Exploration 4.2: Will an Extended-Release Steroid Induce Ulcerative Colitis Remission?

Observational Studies Versus Experiments

LEARNING GOALS

- Identify a study as observational or experimental.
- Explain that random assignment gives us the ability to draw cause–effect conclusions because it ensures that treatment groups have similar characteristics.
- Identify whether a study uses random assignment and/or random sampling and the implications of these design decisions on the conclusions that can be drawn.
- Describe what a block study design is and the benefits of using a blocking variable.

Ulcerative Colitis is an inflammatory bowel disease that involves chronic inflammation of the colon as a result of a person's overactive immune system. Different types of oral steroids are often used to reduce the amount of inflammation in the colon and thus reduce symptoms in the patient. However, steroids can cause many adverse side effects and must by limited in their use. For this reason, having a non-systematic steroid delivered to the location of the inflammation is important. One such steroid is budesonide encased in a coating that releases the drug when the pH is seven or higher, which usually occurs at the beginning of the colon. All of this is in theory until the drug is tested in humans to see how effective it is and whether it has any adverse side effects.

The way these tests are typically done is by giving one group of patients the new drug and another group a placebo (a fake pill). Suppose 24 subjects have been recruited to participate in such a study. They are between 18 and 77 years old. They all have had mild to moderate ulcerative colitis for at least 6 months, have not used any steroids for at least 4 weeks, and have not used any immunosuppressive agents within 8 weeks. This group happens to consist of 8 females and 16 males.

1. One way to design this study would be to assign the 8 females to the placebo group and the 16 males to use drug group. Would this be a reasonable strategy? Why not?

2. One way to deal with the issue from #1 is to assign 4 females and 8 males to each group. This could be done by organizing the names alphabetically and putting the first 4 females in one group and the last 4 in the other group. Do this similarly for the males. What is the proportion of males in each group?

3. Suppose you used the method of assigning subjects described in #2. If you saw a difference in the proportion of patients that achieved remission in their disease, could the difference be explained by the sex of subject? Why or why not? Could the difference be due to other variables, distinct from the drug? Why or why not?

4. While this approach may prevent sex from being a confounding variable, there are still many other potential confounding variable. Identify a better method than #2 for deciding who uses which drug.

Definition
In an ***observational study,*** the groups you compare are "just there," that is, they are defined by what you *see* rather than by what you *do*. In an ***experiment***, you actively create the groups and assign the conditions to be compared. These conditions may be one or more **treatments** or a **control** (a group you do nothing to except perhaps a placebo treatment).

We have been calling subjects in an observational study, observational units. Similarly, subjects in an experiment can be called ***experimental units***.

Definitions
In a ***randomized experiment***, you use a chance device to make the assignments of the experimental units to the groups. The role of the random assignment is to balance out potential confounding variables among the explanatory variable groups, giving us the potential to draw cause-and-effect conclusions.

5. Let's explore the process of random assignment to determine whether it does "work." First, let's focus on the sex (male vs. female) variable. Suppose we put each person's name on a slip, put those slips in a hat and mix them up thoroughly, and then randomly draw out 12 slips for names of people to assign to the drug treatment group. The other 12 will be assigned to the placebo group. What proportion of this treatment group do you expect will be male? What proportion of the placebo group do you expect will be male? Do you think we will always get an 8/8 split (8 males in each treatment group)?

6. To repeat this random assignment a large number of times to observe the long-run behavior, we will use the **Randomizing Subjects** applet. Open the applet and press the **Randomize** button. What proportion of subjects assigned to Group 1 are male? Of Group 2? What is the difference in these two proportions?

You will notice that the difference in proportions of males is shown in the dotplot in the bottom graph. In this graph, each dot represents one repetition of the random assignment process where we are recording the difference in proportions of males between the two groups.

7. Press the **Randomize** button again. Was the difference in proportions of males the same this time?

8. Change the number of replications from 1 to 198 (for 200 total), uncheck the **Animate** option, and press the **Randomize** button. The dotplot will display the difference between the two proportions of males for each of the 200 repetitions of the random assignment process. Where are these values centered?

9. Does random assignment *always* equally distribute/balance the males and females between the two groups? Is there a tendency for there to be a similar proportion of males in the two groups? Explain.

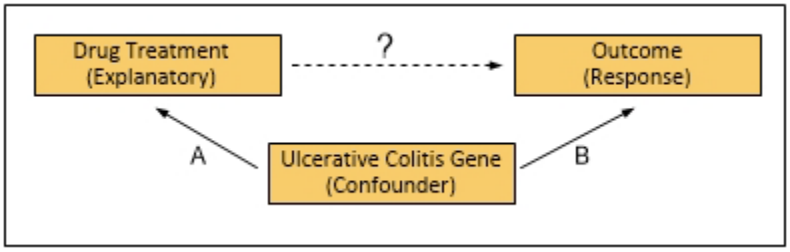
Definition
A comparative study is *balanced with respect to a possible confounding variable* if the distribution of the variable is the same for each group in the study.

10. Prior research has also shown that the likelihood of going into remission is related to variables such as body mass index (BMI), severity of the disease, location of the disease, and prior drug use, so we would like the random assignment to distribute these variables equally between the groups as well. In the applet, click on the **Reveal BMI?** button and then use the pull-down menu to switch from the sex-of-participant variable to the **BMI** variable. The dotplot now displays the differences in *average* BMI between Group 1 and Group 2 for these 200 repetitions. In the long run, does random assignment tend to equally distribute the BMI measurements between the two groups? Explain.

