

Selected Issues in Study Design

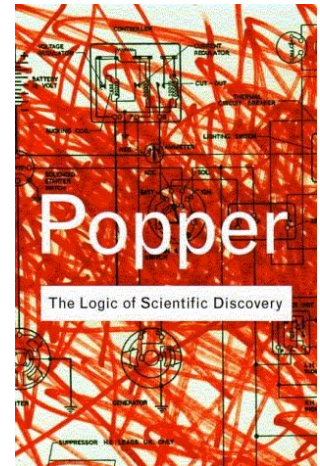
*Most problems in studies are due to poor
design (not poor analysis)*

The Research Question

When I came to practice I was looking for answers like everybody else. For years I asked "what's the right answer?" Now I am learning "What is the right question?"

- Science is the holding of multiple working hypotheses (Thomas Huxley)
- A study is only as good as its hypothesis
- But where do hypothesis come from?
observation + biological understanding + social
understanding + intuition → **causal hypothesis**

*Admittedly, creative action can never
be fully explained. (Popper)*



Hypothesis Refinement

- Research is an ongoing process of hypothesis generation, refutation, refinement, and corroboration
- Results from a single study are seldom definitive (or even clear)
- So how do you know whether a hypothesis is correct?
- Good scientific practice . . . places the emphasis on reasonable scientific judgment and the accumulation of evidence and not dogmatic insistence of the unique validity of a certain procedure (Jerome Cornfield cited in Vandenbroucke & de Craen, 2001)
- There is no such as “proof” (in the mathematical sense in science), but there is “proof” that it “works”:

When you ask people what made the modern West different from other cultures around the world, most of the answers are terribly negative: the disenchantment of the world, the destabilization of the earth, the death of God, the death of the Goddess, nightmare after nightmare. These naysayers tend to overlook the 40 years of life extension that the West has given us, the wonders of modern physics, modern medicine, the abolition of slavery, the rise of democracies, the rise of feminism, and so on. Until we honor both the good and bad news of modernity, we're not going to see our situation clearly. -- Ken Wilber

Beautiful Theory, Ugly Fact

Science is organized common sense where many a beautiful theory is killed by an ugly fact (Thomas Huxley)

- Our job is to draw conclusions based on “ugly fact”
- Illustrative example: “Whole language learning education theory”
 - Educational theorists long pushed the “whole language” approach to teaching reading and talked down the need for breaking words into basic sounds called “phonics.”
 - In 2000, a national panel reviewed ugly facts from 52 randomized studies.
 - Conclusion: no matter what the theory says, phonics is essential in teaching reading.



How do we create a study to gather ugly facts?

- There is no recipe for study design
- However, it helps to know
 - Elements of design
 - Where studies tend to go astray

Selected Elements of Study Design

- Measurement accuracy (variables)
- Effects can only be gauged relative to baseline (provided by a control group)
- Experimental studies differ from non-experimental studies (of course)
- The unit of recorded measure - individual or aggregate (ecological)
- Upstream and downstream causes should be considered
- Measurements may be longitudinal in individuals over time
- Cohort or case-control samples
- Hypothesis testing (“analytic”) or hypothesis generating (“descriptive”) studies
- Is the exposure randomized?
- Are groups comparable at baseline (confounding)
- Will you use prospective or retrospective measurements?
- Incident or prevalent cases?
- Matched or independent samples?
- Will you blind subjects and/or observers?
- Is the study based in an open- or closed-population?
- There are too many design elements to discuss in a single week. We can’t cover them all!

Objectives

- Review basic design aspects of of lab data sets
- Increase understanding of (inevitable) errors in studies

HS267 Variable Types

- Seek to understand and quantify relations between **explanatory variables** and **response variables**
- Classify variables as either **categorical** or **quantitative**
- How our curriculum applies:

Response	Explanatory	
	Categorical	Quantitative
Quantitative	11, 12, 13	14, 15
Categorical	16, 17	not covered

Comparative studies may be classified as:

- I. Experimental - investigator assigns an intervention to see if he or she can influence a response

