16

Marginal *versus* Random-effects Models

16.1 Introduction

The most frequently used models for discrete repeated measurements are of the marginal or random-effects type, and most of them can be viewed as direct extensions of general linear models introduced in Chapter 3 for independent observations to the context of correlated data. Despite the severe similarities between marginal and random-effects model specifications, both families often produce very different results, confusing many statisticians less familiar with these types of models. The aim of the current chapter is therefore to investigate why such strong differences occur in so many applications. In Section 16.2, marginal and random-effects results are compared for the toenail data. Section 16.3 provides some theoretical arguments about the observed differences between both modeling approaches. Finally, the Sections 16.4 and 16.5 will apply these ideas to the toenail and the NTP data, respectively.

16.2 Example: The Toenail Data

Table 16.1 summarizes the parameter estimates and standard errors for a marginal model and a random-effects model, fitted to the toenail data. Both models include linear time-effects, with treatment-specific intercepts and slopes. The marginal model parameter estimates are obtained using generalized estimating equations (GEE1), where a marginal logit function

TABLE 16.1. *Toenail Data. Parameter estimates (standard errors) for a generalized linear mixed model (GLMM) and a marginal model (GEE), as well as the ratio between both sets of parameters.*

	GLMM	GEE	
Parameter	Estimate (s.e.)	Estimate (s.e.)	Ratio
Intercept group A	−1.63 (0.44)	−0.72 (0.17)	2.26
Intercept group B	−1.75 (0.45)	−0.65 (0.17)	2.69
Slope group A	−0.40 (0.05)	−0.14 (0.03)	2.87
Slope group B	−0.57 (0.06)	−0.25 (0.04)	2.22
SD random intercept (τ)	4.02 (0.38)		

is combined with unstructured working assumptions about the association structure. The random-effects model is of the logistic-normal type, with no other random effects than intercepts with variance τ^2, fitted using adaptive Gaussian quadrature with 50 quadrature points. The models are reparameterized versions for the models used earlier in the Chapters 10 and 15, for the same data. Obviously, both analyses produce very different results in the sense that the estimates from the generalized linear mixed model analysis are much bigger in magnitude.

16.3 Parameter Interpretation

The severe differences in results obtained from marginal and random-effects models follow from the fact that the parameters in both models have completely different interpretations. To see the nature of the difference between both model families, consider a binary outcome variable and assume a random-intercepts logistic model with linear predictor $\text{logit}[P(Y_{ij} = 1|b_i)] = \beta_0 + b_i + \beta_1 t$, where t is the time covariate. This model was used in Section 16.2 for each treatment group separately. The conditional means $E(Y_{ij}|b_i)$, as functions of t, are given by

$$E(Y_{ij}|b_i) = \frac{\exp(\beta_0 + b_i + \beta_1 t)}{1 + \exp(\beta_0 + b_i + \beta_1 t)}. \tag{16.1}$$

The model assumes that the conditional means all satisfy a logistic model, with the same slope β_1 but with different intercepts $\beta_0 + b_i$ for all subjects. The marginal average evolution $E(Y_{ij})$ is obtained from averaging (16.1) over the random effects, i.e.,

$$\begin{aligned} E(Y_{ij}) &= E[E(Y_{ij}|b_i)] \\ &= E\left[\frac{\exp(\beta_0 + b_i + \beta_1 t)}{1 + \exp(\beta_0 + b_i + \beta_1 t)}\right] \end{aligned} \tag{16.2}$$

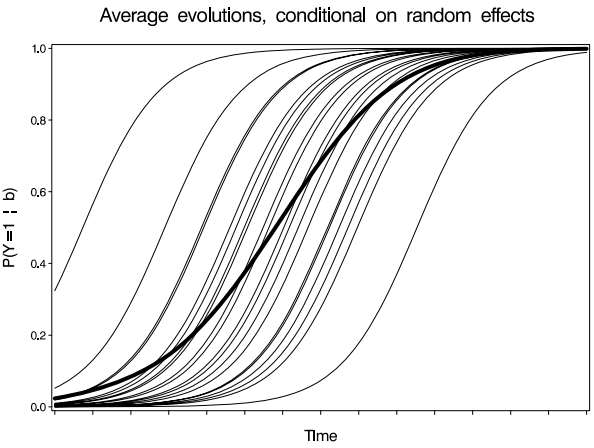

FIGURE 16.1. *Graphical representation of a random-intercepts logistic curve, across a range of levels of the random intercept, together with the corresponding marginal curve.*

