A clinical trial for chronic pulmonary diseases (COPD)

Chronic pulmonary disease (such as Chronic Obstructive Pulmonary Disease – COPD) concerns the development of emphysema. It is a slow progression over many years and the assessment of drug efficacy requires the observation of large numbers of patients for a long period of time. Recently, lung densitometry (measuring the lung density through CT scan) considered for assessing the lung tissue loss over time in patients with emphysema.

A clinical trial with lung densitometry as an endpoint is typically designed as a longitudinal study with repeated measurements at fixed time intervals. Since lung density measurements are closely correlated with lung volume (inspiration level), it is important to include lung volume measurements in statistical analyses as a longitudinal covariate. Lung volume is normally measured at the same time as the lung density is measured.

The clinical efficacy can be assessed by comparing the progression of lung density loss between two treatment groups using a random coefficient model – a longitudinal linear mixed model with a random intercept and slope. In planning the clinical trial with such complex statistical analyses, the calculation of the sample size required to achieve a given power to detect a specified treatment difference is an important, often complex issue.

In this exercise, you are required to simulate outcomes of lung density and lung volume. This can be done using the following model.

$$Y_{ij} = (\beta_0 + b_0) + \beta_1 * TRT + (\beta_2 + b_2) * TIME + \beta_3 * COV_{ij} + \beta_4 * TRT * TIME + \varepsilon_{ij} \quad (1)$$

Here Y_{ij} is the efficacy endpoint (i.e. lung density) measurement for subject *i = 1, 2, …, n, at fixed time point j = 1, 2, …, K.*

- *TRT is an indicator of the i^{th} subject's treatment group (i.e. TRT=1 for active drug; TRT=0 for placebo).*

- *COV_{ij} is a longitudinal covariate (i.e. logarithm of lung volume) for subject i = 1, 2,…, n, at fixed time point j = 1, 2, …, K.*

- *Here b_0 and b_2 are subject-specific random effects for the intercept and slope, respectively, which are from a normal distribution with mean 0 and variance σ_{02} and σ_{02}, respectively.*

- *ε_{ij} is the random error from a normal distribution with mean 0 and variance σ_2 .*

- *The regression parameters β_0, β_1, β_2, β_3, and β_4 are the fixed effects for intercept, treatment, time, covariate and interaction of treatment and time respectively.*

- Here we assume that the benefits can be assessed quantitatively by comparing the slopes of lung density trajectories for the two treatment groups. This quantity is captured by β_4 which is the difference in slope of time between two treatment groups (active vs. placebo).

Assume we know the parameters (*β_0, β_1, β_2, β_3, and β_4,* and *σ_{02} and σ_{02}*) from history data, previous clinical trials or meaningful clinical differences (cf. below). Assuming fixed time intervals (TIME) and fixed equal sample size n for each treatment, the trajectories of efficacy measurement Y_{ij} (i.e. lung density) for the n subjects can be simulated through the model for each treatment group. Then, we could perform a statistical test on *β_4* =0 on the simulated data set.

In order to simulate the trajectories of Y_{ij}, it is necessary to simulate the trajectories of longitudinal covariate COV_{ij}. Assume COV_{ij} is from a linear model regressing against time with a random intercept ($COV_{ij}=(\gamma_0+r_0)+\gamma_1 TIME+\varepsilon_{ij}$ (2)) where γ_0 and γ_1 are the fixed intercept and slope respectively; r_0 and ε_{ij} are from a normal distribution with mean 0 and variance δ_{12} and δ_{22}, respectively. Since we know the parameters (γ_0,γ_1 , δ_{12} and δ_{22}) from history data or previous clinical trials for the study population, it will be simple to simulate the trajectories of the longitudinal covariate COV_{ij} by using some random generating function. Assume there are two treatment groups (active vs. placebo) in a study design. The efficacy endpoint along with the

longitudinal covariate will be measured at *K=2* time points. All corresponding parameters specified in model (1) and (2) could be obtained either through historical data, previous clinical trials or meaningful clinical difference to be tested from clinicians. For the purpose of this project, they are randomly selected and specified as below. Earlier sample size calculations have shown that sample size and power behave as in Table 1 below where α=0.05 and *1-*β=0.80.

N per treatment	Statistical Power (%)
30	62.4
40	76.9
45	**79.9**
50	84.4
60	91.3

Table 1

$\beta_0 = 150, \beta_1 = 5, \beta_2 = -1.8, \beta_3 = -57, \beta_4 = 0.7,$ and $\sigma_0^2 = 280, \sigma_2^2 = 0.4, \sigma^2 = 5;$
$\gamma_0 = 2, \gamma_1 = 0.0007, \delta_0^2 = 0.05, \delta_1^2 = 0.0016.$