16

Contrasting Marginal and Mixed Effects Models

16.1 INTRODUCTION

In this chapter we compare and contrast marginal and mixed effects models for longitudinal data. There are a number of important distinctions between these two broad classes of models that go beyond simple differences in approaches to accounting for the within-subject association. We emphasize that these two classes of models have different targets of inference and therefore address subtly different questions regarding longitudinal change. In this chapter we highlight the main distinctions and discuss the types of scientific questions addressed by each of the two classes of models.

16.2 LINEAR MODELS: A SPECIAL CASE

In Part II we focused on linear models for longitudinal data where the model for the mean response vector can be expressed as

$$E(Y_i) = X_i\beta. \tag{16.1}$$

To account for the positive correlation among the repeated measurements, we described two broad approaches. The first approach is to adopt a covariance pattern model (e.g., autoregressive, Toeplitz) for $\Sigma_i = \text{Cov}(Y_i)$. The second approach is to introduce random effects in the model for the mean response,

$$E(Y_i|b_i) = X_i\beta + Z_ib_i, \tag{16.2}$$

Applied Longitudinal Analysis, Second Edition. By Garrett M. Fitzmaurice, Nan M. Laird, and James H. Ware.

where b_i is a vector of random effects that vary from individual to individual according to a probability distribution (commonly assumed to be multivariate normal). The introduction of random effects induces a random effects covariance structure for Σ_i,

$$\Sigma_i = Z_i G Z_i' + \sigma^2 I_{n_i},$$

where $G = \text{Cov}(b_i)$ and σ^2 is the variance of the measurement or sampling errors.

When discussing these two different approaches for accounting for the within-subject association, issues concerning the interpretation of β in (16.1) and (16.2) simply did not arise because β has the same interpretation in both models. That is, β describes how the mean response in the study population changes with time and how these changes are related to the covariates. This interpretation of β is transparent when (16.1) is considered. However, the linear mixed effects model given by (16.2) implies the exact same model for the marginal mean response when averaged over the distribution of the random effects. That is,

$$\begin{aligned} E(Y_i) &= E\{E(Y_i|b_i)\} \\ &= E(X_i\beta + Z_i b_i) \\ &= X_i\beta + Z_i E(b_i) \\ &= X_i\beta, \end{aligned}$$

since the vector of random effects has mean zero (i.e., $E(b_i) = 0$). Thus, in the linear mixed effects model, β can be interpreted *either* as regression coefficients in the model for the conditional mean of Y_i given X_i and b_i (i.e., conditonal on b_i) *or* as regression coefficients in the model for the population average mean of Y_i given X_i. In the process of taking the expectation or average over the distribution of the random effects, we have implicitly used the property that expectation is a linear operation. This means that the expectation of any linear function of b_i can be easily evaluated. That is,

$$E(X_i\beta + Z_i b_i) = X_i\beta + Z_i E(b_i),$$

for any constants $X_i\beta$ and Z_i. Thus β has the same interpretation in (16.1) and (16.2) because (16.2) is a *linear* mixed effects model (i.e., the right-hand side of (16.2) is a linear function of b_i). Put more simply, in the linear mixed effects model β has a marginal interpretation because the average of the linear rates of change over time for individuals is the same as the linear rate of change over time in the population mean response. However, as we will see in the next section, when evaluating expectations of any *non-linear* functions of b_i we can no longer proceed in this manner. That is, for any non-linear function of b_i, say $h(X_i\beta + Z_i b_i)$,

$$E\{h(X_i\beta + Z_i b_i)\} \neq h\{X_i\beta + Z_i E(b_i)\}.$$

16.3 GENERALIZED LINEAR MODELS

Next we consider the comparison of marginal and mixed effects generalized linear models for longitudinal data. Recall that one of the components in the specification of a generalized linear model is the link function, $g(\mu_i)$, which relates the mean of Y_i to

the linear predictor. In the previous discussion of linear models, the link function was the identity function, $g(\mu_i) = \mu_i$. For the special case of an identity link function, and hence linear models, the regression parameters β have the same interpretation in both marginal and mixed effects models. In this section we focus on non-linear (or non-identity) link functions and compare the regression parameters in marginal and generalized linear mixed effects models.