$$\neq \quad \frac{\exp(\beta_0 + \beta_1 t)}{1 + \exp(\beta_0 + \beta_1 t)}.$$

A graphical representation of both (16.1) and (16.2) is given in Figure 16.1. Obviously, the marginal time trend is much less steep than each of the individual time trends. Intuitively, it is to be expected that this effect strongly depends on the amount of between-subject variability: In case the random-intercepts variability is large, parameters from fitting marginal models and random-effects models will be very different, while equal parameter values hold if the variance of the random-effects equals zero.

Figure 16.1 clearly shows that the regression parameters in marginal and random-effects models have a completely different inerpretation. Therefore, it may be helpful to denote them differently, such as $\boldsymbol{\beta}^{\mathrm{RE}}$ for the parameter vector in the random-effects model, and $\boldsymbol{\beta}^{\mathrm{M}}$ for the parameter vector in the marginal model. The vector $\boldsymbol{\beta}^{\mathrm{RE}}$ models the evolution of each individual subject separately, whereas $\boldsymbol{\beta}^{\mathrm{M}}$ expresses how, on average, the success probability evolves in the population.

This phenomenon holds more generally for any generalized linear mixed model, and there is no straightforward relation between the parameter vector $\boldsymbol{\beta}^{\mathrm{RE}}$ in the random-effects model and the parameter vector $\boldsymbol{\beta}^{\mathrm{M}}$ in the marginal model, except in a few special cases. For example, consider the linear mixed model introduced in Section 4.3, where the random-effects model $\boldsymbol{Y}_i|\boldsymbol{b}_i \sim N(X_i\boldsymbol{\beta} + Z_i\boldsymbol{b}_i, \Sigma_i)$ implies that, marginally, $\boldsymbol{Y}_i$ has mean $E(Y_{ij}) = E[E(Y_{ij}|b_i)] = X_i\boldsymbol{\beta}$, showing that, in this case $\boldsymbol{\beta}^{\mathrm{RE}} = \boldsymbol{\beta}^{\mathrm{M}}$. Another example is the above discussed logistic model with random intercepts,

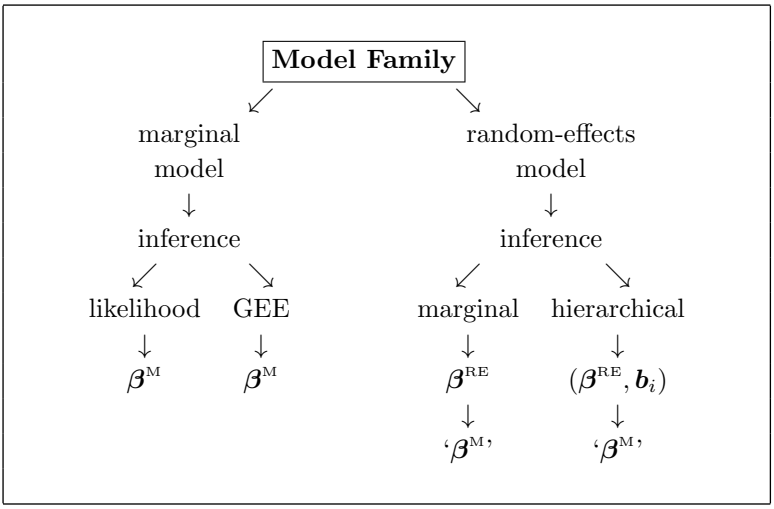

FIGURE 16.2. *Representation of model families and corresponding inferences. A superscript 'M' stands for marginal, 'RE' for random effects. A parameter between quotes indicates that marginal functions but no direct marginal parameters are obtained.*

for which it can be derived that

$$\left|\frac{\boldsymbol{\beta}^{\mathrm{RE}}}{\boldsymbol{\beta}^{\mathrm{M}}}\right| \approx \sqrt{c^2\tau^2 \ + \ 1} \ > \ 1 \tag{16.3}$$

where τ^2 is the variance of the random intercepts and with $c = 16\sqrt{3}/(15\pi)$ (Diggle *et al* 2002, Section 7.4). Note that (16.3) implies our heuristically obtained result that $\boldsymbol{\beta}^{\mathrm{RE}}$ is not smaller than $\boldsymbol{\beta}^{\mathrm{M}}$, with equality when the random-intercepts variance τ^2 is zero.

