping the superscript $*$.

Consequently,

$$\begin{aligned}
&\phi^{c*}_{rd|k_1\ldots k_n}(\boldsymbol{\psi}) \\
&\quad = \quad \phi^{c*}_{rd|k_1\ldots k_d}(\boldsymbol{\psi}) \\
&\quad = \quad \begin{cases} \displaystyle\prod_{t=2}^{d-1} [1 - p_{rt}(H_t, k_t; \boldsymbol{\psi})]\, p_{rd}(H_d, k_d; \boldsymbol{\psi}) & \text{if } D \leq n, \\ \displaystyle\prod_{t=2}^{T} [1 - p_{rt}(H_t, k_t; \boldsymbol{\psi})] & \text{if } D = n+1. \end{cases}
\end{aligned} \tag{29.3}$$

where

$$p_{rd}(H_d, k_d; \boldsymbol{\psi}) = P(D = d | D \geq d, H_d, k_d; W_r; \boldsymbol{\psi}).$$

Here, W_r is a set of covariates, used to model the dropout process. Expression (29.3) is similar to (27.14)–(27.15), used in the context of weighted generalized estimating equations. The difference is that here dropout is allowed to depend on the current, possibly unobserved, measurement.

Molenberghs, Kenward, and Lesaffre (1997) specified the model for the dropout probabilities by logit links, and assuming a linear relationship between the log-odds and the original response. However, the latter is not necessary. For example, non-linear relations and ones involving interactions between the response variables and the covariates could be used. Here, we expect that dropout does not depends on observations preceding k_{d-1}, and thus only depends on k_{d-1} and k_d, but an extension would be straightforward:

$$\text{logit}[p_{rd}(H_d, k_d; \boldsymbol{\psi})] = \psi_0 + \psi_1 k_{d-1} + \psi_2 k_d.$$

This model can also be extended by allowing dependence on covariates W_r. The case $\psi_2 = 0$ corresponds to a MAR dropout process and the case $\psi_1 = \psi_2 = 0$ to a MCAR dropout process.

With dropout occurring, we will not observe $\boldsymbol{Z}^c_r$ but only $\boldsymbol{Z}_r$, a partially classified table, with corresponding probabilities ν_r. The components of ν_r are simple linear functions of the components ν^c_r. This is true for both the cell counts and the cumulative counts.

The multinomial log-likelihood is

$$\ell(\boldsymbol{\beta}, \boldsymbol{\psi}; \boldsymbol{Z}^*) = \ln\left(\frac{1}{\prod_1^N \boldsymbol{Z}_r^*!}\right) + \sum_{r=1}^{N} (\boldsymbol{Z}_r^*)' \ln(\boldsymbol{\nu}_r), \tag{29.4}$$

with the components of ν_r summing to one. The kernel of the log-likelihood is the sum of two contributions. For the complete sequences we have,

$$\begin{aligned}\ell_1(\boldsymbol{\beta}, \boldsymbol{\psi}; \boldsymbol{Z}^*) &= \sum_{r=1}^{N} \sum_{(k_1,\ldots,k_n)} Z^*_{r,n+1,k_1,\ldots,k_n} \\ &\quad \times \log\left\{\mu^*_{rk_1\ldots k_n}(\boldsymbol{\beta}) \prod_{t=2}^{n} [1 - p_{rt}(H_t, k_d; \boldsymbol{\psi})]\right\},\end{aligned}$$

and similarly for the incomplete sequences (say $r = N_1 + 1, \ldots, N = N_1 + N_2$):

$$\begin{aligned}&\ell_2(\boldsymbol{\beta}, \boldsymbol{\psi}; \boldsymbol{Z}^*) \\ &\quad = \sum_{r=1}^{N} \sum_{d=2}^{n} \sum_{(k_1,\ldots,k_{d-1})} Z^*_{rdk_1,\ldots,k_{d-1}} \\ &\qquad \times \ln\left\{\prod_{t=2}^{d-1} [1 - p_{rt}(H_t, k_t; \boldsymbol{\psi})] \sum_{k_d=1}^{c} \mu^*_{rk_1\ldots k_d} p_{rd}(H_d, k_d; \boldsymbol{\psi})\right\}.\end{aligned}$$

We note that, when the probability of dropout does not depend on k_d, i.e., when the dropout process is MAR, the second part of the likelihood partitions into two components, the first for the response process involving $\boldsymbol{\beta}$ only and the second for the dropout process involving $\boldsymbol{\psi}$ only. When the missingness mechanism is MNAR, the resulting likelihood is complex, but the processes of maximization for $\boldsymbol{\beta}$ and for $\boldsymbol{\psi}$ can be separated through the use of the EM algorithm (Dempster, Laird, and Rubin 1977), outlined in Section 28.3. Details are provided in the next section.

29.2.2 Maximization Using the EM Algorithm

We will now show how the likelihood derived in Section 29.2.1 can be maximized using the EM algorithm (Dempster, Laird, and Rubin, 1977; see also Section 28.3), where dropout and response components of the likelihood are maximized separately within each iteration of the algorithm.

Let $(\boldsymbol{\beta}^{(0)}, \boldsymbol{\psi}^{(0)})$ be initial parameters, which can be found from, e.g., a complete case analysis, an available case analysis, or a simple method of imputation. Given current values $(\boldsymbol{\beta}^{(t)}, \boldsymbol{\psi}^{(t)})$ for the parameters, the E step computes the objective function, which is in the case of the missing data

problem equal to the expected value of the observed data log-likelihood, given the observed data and the current parameters:

$$Q\left[(\boldsymbol{\beta},\boldsymbol{\psi})|(\boldsymbol{\beta}^{(t)},\boldsymbol{\psi}^{(t)})\right] = E\left\{\ell\left[(\boldsymbol{\beta},\boldsymbol{\psi})|Z^{c*}_{r,d,k_1\ldots k_n}\right]|Z^{*}_{rdk_1\ldots k_{d-1}},(\boldsymbol{\beta}^{(t)},\boldsymbol{\psi}^{(t)})\right\}.$$

