Class 9: Experiments and Observational Studies (Text: Section 3.1)

A statistical study can tell us

- What is going on (the characteristics of a population)
- Effect of an intervention (taking a drug, using a new fertilizer, implementing a policy to create jobs)

Studies may be of two types:

- **Observational study**: Characteristics of a population are observed, but the population is not disturbed.
- **Experiment**: Treatment is given to a individuals and their responses recorded.

In an experiment, the treatment is generally applied to **randomly selected** subset of the population, and the subset compared to a randomly selected **control group**. This is called a **RCT (randomized controlled trial).** In medicine, RCTs are the gold standard for clinical trials of new drugs.

Ex. To investigate the safety of the anesthetics used in surgery, records from 850,000 operations in 34 major hospitals were examined. The death rates for four common anesthetics are shown in the table.[1]

Anesthetic	A	B	C	D
Death rate	1.7%	1.7%	3.4%	1.7%

The anesthetic used is the **explanatory** variable and the death rate is the **response** variable.
Was this an observational study or an experiment? What can you conclude from it?
The study looked at records of previous operations, so it is an observational study. Anesthetic C looks worse, but we can't say that it caused the death. This anesthetic may be given to more risky cases. There may be confounding/hidden variables.

Ex: If you have to go to hospital, should you choose the one with the lowest mortality rates? (These rates are published.) No. The best hospitals may have the highest mortality rates.

Ex. The Physicians' Health Study (1980-1995) was a longitudinal study of the effect of aspirin and beta carotene on heart attacks.[2] In the study, 21,996 male physicians were randomly divided into four equal groups and given pills or placebos of aspirin and beta carotene (real versions of both, just one, or neither). The result was that 239 of the aspirin placebo group and 139 of the aspirin group had heart attacks.[3] Was this an observational study or an experiment? What can you conclude from it?
This is an experiment with random assignment to the four groups. The numbers of heart attacks in the two groups were sufficiently different that it allowed doctors to conclude that aspirin helps prevent heart attacks. The experiment established the effectiveness of aspirin in reducing heart attacks. The beta carotene, which the body converts into Vitamin A and which was thought might prevent some kinds of cancer, turned out not to have significant health effects.

Ex. A Sept 2006 article[4], "Vitamin D halves the risk of pancreatic cancer," reported on the work scientists at Northwestern and Harvard who examined data from two long-term health studies of 46,771 men and 75,427 women. They found that people who took vitamin D had a 43% lower risk of pancreatic cancer.
Comment. This is an observational study; it maybe that people who take vitamin D have less chance of getting pancreatic can cancer in any case. It is certainly suggestive; it merits an experiment.

1 From *Introduction to the Practice of Statistics*, by D. Moore and G. McCabe 4th edn.
2 Physicians' Heath Study II was launched in 1997 to study the effects of Vitamin E, C and multivitamins.
3 Moore and McCabe, *Introduction to the Practice of Statistics*, 4th edn. "Aspirin in the primary and secondary prevention of vascular disease" published in *The Lancet* on May 30, 2009 provided a more nuanced view of the effectiveness of aspirin.
4 By Halcyon Skinner et al in MedlinePlus at NIH http://www.hpnonline.com/dailyupdates/September_06.html

Why Are Experiments/RCTs Useful? Causation

Even if the sample is well chosen, an observational study cannot easily establish causation. In an observational study there may be **hidden** variables, which explain the effect observed, but are not among the variables studied. We say that two variables are **confounded** if their effects cannot be distinguished from each other. If well designed, experiments *can* provide evidence of causation.

Ex: Children with larger shoe sizes read better.
Age is hidden variable; older children read better. Shoe size is **confounded** with age.

Ex: Countries with more TV sets per person have longer life expectancy.
GDP is **hidden variable**; money enables people to buy TVs *and* to take better care of their health.

Ex. Oklahoma and West Virginia adopted state-funded programs promoting marriage. These are based on studies done in Oklahoma in 1998, from which[5]

> "state economists concluded that being single and being poor were interrelated".

The programs were aimed at poor single mothers:

> "The hope is that marriage will provide these women, and their children, with a route out of poverty and alleviate the burden they place on the state. The thesis is simple: If failure to marry––or divorce—means poverty, marriage must enhance wealth."

Are the 1998 studies observational studies or experiments? Can you conclude that being single causes poverty? Observational studies. Other observers "do not accept that lack of money is a consequence of lack of marriage". Being single and poverty are associated, but there is not necessarily a causal relationship. There may be hidden variables, or poverty may cause people to stay single. In other words, even if causation is involved, it may be in the wrong direction. (That is, reverse causation.)

Ex. On Sept 3, 2002, a New York Times article, "Sorting through the Confusion over Estrogen", reported

> "…..in the decades since its initial approval by the Food and Drug Administration in 1942, estrogen had acquired a reputation as an antidote to many of the illnesses and afflictions of aging. Scores of observational and case studies supported this view……Major new findings from such clinical trials have seriously challenged estrogen's image as a preventive of chronic disease…."

