

Experiments and Observational Studies

Goal: identify causal associations

- Does changing the value of the **explanatory variable(s)** *cause* the value of the **response variable** to change?
 - Explanatory variables: whether or not a child eats peanuts, whether or not a child takes medicines
 - Response variable: whether or not the child develops a peanut allergy

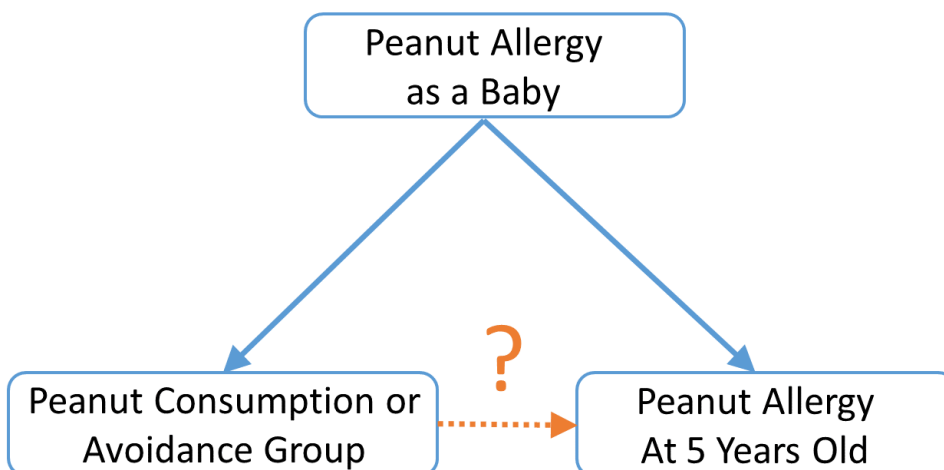
2 Types of Studies

1. **Experiments:** The explanatory variable(s) are controlled by the researcher (the researcher randomly assigns the value of the explanatory variable to each subject).
 - Researcher assigns babies to either consume or avoid peanuts, and to either take or not take medicine
 2. **Observational studies:** the explanatory variable(s) are not manipulated or controlled by the researcher.
 - Babies either end up consuming peanut or avoiding peanuts, taking medicine or not, for reasons outside of the researcher's control.
- To demonstrate a **causal** relationship, need to run an **experiment**.

Confounding

- A **confounder** is a variable that is associated with both the explanatory variable and the response variable, but is not in the 'causal pathway'.

Example: Suppose we let families choose whether their child is in the peanut consumption group or the peanut avoidance group...



Factors and Treatments

- **Factor:** an explanatory variable under the researcher's control (different from R's use of the term factor for any categorical variable!)
 - Factor 1: Peanut consumption (2 Levels: Yes, No)
 - Factor 2: Allergy Medication (2 Levels: No Medication, Medication)
- **Treatment:** A combination of levels of factors for a given subject
 - 4 treatments: (1) Peanut consumption, No medication (2) Peanut consumption, Medication (3) Peanut avoidance, No medication (4) Peanut avoidance, Medication

Four Principles of Experimental Design

Goal: Prevent/Limit Confounding

- **Control:** Control sources of variation other than the factors we are testing by making conditions as similar as possible for all treatment groups.
 - Make sure that there are no other major differences in diet between the groups, other than peanut consumption
- **Randomization:** Subjects/experimental units are assigned to treatments at random to equalize the effects of unknown or uncontrollable sources of variation.
- **Replication:**
 - Within Experiment: Each treatment is applied to more than one subject/experimental unit.
 - Across Experiments: Re-do the entire experiment under different settings
- **Blocking:** Group together subjects/experimental units that are similar in important ways that you cannot control, then randomize the assignment of treatments within each of these groups, or blocks.
 - Block according to initial allergy test results.

What to do in Observational Studies??

We can't do control or randomization or blocking!!

- **Matching:** Find study participants who:
 - have different levels of the explanatory variables of interest, but
 - are similar in ways that are not directly being studied
 - compare values of the response variable between these matched participants.
- Closest we can get to blocking and randomization in an observational study.