NON-LINEAR MIXED MODELS IN A DOSE RESPONSE MODELLING

Madona Yunita Wijaya
Program Studi Matematika, Fakultas Sains dan Teknologi
Universitas Islam Negeri Syarif Hidayatullah Jakarta
mywijaya@gmail.com

Abstract: Study designs in which an outcome is measured repeatedly over time result in longitudinal data. Most of the methodological works have been done in the setting of linear and generalized linear models, where a certain amount of linearity is preserved. This may still be considered a limiting factor and non-linear models is another option offering its flexibility. Non-linear model involves complexity of non-linear dependence on parameters than that in the generalized linear class. It has been utilized in many situations such as growth curves modeling, pharmacokinetic and pharmacodynamic modeling, and dose-response modeling. The latter modeling will be the main interest in this study to construct a dose-response relationship, as a function of time to IBD dataset. Both linear and non-linear models are considered. The results indicate that non-linear models are more flexible than linear models hence able to capture more variability present in the data.

Keywords: IBD, longitudinal, linear mixed model, non-linear mixed model.

1. INTRODUCTION

The classical linear normal model was extended to generalized linear models (GLM) by Nelder and Wedderburn (1972). In LM and GLM settings, a certain amount of linearity is preserved; the former having a linear relationship between the mean response and the linear predictor, and the latter, being linear at the predictor. These models are very flexible and very popular among researchers where they have received much attention in the past few decades as compared to non-linear models. However, there has been increased need for advance modelling technique including non-linear modelling methods. Some of notable reasons are: first, modern statistics are confronted with complex data structures which become increasingly available with modern computing power. Secondly, within GLM framework, of which the LM is a special case with a normal distribution and the identity link, one is limited...
to the choice of distributions only from the exponential family. Hence non-linear models provide a more extensive option. Thirdly, with non-linear models, one captures more shapes with few parameters since few parameters generate a vast majority of shapes. Lastly, linearity is in most cases, just an approximation which may be less meaningful in some situations. Such situations include growth phenomena over sufficiently extended periods, especially when the observational period includes both growth spurts and asymptotic behaviour of growth toward maturation. Dose-response modelling, pharmacokinetic, and pharmacodynamic applications often demand non-linear models as well (Molenberghs and Verbeke, 2005). The non-linearity is not restricted to the fixed effects but sometimes includes the random effects. A flexible framework to accommodate various deviation from the general linear mixed model is the non-linear mixed model of which the generalized linear mixed model is a special case with non-linearity in link function but linear in the predictors.

In this study, we will apply both linear and non-linear models to Inflammatory bowel disease (IBD) dataset that was obtained from a longitudinal observational study. IBD commonly refers to ulcerative colitis (UC) and Crohn disease (CD), which are chronic inflammatory diseases of the GI tract of unknown etiology. Crohn disease usually affects the small intestine, but may be more widespread through the gastrointestinal system. It can present itself in a greater variety of ways than ulcerative colitis, for example it could start with fissures around the anus. Ulcerative Colitis is confined to the colon and rectum (the large bowel). Sometimes the rectum alone is involved. Inflammation and ulceration of the lining of the bowel cause urgent diarrhea which can be very frequent. Both conditions require anti-inflammatory drug treatments, but there are some differences between the two in the choice of drugs. Steroids are often needed either locally in the rectum, or orally to treat acute attacks. Frequently relapsing cases will need immunosuppressive treatments such as azathioprine.

The aim of this study is to construct a dose-response relationship, as a function of time to IBD dataset as described in Section 2. The method that will be used is discussed in Section 3 including linear and non-linear mixed models. The results are presented in Section 4 and finally the discussion and conclusion in Section 5.

2. DATA DESCRIPTION

The dataset comes from a clinical trial with 291 subjects, divided over four treatments: placebo (0), 1000mg (1), 2000mg (2), and 4000 mg(3). Patients are measured during a 7 week period. The outcome of interest is an IBD activity score (IBDSC). Note that not all of the patients were followed up until the end of the study resulting in missing measurements for some time points as can be seen in Table 1. It is also observed that 64.9% of the profiles are complete, while 7.6% exhibit dropouts and the remaining 27.5% have intermittent missing values.