Recall that a marginal model for the mean response vector is given by

$$g(\mu_i) = g\{E(Y_i)\} = X_i\beta, \tag{16.3}$$

where $g(\cdot)$ is an appropriate vector-valued non-linear link function (e.g., logit or log). The regression parameters β in a marginal model have interpretation in terms of changes in the transformed mean response in the study population, and their relation to covariates. For example, when the components of Y_i are binary and a logit link function is adopted, with

$$\text{logit}(\mu_i) = X_i\beta,$$

the regression parameters have interpretation in terms of changes in the log odds of success in the study population. For any known link function $g(\cdot)$, the population means can be expressed in terms of the inverse link function, say $h(\cdot) = g^{-1}(\cdot)$,

$$h\{g(\mu_i)\} = \mu_i = E(Y_i) = h(X_i\beta). \tag{16.4}$$

For example, when the components of Y_i are binary and a logit link function has been adopted, the model for μ_i is

$$\mu_i = h(X_i\beta) = \frac{\exp(X_i\beta)}{1 + \exp(X_i\beta)},$$

where $h(\cdot)$ is the inverse-logit link function. Whether expressed as (16.3) or (16.4), the regression parameters β in a marginal model describe changes in the transformed population mean response vector, μ_i. Note also that μ_i depends on the index i only via fixed and known covariate values.

Next consider generalized linear mixed models where the conditional mean of Y_i, given a vector of random effects b_i, is

$$g\{E(Y_i|b_i)\} = X_i\beta^* + Z_ib_i, \tag{16.5}$$

where the random effects b_i have a distribution with mean zero and covariance matrix G. Here we denote the fixed effects by β^* to clearly distinguish them from the corresponding regression parameters in the marginal model in (16.3). The regression coefficients β^* have subject-specific interpretations in terms of changes in the transformed mean response for any individual. That is, to interpret any component of β^* we must consider a unit change in the corresponding covariate while holding b_i fixed. However, the most natural way to hold b_i fixed at a particular value is to focus on the conditional mean response vector of any given individual. Alternatively, we can compare two individuals who have the same values for b_i but who differ by a single

unit in the corresponding covariate. The former interpretation of any component of β^* is most natural when the covariate is time-varying; the latter interpretation is more natural when the covariate is time-invariant.

Thus, unlike β in marginal models, β^* has interpretation in terms of changes in the transformed mean response for any individual (or the notional comparison of individuals with the same values for b_i). The regression coefficients β^* do not describe changes in the transformed mean response in the study population. The implied model for the marginal means can only be obtained by averaging over the distribution of the random effects. This involves taking an expectation of a non-linear function of b_i,

$$\begin{aligned} \mu_i &= E(Y_i) \\ &= E\{E(Y_i|b_i)\} \\ &= E\{h(X_i\beta^* + Z_ib_i)\}. \end{aligned} \tag{16.6}$$

Note that μ_i depends on the index i only via fixed and known covariate values.

The expression given in (16.6) is the expectation of a non-linear function of b_i. It must be evaluated from the definition of expectation as a weighted average, weighted according to the distribution of the random effects,

$$\mu_i = E(Y_i) = E\{h(X_i\beta^* + Z_ib_i)\} = \int_{-\infty}^{\infty} h(X_i\beta^* + Z_ib_i)f(b_i)db_i, \tag{16.7}$$

where the integration denotes summation or averaging and $f(b_i)$ is the probability density function for b_i (or the "weights" used in the process of averaging). However, the expression for $E(Y_i)$ given by (16.7) does not, in general, have a closed-form, and moreover, as noted in the previous section,

$$E(Y_i) \neq h(X_i\beta),$$

for any β. For example, consider the logistic regression model with a randomly varying intercept,

$$\text{logit}\{E(Y_i|b_i)\} = X_i\beta^* + b_i,$$

where $b_i \sim N(0, \sigma_b^2)$. The implied model for the marginal mean or marginal probability of success is

$$\begin{aligned} \mu_i &= E(Y_i) \\ &= E\{E(Y_i|b_i)\} \\ &= E\left\{\frac{e^{(X_i\beta^*+b_i)}}{1+e^{(X_i\beta^*+b_i)}}\right\} \\ &= \int_{-\infty}^{\infty} \frac{e^{(X_i\beta^*+b_i)}}{1+e^{(X_i\beta^*+b_i)}} \frac{1}{\sqrt{2\pi\sigma_b^2}} e^{-\frac{1}{2}b_i^2/\sigma_b^2}\, db_i. \end{aligned}$$

This expression cannot be evaluated in closed-form, and moreover, is not of the logistic regression form,

$$\frac{e^{(X_i\beta)}}{1+e^{(X_i\beta)}},$$

for any choice of β. That is, marginally (when averaged over the distribution of the random effects) the logit link function is no longer preserved.