The fact that parameters from marginal and random-effects models need to be interpreted completely differently shows that the choice between these model families has important consequences and should be reflected upon very carefully. A schematic display of the possible choices is given in Figure 16.2. Whenever a marginal model is fitted, one directly obtains estimates and inferences for the components in $\boldsymbol{\beta}^{\mathrm{M}}$, the regression vector that models the average trend in the population. Within this class of approaches, fitting and inference can be based on full maximum likelihood principles, or on methods that only require correct specification of a number of moments (GEE and related methods). In case a random-effects model is fitted, one should realize that, even when estimation and inference is based on likelihood principles for the marginal likelihood (14.2) where the random effects have been integrated out, the parameters keep their original random-effects interpretation, such that estimates as well as inferences are obtained for the components in $\boldsymbol{\beta}^{\mathrm{RE}}$ rather than $\boldsymbol{\beta}^{\mathrm{M}}$. Note that, under the random-effects

model family, one can also obtain inferences for the random effects, under the assumption that the hierarchical model formulation was correct, i.e., under the assumption that the correlation between repeated measurements was indeed implied by an underlying random-effects structure. Alternatively, one may consider the generalized linear mixed model as just one approach to construct a marginal likelihood, without having any interests in possible presence of underlying latent variables $\boldsymbol{b}_i$.

Note that, because the random-effects approach results in a marginal likelihood, hereby completely specifying the distribution of $\boldsymbol{Y}_i$, it is possible to derive the marginal average trends in the data. As is indicated in (16.2), this requires averaging the conditional means in (16.1), over the random effects $\boldsymbol{b}_i$. Again, numerical integration methods can be used, but it is often much easier to use numerical averaging by sampling a large number M of random-effects vectors $\boldsymbol{b}_i$ from their fitted distribution $N(\boldsymbol{0}, \widehat{D})$, and to estimate $E(Y_{ij})$ at a specific point t in time by

$$\widehat{E}(Y_{ij}) \quad = \quad \frac{1}{M} \sum_{i}^{M} \frac{\exp(\widehat{\beta}_0^{\mathrm{RE}} + b_i + \widehat{\beta}_1^{\mathrm{RE}} t)}{1 + \exp(\widehat{\beta}_0^{\mathrm{RE}} + b_i + \widehat{\beta}_1^{\mathrm{RE}} t)}.$$

This can be calculated for a fine grid of time-points t, such that a graphical representation for the average trend can be obtained. An example, including SAS code for averaging over the fitted random-effects distribution can be found in Section 19.4. It should be emphasized that, in general, the average trend $E(Y_{ij})$ is not of the same parametric form as the conditional means $E(Y_{ij}|\boldsymbol{b}_i)$. Hence, the averaging over the random effects will not yield formal estimates for the elements in $\boldsymbol{\beta}^{\mathrm{M}}$. They can only provide a plot of the population-averaged trends. This explains why, in Figure 16.2, the marginal trends obtained from the random-effects approach are indicated as '$\boldsymbol{\beta}^{\mathrm{M}}$.'

16.4 Toenail Data: Marginal *versus* Mixed Models

We reconsider the toenail data, with the results from a GEE analysis and a random-effects analysis summarized in Table 16.1. The generalized linear mixed model is logistic with random intercepts only, hence, the approximate relation (16.3) holds and yields as approximate ratio