<table>
<thead>
<tr>
<th>Week 1</th>
<th>Week 2</th>
<th>Week 3</th>
<th>Week 4</th>
<th>Week 5</th>
<th>Week 6</th>
<th>Week 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>#Obs.</td>
<td>291</td>
<td>279</td>
<td>265</td>
<td>247</td>
<td>231</td>
<td>221</td>
</tr>
<tr>
<td>Percentage</td>
<td>100%</td>
<td>96%</td>
<td>91%</td>
<td>85%</td>
<td>79%</td>
<td>76%</td>
</tr>
</tbody>
</table>

3. METHODOLOGY

3.1. Linear Mixed Model

In practice we are often confronted with unbalanced data, therefore many longitudinal data sets cannot be analyzed using multivariate regression techniques (Verbeke and
Non-Linear Mixed Models in a Dosen Response Modelling

Molenberghs, 2000). This leads to a two-stage model where in stage one the evolution of response variables is modeled for each patient or subject using a linear regression model. In stage two, the subject-specific regression coefficients obtained in stage one are used as response variables and is fitted with other known covariates (Verbeke and Molenberghs, 2000). The first stage assumes that \( Y_i \) satisfies the linear regression model:

\[
Y_i = Z_i \beta_i + \varepsilon_i,
\]

where \( Z_i \) is \((n_i \times q)\) matrix of known covariates, modeling how the response evolves over time for the \( i^{th} \) subject, \( \beta_i \) is a \( q \)-dimensional vector of unknown subject specific regression coefficients, and \( \varepsilon_i \) is a vector of residual components \( \varepsilon_{ij} \), \( j=1,2,...,n_i \). In a second stage, a multivariate regression model of the form:

\[
\beta_i = K_i \beta + b_i,
\]

is employed to explain the observed variability between the subjects, with respect to their subject-specific regression coefficients \( \beta_i \), \( K_i \) is a \((q \times p)\) matrix of known covariates, and \( \beta \) is a \( p \)-dimensional vector of unknown regression parameters. In addition, the \( b_i \) are assumed to be independent, following a \( q \)-dimensional normal distribution with mean vector zero and general covariance matrix \( D \) (Verbeke and Molenberghs, 2000).

Due to the extra variability and loss of information experienced in the two-stage analysis, the random-effects models that combines the two steps is applied and is defined as follows (Verbeke and Molenberghs, 2000):

\[
Y_i = X_i \beta + Z_i b_i + \varepsilon_i,
\]

where \( X_i = Z_i K_i \) is a matrix of known covariates and where all other components are as defined earlier with:

\[
\begin{align*}
    b_i &\sim N(0, D), \\
    \varepsilon_i &\sim N(0, \Sigma_i), \\
    b_i \text{ and } \varepsilon_i &\text{ are independent.}
\end{align*}
\]

The selection of a random-effects model is done by selecting the preliminary structures for the mean, for random-effects, and for the residual covariance. In order to find the appropriate random effects one should test the significance of the highest order random effects first in a hierarchical way. Therefore, the Restricted Maximum Likelihood (REML) test will be applied to test the need of the random effects since this will be based on the same mean structure. The classical likelihood-based inference to test for the need of random effects cannot be applied because the corresponding hypothesis is on the boundary of the parameter space. Thus, the asymptotic null distribution of the likelihood ratio test statistic is a mixture of chi-square with equal weights 0.5. The Maximum Likelihood (ML) approach will be used in the mean structure reductions since the REML approach breaks down. This is due to the fact that different mean structures are compared, which will yield incomparable results in case the REML is used.

3.2. Non-linear Mixed Model

In some applications, models are needed in which the mean is no longer modeled as a function of a linear predictor. These are called non-linear mixed models that can take various
forms, but the most common ones involve a conditional distribution of $Y_{ij}$ given $b_i$ belongs to the exponential family, encompassing both normally distributed and non-normal outcomes. In general, the mean is modeled as $E(Y_{ij} | h_i) = h(x_{ij}, \beta, z_{ij}, h_i)$. In agreement with generalized linear mixed models (GLMM), it is customary to assume normally distributed random effects with mean $0$ and covariance matrix $D$, even though other distributions are possible in principle as well. The same approaches can be used to parameter estimation as were developed for GLMM (Serroyen et al, 2009).