Due to the linearity of the complete data log-likelihood, it is natural to consider the expectations in terms of counts of contingency table Z^{c*}_{r}. Consider now the cell count for a particular joint outcome $(k_1,\ldots,k_n)$ with dropout time d, i.e., $Z^{c*}_{rdk_1\ldots k_n}$. The corresponding observed count is $Z^{*}_{rdk_1\ldots k_{d-1}}$. It can be shown that the conditional expectation for this cell count given the history can be written as

$$\begin{aligned} & E(Z^{c*}_{rdk_1\ldots k_d}|Z^{*}_{rdk_1\ldots k_{d-1}},\boldsymbol{\beta},\boldsymbol{\psi}) \\ &= Z^{*}_{rdk_1\ldots k_{d-1}}\frac{\mu^{c*}_{rk_1\ldots k_d}(\boldsymbol{\beta})p_{rd}(H_d,k_d,\boldsymbol{\psi})}{\sum_{k_d}\mu^{c*}_{rk_1\ldots k_{d-1}k_d}(\boldsymbol{\beta})p_{rd}(H_d,k_d,\boldsymbol{\psi})}. \end{aligned} \tag{29.5}$$

Consequently, the maximization step of the EM cycle requires as input only the expectations $E(Z^{c*}_{rdk_1\ldots k_d}|Z_{rdk_1\ldots k_{d-1}},\boldsymbol{\beta},\boldsymbol{\psi})$ for $k_d = 1,\ldots c$. Given this the likelihood can be partitioned into separate components for the response variable and dropout measurements. Each can be maximized separately using conventional likelihood methods.

To summarize, the two steps of the EM algorithm are as follows.

1. *Expectation.* Predict $Z^{c*}_{rdk_1\ldots k_d}$, $k_d = 1,\ldots,c$ for $d < n+1$, given current estimates of $\boldsymbol{\beta}$ and $\boldsymbol{\psi}$, $\left(\boldsymbol{\beta}^{(t)},\boldsymbol{\psi}^{(t)}\right)$:

$$\begin{aligned} & E\left(Z^{c*}_{rdk_1\ldots k_d}|Z^{*}_{rdk_1\ldots k_{d-1}},\boldsymbol{\beta}^{(t)},\boldsymbol{\psi}^{(t)}\right) \\ &= Z^{*}_{rdk_1\ldots k_{d-1}}.\frac{\mu^{c*}_{rk_1\ldots k_d}\left(\boldsymbol{\beta}^{(t)}\right)p_{rd}\left(H_d,k_d,\boldsymbol{\psi}^{(t)}\right)}{\sum_{k_d}\mu^{c*}_{rk_1\ldots k_{d-1}k_d}\left(\boldsymbol{\beta}^{(t)}\right)p_{rd}\left(H_d,k_d,\boldsymbol{\psi}^{(t)}\right)}. \end{aligned}$$

2. *Maximization.* Maximize separately the kernels of the two components of the likelihood corresponding to the response variable and dropout measurements with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\psi}$:

$$\begin{aligned} \ell^{c}(\boldsymbol{\beta},Z^{c*}) &= \sum_{i=1}^{N}\sum_{d=2}^{n+1}\sum_{(k_1,\ldots,k_d)}Z^{c*}_{rdk_1\ldots k_d}\ln\left(\mu^{*}_{rk_1\ldots k_d}(\boldsymbol{\beta})\right), \\ \ell^{c}(\boldsymbol{\psi},Z^{c*}) &= \sum_{i=1}^{N}\sum_{(k_1,\ldots,k_n)}Z^{c*}_{r,n+1,k_1\ldots k_n}\ln\left(\prod_{t=2}^{n}\{1-p_{rt}(H_t,k_t;\boldsymbol{\psi})\}\right) \\ &\quad +\sum_{i=1}^{N}\sum_{d=2}^{n}\sum_{(k_1,\ldots,k_d)}Z^{c*}_{rdk_1\ldots k_d} \end{aligned}$$

$$\times \ln \left(\prod_{t=2}^{d-1} \left[1 - p_{rt}(H_t, k_t; \boldsymbol{\psi}) \right] p_{rd}(H_d, k_d; \boldsymbol{\psi}) \right).$$

The log-likelihood for the measurement model can be maximized using a Fisher scoring algorithm, as discussed in Sections 7.3 and 7.7.

For the dropout portion of the model, one proceeds as follows. By taking each time of measurement and conditioning on the number of units still present at that time, an overall likelihood can be assembled from independent components and, given k_d, this can be seen to be the likelihood of a conventional logistic regression. The maximum likelihood estimate of ψ can then be obtained simply using iteratively reweighted least squares (McCullagh and Nelder 1989, Section 4.4), or any other tool to maximize a logistic regression based likelihood.

Observe that not only the EM algorithm itself is iterative, but that each M step consists of a pair of iterative maximizations. A way to speed up the EM algorithm is to restrict the iterative schemes in the M step to only a few iterations. This yields a so-called generalized EM algorithm (GEM, Dempster, Laird, and Rubin 1977). Rather than fully maximizing the response log-likelihood and the dropout log-likelihood, one can reduce the number of iterations for either or both of the two maximizations, possibly to one.

Two of the main drawbacks of the EM algorithm are its typically very slow rate of convergence and its lack of direct provision of a measure of precision for the maximum likelihood estimates. Several proposals for overcoming these limitations have been made in the literature, and were discussed in some detail in Sections 28.3.3, 28.3.4, and 28.3.5. Molenberghs, Kenward, and Lesaffre (1997) accelerated convergence using a diagonal matrix analogous to the rate matrix introduced by Meng and Rubin (1991, Eq. 2.2.1). Approximations to the observed Fisher information were found through the technique termed EM-aided differentiation by Meilijson (1989). This technique is easy to implement as it requires a negligible amount of extra code. Standard errors and Wald statistics were computed directly from the observed information and score tests are also relatively simple to compute; calculation of the scores being straightforward. Alternatively, inferences can be based on likelihood ratios; the observed data likelihood is not difficult to evaluate in the current multinomial setting.

All computations were carried out in the statistical programming language GAUSS. As a convergence criterion the L_∞ norm of the relative *observed* data score vector was required to be smaller than 10^{-3}.