$$\sqrt{[16\sqrt{3}/(15\pi)]^2 (4.02)^2 + 1} = 2.56$$

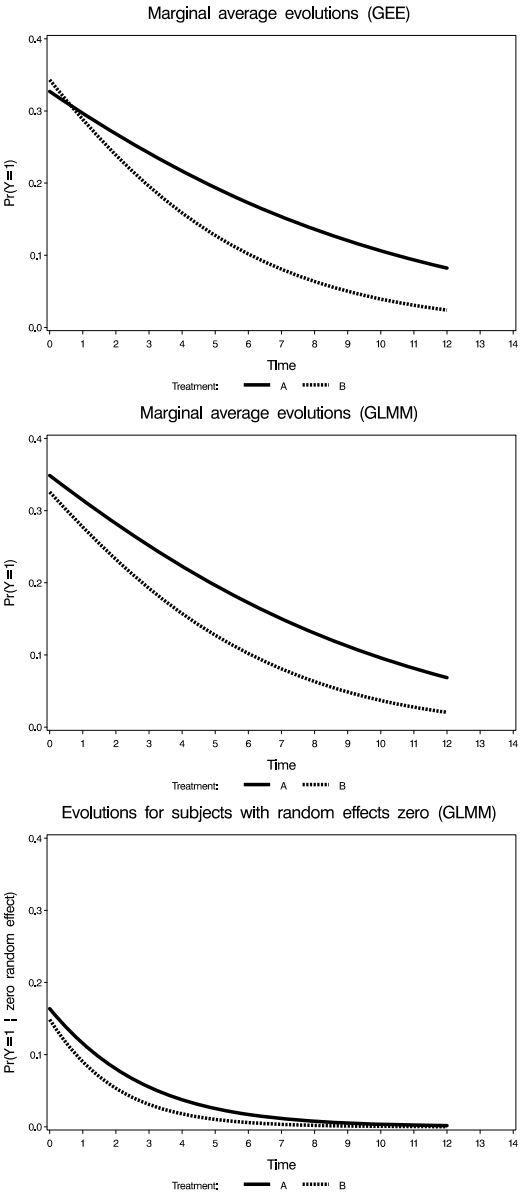

FIGURE 16.3. *Toenail Data. Treatment-arm specific evolutions. (a) Marginal evolutions as obtained from a marginal (GEE) model, (b) marginal evolutions as obtained from integrating out a GLMM, and (c) evolutions for an "average" subject from a GLMM, i.e., a subject with* $b_i = 0$.

which is in line with the observed ratio reported in Table 16.1. The fitted average evolutions, directly obtained from the GEE analysis, are given by

$$P(Y_{ij}=1) \quad = \quad \begin{cases} \dfrac{\exp(-0.72-0.14t)}{1+\exp(-0.72-0.14t)}, & \text{Treatment A} \\ \\ \dfrac{\exp(-0.65-0.25t)}{1+\exp(-0.65-0.25t)}, & \text{Treatment B,} \end{cases}$$

and are shown in the top graph in Figure 16.3. The middle panel of Figure 16.3 shows the marginal trends implied by the mixed model, i.e.,

$$P(Y_{ij}=1) \quad = \quad \begin{cases} E\left[\dfrac{\exp(-1.63+b_i-0.40t)}{1+\exp(-1.63+b_i-0.40t)}\right], & \text{Treatment A} \\ \\ E\left[\dfrac{\exp(-1.75+b_i-0.57t)}{1+\exp(-1.75+b_i-0.57t)}\right], & \text{Treatment B,} \end{cases}$$

where the expectation is taken over the fitted random-effects distribution $N(0, 4.02^2)$. Note that very similar trends are obtained, except maybe early in the study (first 2 months). This may be due to sampling variability, or due to the fact that not all subjects have been followed until the end of the experiment. Indeed, as discussed in Section 2.3, 72 (24%) out of the 298 participants left the study prematurely, due to a variety of, often unknown, reasons. As will be discussed in Part VI, GEE and random-effects analyses make different assumptions about the relation between missingness and the longitudinal response of interest. This may result in (slightly) different fitted average trends. Finally, the bottom plot in Figure 16.3 shows the expected trends for 'average' patients, i.e., for patients with random intercept $b_i = 0$. This again illustrates that, unlike for linear mixed models, the population-averaged trends cannot be obtained by setting random effects in a generalized linear mixed model, equal to zero.