For the special case of the logistic regression model with only a single randomly varying intercept (or subject effect),

$$\text{logit}\{E(Y_i|b_i)\} = X_i\beta^* + b_i,$$

where $b_i \sim N(0, \sigma_b^2)$, the following approximate relationship holds:

$$\text{logit}\{E(Y_i)\} \approx (1 + k^2\sigma_b^2)^{-\frac{1}{2}} X_i\beta^*,$$

where $k = \frac{16\sqrt{3}}{15\pi}$. The derivation of this approximation is not important. What this approximate relationship highlights is how the logistic regression coefficients in the marginal model are attenuated relative to the corresponding fixed effects in the logistic regression model with a randomly varying intercept,

$$\beta \approx \frac{\beta^*}{\sqrt{1 + 0.346\sigma_b^2}},$$

where $k^2 = 0.346$. Thus, when $\text{Var}(b_i) = \sigma_b^2 > 0$, the marginal logistic regression model parameters, β, are smaller in absolute value than the fixed effects, β^*, in the mixed effects model. In addition the discrepancy between β and β^* increases with increasing σ_b^2. For example, if $\sigma_b^2 = 3.5$, then $\beta^* \approx 1.5 \times \beta$; if $\sigma_b^2 = 9$, then $\beta^* \approx 2\,\beta$. For the more general logistic regression model with a vector of random effects b_i, a similar approximate relationship holds and indicates that the marginal logistic regression parameters β are always attenuated toward zero when compared to β^*.

Thus, for non-linear link functions, the fixed effects β^* in generalized linear mixed models are not comparable to the regression parameters β in marginal models. The lack of comparability reflects the distinct targets of inference associated with generalized linear mixed models and marginal models. That is, the fixed effects β^* describe the effects of covariates on changes in an individual's response over time while the regression parameters β describe the effects of covariates on changes in the population mean response over time.

In addition to the special case of an identity function (i.e., linear mixed effects models), there happens to be one exceptional case where β^* and β are almost comparable. When a log link function is adopted, the link function is preserved marginally. In addition the subset of fixed effects β^* for components of X_{ij} that do not overlap with Z_{ij} are directly comparable to the corresponding set of marginal regression parameters β. Specifically, for the log-linear regression model with vector of random effects b_i,

$$\log\{E(Y_{ij}|b_i)\} = X'_{ij}\beta^* + Z'_{ij}b_i,$$

where $b_i \sim N(0, G)$, the implied model for the marginal mean is

$$\begin{aligned}
\mu_{ij} &= E\{E(Y_{ij}|b_i)\} \\
&= E\left\{e^{(X'_{ij}\beta^* + Z'_{ij}b_i)}\right\} \\
&= e^{(X'_{ij}\beta^* + \frac{1}{2}Z'_{ij}GZ_{ij})}.
\end{aligned}$$

Therefore

$$\log\{E(Y_{ij})\} = X'_{ij}\beta^* + \frac{1}{2}Z'_{ij}GZ_{ij},$$

and β^* differs from β only for those covariates that overlap between X_{ij} and Z_{ij}; for all other covariates, the corresponding components of β^* and β are the same. For example, if the model includes only a single randomly varying intercept (or subject effect),

$$\log\{E(Y_{ij}|b_i)\} = X'_{ij}\beta^* + b_i,$$

then

$$\log\{E(Y_{ij})\} = \log(\mu_{ij}) = X'_{ij}\beta^* + g_{11}/2,$$

where g_{11} denotes the variance of the randomly varying intercept b_i. Thus, for the special case of a log link function and a single randomly varying intercept, the fixed effects β^* are directly comparable to the marginal model regression parameters β (with the exception of the intercept).