29.2.3 Analysis of the Fluvoxamine Data

The data were introduced in Section 2.4 and analyzed before in Sections 6.5, 7.2.4, and 7.11. Analyses of the data, assuming MAR, are described in

TABLE 29.1. *Fluvoxamine Trial. Summary of the ordinal therapeutic outcomes at three follow-up times. (For example, category 241 corresponds to a classification of 2 on the first visit, 4 on the second visit, and 1 on the third visit; a $*$ in one of the positions indicates dropout.)*

Cat	#	Cat	#	Cat	#	Cat	#
Completers							
111	10	211	32	311	12	411	1
112		212	1	312	1	412	
113		213		313		413	
114		214	1	314		414	
121	1	221	13	321	35	421	5
122		222	16	322	14	422	5
123	1	223	1	323	1	423	
124		224	3	324	1	424	1
131		231	1	331	6	431	13
132		232	2	332	5	432	13
133		233	2	333	3	433	5
134		234		334	1	434	
141		241	1	341	1	441	4
142		242		342	2	442	2
143		243	1	343		443	4
144		244		344		444	3
Dropout after 2nd visit							
11*	3	21*	3	31*		41*	
12*		22*	7	32*	7	42*	2
13*		23*	3	33*	3	43*	5
14*		24*	2	34*	1	44*	8
Dropout after 1st visit							
1**	4	2**	6	3**	9	4**	12

Molenberghs and Lesaffre (1994) and Kenward, Lesaffre, and Molenberghs (1994).

From the initially recruited subjects, 14 were not observed at all after the start, 31 and 44 patients, respectively, were observed on the first only and first and second occasions and 224 had complete observations. We omit from the current analyses two patients with non-monotone missing values, leaving 299 in the current analyses. We summarize the therapeutic and side effects results in two sets of contingency tables, Tables 29.1 and 29.2.

TABLE 29.2. *Fluvoxamine Trial. Summary of the ordinal side effects outcomes at three follow-up times. (For example, category 241 corresponds to a classification of 2 on the first visit, 4 on the second visit, and 1 on the third visit; a $*$ in one of the positions indicates dropout.)*

Cat	#	Cat	#	Cat	#	Cat	#
Completers							
111	86	211	25	311	1	411	2
112	5	212	6	312		412	1
113	1	213		313		413	
114		214		314		414	
121	3	221	28	321	1	421	
122		222	39	322	5	422	1
123	7	223	4	323		423	
124		224		324		424	
131		231		331		431	
132		232	4	332	3	432	
133		233		333	2	433	
134		234		334		434	
141		241		341		441	
142		242		342		442	
143		243		343		443	
144		244		344		444	
Dropout after 2nd visit							
11*	13	21*	3	31*	1	41*	
12*	4	22*	9	32*	1	42*	
13*		23*	3	33*	5	43*	
14*		24*	1	34*	2	44*	2
Dropout after 1st visit							
1**	9	2**	6	3**	7	4**	9

For the data on therapeutic effect as well as on side effects we present four sets of parameter estimates. Each set is the result of fitting a marginal proportional odds model to the response and, for non-ignorable models, a logistic regression model to the dropout process. In the first set, the response model alone is fitted to the data from those subjects with complete records. Such an analysis will be consistent with an analysis of the full data set if the dropout process is completely random. The remaining three sets of estimates are obtained from fitting models with non-random, random,

TABLE 29.3. *Fluvoxamine Trial. Maximum likelihood estimates (standard errors) for side effects.*

Parameter	Completers	MCAR	MAR	MNAR
		Measurement model		
intercept 1	1.38(1.00)	-0.60(0.82)	-0.60(0.82)	-0.78(0.79)
intercept 2	4.42(1.04)	1.59(0.83)	1.59(0.83)	1.31(0.80)
intercept 3	6.32(1.14)	2.90(0.85)	2.90(0.85)	2.51(0.82)
age	-0.22(0.08)	-0.20(0.07)	-0.20(0.07)	-0.19(0.07)
sex	-0.35(0.25)	-0.03(0.22)	-0.03(0.22)	0.00(0.21)
duration (visit 1)	-0.05(0.08)	-0.13(0.05)	-0.13(0.05)	-0.12(0.05)
duration (visit 2)	-0.10(0.08)	-0.20(0.06)	-0.20(0.06)	-0.21(0.05)
duration(visit 3)	-0.13(0.08)	-0.19(0.07)	-0.19(0.07)	-0.23(0.06)
severity (visit 1)	0.00(0.16)	0.26(0.13)	0.26(0.13)	0.28(0.12)
severity (visit 2)	0.09(0.16)	0.33(0.13)	0.33(0.13)	0.34(0.13)
severity (visit 3)	0.17(0.16)	0.41(0.13)	0.41(0.13)	0.40(0.13)
Association				
visits 1 and 2	2.89(0.33)	3.12(0.30)	3.12(0.30)	3.26(0.29)
visits 1 and 3	2.06(0.32)	2.33(0.35)	2.33(0.35)	2.30(0.32)
visits 2 and 3	2.86(0.34)	3.16(0.37)	3.16(0.37)	3.18(0.36)
visits 1, 2, and 3	0.45(0.76)	0.48(0.79)	0.48(0.79)	0.61(0.71)
		Dropout model		
ψ_0		-1.90(0.13)	-3.68(0.34)	-4.26(0.48)
ψ_1				1.08(0.54)
ψ_2			0.94(0.15)	0.18(0.45)
-2 log-likelihood		1631.97	1591.98	1587.72

and completely random dropout, defined in terms of constraints on the ψ parameters.

We consider first the analysis of the side-effects data, Table 29.3. Covariates have been included in the response component of the model. The relationships with two covariates, sex and age, have been held constant across visits, the relationships with the other two covariates, duration and severity, have been allowed to differ among visits.

Conditional on acceptance of the validity of the overall model we can, by examining the statistical significance of the parameters in the dropout model, test for different types of dropout process. Three statistics, likelihood-ratio, Wald, and score can be computed for each null hypothesis, and we present each in Table 29.4 for comparisons of (1) MNAR *versus* MAR and

TABLE 29.4. *Fluvoxamine Trial. Side effects. Test statistics for dropout mechanism.*

	MNAR *vs* MAR		MAR *vs* MCAR	
Wald	4.02	$(p = 0.045)$	38.91	$(p < 0.001)$
LR	4.26	$(p = 0.039)$	39.99	$(p < 0.001)$
score	4.24	$(p = 0.040)$	45.91	$(p < 0.001)$

of (2) MAR *versus* MCAR. In line with Diggle and Kenward (1994) and Molenberghs, Kenward, and Lesaffre (1997), it is tempting to assume both statistics follow a null asymptotic χ^2_1 distribution. Jansen *et al* (2005) show that great care has to be taken with the test for MNAR against MAR (see Chapter 31).

