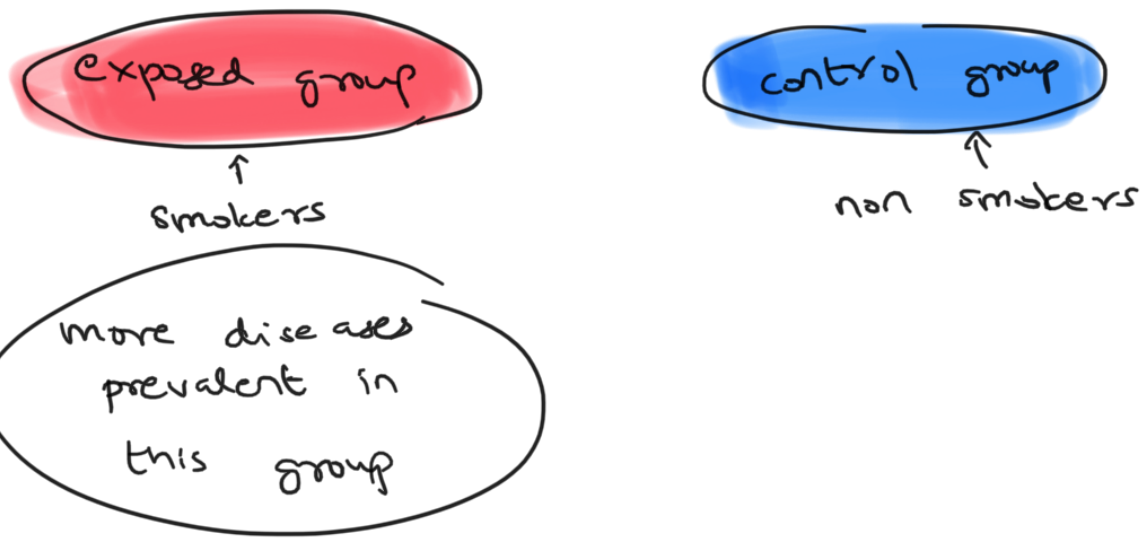
Day 2 – Observational studies

control: subject who didn't get treatment

controlled experiment: Investigators decide which group a person goes to

Observational studies - subjects assign themselves to their groups

(eg) studies on smoking



Conclusion

→ **association** bet smoking and disease.

→ [smoking ⇒ disease] **explains** the association

↑

Causality: one candidate explanation

circumstantial evidence only as might be Confounding factors

more observational studies run on proposed confunding factors → to make strong case that

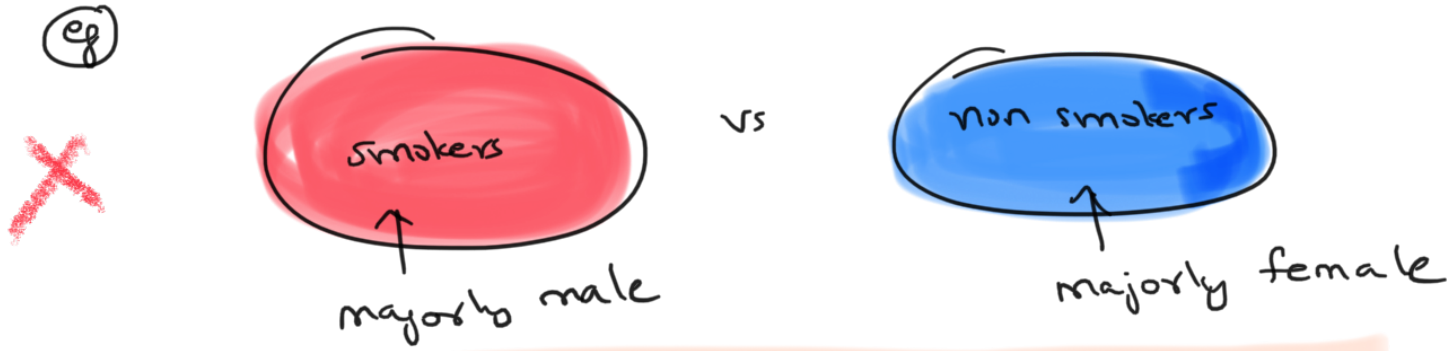
smoking ⇒ disease

To make observational studies better:

