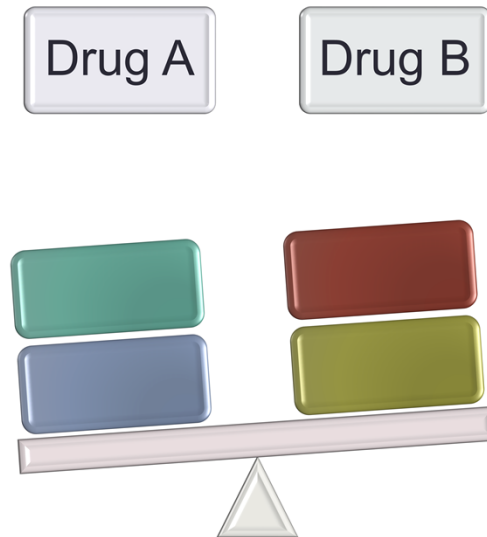# CAUSATION AND OBSERVATIONAL STUDIES

Unit 2: Producing Data

In the case of observational studies, we must be extremely careful about lurking variables, especially when we wish to show a causal relationship between our explanatory variable and the response variable.

Suppose we conduct an observational study in which we are interested in comparing two drugs (A and B) that are currently on the market for over-the-counter use.

Further suppose we are able to select a perfect random sample and get everyone to selected to participate.

We collect our data and begin to compare the results of the outcome between the two drugs, A and B.

In our data analysis, we find that individuals using drug A report better results than individuals using drug B.

Hopefully our discussion on lurking variables makes you immediately question any causal link between drug and response.

Maybe it is the case that drug A is better but here is another possibility.
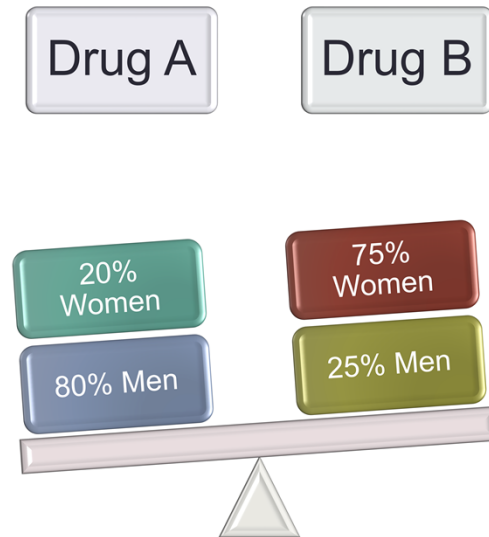
Marketing campaigns for the two drugs tend to target different genders – Drug A targets men most often and Drug B targets women.

In our observational study – suppose we look at the results carefully and we see that 80% of those using Drug A are men but only 25% of those using Drug B are men.

What is the potential problem with this study?

The answer: Without accounting for gender, what we see is the combined effect of gender and drug on the results. We cannot be certain – without further analysis – whether the result we see is due to Drug or Gender or some combination.

The variables of gender and drug are confounded in this analysis!

To investigate – we can stratify our analysis and look at the effect of drug within each gender.
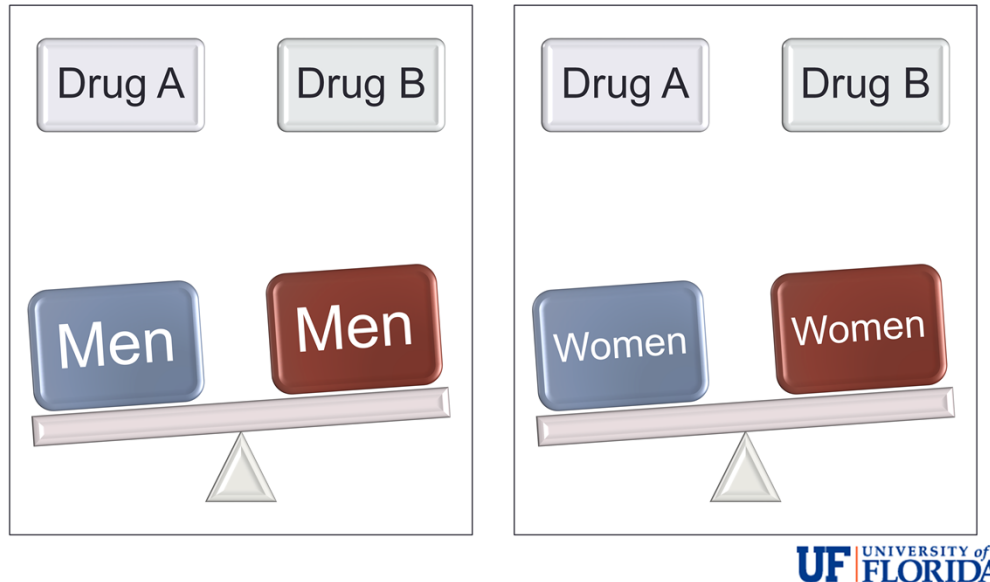
If when we split the data by gender and conduct an appropriate analysis and we still see that Drug A is better for both groups – we have moved a step closer to proving that Drug A is better.