Based on statistical literature review, the following dose time-response model (generalized diffusion function) is proposed to fit to the IBD dataset (Woodruff, 2001; Zhu 2001, 2003):

$$Y_{ij} = A_0 + r_{ij} + \frac{(B_0 + r_2 + B_1 \times \text{dose}_i) \times \text{week}_{ij}}{1 + \exp(C_1 \times \text{dose}_i)} + \epsilon_{ij}$$

(5)

where $r_{ij} \sim N(0, \sigma_{r_i}^2)$, $r_2 \sim N(0, \sigma_{r_2}^2)$, and $\epsilon_{ij} \sim N(0, \sigma^2)$. The term $(B_0 + B_1 \times \text{dose}_i) \exp(C_1 \times \text{dose}_i)$ is the dose-dependent asymptote in time and $\exp(C \times \text{dose}_i)$ is the dose-dependent half life time. As such it prescribes a trajectory of persistent dose-effects, i.e. dose effects sustain at a constant level and do not show any recovery from the dose effects within the experimental time course.\[8\]. Due to computational difficulty, only two subject specific or random effects are considered, i.e. attaching $r_1$ and $r_2$ to $A_0$ and $B_0$, respectively, and they are assumed to be independent.

4. RESULTS

4.1. Exploratory Data Analysis (EDA)

The individual profiles from 40 randomly selected patients for each treatment groups are displayed in Figure 1. It can be observed that there are much variability between and within patients suggesting the need of random intercept and random slope in the model.

It is also confirmed by the non-constant variance structure across the time (Figure 2, right panel). The variances seem to increase with time and this could be due to attrition. This suggests that caution should be used with incomplete data. In addition, mean structure for each treatment indicates the need of non-linear function to describe the evolution. They seem to follow a quadratic evolution over time (Figure 2, left panel) and this can be considered to construct our initial model for linear mixed model in Section 4.2.

4.2. Linear Mixed Model

The adequacy of the first stage model based on graphical exploration in Section 4.1 was explored by testing for the need of a model extension. There is strong evidence in favor using the quartic first stage model results in $F_{\text{meta}} = 1.105$ which is not significant (p-value = 0.2402) when compared to an $F$-distribution with 224 and 189 degrees of freedoms. In the second stage model, the subject-specific intercepts and time effects are related to the dose levels as a continuous rather than as a categorical variable assuming the placebo group received dose level = 0. This was considered since no substantial differences obtained from the model by treating the dose as continuous or categorical. Thus by taking continuous dose level, we do not lose information. Further efficiency is achieved by estimating less parameter.

The preliminary linear mixed effects model is suggested by combining the first and second stage models with random intercept and random slopes with unstructured working assumption. Likelihood Ratio Test (LRT) statistics with the correct null distribution was
Non-Linear Mixed Models in a Dosen Response Modelling

performed to test the need of random effects. The model is then reduced to a more parsimonious model by deleting insignificant terms in a hierarchical way. The results for the final model based on the observed data (direct likelihood) are shown in Table 2 assuming the missingness in the data is at random. The final model was also fitted after generating 5 multiple imputations. Both models yielded similar parameter estimates and led to the same inference.

Figure 1. Individual profiles per treatment groups.