All tests provide weak evidence for MNAR in the context of the assumed model. They also strongly support MAR over MCAR. But again, one has to be very cautious with such conclusions. Section 31.3 will study sensitivity of the MNAR model to the model assumptions made. Further detail on the precise nature of sensitivity can be found in Jansen *et al* (2005).

The estimated dropout model is, with simplified notation:

$$\text{logit}[P(\text{dropout})] = -4.26 + 1.08Y_c + 0.18Y_{pr}$$

for Y_{pr} and Y_c the previous and current observations, respectively. It is instructive to rewrite this in terms of the increment and sum of the successive measurements. Standard errors of the estimated parameters have been added in square brackets.

$$\text{logit}[P(\text{dropout})] = -4.26 + 0.63[0.08](Y_c + Y_{pr}) + 0.45[0.49](Y_c - Y_{pr}).$$

It can be seen that the estimated probability of dropout increases greatly with large side effects. The corresponding standard error is comparatively small. Although the coefficient of the increment does not appear negligible in terms of its absolute size, in the light of its standard error it cannot be said to be significantly different from zero. This reflects the lack of information in these data on the coefficient of the increment in the dropout model.

Although the evidence of dependence of the dropout process on previous observation is overwhelming, that for MNAR is borderline.

It is worth noting that there are substantial differences between the analyses of the completers only and full datasets with respect to the parameter estimates of the response model. In the presence of an MAR and MNAR process, the former analysis produces inconsistent estimators. Given the clear association of side-effect occurrence and the covariates age, duration, and severity, we investigated the relationship between these and dropout, but found only marginal evidence for a dependence on sex and severity.

TABLE 29.5. *Fluvoxamine Trial. Maximum likelihood estimates (standard errors) for therapeutic effect.*

Parameter	Completers	MCAR	MAR	MNAR
		Measurement model		
intercept 1	-2.36(0.17)	-2.32(0.15)	-2.32(0.15)	-2.33(0.14)
intercept 2	-0.53(0.13)	-0.53(0.12)	-0.53(0.11)	-0.52(0.10)
intercept 3	1.03(0.14)	0.90(0.11)	0.90(0.12)	0.90(0.09)
visit 2 - visit 1	1.38(0.12)	1.22(0.10)	1.22(0.10)	1.32(0.11)
visit 3 - visit 1	2.70(0.19)	2.58(0.18)	2.58(0.18)	2.83(0.19)
association				
visits 1 and 2	2.58(0.24)	2.57(0.22)	2.57(0.22)	2.46(0.20)
visits 1 and 3	0.85(0.23)	0.86(0.24)	0.86(0.24)	0.77(0.19)
visits 2 and 3	1.79(0.25)	1.79(0.25)	1.79(0.25)	1.59(0.20)
visits 1, 2 and 3	0.39(0.52)	0.27(0.52)	0.27(0.52)	0.22(0.23)
		Dropout model		
ψ_0		-1.88 (0.13)	-2.56(0.37)	-2.00(0.48)
ψ_1				-1.11(0.42)
ψ_2			0.26(0.13)	0.77(0.19)
-2 log-likelihood		2156.91	2152.87	2145.93

In Table 29.5, the results from the analyses of the therapeutic effect are presented. Here, apart from overall effects of time, no covariates are included because all showed negligible association with the response. Interestingly the comparison of the three dropout models (Table 29.6) produces somewhat different conclusions about the dropout mechanism, when compared to those of the side-effects analysis (Table 29.4).

Here, the three classes of tests again behave consistently. The evidence for MNAR is strong, but the same warnings about the sensitivity of the MNAR model to modeling assumptions apply here. The tests comparing the MAR and MCAR processes show only moderate evidence of a difference. The latter tests are not strictly valid however in the presence of MNAR missingness. It is interesting that a comparison of the MCAR and MAR models, which is much easier to accomplish than the comparison of MAR and MNAR, gives little suggestion that such a relationship might exist between dropout and response. This is partly a consequence of the nature of the dropout relationship in this example. With the side-effects the association between dropout and response was dominated by the average response. With the therapeutic observations however dependence of dropout probability is largely on the measurement increment, also a fea-

TABLE 29.6. *Fluvoxamine Trial. Therapeutic effect. Test statistics for dropout mechanism.*

	MNAR *vs* MAR		MAR *vs* MCAR	
Wald	6.98	$(P = 0.008)$	3.98	$(P = 0.046)$
LR	6.94	$(P = 0.008)$	4.03	$(P = 0.044)$
score	9.31	$(P = 0.002)$	4.02	$(P = 0.045)$

ture of the analyses in Diggle and Kenward (1994). From the fitted MNAR model we have:

$$\begin{aligned}\text{logit}\{P(\text{dropout})\} &= -2.00 - 1.11Y_{\text{c}} + 0.77Y_{\text{pr}} \\ &= -2.00 - 0.17[0.17](Y_{\text{c}} + Y_{\text{pr}}) - 0.94[0.28](Y_{\text{c}} - Y_{\text{pr}}).\end{aligned}$$

A plausible interpretation would be that dropout decreases when there is a favorable change in therapeutic effect, and increases only comparatively slightly when there is little therapeutic effect. Larger differences can also be seen among the parameter estimates of the response component, between the MCAR and MAR models on one hand and the non-random dropout model on the other, than are apparent in the analysis of the side effects. The estimated differences between visits are greater in the MNAR model; in the MAR analysis no account is taken of the dependence of dropout on increment, so the sizes of the changes between visits is biased downwards. These differences are however of little practical importance given the sizes of the associated standard errors. Similarly, the statistical dependence between repeated measurements as measured by the log odds-ratios is smaller under the MNAR model, possibly because of the effect of selection under the MAR model.

29.3 A Model for Non-monotone Missingness

In Section 29.2, we presented a model for ordinal data but confined missingness to the dropout type. Here, general missingness will be studied, in the specific context of a bivariate binary outcome.

Baker, Rosenberger, and DerSimonian (1992) considered a log-linear type of model for two binary outcomes subject to incompleteness. A main advantage of this method is that it can easily deal with non-monotone missingness.