As a summary and conclusion, we now compare the results from various models and estimation techniques applied to the toenail data. Table 16.2 summarizes the results from the marginal model and the random-effects model, considered earlier in Section 16.2: Both models include linear time-effects, with treatment-specific intercepts and slopes. The marginal model parameter estimates are obtained using generalized estimating equations (GEE1), where a marginal logit function is combined with unstructured working assumptions about the association structure. The random-effects model is of the logistic-normal type, with no other random effects than intercepts with variance τ^2. The mixed model has been fitted using MQL and PQL (both with REML for fitting the linear mixed models to the pseudo-data), as well as with adaptive Gaussian quadrature with 50 quadrature points. A selection of the results was shown before in Table 16.1. We now clearly observe that the estimates obtained from PQL and MQL are situated somewhat in between the estimates obtained from QUAD and GEE,

TABLE 16.2. *Toenail Data. Parameter estimates (standard errors) for a generalized linear mixed model and a marginal model (GEE). The mixed model has been fitted using MQL and PQL (both with REML for fitting the linear mixed models to the pseudo-data), as well as with adaptive Gaussian quadrature with 50 quadrature points (QUAD).*

Parameter	QUAD	PQL
Intercept group A	−1.63 (0.44)	−0.72 (0.24)
Intercept group B	−1.75 (0.45)	−0.72 (0.24)
Slope group A	−0.40 (0.05)	−0.29 (0.03)
Slope group B	−0.57 (0.06)	−0.40 (0.04)
Var. random intercepts (τ^2)	15.99 (3.02)	4.71 (0.60)
Parameter	**MQL**	**GEE**
Intercept group A	−0.56 (0.17)	−0.72 (0.17)
Intercept group B	−0.53 (0.17)	−0.65 (0.17)
Slope group A	−0.17 (0.02)	−0.14 (0.03)
Slope group B	−0.26 (0.03)	−0.25 (0.04)
Var. random intercepts (τ^2)	2.49 (0.29)	

where MQL is closest to GEE. As has been discussed in Section 14.4, MQL is based on a Taylor series expansion of the mean μ_{ij} around current estimates of the fixed effects and around random effects equal to zero. Therefore it produces estimates relatively close to those from marginal models, which do not contain any random effects at all (i.e., which have all $\boldsymbol{b}_i \equiv \mathbf{0}$). PQL, on the other hand, explicitly accounts for the random effects in its Taylor series expansion and therefore yields estimates closer to those obtained under Gaussian quadrature.

16.5 Analysis of the NTP Data

As discussed in Chapter 13, the generalized linear mixed model (GLMM) is not the only model in the class of random-effects models. An alternative model is the beta-binomial model, introduced in Section 13.4.2. We will now fit a beta-binomial model and compare it to the results obtained from previous analyses. It will be assumed that the success probability π_i and the within-cluster correlation ρ_i satisfy

$$\ln\left(\frac{\pi_i}{1-\pi_i}\right) = \beta_0 + \beta_d d_i \tag{16.4}$$

$$\ln\left(\frac{1+\rho_i}{1-\rho_i}\right) = \beta_a, \tag{16.5}$$

TABLE 16.3. *NTP Data. Parameter estimates (standard errors) for the beta-binomial model, fitted to various outcomes in three studies.* β_0 *and* β_d *are the marginal intercept and dose effect, respectively;* β_a *is the Fisher z transformed correlation;* ρ *is the correlation.*

Outcome	Parameter	DEHP	EG	DYME
External	β_0	-4.91(0.42)	-5.32(0.71)	-7.27(0.74)
	β_d	5.20(0.59)	2.78(0.81)	8.01(0.82)
	β_a	0.21(0.09)	0.28(0.14)	0.21(0.12)
	ρ	0.10(0.04)	0.14(0.07)	0.10(0.06)
Visceral	β_0	-4.38(0.36)	-7.45(1.17)	-6.21(0.83)
	β_d	4.42(0.54)	4.33(1.26)	4.94(0.90)
	β_a	0.22(0.09)	0.04(0.09)	0.45(0.21)
	ρ	0.11(0.04)	0.02(0.04)	0.22(0.10)
Skeletal	β_0	-4.88(0.44)	-2.89(0.27)	-5.15(0.47)
	β_d	4.92(0.63)	3.42(0.40)	6.99(0.71)
	β_a	0.27(0.11)	0.54(0.09)	0.61(0.14)
	ρ	0.13(0.05)	0.26(0.04)	0.30(0.06)
Collapsed	β_0	-3.83(0.31)	-2.51(0.09)	-5.42(0.45)
	β_d	5.59(0.56)	3.05(0.17)	8.29(0.79)
	β_a	0.32(0.10)	0.28(0.02)	0.33(0.10)
	ρ	0.16(0.05)	0.14(0.01)	0.16(0.05)

