Department of Statistics, National Taipei University
Instructor: Yi-Ting Hwang

A common way of treating patients with cardiovascular disease is by surgical intervention. In particular, such patients may arrive at a hospital with symptoms such as unstable angina or suspected myocardial infarction (heart attack), requiring that physicians perform an invasive procedure called a percutaneous coronary intervention (PCI) to investigate the extent to which coronary arteries might be blocked. During such an investigation, the blockage may be treated using a balloon to dislodge the blockage and widen the artery (balloon angioplasty); in addition, a device known as a stent may be inserted to prop the artery open.

When such PCI procedures are performed, it is necessary for the subject to be treated with a drug that inhibits the aggregation of platelets in the blood. Informally, platelets are a blood constituent involved in clotting of the blood; clotting occurs when the platelets aggregate together in clumps. To ensure that clotting does not interfere with the procedure, inhibition of the clotting mechanism is desirable; clotting during the procedure can lead to complications such as stroke or heart attack. A long-standing issue has been to determine which of two popular drugs elicits the most desirable pattern of inhibition of platelet aggregation.

Accordingly, an experiment was conducted to compare the platelet aggregation patterns of the drugs in such subjects under controlled conditions. Subjects arriving at a major medical center with symptoms of unstable angina or myocardial infarction who were judged to require a PCI procedure were randomized into two groups, one for each of the drugs, with 200 subjects per group. For each subject, at time 0, the assigned drug was administered according to the manufacturer's recommended dosage; for each drug, this involved giving the subject a large dose by injection to start inhibition of platelet aggregation immediately and simultaneously giving the subject a smaller dose of the drug intravenously at a constant rate over several hours, a method of administration known as an infusion. The purpose of the infusion was to keep platelet aggregation inhibited over at least a 12 hour period, so that clotting would be minimized during the PCI procedure and subsequent recovery for the subject. For each subject, blood samples were to be taken at 0.5, 2.0, 3.5, 5.0, 8.0, 11.0, and 12.5 hours. Each sample was to be analyzed for degree of platelet inhibition, characterized by the response percent inhibition, a value between 0 and 100 representing the percentage of inhibition relative to that of an untreated sample (in units of %μM). Also recorded for each subject was whether the subject had experienced a previous myocardial infarction before the current hospitalization (0=no, 1=yes) and gender (0=female, 1=male).

The data from the study may be found on the class web page in the file longdata110sp5.dat. Each record in the file corresponds to a single observation, and the columns are

1. subject id number (1 – 400),
2. previous myocardial infarction indicator (0 = no, 1 = yes),
3. gender indicator (0 = female, 1 = male),
4. time (hours, measured since administration of drug),
5. percent inhibition,
6. drug group indicator (1 or 2).

Note that for some subjects, the response is not available at all intended time points; some samples were mishandled and in some instances study personnel did not follow the instructions and neglected to obtain samples. It was determined that the reasons for the missing values had nothing to do with the drugs or the patterns of inhibition.

The time plot shows that over the period of the study, platelet inhibition appears for most subjects to follow a rough straight-line trajectory that either stays relatively flat or rises, although a few profiles seem to decrease. To represent this, the investigators proposed the following model. Because the investigators were particularly interested in the time point 0.5 hours post-administration, as we will see shortly, they defined time in the model so that $t = 0$ corresponds to 0.5 hours after administration of the drug. That is, they let t_{ij}, the time of the jth platelet inhibition response on subject i, be defined as

$$t_{ij} = s_{ij} - 0.5,$$

where s_{ij} = time of the jth response on subject i measured from administration of the drug (so s_{ij} equals the time value given in the data file). Letting y_{ij} be the corresponding platelet inhibition response for subject i at the jth time, the model is

$$y_{ij} = \beta_0 + \beta_1 t_{ij} + b_{0i} + b_{1i} t_{ij} + e_{ij}, \tag{1}$$

This model thus allows the pattern after 0.5 hours to follow a straight line for each subject.

(a) Using your favourite graphical software, make separate spaghetti plots of the raw longitudinal profiles for each group separately. Give your general impressions on the overall observed patterns and how they might compare across groups.

(b) The investigators initially wished to assume that, for model (1), mean platelet inhibition at 0.5 hours following administration of drug has the following features within each drug group:

- is associated with whether the subject has had a previous myocardial infarction
- is associated with whether the subject is male or female
- the way in which it is associated with whether the subject is male or female is the different depending on whether the subject has had a previous myocardial infarction. Because the subjects had been on the drugs for 0.5 hours, the investigators assumed that mean platelet inhibition at 0.5 hours and the way the above features occur is different for the two drugs.

The investigators also wished to assume that the typical or mean rate of change of platelet inhibition over the study period also has these features, and they wished to allow for the possibility that the mean rate of change of platelet inhibition and its association with prior myocardial infarction and gender could be different for each drug. This would allow the possibility that the drug that is received is associated with the pattern of change of platelet inhibition in different ways for subjects of different genders and prior history of myocardial infarction.

Let m_i = 0 if subject i has not had a previous myocardial infarction and m_i = 1 if s/he has, and let g_i = 0 if i is female, and g_i = 1 if i is male. Given these beliefs, write down expressions for fixed effects for subject i taking drug k, k = 1, 2. Be sure to define and fully describe all additional symbols you use.