What might have caused the results to be so different?
The article says "….*there is always a chance that factors not considered* [in the observational studies] *could have influenced the results, especially since the women who choose to use hormones tend to have healthier habits over all and are likely to be followed more closely by their physicians."* The results of the observational studies may have been confounded with the effect of the patient's overall health. The clinical studies suggest that with the exception of a few illnesses, estrogen not an antidote to the illnesses of aging that it was once thought to be.

Multiple Regression is a primary tool for establishing causation. Also **natural experiments**, when available, as in the next example.

Ex: The Harlem Children's Zone is an NGO (non-governmental organization) providing social services for poor children in New York City. It runs a charter school, Promise Academy, which accepts students by lottery. A study compared students who went to the school to comparable students who went to other schools.[6] Students in the school gained 1.3 or 1.4 standard deviations (usual gains are about 0.2 SD.)
What kind of study is this? Because of random selection of students in the school, this is a RCT: It is a natural experiment.

5 From "Wedded to the Value of Marriage", BBC News, January 15, 2004.
6 "The Harlem Miracle", May 7, 2009, http://www.nytimes.com/2009/05/08/opinion/08brooks.html

Experimental Methods: Good Design

Experiments should be designed to avoid confounding and Hawthorne and experimenter effects. Confounding is made less likely by having a

- **Control group** and
- **Random assignment** to groups.

Ex: "Learning in an online format versus and in-class format: an experimental study"[7] reported that in Nova Southeastern University, FL, online students did as well as those who learned in class. What might you want to ask about the experiment before investing in online learning?
Was there a control group? Yes, the in-class group.
Were students assigned randomly to the two groups? No. So bias introduced by lack of randomization.
Were the two groups similar? No, the online group had better scores at the start of the course. The fact they only did as well as the in-class group was worrying.

Ex: A 2005 experiment[8] claimed that a drink containing pig whipworm eggs reduced the symptoms of Crohn's disease (a disease of the bowel). What might you ask?
Was there a control group*?* No. So the result could have been a placebo effect.
Was the reduction medically significant? That is, big enough it brought relief to the patients.

Experimenter and Hawthorne effects

The experimenter effect occurs when the experimenter influences the result of the experiment (by accident). The Hawthorne effect occurs when the subjects respond differently just because they are part of an experiment. Both can be eliminated by using a **placebo.**

Ex: Crohn's disease experiment: Give the control group a drink that looks the same.

To make sure the experimenters' or subjects' behavior or beliefs do not unconsciously affect the outcome, use **double blinding** (neither experimenter nor subject knows who is in the control group) or **single blinding** (one knows; the other does not).

Ex: The Physicians' Health Study experiment
Used random assignment of participants to groups, placebos, and double-blinding.

Ex: In 2007, doctors in Germany showed that people who ate a small amount of dark chocolate each day had lower blood pressure than those who ate white chocolate.[9] What kind of blinding was possible? This experiment was single (investigator) blinded. The subjects knew what color chocolate they were eating!

Ex: Are some "natural remedies" placebos? Possibly

7 Allan Schulman and Randi Sims, *T.H.E Journal* June1999.
8 R. Summers at al, "Trichuris suis therapy in Crohn's disease" in *Gut* v. 54, 2005.
9 www.medicinenet.com/script/main/art.asp?articlekey=82299 "Dark Chocolate May Help Blood Pressure" July 3, 2007

The MMR Vaccine Controversy

Children in several countries are routinely vaccinated against measles, mumps, and rubella by the MMR vaccine. As a result, measles was close to eradication. However, a 1998 paper[10] linking the MMR vaccine to autism caused a considerable reduction in vaccination levels, triggering new outbreaks of measles in the US, Europe,[11] Israel, Japan. The 1998 paper suggested the MMR vaccine led to bowl infections and to autism. What statistical questions would help you evaluate the results of the paper?[12]

- What was the research question? Not stated but would have been that the MMR vaccine led to an increase in bowl problems and autism.
- Was this an observational study or an experiment? Observational
- How big was the sample? 12 children; small sample.
- How was the sample chosen? Children referred for gastrointestinal problems and symptoms of autism—selected to have both symptoms.
- Was there a control group? No. When a randomized study is not possible, a matched sample, equivalent in every way except for the possible causal agent (here the MMR vaccine) can help establish causation.
- Was the study blinded? No; the experimenters knew which children had had the MMR vaccine. Parental memory used to establish date of vaccine and of onset of symptoms.
- How long was study? Short; children followed for days or week not years.
- How was the study funded? Study claimed funding was from the hospital where the experimenters worked, plus a children's charity. Later it was discovered the charity was defunct and additional funding came from a lawyer who was putting together a case against the MMR vaccine and in favor of single vaccines. The children and their parents were part of his legal case.
- What has later research shown?[13] No later studies have replicated Wakefield's results. Alarmingly, the level of vaccination has not returned to previous levels even though the result in the paper is now thought to be false.
- What legal action has been taken?[14] On Feb 12, 2009, the US federal Vaccine Injury Compensation Program (VICP) rules that there was no evidence of a link between autism and the MMR vaccine or thimerosal, a preservative used in vaccines. In Feb 2010, The Lancet fully retracted the article. In May 2010, Wakefield was struck from the medical register. He was working in Texas.
- British Medical Journal, 5 Jan 2011 article:[15]
  "Authored by Andrew Wakefield and 12 others, the paper's scientific limitations were clear when it appeared in 1998. As the ensuing vaccine scare took off, critics quickly pointed out that the paper was a small case series with no controls, linked three common conditions, and relied on parental recall and beliefs. Over the following decade, epidemiological studies consistently found no evidence of a link between the MMR vaccine and autism. By the time the paper was finally retracted 12 years later, after forensic dissection at the General Medical Council's (GMC) longest ever fitness to practise hearing, few people could deny that it was fatally flawed both scientifically and ethically. ***But it has taken the diligent scepticism of one man, standing outside medicine and science, to show that the paper was in fact an elaborate fraud."***