where d_i is the dose administered to the ith cluster. Note that this is the same parameterization as was used before for the Bahadur model in Section 7.2.3. Table 16.3 shows the results for the three NTP studies, and for the four different outcomes.

For the sake of comparison, we will focus on the outcome 'External malformations' in the DEHP study. Table 16.4 summarizes the results from analyses based on marginal models (Chapters 7 and 8) conditional models (Chapters 11 and 12), and random-effects models (Chapters 14 and 16). As has been indicated in Section 11.4, estimates for the conditional models are typically considerably smaller than their marginal counterparts, due to the fundamental difference in interpretation. Indeed, conditional-model parameters describe the conditional logit and log odds ratios of outcomes, given other outcomes, whereas in marginal models no such conditioning takes place. A similar argument explains the differences between marginal and random-effects models (Section 16.3).

The results from conditional models are very similar, whatever estimation method is used (maximum likelihood or pseudo-likelihood). The various marginal modeling approaches (Bahadur, various forms of GEE, ALR) provide very similar inferences as well, even though some subtle differ-

TABLE 16.4. *NTP Data. External malformations in the DEHP study. Parameter estimates (standard errors) from analyses based on marginal models (Chapters 7 and 8), conditional models (Chapters 11 and 12), and random-effects models (Chapters 14 and 16).* β_0 *and* β_d *are the intercept and dose effect, respectively; the association parameter varies between models.*

Model	β_0	β_d	Association	
Conditional models				
Quadr. loglin. (ML)	-2.81(0.58)	3.07(0.65)	LOG OR	0.18(0.04)
Quadr. loglin. (PL)	-2.85(0.53)	3.24(0.60)	LOG OR	0.18(0.04)
Marginal models				
Lik. Bahadur	-4.93(0.39)	5.15(0.56)	β_a	0.11(0.03)
St. GEE1 (exch)	-4.98(0.37)	5.33(0.55)	ρ	0.11
St. GEE1 (ind)	-5.06(0.38)	5.31(0.57)		
Prent. GEE1 (exch)	-4.99(0.37)	5.32(0.55)	ρ	0.11 (0.04)
Prent. GEE1 (ind)	-5.06(0.38)	5.31(0.57)		
Lin. based (exch)	-5.00(0.37)	5.32(0.55)	ρ	0.06
Lin. based (ind)	-5.06(0.38)	5.31(0.57)		
GEE2	-4.98(0.37)	5.29(0.55)	β_a	0.15(0.05)
ALR	-.516(0.35)	5.64(0.52)	β_a	0.96(0.30)
Random-effects models				
Beta-binomial	-4.91(0.42)	5.20(0.59)	β_a	0.21(0.09)
GLLM (MQL)	-5.18(0.40)	5.70(0.66)	Int. var τ^2	1.20(0.53)
GLMM (PQL)	-5.32(0.40)	5.73(0.65)	Int. var τ^2	0.95(0.40)
GLMM (QUAD)	-5.97(0.57)	6.45(0.84)	Int. var τ^2	1.27(0.62)

ences exist, as was explained throughout the various analyses conducted in Chapter 8. More severe discrepancies are observed when the various random-effects analyses are compared. The differences between MQL, PQL and Gaussian quadrature have been observed and explained before in Section 16.4. However, note how the results from the beta-binomial model are closer to those from the marginal models than to those from the GLMM model under Gaussian quadrature. This can be explained as follows. It follows from (13.4) and (13.5) that the parameters π_i and ρ_i modeled in (16.4) and (16.5) have marginal interpretations. Hence, although the beta-binomial model has a random-effects genesis, the regression coefficients need to be interpreted marginally.

Part V

Case Studies and Extensions