Observational versus Randomized Studies

Leg Injury in Motor Vehicle Collisions

Dischinger P.C. et al., Advan. of Automotive Med., 1992.

'In a study of trauma-center population in Maryland, it was found that:
(a) there was a higher incidence of lower extremity injury in frontal collisions,
(b) seatbelt use was not effective in preventing lower extremity fractures, and
(c) there was a higher incidence of lower extremity fracture among women.'

Dischinger P.C. et al., Accid. Analysis and Prev., 1995.

'For both men and women, results indicate an association between driver height and the incidence of lower extremity fractures. The incidence of injuries increased among shorter drivers, most of whom were women.'

Lurking Variable: Height  Gender, Leg injury

Removing kidney stones

	Success	< 2 cm	≥ 2 cm
Open surgery	**78%**	93%	73%
PCNL	**83%**	87%	69%

Percent of Sample

Open Surgery	25%	75%
PCNL	77%	23%

Simpson's Paradox:
Beware of comparing combined scores

Simpson's Paradox: (Comparing aggregate scores)

Basketball

	Shot %	FT	3-pt
Me	**.540**	.600	.300
You	**.460**	.700	.400

Percentage of population

Me	80%	20%
You	20%	80%

Gender Bias in Graduate Admissions

UC Berkeley Admitted:
44% of 8442 men, 35% of 4321 women

Q: Which major discriminated against women?

	Men		Women	
Major	Number of applicants	Percent admitted	Number of applicants	Percent admitted
A	825	62	108	82
B	560	63	25	68
C	325	37	593	34
D	417	33	375	35
E	191	28	393	24
F	373	6	341	7
:	:	:	:	:
:	:	:	:	:
All	8442	44	4321	35

A Common Cause

Driver Leg Injury Study

Men	Women

-Height: Not Balanced

Kidney Stones

Open Surgery	PCNL

Severity of disease: Not Balanced

Gender Bias in Graduate Admissions

Men	Women

Application rate to majors: Not Balanced

Q: Are there situations where we can (safely) jump to conclusions?
A: Yes.

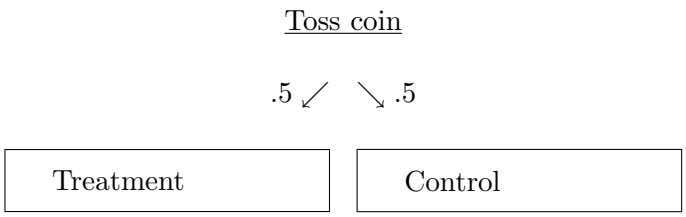
Soundbite:

The Treatment and Control groups should be the same in all aspects except for the treatment.

Silver Bullet: *Randomization*

Ex. Clinical study: "Lipitor lowers cholesterol."

Design: Take 1000 patients, randomize into Treatment and Control groups.



Age profile of the two groups? Diet? Exercise? Attitudes? Balanced!

Q: Why not do this for all studies?
A: Often cannot randomize. Group membership like gender or race or smoking is *observed* rather than *assigned*.

Observational studies
versus *Randomized Controlled* studies

Observational studies have a problem
⇓
Apples and Oranges

Men are different from women in other ways than the variable of interest (e.g. they have different heights, and they prefer different majors)

Smokers are different from nonsmokers in other ways besides smoking! (e.g. Smokers drink more coffee, stand out in the cold more, etc.) If average insurance claims for the two groups are different, is this due to smoking or diet or lifestyle?

Smoking effects are "confounded" with Diet effects, Lifestyle effects, etc. (a lot of lurking 3rd variables here).

Clofibrate Study

Clofibrate is a cholesterol lowering drug. The study looked at mortality rates of patients with heart disease over a period of 5 years.

	Clofibrate		Placebo	
	Number	Deaths	Number	Deaths
Total	1103	20%	2789	21%
Adherers	708	15%	1813	15%
Non-adherers	357	25%	882	28%

Conclusions?

1. Clofibrate is not effective
2. Adherers are different from non-adherers

Remember:

1. Anecdotal studies: No comparison of rates or averages
2. Observational studies: Allows comparison, but beware of confounded effects. (Epidemiological studies and social science research are typically observational.)
3. Randomized studies: Best way to avoid confounded effects, but not always possible. (This is the gold standard in clinical studies.)