10 "Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children" by A J Wakefield, S H Murch, A Anthony, J Linnell, D M Casson, M Malik, M Berelowitz, A P Dhillon, M A Thomson, P Harvey, A Valentine, S E Davies, J A Walker-Smith in *The Lancet*, Vol 351, No. 9103 , 28 Feb 1998

11 http://news.bbc.co.uk/2/hi/health/7813149.stm and http://news.bbc.co.uk/2/hi/health/8293149.stm

12 http://briandeer.com/mmr/lancet-greenhalgh.htm and www.timesonline.co.uk/tol/life_and_style/health/article5683671.ece

13 www.nytimes.com/2008/08/24/opinion/24sun2.html and www.nytimes.com/2008/09/09/opinion/09tue3.html http://www.csicop.org/si/show/autism-vaccine_link_researcher_andrew_wakefield_accused_of_faking_his_data/

14 http://jolt.unc.edu/blog/2009/02/18/no-link-between-mmr-vaccine-and-autism and http://news.bbc.co.uk/2/hi/health/8695267.stm

15 http://www.bmj.com/content/342/bmj.c7452.full

Ex: In March 2009 in Minneapolis, 25% of the children in preschool classes for children with severe autism were Somali, while 6% of the school enrollment was Somali.[16] What can you conclude? "…public health experts say it is hard to tell whether the apparent surge of cases is an actual outbreak, with a cause that can be addressed, or just a statistical fluke." *New York Times,* March 17, 2009

Ex: On September 22 2009, a study[17] compared adult and child autism rates in UK. What might you expect to see? What light could this shed on the safety of the MMR vaccine? MMR vaccine has existed since 1990, so if it has an effect on autism, we'd expect autism prevalence to be higher in children than in adults. In fact the rates were the same—1% in both.

What is an Institutional Review Board? (IRB) Why were they created?

Tuskegee Study
"…. study that was conducted between 1932 and 1972 in Tuskegee, Alabama. That study included nearly 400 poor African-American men with preexisting syphilis whose disease was allowed to progress without treatment. Researchers did not infect the subjects, but they did not tell them they had the disease either."[18] The men thought they were receiving free health care from the US government, but they were not given treatment even after penicillin was discovered to be a cure for syphilis in the 1940s." As a result, some died of syphilis, wives got syphilis and children were born with congenital syphilis.
Study was brought to light and stopped in 1972.

Guatemala Study
"….a 1946-1948 research study in which people in Guatemala were intentionally infected with sexually transmitted diseases…… the U.S. Public Health Service Sexually Transmitted Disease Inoculation Study of 1946-1948, aimed at determining the effectiveness of penicillin in treating or preventing syphilis after subjects were exposed to the disease. Gonorrhea and chancres were also studied. Penicillin was a relatively new drug at the time. The tests were carried out on female commercial sex workers, prisoners in the national penitentiary, patients in the national mental hospital and soldiers. According to the study, more than 1,600 people were infected: 696 with syphilis, 772 with gonorrhea and 142 with chancres."[19]

An IRB must give approval to all research involving human subjects in its institution. The IRB considers:

- Informed consent
- No harm to participants
- Stopping study if a treatment appears to be effective and providing treatment to everyone.

Can an observational study require IRB supervision? Possibly yes. Traditionally observational studies have been considered less of a danger to participants. However, the information gained from the study may itself be a risk.
For example, an HIV study attempting to establish the prevalence of HIV in Tanzania was cancelled because Yale's IRB insisted that participants were informed if they were HIV positive. The government of Tanzania did not want this because they would be overwhelmed with requests for care that they could not provide.

16 http://www.nytimes.com/2009/03/17/health/17auti.html?pagewanted=1&_r=2&hp
17 "Autism Spectrum Disorders in adults living in households throughout England - report from the Adult Psychiatric Morbidity Survey 2007" Sept 22, 2009, The NHS Information Centre
18 http://edition.cnn.com/2010/WORLD/americas/10/01/us.guatemala.apology/index.html?iref=obinsite
19 http://edition.cnn.com/2010/WORLD/americas/10/01/us.guatemala.apology/index.html?iref=obinsite