Finally, although it is possible, in principle, to obtain estimates of the marginal means from a generalized linear mixed effects model, the assumed form for the regression model for the conditional means given b_i (e.g., logistic or log-linear) no longer holds for the resulting marginal means when averaged over the distribution of the random effects; moreover any misspecification of the model can yield biased estimates of the implied population average means. As a result a set of regression parameters for μ_i, describing the dependence of the population mean response on the covariates, is not immediately available from a generalized linear mixed effects model, even after averaging over the distribution of the random effects.[1] The practical consequence is that it is not possible to describe parsimoniously the effects of covariates on the population means in terms of regression coefficients. This may not be so problematic in the setting of a randomized longitudinal clinical trial where the parameter of interest is often a simple difference or contrast of treatment means (or changes in treatment means from baseline). In the latter setting there is only a single covariate of interest (e.g., treatment group) and it is discrete; furthermore a suitable contrast of the *marginal* mean response profiles in the treatment groups can be estimated. However, when one or more of the covariates of interest is quantitative and/or when there are potential confounding variables that need to be controlled for in the analysis, no simple summaries of the effects of covariates on μ_i are readily available from generalized linear mixed effects models.

[1]Note that marginal regression parameters can be estimated directly from models with random effects only when a non-standard relationship is assumed for the conditional mean given the random effects. Alternatively, certain non-standard random effects distributions preserve the same link function relating the marginal and conditional means to the covariates. This class of models are known as *marginalized* mixed effects models.

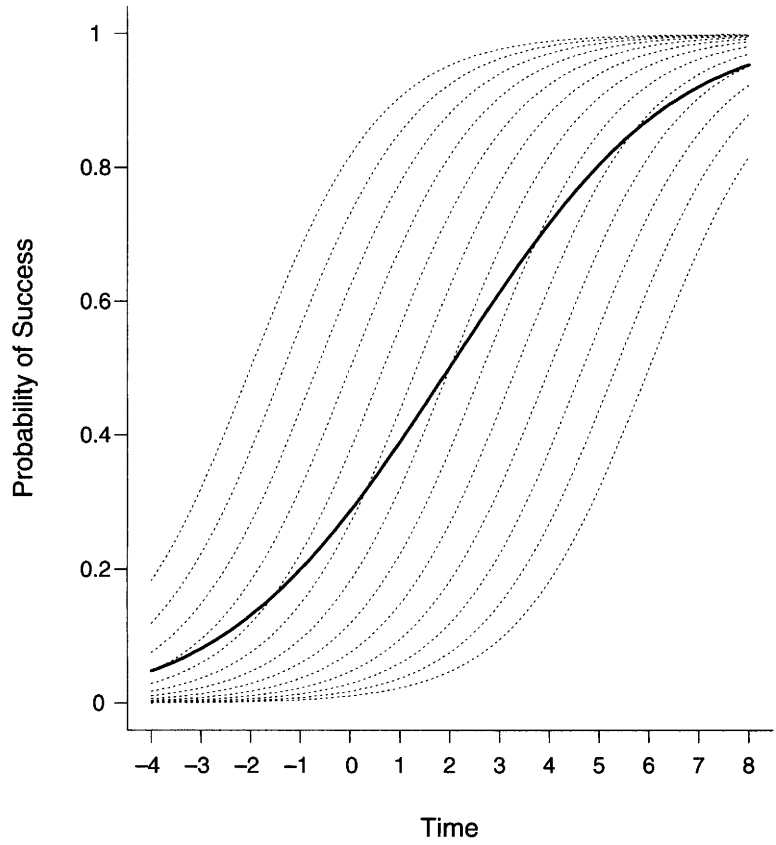

Fig. 16.1 Comparison of conditional probabilities of success (dotted lines) and marginal probability of success (solid line), averaged over the distribution of the random effects.

16.4 SIMPLE NUMERICAL ILLUSTRATION

To reinforce the distinctions made in the previous section, consider the following simple numerical illustration. Suppose that Y_i is a vector of binary responses and it is of interest to describe changes in the log odds of success over time. For simplicity we assume that there are no covariates other than the times of measurement. A logistic regression model, with randomly varying intercepts, is given by

$$\text{logit}\{E(Y_{ij}|b_i)\} = \beta_1^* + \beta_2^* t_{ij} + b_i,$$

where b_i is assumed to have a normal distribution with zero mean and variance $g_{11} = \text{Var}(b_i)$. Figure 16.1 displays a plot of $E(Y_{ij}|b_i)$ versus t_{ij} for a random

Table 16.1 Data on whether an electrocardiogram (ECG) was normal (0) or abnormal (1) from a two-period crossover trial comparing the effects of active drug to placebo.

