Class project

BSTA 519 Applied Longitudinal Data Analysis

Fall 2021

Presentation – Monday 12/06/2021 3:00 to 6:00pm

Report due – Wednesday 12/08/2021 Midnight

The Childhood Asthma Management Program study

The Childhood Asthma Management Program (CAMP) was a clinical trial carried out in children with asthma. The trial was designed to determine the long-term effects of 3 treatments (budesonide, nedocromil, or placebo) on pulmonary function. The design of CAMP was a multicenter, masked, placebo-controlled, randomized trial. A total of 1041 children aged 5-12 years were enrolled between December of 1993 and September of 1995. In this analysis, we will look at a subset of 695 patients.

The primary outcome of the trial was lung function as measured by FFratio, the ratio of FEV1 to forced vital capacity (FVC) where FEV1 is the Forced Expiratory Volume at 1 second. The normal value for the FEV1/FVC ratio is 70%, and a lower measured value corresponds to a more severe lung abnormality. Additional patient characteristics that might impact the outcome variable are also collected. The data for this analysis are provided in the file ***camp_2021 class project.xlsx***.

Dataset Variables

id	Randomized participant ID
Trtment	Treatment group: A= budesonide, B= nedocromil, C= placebo
Age_rz	Age in years at Randomization
Gender	m=male, f=female
Ethnic	w=white, b=black, h=hispanic, o=other
FFratio	FEV1/FVC ratio %
Anypet	Any pets, 1=Yes 2=No
Woodstove	Used wood stove for heating/cooking, 1=Yes 2=No
Dehumid	Use a dehumidifier, 1=Yes 2=No 3=Don't Know
Parent_smokes	Either Parent/partner smokes in home, 1=Yes 2=No
Visitc	Followup visit (mos)
Fdays	Days since randomization
Fyears	Years since randomization

Note that information for a few variables (Anypet, Woodstove, Dehumid and Parent_smokes) is not available at every visit. For visits with missing values, replace the missing values with the value from the previous visit with information collected.

Question 1: Is there change over time in FFratio, and are budesonide and nedocromil effective in preventing the decline of FFratio? Is there a difference in effectiveness between budesonide and nedocromil?

Class project

BSTA 519 Applied Longitudinal Data Analysis

Fall 2021

Presentation – Monday 12/06/2021 3:00 to 6:00pm

Report due – Wednesday 12/08/2021 Midnight

- Fit an appropriate model to answer the above research question.
    - Provide exploratory plots that you think are helpful for presenting the data.
    - Justify your choice of model and the approaches you choose to model the mean response and the correlation among repeated measures.
    - Perform relevant model diagnostics to check model assumptions and outlying/influential data points (subjects).

- It is also of interest to examine whether patient characteristics are associated with FFratio and what the treatment effects are after adjusting for patient characteristics. Fit an appropriate model to answer this question.

Question 2: Define abnormal FFratio as FFratio (%) < 70. Is there change over time in abnormal FFratio, and are budesonide and nedocromil effective in preventing abnormal FFratio? Is there a difference in effectiveness between budesonide and nedocromil?

- Fit an appropriate model to answer the above research question.
    - Provide exploratory plots that you think are helpful for presenting the data.
    - Justify your choice of model and the approaches you choose to model the mean response and the correlation among repeated measures.

- It is also of interest to examine whether patient characteristics are associated with abnormal FFratio and what the treatment effects are after adjusting for patient characteristics. Fit an appropriate model to answer this question.

Respiratory infection in preschool children

Respiratory infections are infections of parts of the body involved in breathing, such as the sinuses, throat, airways or lungs. Two hundred and seventy-five preschool children in Indonesia were examined for up to six consecutive quarters for the presence of respiratory infection. It is of interest to evaluate the change in the prevalence of respiratory infection with age, and whether the prevalence of respiratory infection is higher among children who suffer xerophthalmia, an ocular manifestation of chronic vitamin A deficiency. The data set you will work with is a subset of the original bigger cohort study (Sommer et al. 1984 ***Am J Clin Nutr.*** Nov; 40(5):1090-5.)

Class project

BSTA 519 Applied Longitudinal Data Analysis

Fall 2021

Presentation – Monday 12/06/2021 3:00 to 6:00pm

Report due – Wednesday 12/08/2021 Midnight

The data set is ***Respiratory infection.xlsx*** and the variables include:

(1) ID, identification number of each participant
(2) Respiratory infection, the outcome variable (1 = Yes, 0 = No)
(3) Age in months (Centered at 36 months)
(4) Xerophthalmia (1 = Yes, 0 = No)
(5) Sex (1 = female, 0 = male)
(6) Height for age, a percentage of US National Center for Health Statistics standard
(7) Visit
(8) Baseline age in months (Centered at 36 months)
(9) Season (1 = summer, 2 = autumn, 3 = winter, 4 = spring)

Question 3: How does the prevalence of respiratory infection change with age, and whether change in the prevalence of respiratory infection differs by xerophthalmia status?

Based on this dataset,

- The change in the prevalence of respiratory infection with age could be evaluated both cross-sectionally and longitudinally. Fit an appropriate model to evaluate both cross-sectional and longitudinal associations between respiratory infection with age.
    - Provide exploratory plots that you think are helpful for presenting the data.
    - Justify your choice of model and the approaches you choose to model the mean response and the correlation among repeated measures, and examine whether cross-sectional and longitudinal associations are the same.
    - Adjust other variables (Sex, Height for age, Xerophthalmia, and Season) in the model.

- Fit an appropriate model to evaluate whether the longitudinal change in the prevalence of respiratory infection differs by xerophthalmia status.
    - Provide exploratory plots that you think are helpful for presenting the data.
    - Justify your choice of model and the approaches you choose to model the mean response and the correlation among repeated measures.
    - Adjust other variables (Sex, Height for age, Baseline age in months and Season) in the model.

For each question,

1. Summarize the results in a report up to 5 pages, and organize your report as abstract, introduction with the research question, statistical model, analysis methods, results, discussion and conclusions.

Class project

BSTA 519 Applied Longitudinal Data Analysis

Fall 2021

Presentation – Monday 12/06/2021 3:00 to 6:00pm

Report due – Wednesday 12/08/2021 Midnight

---

   i. Page limit is strictly enforced. Can be single spaced, no less than 0.8" margins.
   ii. Supplement your introduction section with a ***simple*** literature search on the background.
   iii. In the beginning of results section, provide a table to summarize patient characteristics. It is also helpful to provide descriptive statistics of the outcome at each time point.
   iv. Remember to provide justification of your model, and you may include up to two plots to illustrate your data and results.
   v. Make sure to provide interpretations of the model coefficients and parameters in the context of the research question.
   vi. In the conclusion, also communicate your results in a way that could be understood by the general public with limited statistical knowledge.
   vii. Present your results in a table if necessary.
2. Provide an appendix of your code, relevant output of your final models, as well as the data file. Be sure that coding and data management is clearly documented to ensure the reproducibility of your analysis.
3. Provide a written description of role of each group member in completing the class project.
4. Create a 20 minute oral presentation to present your results to the class. (Think about what are the most important results/message to present.) Submit your power point slides after the presentation.

Group assignments

It is up to you to find your group members and pick a question to work on. Each group will have 3 students, except that one group will have 2 students. I will send out a separate email to record your group assignments.