- Randomized experiments**

- Non-randomized experiments**

- II. Observational – no investigator intervention *per se*

- Cohort**

- Case-Control**

- Cross-sectional**

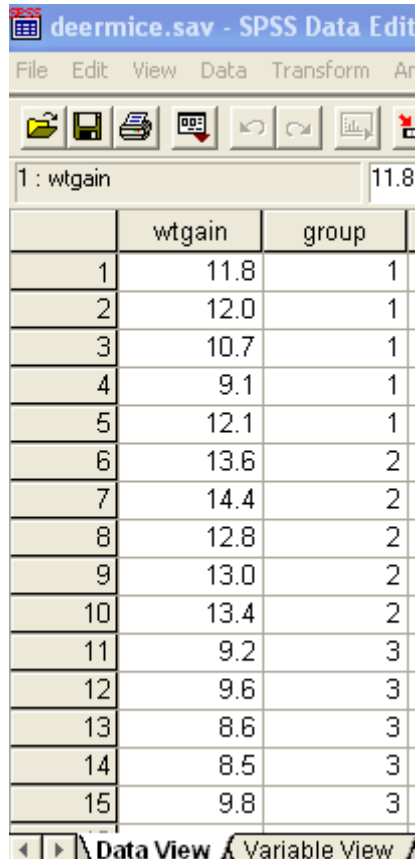
- Ecological**

Weight Gain on Different Diets

deermice.sav (Labs 2 & 3)

Explanatory variable = diet group (1=standard, 2=junk, 3=health)

Response variable = weight gain (grams)



	wtgain	group
1	11.8	1
2	12.0	1
3	10.7	1
4	9.1	1
5	12.1	1
6	13.6	2
7	14.4	2
8	12.8	2
9	13.0	2
10	13.4	2
11	9.2	3
12	9.6	3
13	8.6	3
14	8.5	3
15	9.8	3

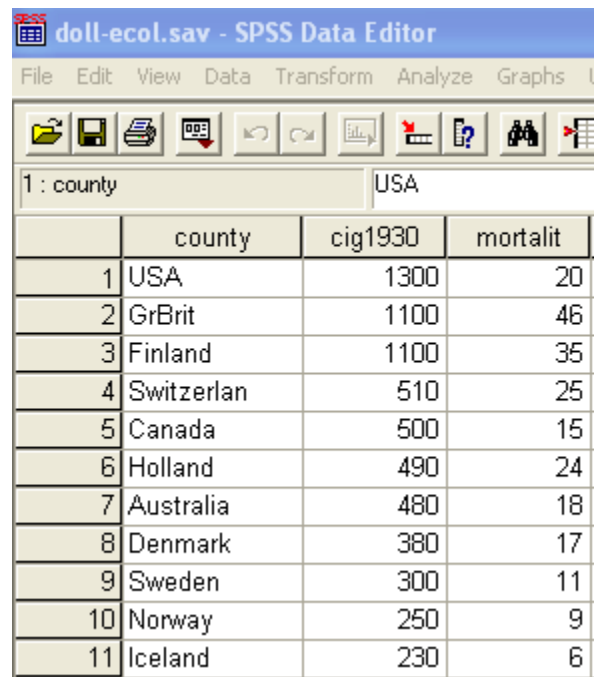
*Data are **experimental** because the investigator assigned the explanatory variable*

Cigarettes and Lung Cancer Mortality

doll-ecol.sav (Chap 12 and 13 labs)

Explanatory var = per capita cigarette consumption (cig1930)

Response var = lung cancer mortality per 100,000 (mortalit)



	county	cig1930	mortalit
1	USA	1300	20
2	GrBrit	1100	46
3	Finland	1100	35
4	Switzerlan	510	25
5	Canada	500	15
6	Holland	490	24
7	Australia	480	18
8	Denmark	380	17
9	Sweden	300	11
10	Norway	250	9
11	Iceland	230	6

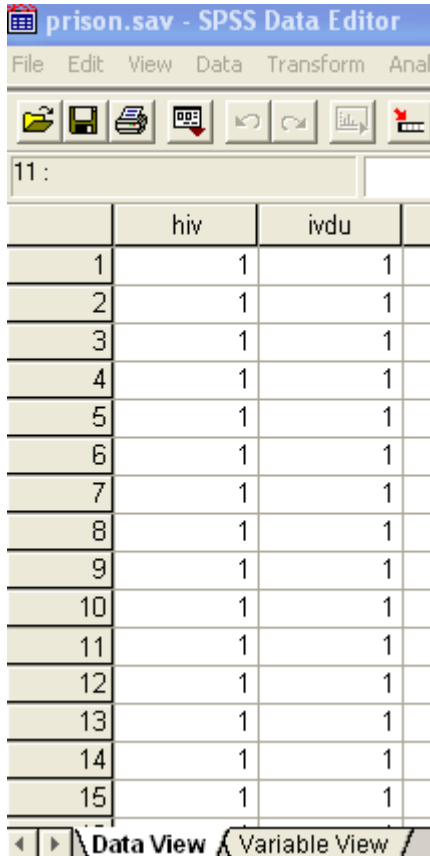
Data are observational with data on aggregate-level. This is an ecological study