	Response (Period 1, Period 2)			
Sequence	(1, 1)	(1, 0)	(0, 1)	(0, 0)
Sequence 1 (P → A)	6	0	6	22
Sequence 2 (A → P)	9	4	2	18

Source: Reprinted with permission from Table 3.1 of Jones and Kenward (1989).
Note: P: Placebo; A: Active drug.

sample of b_i from a normal distribution with zero mean and variance $g_{11} = 4$ (with $\beta_1^* = -1.5$ and $\beta_2^* = 0.75$; $t_{ij} \in [-4, 8]$). Also displayed in Figure 16.1 is a plot of the marginal probability of success, averaged over the distribution of b_i. When the subject-specific logistic curves are compared to the population average curve, it is apparent that the slopes of the former (determined by β_2^*) are steeper than the slope of the latter. Notice that in the range of probabilities from 0.3 to 0.7, where the logistic curves are approximately linear, the slopes for the subject-specific curves rise faster than the slope for the marginal success probabilities. This reinforces the notion that β^* does not characterize aspects of the population log odds of response, but instead describes changes in the log odds of success for an individual from the population.

16.5 CASE STUDY

This section highlights aspects of interpretation of the regression coefficients in marginal and generalized linear mixed effects models using safety data from a crossover trial on the disease cerebrovascular deficiency. The variable we analyze is not a trial endpoint per se but rather a potential side effect. In this two-period crossover trial, comparing the effects of active drug to placebo, 67 patients were randomly allocated to the two treatment sequences, with 34 patients receiving placebo → active, and 33 patients receiving active → placebo. The response variable is binary, indicating whether an electrocardiogram (ECG) was abnormal ($Y = 1$) or normal ($Y = 0$). Thus each patient has a bivariate binary response vector, $Y_i = (Y_{i1}, Y_{i2})'$, where Y_{ij} denotes the response for the i^{th} subject in the j^{th} period (for $i = 1, ..., 67; j = 1, 2$). In Table 16.1 the data are summarized in the form of a 2×4 contingency table.

First, we analyze these data using a marginal model. The marginal mean of the response (or probability of an abnormal ECG) is modeled as a logistic function of the

Table 16.2 Parameter estimates and standard errors from marginal logistic regression model for the ECG data.

Variable	Estimate	SE	Z
Intercept	−1.2433	0.2999	−4.15
Treatment	0.5689	0.2335	2.44
Period	0.2951	0.2319	1.27
log OR(α)	3.5617	0.8148	4.37

covariates, Treatment_{ij} (0 = Placebo, 1 = Active drug) and Period_{ij} (0 = Period 1, 1 = Period 2),

$$\text{logit}(\mu_{ij}) = \beta_1 + \beta_2 \,\text{Treatment}_{ij} + \beta_3 \,\text{Period}_{ij}.$$

The within subject association between the two responses is modeled in terms of a common log odds ratio, α,

$$\log \text{OR}(Y_{i1}, Y_{i2}) = \log \left\{ \frac{\Pr(Y_{i1} = 1, Y_{i2} = 1) \Pr(Y_{i1} = 0, Y_{i2} = 0)}{\Pr(Y_{i1} = 1, Y_{i2} = 0) \Pr(Y_{i1} = 0, Y_{i2} = 1)} \right\} = \alpha.$$

The results, obtained using the GEE approach, are presented in Table 16.2. These results indicate that treatment with the active drug is harmful, increasing the rates of abnormal electrocardiograms. The odds of an abnormal electrocardiograms is 1.77 (or $e^{0.57}$) times higher when treated with active drug versus placebo. The estimate of the within-subject association is $\widehat{\alpha} = 3.56$, indicating very strong positive association. That is, the odds of an abnormal electrocardiogram at the second occasion are approximately 35 times higher if the electrocardiogram at the first occasion is abnormal rather than normal.