Figure 2. Mean and variance structures for the different treatment groups
Table 2. The estimated parameter (standard error) for the linear mixed model.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Direct Likelihood</th>
<th></th>
<th>Multiple Imputation</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate (s.e.)</td>
<td>P-value</td>
<td>Estimate (s.e.)</td>
<td>P-value</td>
</tr>
<tr>
<td>Intercept</td>
<td>150.150 (5.394)</td>
<td>&lt;.0001</td>
<td>149.904 (5.899)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Dose</td>
<td>2.863 (1.132)</td>
<td>0.012</td>
<td>2.757 (1.172)</td>
<td>0.0193</td>
</tr>
<tr>
<td>Week</td>
<td>-68.040 (7.665)</td>
<td>&lt;.0001</td>
<td>-67.333 (8.704)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Week²</td>
<td>26.152 (3.500)</td>
<td>&lt;.0001</td>
<td>25.682 (4.192)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Week³</td>
<td>-4.309 (0.638)</td>
<td>&lt;.0001</td>
<td>-4.204 (0.794)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Week⁴</td>
<td>0.252 (0.040)</td>
<td>&lt;.0001</td>
<td>0.245 (0.051)</td>
<td>0.0001</td>
</tr>
<tr>
<td>Dose × week</td>
<td>-2.177 (0.721)</td>
<td>0.003</td>
<td>-2.143 (0.789)</td>
<td>0.0087</td>
</tr>
<tr>
<td>Dose × week²</td>
<td>0.233 (0.083)</td>
<td>0.005</td>
<td>0.231 (0.091)</td>
<td>0.0145</td>
</tr>
</tbody>
</table>

*dosage unit is g instead of mg

4.3. Non-linear Mixed Model

Non-linear mixed model can be considered as an alternative to linear mixed model to incorporate a more flexible function that can approximate the observed mean profile. Parameter estimates and standard errors obtained from fitting model (5) are presented in Table 3. It can be observed that all parameters are significant. The model was tried to simplify by removing the random $r_2$ effect but it was not possible since the likelihood ratio equals 151 on 0 and 1 degree of freedom ($p < 0.0001$). Note that this model is non-linear in the fixed-effects parameters, but linear in the random effect, simplifying the calculation of the marginal mean over the random-effects distribution.

Table 3. The estimated parameter for the non-linear mixed model.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_0$</td>
<td>141.620</td>
<td>6.8184</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>$B_0$</td>
<td>-77.617</td>
<td>12.671</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>$B_1$</td>
<td>3.907</td>
<td>1.9592</td>
<td>0.0471</td>
</tr>
<tr>
<td>$C_1$</td>
<td>-0.079</td>
<td>0.0278</td>
<td>0.0046</td>
</tr>
<tr>
<td>$D_0$</td>
<td>1.045</td>
<td>0.0355</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>$Var(r_1)$</td>
<td>39.650</td>
<td>16.293</td>
<td>0.0156</td>
</tr>
<tr>
<td>$Var(r_2)$</td>
<td>789.67</td>
<td>111.65</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>16.450</td>
<td>0.3079</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

4.4. Model Comparison

The marginal profiles obtained from linear as well as non-linear mixed models are depicted in Figure 3 for each doselevel/treatment group. The linear mixed model fits the data well at the beginning and at the end of the study and seems to overfit in the middle of the study for the four treatment groups, especially for 1000mg and 4000mg. The same picture is delineated from fitting the non-linear mixed model. In addition, less parameter is needed to be estimated under nonlinear mixed model compare to linear mixed model. Thus, we achieve higher efficiency by fitting non-linear mixed model to come up with excellent results. The four different dose levels show similar trend, i.e. higher IBD score in week 1 and the score decreases as time increases. Higher dose level implies lower IBD score.
5. DISCUSSION AND CONCLUSION

In this paper, it is of interest to explore the dose-response relationship as a function of time. When conventional linear models may be insufficient, non-linear models can be used as an alternative to describe the evolution of the profiles in a longitudinal setting. Linear and non-linear mixed models are used to fit to the IBD dataset and both approaches seem to fit very well although they seem to overfit in the middle of the study as seen in Figure 3. This could be due to intermittent missing value and it was observed that 27.5% of the patients have non-monotone pattern. However, linear mixed model might be inadequate in this case since taking a higher order polynomial can impose multiple peaks or valley on the dose-time-response, and is difficult to interpret [8]. Meanwhile, non-linear mixed model offers efficiency; only need less parameter than linear mixed model. The user also has the flexibility to define the nonlinear model that applies to a particular data set or response although it is not an easy task. But it will be very handy if we have a theoretical background.

REFERENCES