(c) In terms of your model in (b)

(i) Give an expression that represents the typical value of platelet inhibition at 0.5 hours after drug administration for male subjects with a previous myocardial infarction taking drug 2.

(ii) Give an expression for mean platelet inhibition for female subjects with no previous myocardial infarction at 12 hours following administration of drug 1.

(d) Let $\boldsymbol{e}_i = \boldsymbol{e}_{1i} + \boldsymbol{e}_{2i}$. The investigators were all willing to believe that

(i) the assay used to measure platelet inhibition for both drug groups exhibits constant variation regardless of the true value of platelet inhibition being ascertained, i.e., $\text{Var}(\boldsymbol{e}_{2i}) = \sigma_1^2 \boldsymbol{I}_{n_i}$,

(ii) within-subject local fluctuations in platelet inhibition are of similar magnitude for both drugs and across time for all subjects, $\text{Var}(\boldsymbol{b}_i) = \boldsymbol{G}$,

(iii) variation in inherent true platelet inhibition at 0.5 hours is similar for patients in both drug groups, as is variation in the inherent rates of change of platelet inhibition over the study period and the way these quantities covary, i.e., $\text{Var}(\boldsymbol{e}_{1i}) = \sigma_1^2 \boldsymbol{\Gamma}_{n_i}$, where $\boldsymbol{\Gamma}_{n_i}$ is a correlation matrix that allows the correlation among elements to fall off as they get farther apart in time.

For example, $\mathbf{\Gamma}_{n_i}$ could be an exponential covariance matrix or a Gaussian covariance matrix.

One of the investigators was concerned, however, that the time points at which platelet inhibition was measured were not sufficiently far apart in time to ensure that measurements within a subject are uncorrelated. He was willing to believe that, if such correlation is present, it falls off as the time points get farther apart, but he insisted that an analysis be done to resolve this issue.

Give two different sets of assumptions on the e_{ij}, $i = 1, \ldots, n_i$, and random effects corresponding to b_{0i} and b_{1i} in (1) that incorporate (i) – (iii). The first set of assumptions should incorporate the investigator's concern; the second set should represent the case where the investigator's concern is unwarranted.

Fit the overall model (1) along with your model for β_0, β_1, b_{0i} and b_{1i} in (b) using SAS proc mixed and the method of restricted maximum likelihood. Which set of assumptions is best supported?

(e) From the output for the fit of the model you preferred in (d), write down an estimate of the variance associated with among-subject variation in true platelet inhibition in the population of male subjects with no previous myocardial infarction receiving drug 2 at 0.5 hours post-administration.

(f) Previous research has suggested that the way in which platelet inhibition occurs for both drugs over this period may be associated with whether a subject has had a previous myocardial infarction, but there is no evidence to suggest that it is associated with gender in any way. Thus, the investigators planned to base their subsequent analyses not on the model you developed in (a) but on a model that includes no effect of gender either in the representation of mean platelet inhibition at 0.5 hours or in the representation of the typical rate of change of platelet inhibition over the study period. Write down this simpler model and fit it using ML and your preferred covariance structure from (d). Based on your preferred fit in (d) and this fit, is there any evidence against doing this?

(g) For the rest of the problem, consider the simpler model in (f) with no gender effects. The reason that the investigators were so interested in 0.5 hours post-administration is because another research team had recently published a paper receiving a lot of press, which claimed that the 2 drugs exhibit the same mean platelet inhibition and that, furthermore, mean platelet inhibition on the two drugs is the same for subjects with or without a previous myocardial infarction. This team based their finding on comparing platelet inhibition levels 0.5 hours post-administration. Our investigators felt that comparing platelet inhibition at a single time point, particularly one so soon after administration, was not very informative. Thus, their first goal was to examine whether the data from the current study offer evidence refuting the claim of their rival investigators. Write

down a set of hypotheses that addresses the issue of interest to the investigators in terms of the model in (f), and express your null hypothesis in terms of a linear function $\boldsymbol{L}$, defining $\boldsymbol{L}$. Using Wald methods, carry out the test at level of significance 0.05 based on a REML fit. State your conclusion as a meaningful sentence.

(h) The investigators' second goal was to make the point that comparing platelet inhibition at a single point does not tell the whole story. Thus, regardless of how the test in (g) turned out, they wanted to investigate longer time periods and the rate of change of platelet inhibition over them. The first question along these lines was whether the way "typical" rate of change differs between subjects who have had a previous myocardial infarction and those who have not is different for the two drugs. Write down a set of hypotheses that addresses the issue of interest to the investigators in terms of the model in (f), and express your null hypothesis in terms of a linear function $\boldsymbol{L}$, defining $\boldsymbol{L}$. Using Wald methods, carry out the test at level of significance 0.05 based on a REML fit. State your conclusion as a meaningful sentence.

(i) Consider also the model you specified in part (f). Using SAS proc mixed, fit the model using restricted maximum likelihood. This time, have your program print out both the estimates of the fixed parameters in $\boldsymbol{\beta}$ and the approximate best linear unbiased predictors (BLUPs) of the random effects $\boldsymbol{b}_i$. In your program, create 4 data sets and have the program print out their contents:

(i) A data set containing only the data for subject 4 in drug group 1

(ii) A data set containing only the data for subject 100 in drug group 1

(iiii) A data set containing only the data for subject 230 in drug group 2

Note

- Attach your SAS program as an appendix.
- **You should include appropriate outcomes in your report to support your findings.**