Next we analyze these data using a generalized linear mixed model. In particular, we consider the following logistic regression model for the conditional mean of the response, given a random patient effect:

$$\text{logit}\{E(Y_{ij}|b_i)\} = \beta_1^* + \beta_2^* \,\text{Treatment}_{ij} + \beta_3^* \,\text{Period}_{ij} + b_i,$$

where the random effect b_i is assumed to have a normal distribution with zero mean and variance, $\text{Var}(b_i) = g_{11}$. In this model each patient is assumed to have some underlying propensity for an abnormal electrocardiogram given by b_i. Then a patient's odds of an abnormal electrocardiogram is multiplied by a common factor $e^{\beta_2^*}$ if the patient is treated with the active drug, regardless of that patient's underlying propensity. Thus $e^{\beta_2^*}$ has interpretation in terms of the ratio of a patient's odds of an abnormal electrocardiogram, when treated with the active drug versus placebo.

Table 16.3 Parameter estimates and standard errors from mixed effects logistic regression model for the ECG data.

Variable	Estimate	SE	Z
Intercept	−4.0816	1.6710	−2.44
Treatment	1.8631	0.9269	2.01
Period	1.0376	0.8189	1.27
$g_{11} = \text{Var}(b_i)$	24.4345	18.8474	

Note: ML estimation is based on 100-point adaptive Gaussian quadrature.

The ML estimates of the fixed effects and variance component are presented in Table 16.3. These results also indicate that treatment with the active drug is harmful, increasing the patient-specific rates of abnormal electrocardiograms. In particular, a patient's odds of an abnormal electrocardiogram is 6.4 (or $e^{1.86}$) times higher when treated with active drug than when treated with the placebo. This common treatment effect is the same, regardless of the patient's underlying propensity for an abnormal electrocardiogram. The estimate of the variance of b_i, $\widehat{g}_{11} = 24.4$, indicates that there is very substantial between-patient variability in the propensity for abnormal electrocardiograms, and this is consistent with the very strong within-subject association found in the results from the marginal model reported in Table 16.2. Note, however, that the variance of b_i is poorly estimated (as evidenced by the very large standard error) since there are only two observations per patient. The estimates of the fixed effects presented in Table 16.3 are all substantially larger than the corresponding estimates in Table 16.2 (approximately three to four times larger), but so too are the standard errors (approximately four times larger). As a result Wald test statistics for null hypotheses concerning the fixed effects are often quite similar in magnitude for the two classes of models. In general, it can be shown that the discrepancy between β^* from the logistic regression model with random subject effect and β from a marginal model (with corresponding fixed effects) increases with the variance of the random subject effect. That is, the greater the underlying heterogeneity among subjects, the greater is the discrepancy between β^* and β, with $|\beta^*| > |\beta|$ for $\text{Var}(b_i) > 0$.

Comparison of the two estimates of treatment, $e^{\widehat{\beta_2}} = 1.8$ and $e^{\widehat{\beta_2^*}} = 6.4$, from the marginal and mixed effects logistic regression models highlights the distinction between these two analytic approaches. The estimated treatment effect from the marginal model describes how the average rates (expressed in terms of odds) of abnormal electrocardiograms would increase in the study population if patients were treated with the active drug. In contrast, the estimated treatment effect from the mixed effects model describes how the odds of an abnormal electrocardiograms increases for

any patient treated with the active drug. As a result the answer to the question "how harmful is the active drug" will depend on whether scientific interest is in its impact on the study population or on an individual drawn at random from that population.

We conclude this section by considering the sensitivity of the results reported in Tables 16.2 and 16.3 to sampling zeros and the distributional assumption for the random effects. Note that the first row of Table 16.1 contains a sampling zero. For patients receiving the first sequence of treatments (P → A), the response pattern (1, 0) happens not to have occurred; we refer to this as a "sampling zero" because the response pattern is assumed to be unobserved due to the limited size of the sample (34 patients receiving P → A). Sampling zeros, an extreme case of sparseness, are known to potentially cause problems for estimation of certain parameters. To assess whether the sampling zero in Table 16.1 has an inordinate influence on the estimate of the treatment effect, a small constant was added to that cell of the table. Specifically, we added $\frac{1}{4}$ to the cell with the sampling zero and repeated the two analyses reported in Tables 16.2 and 16.3. The estimated treatment effect from the marginal model was $\widehat{\beta}_2 = 0.55$ (SE = 0.236), almost identical to the result found in Table 16.2. The estimated treatment effect from the mixed effects model was $\widehat{\beta}_2^* = 1.64$ (SE = 0.821), somewhat smaller than the estimated treatment effect in Table 16.3; the estimated variance of b_i, $\widehat{g}_{11} = 20.7$, was also smaller. However, from a subject-matter point of view, the substantive conclusions do not change and the difference between the subject-specific and population-averaged effects of treatment is of the same order of magnitude as reported in Tables 16.2 and 16.3.