As in Section 29.2, let $r = 1, \ldots, N$ index distinct covariate levels. In this section, the index r will be suppressed from notation. Let $j, k = 1, 2$ correspond to the outcome categories of the first and second measurement, respectively and let $r_1, r_2 = 0, 1$ correspond to the missingness indicators

(1 for an observed and 0 for a missing measurement). Such a setup leads to a four-way classification. The complete data and observed data cell probabilities $\pi_{r_1r_2,jk}$ for this setting are presented in Figure 29.1.

To accommodate (possibly continuous) covariates, as proposed by Jansen *et al* (2003), we will use a parameterization, different from and extending the original one, which belongs to the selection model family (Little 1994):

$$\pi_{r_1r_2,jk} = p_{jk} q_{r_1r_2|jk}, \tag{29.6}$$

where p_{jk} parameterizes the measurement process and $q_{r_1r_2|jk}$ describes the missingness mechanism, conditional on the measurements. In particular, we will assume

$$p_{jk} = \frac{\exp(\boldsymbol{\theta}_{jk})}{\sum_{j,k=1}^{2} \exp(\boldsymbol{\theta}_{jk})}, \tag{29.7}$$

$$q_{r_1r_2|jk} = \frac{\exp[\beta_{jk}(1-r_2) + \alpha_{jk}(1-r_1) + \gamma(1-r_1)(1-r_2)]}{1 + \exp(\beta_{jk}) + \exp(\alpha_{jk}) + \exp(\beta_{jk} + \alpha_{jk} + \gamma)}, \tag{29.8}$$

for unknown parameters θ_{jk}, β_{jk}, α_{jk}, and γ. *A priori,* no ordering is imposed on the outcomes. The advantage is that genuine multivariate settings (e.g., several questions in a survey) can be handled as well. When deemed necessary, the implications of ordering can be imposed by considering specific models and leaving out others. For example, one may want to avoid missingness on future observations. In the current bivariate case, the index k would have to be removed from α in the above model. To identify the model, we set $\theta_{22} = \mathbf{0}$ and further $\boldsymbol{\theta}_{jk} = X_{jk}\boldsymbol{\eta}$. This allows the inclusion of covariate effects that, together with (29.7), is similar in spirit to the multigroup logistic model (Albert and Lesaffre 1986). Even though the parameters $\boldsymbol{\eta}$ are conditional in nature and therefore somewhat difficult to directly interpret in case planned sequences are of unequal length (but not in the case considered here), (29.7) allows easy calculation of the joint probabilities. Such computational advantages become increasingly important as the length of the response vector grows. If necessary, specific functions of interest, such as a marginal treatment effect, can be derived. They will typically take the form of non-linear functions. Arguably, a model of the type here can be most useful as a component of a sensitivity analysis, in conjunction with the use of different (e.g., marginal) models.

In many examples, the design matrices X_{jk} will be equal to each other. Stacking all parameters will lead to the following design:

$$\boldsymbol{\theta} = X\boldsymbol{\eta}. \tag{29.9}$$

Likewise, a design can be constructed for the non-response model parameters:

$$\boldsymbol{\delta} = W\boldsymbol{\psi}, \tag{29.10}$$

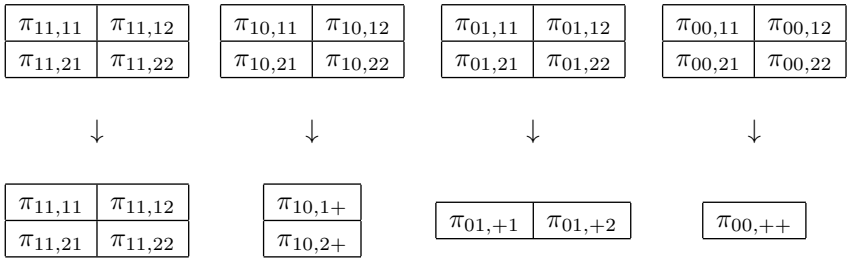

FIGURE 29.1. *Theoretical distribution over complete and observed cells of a bivariate binary outcome. Tables correspond to completely observed subjects and subjects with the second, the first, and both measurements missing, respectively.*

where the vector $\boldsymbol{\delta}$ stacks the β_{jk}, α_{jk} and γ and W is an appropriate design matrix. The vector $\boldsymbol{\psi}$ groups the parameters of interest. For example, if MCAR would be considered, the α and β parameters do not depend on neither j nor k and then $\boldsymbol{\psi}' = (\alpha, \beta, \gamma)$. Both designs (29.9) and (29.10) can be combined into one, using $\boldsymbol{\xi} = (\boldsymbol{\theta}', \boldsymbol{\delta}')'$,

$$T = \begin{pmatrix} X & 0 \\ 0 & W \end{pmatrix},$$

and

$$\boldsymbol{\phi} = (\boldsymbol{\eta}', \boldsymbol{\psi}')'. \tag{29.11}$$

The corresponding log-likelihood function can be written as:

$$\begin{aligned}
\ell &= \sum_{j,k=1}^{2} y_{11jk} \ln \pi_{11jk} + \sum_{j=1}^{2} y_{10j+} \ln(\pi_{10j1} + \pi_{10j2}) \\
&\quad + \sum_{k=1}^{2} y_{01+k} \ln(\pi_{011k} + \pi_{012k}) \\
&\quad + y_{00++} \ln(\pi_{0011} + \pi_{0012} + \pi_{0021} + \pi_{0022}) \\
&= \sum_{j,k=1}^{2} \sum_{s=1}^{y_{11jk}} \ln \pi_{11jk} + \sum_{j=1}^{2} \sum_{s=1}^{y_{10j+}} \ln \pi_{10j+} \\
&\quad + \sum_{k=1}^{2} \sum_{s=1}^{y_{01+k}} \ln \pi_{01+k} + \sum_{s=1}^{y_{00++}} \ln \pi_{00++}.
\end{aligned}$$

Computation of derivatives, needed for optimization and for the calculation of influence measures, is straightforward. A technical report can be obtained from the authors upon request.