But what about other variables we haven't considered? There are often many potential lurking variables in any one problem.

CAUSATION AND OBSERVATIONAL STUDIES

Unit 2: Producing Data

UF UNIVERSITY of FLORIDA
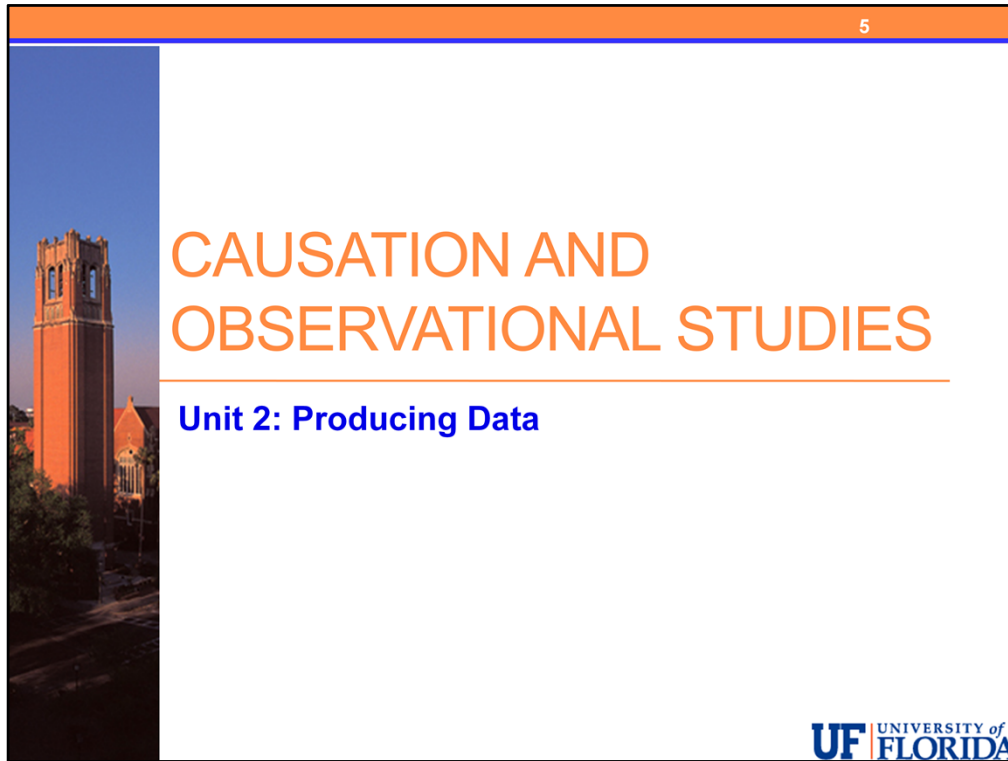
The key to establishing causation is to rule out the possibility of any lurking variable.

To do this we need to ensure that individuals differ only with respect to the values of the explanatory variable.

This is extremely difficult, if not impossible to accomplish in an observational study.

Given what we will learn in this course – our only possible recourse is to STRATIFY our analysis by the potential lurking variable and to investigate if the relationship we see between the explanatory and response variable still holds true within the levels of the lurking or confounding variable.