Next, we consider the validity of the distributional assumption for b_i. In particular, does the large estimated variance of b_i accurately reflect the true between-patient variability in the risk of an abnormal ECG? Because each patient has two responses, there are only four possible patterns of response: (0, 0), (0, 1), (1, 0) and (1, 1). The "sufficient statistic" for b_i is each patient's total number of abnormal ECGs, say $S_i = \sum_{j=1}^{2} Y_{ij}$; moreover S_i can take on only three possible values (0, 1, 2). Therefore for patients randomized to each of the two possible treatment sequences, b_i can take on only three distinct values ordered from smallest to largest according to S_i. Furthermore, note from Table 16.1 that the distribution of S_i is far from symmetric, with 22% of patients having no abnormal ECGs, 18% having 1 abnormal ECG, and 60% having 2 abnormal ECGs. This feature of the distribution of S_i can only be captured by a normal distribution for the b_i, centered at zero, that has a relatively large variance. To assess the sensitivity of the results to the normal assumption for b_i, we estimated the treatment effect using *conditional* maximum likelihood estimation (see Section 14.6). That is, we fitted the following *fixed effects* (or *conditional*) logistic regression model,

$$\text{logit}\{E(Y_{ij}|\alpha_i)\} = \beta_1^* + \beta_2^* \,\text{Treatment}_{ij} + \beta_3^* \,\text{Period}_{ij} + \alpha_i,$$

where the α_i denote *fixed effects* representing time-invariant characteristics of the patients that are not otherwise accounted for by the covariates in the model. The conditional ML estimate of the treatment effect, after we added $\frac{1}{4}$ to the cell with the sampling zero, was $\widehat{\beta}_2^* = 1.94$ (SE = 1.109), larger than the estimated treatment effect,

$\widehat{\beta}_2^* = 1.64$, from the mixed effects model reported earlier. However, this difference in the estimates of β_2^* is well within the sampling variability of the estimates. A statistical test of the difference between the two estimates for the treatment effect, using the analogue of the "Hausman test" developed for GLMMs by Tchetgen and Coull (2006) (see Section 14.6), yielded $Z = 0.402, (p > 0.65)$. Thus the results presented for these data do not appear to be very sensitive to the normal distribution assumption for b_i; however, this cannot be expected in general. In particular, when the number of repeated binary responses is small, and there is a large proportion of subjects with positive (or negative) responses at all occasions, the assumption of a symmetric, normal distribution for the b_i is questionable.

16.6 CONCLUSION

In Chapters 12 through 15 we considered two types of extensions of generalized linear models to longitudinal data: marginal models and generalized linear mixed models. These two quite different analytic approaches arise from different specifications of, or assumptions about, the joint distribution of Y_i and the source of the correlation among the repeated measures on the same individual. Marginal models merely acknowledge the correlation among repeated measures when estimating the regression parameters; in contrast, generalized linear mixed models provide an explanation for the source of the correlation. Unlike the linear models for continuous responses considered in Part II, with generalized linear models (and non-linear link functions) for discrete responses, different assumptions about the correlation can lead to regression coefficients with quite distinct interpretations.

The basic premise of marginal models is to make inferences about population means, and comparisons of sub-population means, albeit on a transformed scale (e.g., logit or log). The term "marginal" is used to emphasize that the mean response modeled is conditional only on the covariates and not on unobserved random effects or on previous responses. A distinctive feature of marginal models is that the regression models for the mean response and the models for the within-subject association are specified separately. This separation of the model for the mean response from the model for the within-subject association ensures that the marginal model regression coefficients have interpretation that does not depend on the assumptions made about the within-subject association. Specifically, the regression coefficients in marginal models describe the effects of covariates on the population mean response.

