

Chapter 12: Reading Biostatistical Research Part I - Experiments

Reading Experimental Studies



- Identify the **response variable(s)** being targeted in the study.
 - Is there just one outcome variable of interest, or several that the researchers are using to compare treatment factors?
 - Is the response variable categorical or numeric?
 - If numeric, the researchers are likely comparing means/medians
 - If categorical, the researchers are likely comparing proportions/risk
- Identify the **causal factor(s)** being explored in the study.
 - What are the various **treatment/control levels** being compared in this study? *These are typically categorical in experiments, but be on the lookout for the possibility of it being numeric!*
 - Is there more than one causal variable being considered? Are the researchers studying an **interaction effect** between these predictors?
 - For example: how does the fertilizer and sunlight level affect the growth of this plant? Does the effectiveness of the fertilizer increase when sunlight levels are higher?
- What **kind** of experiment did they run?
 - Is this a randomized design? (it usually will be!)
 - Are there any other experimental features, such as cross-over trials or repeated measures?
- What **population** is this study trying to generalize to, and what **sample** have they used to represent this population?
 - Is the population all people with a particular disease/condition? Only those with a mild form or severe form? All ages or select ages?
 - How large is the sample? What are the sampling units? (e.g., People, Bacteria cultures, Plots of land, etc.)
- What did the researchers **find**, and **what statistical evidence** are they using to make these claims?
 - For categorical response variables (e.g., having or not having a disease), researchers may report a p-value from Z-tests or confidence intervals for a difference in absolute risk, relative risk, or odds ratios.
 - In some cases, researchers may report a hazard ratio (and Kaplan-meier plot) when comparing time until an adverse event occurs.
 - For numeric response variables (e.g., time to recovery), researchers will commonly report a p-value from a z-test or t-test for independent sample means, or confidence intervals for a difference in means. *In non-parametric tests, it may be a difference in medians!*
- Any threats to the **internal validity** (causality argument) being made?
 - Threats to group selection, setting differences, timing differences, test familiarity, attrition differences, or independence of the units in each group
- Any threats to the **external validity** (generalizability argument) being made?
 - Threats to participant selection, setting, or historical sustainability?

Letermovir vs Valganciclovir for Prophylaxis of Cytomegalovirus in High-Risk Kidney Transplant Recipients

<https://jamanetwork.com/journals/jama/fullarticle/2805945>



1. What is the motivation for this study?

2. What is the response variable in this study? Is it a numeric or categorical outcome? Are there any other response outcomes that were used?

3. What is the primary treatment factor being examined in this study? What experimental levels are being compared?

4. Is this an experiment? What kind of experiment is it, or what experimental features do you recognize?

5. Describe the population of interest and describe the sample being examined

6. What did they find, and what statistical evidence do they use to support those claims?

7. In the methods, the paper states: "All study drugs had a matching placebo. Participants, investigators, study staff, and sponsor personnel involved in study drug administration and clinical evaluation were masked to study drug assignments." What does that mean, and is that an advantage or limitation to this study?

Chapter 12 Additional Practice

Effects of Brief Sodium Fluoride Treatments on the Growth of Early and Mature Cariogenic Biofilms

<https://www.nature.com/articles/s41598-021-97905-0>

1. 2. What is the primary treatment factor being examined in this study? What experimental levels are being compared?
2. Identify at least 5 different response variables that the researcher is examining. *It might be clearest to identify by looking through the figures in the paper!*
3. Describe the sample being examined in this study and identify the unit of observation.
4. What is the population that this researcher is (ideally) trying to represent with this sample?
5. What did the researcher find, and what statistical evidence is used to support those claims? *You will need to scroll down into the findings to see what statistical analysis was done!*

Chapter 12 Learning Goals

After this chapter, you should be able to...

- Read a structured abstract and visualizations/tables from a biomedical research paper documenting an experiment and identify the following:
 - The response variable(s) (and variable type(s))
 - The causal factor(s) being examined, and the various treatment/control levels that the researchers compared
 - The type of experiment used and other experimental features the researchers designed into the study (e.g., repeated measures, cross-over trials, pre-randomization blocking)
 - The population being generalized to and details about the sample used
 - The main findings and the statistical evidence supporting those findings
- Based on selected passages from the article...
 - Identify any threats to the internal validity (causality argument)
 - Identify any threats to the external validity (generalizability argument)
- Think about the response variable measure in context to assess the practical significance of a finding
 - For example, recognizing low p-values as an insufficient measure for determining whether a difference is particularly large or meaningful.