Learning Goals:

- Distinguish statistical significance from practical importance.
- From a description of a study, evaluate whether the conclusion of a hypothesis test is reasonable.

Previously, we discussed Type I and Type II errors, which arise naturally when we are randomly sampling and setting significance levels. There is always the possibility that a Type I or Type II error has occurred and statisticians can estimate the probability of each type of error.

But there are other more serious problems with hypothesis testing that are also more prevalent. In this activity we will examine a variety of real studies that highlight potential problems with statistical studies.

1) The following is an excerpt from *Statistics* by Freedman, Pisani, and Purves.

   In some cases of cirrhosis of the liver, the patient may start to hemorrhage and bleed to death. One treatment involves surgery to redirect the flow of blood through a portacaval shunt. The operation to create the shunt is long and hazardous. Do the benefits outweigh the risks?

   In 51 studies of the portacaval shunt, 34 (66%) reported results that were characterized as "strongly enthusiastic" about the shunt. Only 6 studies (12%) were characterized as "no enthusiasm" for the shunt.

   Here is an analysis of these studies as published in *Gastroenterology*

	Degree of enthusiasm for the shunt by researchers		
	Strong	Moderate	None
No controls	24	7	1
Controls, but not randomized	10	3	2
Randomized controlled	0	1	3

   a) What is the purpose of a control group in an experiment? Give an example of a control group that could be used in a portacaval shunt experiment.

   b) What is the purpose of randomization in an experiment?

   c) Based on these studies, do you think that doctors should use portacaval shunts? Why or why not?

2) Researchers performed a randomized controlled double blind experiment to investigate the use of cholestyramine in reducing blood cholesterol levels and

thus reducing heart attacks. There were 3,806 subjects, who were all middle aged men at high risk of heart attack; 1906 were chosen at random for the treatment group (cholestyramine) and the remaining 1900 were assigned to a control group (placebo).
Source: "The Lipid Research Clinics Primary Prevention Trial Results,"Journal of the American Medical Associations vol.251 (1984) pp 351-64.

The subjects were followed for 7 years. The drug reduced cholesterol levels by about 8% . Futhermore, there was a lower incidence of heart attacks in the treatment group (155 or 8.1%) vs. the control group (187 or 9.8%).

This 1.7% difference was statistically significant with a P-value of about 0.035 for a one-tailed test. The article was published in the *Journal of American Medical Association,* whose editors are quite strict about using significance levels of 0.05.

a) Would the experiment's results have been significant if a two-tailed test was conducted? How do you know?

b) The researchers said this was "strong evidence" that cholestryamine reduces heart attacks by reducing cholesterol levels. Why might a skeptical reader think that the researchers are overstating their results?

c) Cholestryamine can have the following side effects: constipation, stomach/abdominal pain, gas, nausea, and vomiting. The cost per year is approximately $1,350. If you were at high risk of a heart attack and also had high cholesterol, would you take cholestryamine based on what you know at this point? Why or why not?

3) This is a true story … In a meeting at a community college on the impact of tutoring, a researcher said that tutoring had a statistically significant impact on

improving GPA. His sample size was over 2,000 students. A math professor in the audience asked how much GPAs improved. The response: 0.05 GPA points.

a) How can such a small increase in GPA be statistically significant?

b) Should the college increase tutoring funds, at the expense of funding other programs, based on these results? Why or why not?

4) Read through these summaries (footnotes on the next page)

- A 1996 study on the effects of nicotine on cognitive performance revealed that findings of nicotine or smoking improving performance were more likely to be published by scientists who acknowledged tobacco industry support.[1]
- A 2003 study of published research on antidepressants found that studies sponsored by manufacturers of selective serotonin reuptake inhibitors (SSRI) and newer antidepressants tended to favor their products over alternatives when compared to non-industry-funded studies. Also, modelling studies funded by industry were more favorable to industry than studies funded by non-industry sponsors.[2] In general, studies funded by drug companies are four times more likely to favor the drug under trial than studies funded by other sponsors. [3]
- A 2006 review of experimental studies examining the health effects of cell phone use found that studies funded exclusively by industry were least likely to report a statistically significant result.[4]
- The US Food and Drug Administration (FDA) determined in 2008 that the bisphenol A (BPA) in plastic containers is safe when leeched into food, citing chemical industry studies. Independent research studies reached different conclusions[5] with over 90 percent of them finding health effects from low doses of BPA.[6]
- Two opposing commercial sponsors can be at odds with the published findings of research they sponsor. A 2008 Duke University study on rats, funded by the Sugar Association, found adverse effects of consuming the artificial sweetener Splenda. The manufacturer, Johnson & Johnson subsidiary McNeil Nutritionals LLC, responded by sponsoring its own team of experts to refute the study.[7]
- A 2012 analysis of outcomes of studies pertaining to drugs and medical devices revealed that manufacturing company sponsorship "leads to more favorable results and conclusions than sponsorship by other sources."[8]

a) What is the point being made by these reviews of statistical studies?

b) Your instructor will assign your group to read one or more sections of Hilda Bastin's article, *'They would say that, wouldn't they?' A reader's guide to*

author and sponsor biases in clinical research.

Source: Hilda Bastian (December 2006). "'They would say that, wouldn't they?' A reader's guide to author and sponsor biases in clinical research". J R Soc Med. ***99*** *(12): 611–614. doi:10.1258/jrsm.99.12.611. PMC 1676333. PMID 17139062.*

Prepare a summary of your section for class discussion.

1. *Christina Turner; George J Spilich (1997). "Research into smoking or nicotine and human cognitive performance: does the source of funding make a difference?". Addiction.* ***92*** *(11): 1423–1426. doi:10.1111/j.1360-0443.1997.tb02863.x. PMID 9519485.*
2. *C. Bruce Baker; Michael T. Johnsrud; M. Lynn Crismon; Robert A. Rosenheck; Scott W. Woods (2003). "Quantitative analysis of sponsorship bias in economic studies of antidepressants". The British Journal of Psychiatry.* ***183*** *(6): 498–506. doi:10.1192/bjp.183.6.498.PMID 14645020.*
3. *Becker-Brüser W (2010). "Research in the pharmaceutical industry cannot be objective". Z Evid Fortbild Qual Gesundhwes.* ***104*** *(3): 183–9. PMID 20608245.*
4. *Anke Huss; Matthias Egger; Kerstin Hug; Karin Huwiler-Müntener; Martin Röösli (2006-09-15). "Source of Funding and Results of Studies of Health Effects of Mobile Phone Use: Systematic Review of Experimental Studies". Environmental Health Perspectives.* ***115*** *(1): 1–4.doi:10.1289/ehp.9149. PMC 1797826. PMID 17366811.*
5. *vom Saal FS, Myers JP (2008). "Bisphenol A and Risk of Metabolic Disorders". JAMA.* ***300*** *(11): 1353–5. doi:10.1001/jama.300.11.1353. PMID 18799451.*
6. *David Michaels (2008-07-15). "It's Not the Answers That Are Biased, It's the Questions". The Washington Post.*
7. *Stephen Daniells (2009-09-25). "Splenda study: Industry and academia respond". Foodnavigator.com.*
8. *Lundh, A; Sismondo, S; Lexchin, J; Busuioc, OA; Bero, L (Dec 12, 2012). "Industry sponsorship and research outcome.". The Cochrane database of systematic reviews.* ***12****: MR000033. doi:10.1002/14651858.mr000033.pub2. PMID 23235689.*