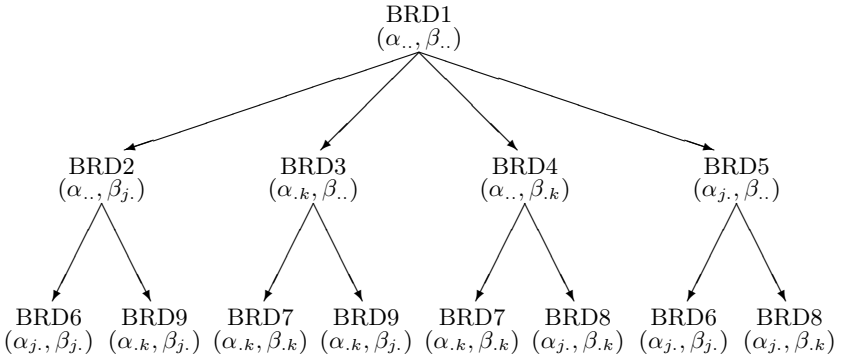

FIGURE 29.2. *Graphical representation of the BRD model nesting structure.*

To include covariates, the design level $r = 1, \ldots, N$ needs to be introduced again. In particular, with subject-specific covariates, it may be sensible to use $i = 1, \ldots, N$ to index individuals.

Baker, Rosenberger, and DerSimonian (1992, BRD) consider nine identifiable models, based on setting α_{jk} and β_{jk} constant in one or more indices. An overview, together with the nesting structure, is given in Figure 29.2.

Whereas these authors considered the nine models in terms of the original parameterization, they do carry over to parameterization (29.8). Interpretation is straightforward. For example, BRD1 is MCAR, in BRD4 missingness in the first variable is constant, while missingness in the second variable depends on its value. Two of the main advantages of this family are ease of computation in general, and the existence of a closed-form solution for several of its members (BRD2 to BRD9).

29.3.1 Analysis of the Fluvoxamine Data

In the analysis, all patients with known duration level are considered, leaving a total of 310 out of 315 subjects in the study. In the measurement model, the effect of duration is held constant over both visits. Regarding the missingness model, an effect of duration is assumed in both the α and the β parameters. Each of the 9 models is represented by a specific choice for the design. For example, for BRD1, and using the index i for individual, we obtain:

$$\boldsymbol{\phi} = (\eta_1, \eta_2, \eta_3, \eta_4, \alpha, \alpha_{\text{dur}}, \beta, \beta_{\text{dur}}, \gamma)',$$

$$X_i = \begin{pmatrix} 1 & 0 & 0 & \text{duration}_i \\ 0 & 1 & 0 & \text{duration}_i \\ 0 & 0 & 1 & \text{duration}_i \end{pmatrix}$$

TABLE 29.7. *Fluvoxamine Trial. Maximum likelihood estimates and standard errors of BRD models. All observations included. No covariates. Part I.*

Effect	BRD1	BRD2	BRD3	BRD4	BRD5
Measurement model					
Intercept_{11}	0.22(0.15)	0.20(0.15)	0.28(0.15)	0.03(0.17)	0.32(0.15)
Intercept_{12}	-1.72(0.30)	-1.74(0.30)	-1.72(0.30)	-1.61(0.30)	-1.62(0.30)
Intercept_{21}	-0.12(0.18)	-0.12(0.18)	-0.05(0.18)	-0.42(0.23)	-0.13(0.18)
Dropout model					
α	-4.72(0.71)	-4.72(0.71)		-4.72(0.71)	
$\alpha_{1.}$					-3.87(0.71)
$\alpha_{2.}$					$-\infty$
$\alpha_{.1}$			-4.27(0.71)		
$\alpha_{.2}$			$-\infty$		
β	-1.09(0.13)		-1.09(0.13)		-1.09(0.13)
$\beta_{1.}$		-1.37(0.22)			
$\beta_{2.}$		-0.91(0.17)			
$\beta_{.1}$				-1.57(0.38)	
$\beta_{.2}$				-0.55(0.29)	
γ	3.04(0.77)	3.04(0.77)	3.04(0.77)	3.04(0.77)	3.04(0.77)
- loglik	565.96	564.55	565.07	564.55	565.34

and

$$W_i = \begin{pmatrix} 1 & \text{duration}_i & 0 & 0 & 0 \\ 1 & \text{duration}_i & 0 & 0 & 0 \\ 1 & \text{duration}_i & 0 & 0 & 0 \\ 1 & \text{duration}_i & 0 & 0 & 0 \\ 0 & 0 & 1 & \text{duration}_i & 0 \\ 0 & 0 & 1 & \text{duration}_i & 0 \\ 0 & 0 & 1 & \text{duration}_i & 0 \\ 0 & 0 & 1 & \text{duration}_i & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

The matrix X_i includes a time dependent intercept and a time independent effect of duration. The W_i matrix indicates which of the nine BRD models is considered; changing the model also changes the vector $\boldsymbol{\psi}$.

We will consider three sets of BRD models in some detail. Tables 29.7 and 29.8 presents models (parameter estimates, standard errors, negative log-likelihoods) without duration. In Tables 29.9 and 29.10, duration is added as a covariate to the measurement model but not yet to the missingness model, whereas in the final set (Tables 29.11 and 29.12) the effect of duration is included in both measurement and missingness parts of the

TABLE 29.8. *Fluvoxamine Trial. Maximum likelihood estimates and standard errors of BRD models. All observations included. No covariates. Part II.*

Effect	BRD6	BRD7	BRD8	BRD9
Measurement model				
Intercept_{11}	0.32(0.15)	0.14(0.16)	0.16(0.17)	0.27(0.15)
Intercept_{12}	-1.62(0.30)	-1.61(0.30)	-1.44(0.32)	-1.72(0.30)
Intercept_{21}	-0.13(0.18)	-0.31(0.21)	-0.39(0.22)	-0.04(0.17)
Dropout model				
α				
$\alpha_{1.}$	-3.93(0.71)		-3.93(0.71)	
$\alpha_{2.}$	-∞		-∞	
$\alpha_{.1}$		-4.29(0.71)		-4.29(0.71)
$\alpha_{.2}$		-∞		-∞
β				
$\beta_{1.}$	-1.37(0.22)			-1.37(0.22)
$\beta_{2.}$	-0.91(0.17)			-0.91(0.17)
$\beta_{.1}$		-1.57(0.38)	-1.56(0.37)	
$\beta_{.2}$		-0.56(0.29)	-0.56(0.29)	
γ	3.31(0.79)	3.51(0.84)	3.31(0.79)	3.11(0.77)
- loglik	563.97	563.70	563.97	563.70

model. Sampling zeroes in some of the cells forces some parameters to lie on the boundary of their corresponding parameter space which, due to the parameterization, is equal to ∞. This should not be seen as a disadvantage of our model, as boundary solutions are a well-known feature of MNAR models (Rubin 1996). The advantage of our parameterization is that either an interior or a boundary solution is obtained, and never an invalid solution.