In contrast, the basic premise of generalized linear mixed effects models is that there is natural heterogeneity across individuals in the study population in a subset of the regression parameters. That is, a subset of the regression parameters (e.g. intercepts and slopes) is assumed to vary across individuals according to some underlying distribution. But, conditional on the random effects, it is assumed that the repeated measurements for any given individuals are independent observations. Generalized linear mixed models extend the conceptual approach of the linear mixed effects model in a very natural way. The correlation among repeated measurements arises from their sharing of common random effects. Unlike the linear mixed effects model, the re-

gression parameters in generalized linear mixed models have subject-specific, but not population-averaged, interpretations. That is, due to the non-linear link functions that are usually adopted for discrete responses, the fixed effects do not describe changes in the mean response in the study population. Instead, they describe how changes in an individual's mean response are related to within-individual changes in the covariates. As a result generalized linear mixed models are most useful when the scientific objective is to make inferences about individuals rather than the study population. For example, the regression parameters in a logistic mixed effects model describe how the log odds of response changes over time, and how these changes relate to within-individual changes in covariates. Unlike marginal models, they do not compare changes in the log odds of response across sub-populations of individuls defined by values of the covariates. In summary, with generalized linear mixed models, the main focus is on inferences about the individual, while with marginal models, the main focus is on inferences about the study population.

The choice between marginal and generalized linear mixed models for longitudinal data can only be made on subject-matter grounds. We have emphasized the different targets of inference for these two classes of models. For any given longitudinal study, different scientific questions will usually demand different analytic models. For example, a physician considering the potential benefits of a novel treatment for one of her patients might be more interested in the subject-specific effect of treatment. On the other hand, public health researchers or health insurance assessors considering the potential reduction in morbidity or mortality in the population if patients receive the novel treatment would be more interested in the population-averaged effect of treatment. When the answers to both of these questions are of interest, there is no contradiction in reporting estimates of both the subject-specific and population-averaged effects.

In summary, we do not prescribe (or proscribe for that matter) one class of models over another. While there has been much debate in the statistical literature concerning the appropriateness of these two classes of models for analyzing longitudinal data, much of the discussion has generated more heat than light. From a purely probabilistic point of view, generalized linear mixed models might appear to have a distinct advantage over marginal models since the marginal distribution of Y_i, the target of inference for marginal models, can, in principle, be derived from the generalized linear mixed effects model by averaging over the distribution of the random effects. However, this apparent advantage is somewhat illusory because the induced marginal model does not, in general, retain the same form. For example, consider a logistic regression model with random effects. The implied model for the marginal mean, averaged over the distribution of the random effects, cannot be a logistic regression model; that is, a logistic regression model with random effect is simply not compatible with a logistic regression model for the marginal means (when averaged over the distribution of the random effects). As a result regression coefficients that parsimoniously summarize the covariate effects of interest are not readily available. In addition any misspecification of the generalized linear mixed effects model can yield biased estimates of the implied marginal means. If the goal is to make an inference about the population average mean of Y_i, a marginal model should be adopted, thereby avoiding the

aforementioned problems, the need to correctly specify the conditional distribution of Y_i given b_i and the marginal distribution of b_i, and the computational demands of integrating over the distribution of the random effects. Thus we find ourselves in substantial agreement with Drum and McCullagh (1993, p. 300) when they comment that:

> "...the megalomaniacal strategy of fitting a grand unified model, supposedly capable of answering any conceivable question that might be posed, is, in our view, dangerous, unnecessary and counterproductive."

The answers to different scientific questions concerning longitudinal change will invariably demand that different statistical models have to be applied to the data at hand. In short, one size does not fit all.

16.7 FURTHER READING

A useful discussion of the distinct interpretations of the regression parameters in marginal and mixed effects models for binary data can be found in Section 12.2 of Agresti (2002); also see Chapter 7 of Fitzmaurice et al. (2009) for a detailed discussion of the distinct targets of inference for marginal and mixed effects models. Gardiner et al. (2009) compare and contrast the assumptions that underlie fixed effects, random effects, and marginal models.

Bibliographic Notes

Neuhaus et al. (1991) compare marginal and mixed effects models for analyzing correlated binary data; also, see Zeger et al. (1988), Graubard and Korn (1994), and Section 7.4 of Diggle et al. (2002). Hubbard et al. (2010) give a remarkably lucid description of the assumptions behind, and the type of inference provided by, marginal and mixed effects models in the cluster-correlated data setting.