Consulting Project

Jiaqi Sun

2022-12-02

Building Model

We structured a multilevel model for this pilot trail study, the variables we use for multilevel are test (test1 - test n), group, patient_id and score. For the random effect is $(1|patient_{ID})$ and $(0 + test|patient_{ID})$.

Column	Description
Y	Survey Score
i	Patient_id
j	Group
t_k	Test
α_j	intercept of random effect of different patients
δ_{ij}	intercept of random effect of different patients
β_c	control group slope, we use 0.08 here
γ_t	treatment group effect size, we use 0.2 here
μ_i	slope of random effect of different patients
e_{ijk}	error

$$
\begin{aligned}
e_{ijk} &\overset{\text{iid}}{\sim} N(0, \sigma_y^2) \\
\delta_{ij} &\overset{\text{iid}}{\sim} N(0, \sigma_\delta^2) \\
\mu &\overset{\text{iid}}{\sim} N(0, \sigma_\mu^2) \\
Y_{ijk} =& \alpha_j + \delta_{ij} + (\beta_c + \gamma_t + \mu_i) \cdot t_k + e_{ijk}
\end{aligned}
$$

$$
score_{ijk} = 3 + \delta_{ij} + (0.08 + 0.2 + \mu_i) \cdot test + e_{ijk}
$$

where δ_{ij} and μ_i are random effects: $(1|patient_ID)$, $(0 + test|patient_ID)$

Suggestions:

Firstly, this is definitely an under power study because the questions are similar and it is the same person answering the questions. As people answering the questions in the sample questionnaire, the answers will be correlated. If we pretend each one is independent, then we underestimate the variability. So the sample size should be larger than the one we gave on the plot.

Secondly, there are four repeated measurements in this experiment, some of the surveys are too long and contain similar questions. It's hard for normal people to finish fulfilling in 15 minutes, not to mention patients with hand joint pain. The patients maybe take the survey seriously at the beginning and then just give random answers. So we suggest to shorten the questionnaire, or pick some of them to represent.