From Tables 29.7 and 29.8, we learn that likelihood ratio tests fail to reject BRD1 in favor of a more complex model, implying the simplest mechanism, MCAR would be adequate. However, this conclusion changes when duration is included in the measurement model (Tables 29.9 and 29.10). The effect of duration is highly significant, whichever of the nine BRD models is chosen to conduct a likelihood ratio test. In addition, within Tables 29.9 and 29.10, not BRD1 but rather BRD4 provides the most adequate description. The likelihood ratio test statistic for comparing BRD1–4 equals 7.10, while those for BRD4–7 and BRD4–8 are 2.10 and 1.52, respectively. Thus, from this set of models, one observes that duration improves the fit and, moreover, one would be inclined to believe duration, included in the measurement model, has the effect of changing the nature of the missingness mechanism, by making it more complex, even though it is often

TABLE 29.9. *Fluvoxamine Trial. Maximum likelihood estimates and standard errors of BRD models. All observations included. Duration as covariate in the measurement model. Part I.*

Effect	BRD1	BRD2	BRD3	BRD4	BRD5
Measurement model					
Intercept_{11}	0.46(0.17)	0.45(0.17)	0.53(0.17)	0.23(0.20)	0.57(0.17)
Intercept_{12}	-1.46(0.31)	-1.48(0.31)	-1.46(0.31)	-1.26(0.32)	-1.37(0.31)
Intercept_{21}	0.10(0.20)	0.10(0.19)	0.17(0.20)	-0.25(0.23)	0.09(0.21)
Duration	-0.02(0.01)	-0.02(0.01)	-0.02(0.01)	-0.02(0.01)	-0.02(0.01)
Dropout model					
α	-4.71(0.71)	-4.71(0.71)		-4.71(0.71)	
$\alpha_{1.}$					-3.85(0.71)
$\alpha_{2.}$					$-\infty$
$\alpha_{.1}$			-4.24(0.71)		
$\alpha_{.2}$			$-\infty$		
β	-1.11(0.13)		-1.11(0.13)		-1.11(0.13)
$\beta_{1.}$		-1.44(0.23)			
$\beta_{2.}$		-0.90(0.17)			
$\beta_{.1}$				-1.86(0.45)	
$\beta_{.2}$				-0.43(0.25)	
γ	2.98(0.77)	2.98(0.77)	2.98(0.77)	2.98(0.77)	2.98(0.77)
- loglik	550.15	548.31	549.12	546.60	549.39

believed that including explanatory variables (either in the model for the outcomes or in the missingness model) may help to explain structure in the missingness mechanism. BRD4 states that missingness at the second occasion depends on the (possibly unobserved) value at that same occasion, a so-called type I model, in the typology of Baker (2000), in contrast to type II models, where missingness in a variable depends at least also on other, possibly incomplete, assessments. Obviously, such models are particularly vulnerable to assumptions made.

A key conclusion is that, up to this point, no covariate effects have been considered on the missingness parameters. An analysis including duration in the missingness part of the model should be entertained and examined carefully. When switching to Tables 29.11 and 29.12, the conclusions do change drastically. First, all evidence for non-MCAR missingness disappears as, based on likelihood ratio tests, BRD1 comes out as the most adequate description of all nine models. Second, comparing corresponding BRD models between Tables 29.9 and 29.10 on the one hand and Tables 29.11 and 29.12 (p-values in bottom line of Tables 29.11 and 29.12),

TABLE 29.10. *Fluvoxamine Trial. Maximum likelihood estimates and standard errors of BRD models. All observations included. Duration as covariate in the measurement model. Part II.*

Effect	BRD6	BRD7	BRD8	BRD9
Measurement model				
Intercept_{11}	0.57(0.17)	0.35(0.18)	0.36(0.19)	0.52(0.18)
Intercept_{12}	-1.37(0.31)	-1.26(0.32)	-1.06(0.33)	-1.46(0.31)
Intercept_{21}	0.09(0.20)	-0.13(0.21)	-0.21(0.22)	0.18(0.20)
Duration	-0.02(0.01)	-0.02(0.01)	-0.02(0.01)	-0.02(0.01)
Dropout model				
α				
$\alpha_{1.}$	-3.92(0.71)		-3.94(0.71)	
$\alpha_{2.}$	-∞		-∞	
$\alpha_{.1}$		-4.28(0.71)		-4.26(0.71)
$\alpha_{.2}$		-∞		-∞
β				
$\beta_{1.}$	-1.44(0.23)			-1.44(0.23)
$\beta_{2.}$	-0.90(0.17)			-0.90(0.17)
$\beta_{.1}$		-1.87(0.46)	-1.86(0.45)	
$\beta_{.2}$		-0.43(0.25)	-0.43(0.25)	
γ	3.31(0.79)	3.74(0.89)	3.39(0.79)	3.07(0.77)
- loglik	547.57	545.55	545.84	547.30

it is clear that the effect of duration on the missingness model cannot be neglected.

Important modeling and data analytic conclusions can be drawn from this. First, it clearly does not suffice to consider covariate effects on the measurement model, but one has to carefully contemplate such effects on the missingness model as well. Therefore, the models in Tables 29.11 and 29.12, should be regarded as the ones of primary interest. Second, it is found that a longer duration implies a less favorable side-effects outcome, as well as an increased change of missing visits. Obviously, duration acts as a confounding variable which, unless included in both parts of the model, may suggest a relationship between the measurement and missingness models and thus one may erroneously be led to believe that the missing data are MNAR. Third, it should be noted that the parameter estimates of duration are remarkably stable. This implies that, in case one is primarily interested in the effect of duration on the occurrence of side effects all 18 models containing this effect provide very similar evidence. Although this need not be the case in general, it is a comforting aspect of this particular data analysis.