HIV in a Women's Prison

Recall prison.sav (Chap 16 Lab)

Explanatory var = IV drug use (1 = users, 2 = non-user)

Response var = HIV serology (1 = positive, 2 = negative)



	hiv	ivdu	
1	1	1	
2	1	1	
3	1	1	
4	1	1	
5	1	1	
6	1	1	
7	1	1	
8	1	1	
9	1	1	
10	1	1	
11	1	1	
12	1	1	
13	1	1	
14	1	1	
15	1	1	

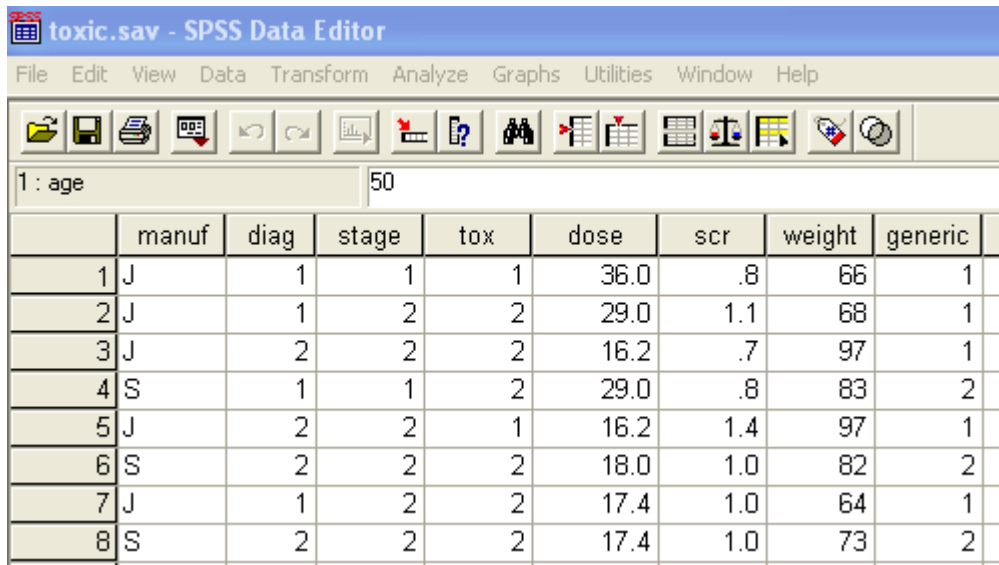
*Data are observational on the individual-level. But onset data cannot be unraveled. Thus, data are **cross-sectional***

Toxicity in Cancer Patients

toxic.sav (Chap 16 illustrative):

Explanatory variable = generic drug use (generic: 1 = yes, 2 = no)

Response variable = cerebellar toxicity (tox: 1 = yes, 2 = no)



	manuf	diag	stage	tox	dose	scr	weight	generic
1	J	1	1	1	36.0	.8	66	1
2	J	1	2	2	29.0	1.1	68	1
3	J	2	2	2	16.2	.7	97	1
4	S	1	1	2	29.0	.8	83	2
5	J	2	2	1	16.2	1.4	97	1
6	S	2	2	2	18.0	1.0	82	2
7	J	1	2	2	17.4	1.0	64	1
8	S	2	2	2	17.4	1.0	73	2

*Data are observational,
individual-level,
longitudinal, with all
individuals followed over
time. Thus, data are **cohort**.*

Comment: This is a *retrospective* cohort based on data abstracted data from medical records.

Esophageal Cancer and Alcohol Consumption

bd1.sav (Chap 17 illustrative)

Explanatory var = alcohol consumption (alc2: 1 = high, 2 = low)

Response var = esophageal cancer (case: 1 = case, 2 = control)

1: case 2

	case	agegrp	alc	tob	alc2	tob2
1	2	1	1	1	2	2
2	2	1	1	1	2	2
3	2	1	1	1	2	2
4	2	1	1	1	2	2
5	2	1	1	1	2	2

*Data are observational, individual-level, with study of all population cases but only a sample of non-cases. Thus, data are **case-control**.*

Error in Research

- All research has errors
- Two types of errors
 - Random error
 - Systematic error
- We will continue the lecture using slides from Chapter 12 in *Epi Kept Simple*