

Stat 204

Quiz 1 1/21/20 Name: _____

Table 1: The randomized controlled double-blind experiment

	Size	Rate
Treatment	200,000	28
Control	200,000	71
No consent	350,000	46

Table 2: The NFIP study

	Size	Rate
Treatment	225,000	25
Control	725,000	54
No consent	125,000	44

1. The tables above summarize the results of the 1954 Salk vaccine trial. The rates (cases per 100,000 children) for the treatment and no consent groups were about the same, but the polio rates for the control groups were quite different. Explain.

The NFIP control group consisted of 1st and 3rd graders whose parents might or might not have consented to treatment, while the other control group consisted of 2nd graders whose parents did consent. Students whose parents would not consent to treatment were less susceptible to polio than those who would have gotten consent, so the mix of students in the NFIP control group were less susceptible to polio overall than the control group in the randomized controlled trial.

2. About 60% of a child's growth hormone is secreted during sleep, so it is believed that a lack of sleep in children might stunt growth.
 - (a) Identify the explanatory variable and the response variable.

The amount of sleep is the explanatory variable; the child's growth is the response variable.

- (b) Would an experiment or observational study be used to investigate this association? Why?

This would have to be investigated with an *observational study*, as it would be profoundly unethical to deprive children of sleep to understand its effect on growth!

3. 40 insomniacs were randomly divided into two groups; 20 received a 1-hour cognitive behavioral therapy session, while the other 20 received no treatment. Three months later, 14 of those in the treatment group reported sleep improvements, and 3 people from the other group reported improvements.

(a) What are the cases in this study?

The cases are the 40 insomniacs.

(b) What are the variables? Identify each as categorical or quantitative.

For each case, the experimental group (*treatment* or *control*, i.e., *therapy* or *no therapy*) is identified and whether sleep improved or not is recorded. Both of these variables are categorical and binary.

(c) How many rows and columns would the corresponding dataset have? What would they contain?

There would be 40 rows, one per case, and 2 columns, one for each variable.