11. Ulcerative colitis does run in certain families. Because it appears heredity plays a role in the disease, suppose there is an "ulcerative colitis gene" that is related to people's ability to both get the disease and to the severity of the disease. We didn't know about this gene ahead of time, but if you select the **Reveal gene?** button and then select **gene** from the pull-down menu, the applet shows you this gene information for each subject and also how the proportions with the gene differ in the two groups. Does this variable tend to be equalized between the two groups in the long run? Explain.

Key idea
Randomly assigning experimental units to groups tends to balance out all other variables between the groups. Any variables that could have an effect on the response should be equalized between the two explanatory variable groups and therefore should not be confounding.

The diagram below shows the possible confounding effect if the study design had not been randomized. We want to know whether differences in the explanatory variable cause differences in the response (dotted arrow). However, in this case, a confounding variable (ulcerative colitis gene) is associated with both the explanatory and response variables, making it impossible to separate the effects of explanatory and confounding variables on the response.



12. How does random assignment remove the effect of potential confounding variables? Does it eliminate the association between the potential confounding variable and the explanatory variable? Between the potential confounding variable and the response variable? Both? Similarly, what arrows in the diagram above will be eliminated through random assignment, A, B, or both?

13. Suppose this study finds a statistically significant difference between the two treatment groups. What conclusion would you draw? Can you conclude there is a cause-and-effect relationship? For what population? What additional information would you need to know?

It was mentioned earlier that one group of subjects took a placebo (an inert pill) instead of the medicine. A "placebo effect" has been found in numerous studies. This is when subjects are told something good is going to happen, they often have a positive response even if nothing is actually done to them. For this reason, subjects are often kept "blind" as to which treatment group they are placed in, for example by giving one group the actual treatment and the other group a "fake" treatment, called a placebo. Placebo treatments have even been used in studies about knee surgery!

14. Explain why it is also important that the person responsible for determining whether or not the patient is in remission should be "blind" to which treatment the patient received.

Definition
In ***double-blind*** studies, neither the subjects nor those evaluating the response variable know which treatment group the subject is in.

15. Many research papers where random assignment is used will present a table, similar to the one shown below (adapted from Sandborn WJ et al., 2012) that gives baseline demographics and clinical characteristics of the subjects after they were randomly assigned to the two groups but *before* the treatment began. These tables will often have an additional column of p-values comparing these characteristics. Do the researchers hope these p-values are large (greater than 0.05) or small (less than 0.05)? Explain.

Parameter	**Placebo** (*n* = 121)	**Budesonide** (*n* = 123)
Mean age, years (range)	39 (18–77)	42 (19–68)
Median number of flares in past 2 years	2	2
Median baseline UCDAI score (activity index for the disease)	7.0	7.0
	Count (%)	**Count (%)**
Male	68 (56.2)	77 (62.6)
Female	53 (43.8)	46 (37.4)
Mild severity of last flare	30 (24.8)	31 (25.2)
Moderate severity of last flare	79 (65.3)	82 (66.7)
Proctosigmoiditis (spread of disease beyond the colon)	47 (33.9)	34 (27.6)
Left-sided colitis (location in the colon)	34 (28.1)	32 (26.0)
Prior mesalamine use (non-steroid anti-inflammatory drug)	74 (62.2)	58 (47.2)

Exploring Further: Blocking

Recall at the beginning of the exploration when we explicitly controlled the number of females to be the same in each group and the number of males to be the same in each group by randomly assigning 4 females to each treatment group and then randomly assigning 8 males to each treatment group. This is related to a technique called ***blocking***.

Definition

A ***block study design*** creates ***blocks*** of experimental units that are similar to each other, randomly assigns the treatments within each block, and then analyzes the data in a way which accounts for block-to-block variations. When there are only two treatment groups being compared, a block study design is called a *matched pairs design*. The term block comes from the first block designs which were agricultural experiments in large fields where separate parts of the field were called "blocks."

16. In the **Randomizing Subjects** applet, click on the **Reset** button at the bottom and then select the **Randomized block** option on the left. Leave **Animate** checked and make sure that **sex** is the variable selected. With the number of **Repetitions** set as 1, click on the **Randomize** button.

    a. What do you notice about the animated shuffling of cards to the two treatments?

    b. Are the proportions of males the same in each group? Are the proportions of females the same in each group?

    c. Uncheck **Animate** and do 999 more random shuffles. Describe the shape, center, and variability of the graph displaying the differences in the proportions of males in each group for each of the 1,000 random shuffles. Why does this graph look like it does?

17. We can see that blocking on sex has completely controlled sex as a confounding variable, that is to say that there will always be the same proportions of males in each treatment and of females in each treatment, so sex could never account for any of the differences in the treatment groups. Might there be added benefits to controlling confounding variables when blocking is used?

    a. Go back to the applet and make sure the **Reveal BMI?** button is on and use the pull-down menu to select the variable **BMI**. Report the SD of the distribution of differences in average BMIs between the two groups.

    b. Now select the **Completely Randomized** button to approximate the SD of the distribution of differences in average BMIs when the proportion of males in the treatment groups could differ by chance. How does this SD compare to the SD of the distribution of differences in average BMI when blocking on the variable sex?

Think about it

Why is the variability in the differences in average BMIs reduced in the block design compared to the completely randomized design?

When other variables (like BMI) are associated with the blocking variable, we see the added benefit that these variables are more evenly distributed across the treatments as well. This will not be the case if the other variable (like gene) is not associated with the blocking variable. Check it out in the applet!

Reference

Sandborn WJ, Travis S, Moro L, Jones R, Gautille T, Bagin R, Huang M, Yeung P, Ballard ED 2nd Once-daily budesonide MMX® extended-release tablets induce remission in patients with mild to moderate ulcerative colitis: results from the CORE I study. *Gastroenterology* 2012 Nov;143(5):1218-26