March 2021
Joy Shi

DAG Exercises

A Selecting Controls in Case-Control Studies

Suppose you are interested in the effect of environmental tobacco smoke on lung cancer incidence among people who have never smoked. To assess environmental tobacco smoke, you ask the participants in your study whether or not their partner is a smoker.

To conduct this study, you decide to use a case-control design. Cases were identified using a local lung cancer registry, whereas controls were selected from people who attended local fitness and recreational centers. Assume that people who attended these fitness and recreational centers were more likely to live in densely populated neighborhoods. Furthermore, people who live in densely populated neighborhoods are more likely to have partners who smoke. Assume that population density of the neighborhood is otherwise unrelated to lung cancer incidence.

1. Based on this scenario, draw a DAG for the effect of having a partner who smokes on lung cancer incidence.
2. Identify the open paths from the exposure (having a partner who smokes) and the outcome (lung cancer incidence).
3. Are these paths causal or non-causal? Is the marginal association between having a partner who smokes and lung cancer incidence an unbiased estimate of the causal effect?
4. If you said no in question 3, what steps could you take in the analysis of the case-control study data to block these non-causal paths.

B Birth Weight Paradox

Observational studies have shown that smoking among pregnant people is positively associated with higher infant mortality. However, when these studies adjust for infant low birth weight (an established predictor of infant mortality), the "birth weight paradox" arises: among low birth weight infants, smoking during pregnancy appears to be *protective* against infant mortality. Let's draw a causal diagram to discern what is going on. The following steps will help guide you through drawing this DAG:

1. Start by including the exposure (smoking during pregnancy) and the outcome (infant mortality) to the causal diagram. Given the variables currently in your DAG, add any relevant arrows.
2. Add infant low birth weight to the causal diagram. Consider the timing of this variable (i.e. does it occur before or after the exposure? Before or after the outcome?) Add any relevant arrows to your DAG.
3. Do you think there are any potential confounders for the relationship between infant low birth weight and infant mortality relationship? If so, add this to the DAG. *(Note: you don't have to come up with a specific variable - you can notate this using the letter U)*
4. Recall that observational studies assess the association between smoking during pregnancy and infant mortality among low birth weight infants. Make sure this is appropriately reflected in your DAG.
5. Does the association between smoking during pregnancy and infant mortality reflect a causal relationship? Why or why not?

Joy Shi