→ controlling for confunding factors in observational study needed

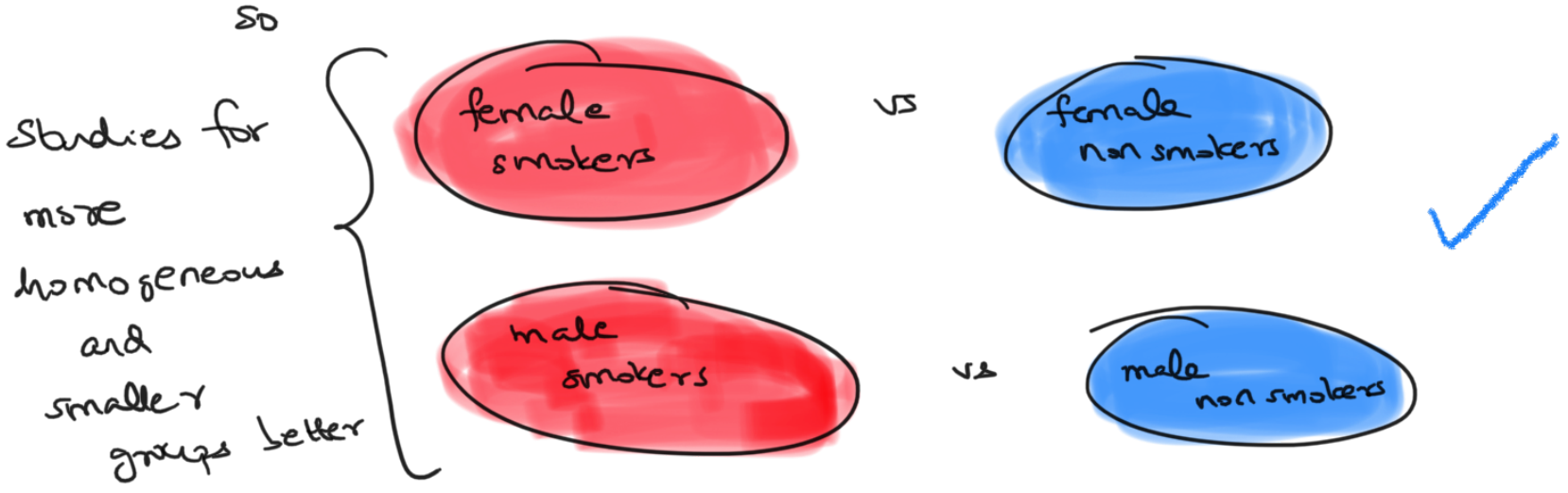
→ ie, make control group ~ exposed group modulo exposure

eg



@ gender difference → possible confunding factor

So



ⓑ age → another confunding factor

Clofibrate trial (observational study)

Goal: Evaluate five drugs for prevention of heart disease

Subjects – middle aged men with heart trouble

8341 people

→ 5552 drug group
one drug - clofibrate
reduce cholestrol in blood level

→ 2789 control group
(placebo - lactose)

After 5 years

20% of clofibrate group died

21% of control group died

Reason proposed: Subjects didn't take medicine (for placebo) regularly

People who did : adherers

	Clofibrate		Placebo	
	number	death	number	death
Adherers	708	15%	1813	15%
non adherers	357	25%	882	28%
Total	1103	20%	2789	21%

Conclusion: (1) drug was not useful

(2) adherers are "different" than non-adherers

perhaps more health conscious

more examples

① Pellagra disease in 18th 19th century

↳ flies (Simulium)

→ disease was there, flies were there

→ disease mostly in spring, flies were most active
developed in spring

etc

∞ infectious?

Any: 1914, Goldberger showed pellagra caused by poor diet. Flies were indicative of poverty, (little niacin)
not pellagra CAUSATION ≠ ASSOCIATION

② Ultrasound ⇒ lower weight in babies?

Any: Pregnancy issues ⇒ ultrasound
doctors suggest

↓

lower birth weights

Sex bias in grad admissions (observational study)

Applicants to grad school to Berkeley

8442 men → 44% admitted

4321 women → 35% admitted

University prefers men to women?

But each major → no bias against women!

(Some major → men, some major → women)
(if any bias, seemed to be against men)

How?

100 majors, 6 top majors → 1/3 of total applicants

same pattern for rest of majors

	men		women	
	Applicants	% admitted	Applicants	% 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
	T_M	P_M	T_W	P_W

easy to get into

harder to get into

more women bias?

more men bias

> 90% of women applied to harder to get into majors.

> 58% of men applied to easier to get into majors

Simpson's paradox

Relationships bet % in subgroups can be **reversed** when subgroups are combined

(major wise, looks like women > men
but over all men > women)
preferred
slight preference

weighted averages

earlier

men
% admitted :

T_m = total men

$$62\% \wedge \frac{825}{T_m} + 63\% \wedge \frac{560}{T_m} + 37\% \wedge \frac{325}{T_m} + 33\% \wedge \frac{417}{T_m} + 28\% \wedge \frac{191}{T_m} + 6\% \wedge \frac{373}{T_m}$$

earlier wt was
men admitted for that course / Total men applicants for that course

$\times 100$

weighted

$$.62 \times \frac{933}{4526} + .63 \times \frac{585}{4526} + .37 \times \frac{918}{4526} + .33 \times \frac{792}{4526} + .28 \times \frac{584}{4526} + 0.06 \times \frac{714}{4526}$$

now wt is applicants for course / total number of applicants

Confounding vs confounders

Confounding → diff bet treatment and control groups other than treatment

confounder → third variable associated with exposure and with disease

(eg) If gene → causes one to smoke AND → increases risk to cancer

is a confounder

If gene → only increases risk to cancer)

this doesn't influence smoking — cancer association

Confounder different from confounding factor