

Homework #1

BSTA 519 Applied Longitudinal Data Analysis - Fall 2021

Due - Monday 10/11/2021 by midnight

The Toenail Data

A randomized, double-blinded parallel group study was conducted for the comparison of two oral treatments (group A: Lamisil; group B: Itraconazol) for toenail dermatophyte onychomycosis (TDO). TDO is a common toenail infection and difficult to treat. The aim of the present study was to compare the efficacy and safety of 3 months of continuous therapy for the two treatments.

In total, 2×189 patients were originally randomized and subjects were followed up to a total of 12 months. Measurements were taken at baseline, every month during treatment, and every 3 months afterward, resulting in a maximum of seven measurements per subject. For the purpose of this homework, we will consider one of the secondary outcomes, unaffected nail length as our outcome variable of interest based on available data.

In Sakai Under submission → Homework 1, you will find a dataset in Excel called *Toenail.xlsx*. Import the file into the statistical software of your choice, and do the following:

Exploratory Data analysis

- a) Create a profile plot of unaffected nail length vs. time for 50 group A subjects (you could use 50 randomly selected group A subjects, or first 50 group A subjects, or any 50 group A subjects of your choice) and based on the plot, comment on
 - 1) How the mean unaffected nail length changes over time;
 - 2) How the variation in unaffected nail length changes over time;
 - 3) Whether there is any outlier in the data of your 50 subjects.
- b) Create a profile plot of mean unaffected nail length vs. time for group A and Group B. Based on the plot, comment on
 - 1) How the mean unaffected nail length changes over time for each group and whether there is any obvious pattern.
 - 2) Whether the trend of mean unaffected nail length over time is the same for the two groups.
- c) Create a scatter plot of unaffected nail length vs. time and add a lowess curve for group A and group B, separately and comment on the trend of mean unaffected nail length over time for each group.
- d) Create a scatter plot matrix of the repeated measurements over time and comment on how the correlation among the repeated measurements changes over time.

Simple longitudinal data analysis looking at data at baseline and month 1 only

Homework #1

BSTA 519 Applied Longitudinal Data Analysis - Fall 2021

Due - Monday 10/11/2021 by midnight

- e) Fit an appropriate model to answer the following questions and interpret your results. Write out the model and specify the null vs. alternative hypotheses for each question.
- 1) Does the change of unaffected nail length from baseline to month 1 equal to zero for the Itraconazol group?
 - 2) Does the change of unaffected nail length from baseline to month 1 differ between the two groups?
- f) Fit an appropriate model to answer the following questions and interpret your results. Write out the model and specify the null vs. alternative hypotheses for each question.
- 1) Does the unaffected nail length at month 1 differ between the two groups after adjusting for baseline difference in unaffected nail length?
 - 2) Is the unaffected nail length at month 1 related to the baseline unaffected nail length after adjusting for group difference?

Note: Make sure to answer the questions AND interpret your results including providing basis (such as results of hypothesis testing, estimates of difference or change and its 95% CI) for your answers, and attach appropriate output if you want.

Lord's paradox

As we discussed in the class, ANCOVA model should be used carefully in observational studies. Read the following two papers:

- **Wright, 2006.** Comparing groups in a before-after design: when t test and ANCOVA produce different results. Br J Educ Psychol. 2006 Sep;76:663-75.
- **Glymour et al. 2005.** When is baseline adjustment useful in analyses of change? An example with education and cognitive change. Am J Epidemiol. 2005 Aug 1;162(3):267-78.

And answer the following questions:

- 1) What is Lord's paradox? Why could the paradox occur?
- 2) Under what conditions is ANCOVA model appropriate to use? Under what conditions is ANCOVA model NOT appropriate to use?