TABLE 29.11. *The Fluvoxamine Trial. Maximum likelihood estimates and standard errors of BRD models. All observations included. Duration as covariate in both measurement and missingness model. Part I.*

Effect	BRD1	BRD2	BRD3	BRD4	BRD5
Measurement model					
Intercept_{11}	0.46(0.18)	0.45(0.17)	0.53(0.18)	0.30(0.20)	0.57(0.17)
Intercept_{12}	-1.46(0.31)	-1.48(0.31)	-1.46(0.31)	-1.37(0.31)	-1.37(0.31)
Intercept_{21}	0.10(0.20)	0.10(0.20)	0.17(0.20)	-0.15(0.24)	0.09(0.20)
Duration	-0.02(0.01)	-0.02(0.01)	-0.02(0.01)	-0.02(0.01)	-0.02(0.01)
Dropout model					
$\alpha_{..}$	-4.57(0.72)	-4.57(0.72)		-4.57(0.72)	
$\alpha_{1.}$					-3.82(0.73)
$\alpha_{2.}$					$-\infty$
$\alpha_{.1}$			-4.20(0.72)		
$\alpha_{.2}$			$-\infty$		
α_{dur}	-0.02(0.02)	-0.02(0.02)	-0.01(0.02)	-0.02(0.02)	-0.01(0.02)
$\beta_{..}$	-1.40(0.16)		-1.40(0.16)		-1.40(0.16)
$\beta_{1.}$		-1.63(0.24)			
$\beta_{2.}$		-1.22(0.20)			
$\beta_{.1}$				-1.79(0.36)	
$\beta_{.2}$				-0.87(0.33)	
β_{dur}	0.02(0.01)	0.02(0.01)	0.02(0.01)	0.02(0.01)	0.02(0.01)
γ	3.10(0.78)	3.10(0.78)	3.10(0.77)	3.10(0.78)	3.09(0.78)
- loglik	543.78	542.74	542.86	542.63	543.14
$p^{\dagger}$	0.0017	0.0038	0.0019	0.0189	0.0019

† p-value for the comparison with the corresponding BRD model in Table 29.9, to test the null hypothesis of no effect of duration in the missingness model.

However, though we have reached plausible conclusions, one should still exercise caution, as non-random missingness models heavily rely on untestable assumptions (Verbeke and Molenberghs 2000). Therefore, it is important to search for observations that may drive these conclusions. This naturally leads to the concept of sensitivity analysis. In Sections 31.4 and 31.5, sensitivity analysis tools applicable to the BRD model, or its extension to covariates used here, will be introduced.

TABLE 29.12. *Fluvoxamine Trial. Maximum likelihood estimates and standard errors of BRD models. All observations included. Duration as covariate in both measurement and missingness model. Part II.*

Effect	BRD6	BRD7	BRD8	BRD9
Measurement model				
Intercept_{11}	0.57(0.17)	0.41(0.18)	0.43(0.19)	0.52(0.18)
Intercept_{12}	-1.37(0.31)	-1.37(0.31)	-1.22(0.33)	-1.46(0.31)
Intercept_{21}	0.09(0.21)	-0.04(0.22)	-0.13(0.23)	0.18(0.20)
Duration	-0.02(0.01)	-0.02(0.01)	-0.02(0.01)	-0.02(0.01)
Dropout model				
$\alpha_{..}$				
$\alpha_{1.}$	-3.87(0.73)		-3.88(0.73)	
$\alpha_{2.}$	-∞		-∞	
$\alpha_{.1}$		-4.23(0.73)		-4.22(0.72)
$\alpha_{.2}$		-∞		-∞
α_{dur}	-0.01(0.02)	-0.01(0.02)	-0.00(0.02)	-0.01(0.02)
$\beta_{..}$				
$\beta_{1.}$	-1.63(0.24)			-1.63(0.24)
$\beta_{2.}$	-1.22(0.20)			-1.22(0.20)
$\beta_{.1}$		-1.79(0.36)	-1.77(0.35)	
$\beta_{.2}$		-0.88(0.33)	-0.88(0.33)	
β_{dur}	0.02(0.01)	0.02(0.01)	0.02(0.01)	0.02(0.01)
γ	3.33(0.79)	3.50(0.84)	3.32(0.79)	3.16(0.78)
- loglik	542.14	541.77	542.05	541.86
$p^{\dagger}$	0.0044	0.0228	0.0226	0.0043

† p-value for the comparison with the corresponding BRD model in Table 29.10, to test the null hypothesis of no effect of duration in the missingness model.

29.4 Concluding Remarks

In Section 29.2, a modeling approach for incomplete ordinal outcomes with dropout was presented. The approach is very general and any measurement model can be used. In fact, it is easy enough to adapt the method to any type of outcome. Not only marginal models, also random-effects models can be used by way of measurement model. In Section 29.3, a model specifically for binary data, but then with general missingness patterns, has been presented. The one limitation of the model in Section 29.2 is its suitability to dropout only. Several extensions to general missingness have been studied in the literature. Troxel, Harrington, and Lipsitz (1998) presented methods for non-ignorable non-monotone missingness. Baker (1995) pre-

sented a modification of Diggle and Kenward (1994) to accommodate non-monotone missingness. Jansen and Molenberghs (2005) modify the model of Section 29.2 to account for non-monotone missingness by replacing the logistic regressions for dropout with a second multivariate Dale model to describe the vector of missingness indicators, given the outcomes.

Thus, a wide variety of selection models is available for incomplete longitudinal data, under MNAR and possibly also with non-monotone missingness. Nevertheless, care has to be taken with such models. As with all model fitting the conclusions drawn are conditional on the appropriateness of the assumed model. Especially here, there are aspects of the model that are in a fundamental sense not testable, namely the relationship between dropout and the missing observations. It is assumed in the modeling approach taken here that the relationships among the measurements from a subject are the same whether or not some of these measurements are unobserved due to dropout. It is this assumption, combined with the adoption of an explicit model linking outcome and dropout probability, that allows us to infer something about the MNAR nature of the dropout process. Given the dependence of the inferences on untestable assumptions, care is needed in the interpretation of the analysis.

The absence of evidence for non-random dropout may simply mean that a non-random dropout process is operating in a quite different manner, and in practice it is likely that many such processes are operating simultaneously.

Thus, the sensitivity of the posited model to modeling assumption needs to be addressed with great caution. Verbeke and Molenberghs (2000, Chapter 19 and 20) discussed ways to assess such sensitivities with continuous longitudinal data. We refer to Chapter 31 for a discussion of sensitivity analysis in the non-Gaussian setting.