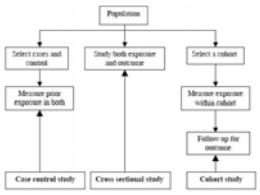
Epidemiology Kept Simple

Ch 11: Observational Study Designs

Observational Designs

- Cross-sectional and ecological (§11.2 & §11.3)
- Cohort (§11.4)
- Case-control (§11.5)



Ecological & Cross-Sectional

- *Cross-sectional* = *no* follow-up of individual experience over time
- Cross-sectional data on individuals = cross-sectional survey
- Cross-sectional data aggregates = ecological design

Example: Ecological Data

doll-ecol.sav - SPSS Data Editor

File Edit View Data Transform Analyze Graphs

1 : county USA

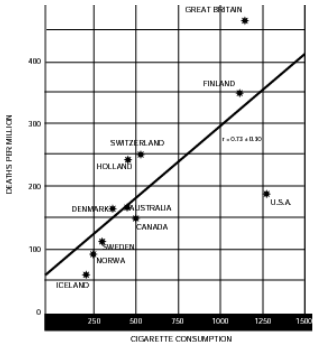
	county	cig1930	mortalit
1	USA	1300	20
2	GrBrit	1100	46
3	Finland	1100	35
4	Switzerlan	510	25
5	Canada	500	15
6	Holland	490	24
7	Australia	480	18
8	Denmark	380	17
9	Sweden	300	11
10	Norway	250	9
11	Iceland	230	6

Unit of observation = geographic region

Cig1930 = cigarettes per capita, 1930

Mortal = lung cancer mortality per 100,000 p-yrs, 1950

Ecological Data



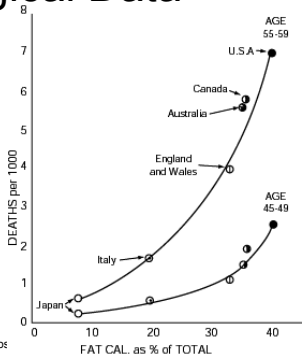
per capita cig consumption & lung cancer

r = 0.74

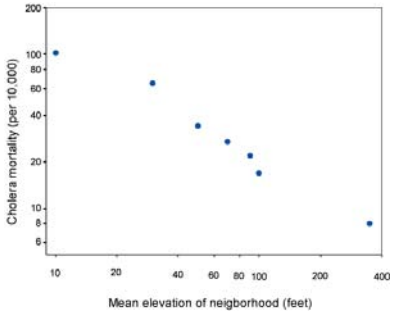
Ecological Data

Calories from fat and cardiovascular disease

Ecological design: no individual data; aggregate unit of observation



Farr's (1852) Faux Pas



Elevation & cholera mortality:

(supports miasma theory)

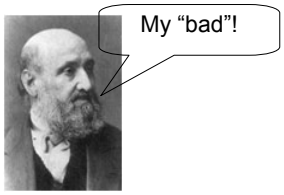
Farr's faux pas

Confounded!

Exposure = elevation

Disease = cholera

Confounder ≡ proximity to contaminated water sources



SES & Mental Disorders

Hollingshead & Redlich (1964)

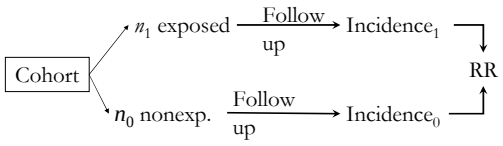
	Prevalence per 100,000	
Social class	Psychosis	Neurosis
High	188	349
Moderate	291	250
Low	518	114
Very low	1505	97

Biases in Hollingshead

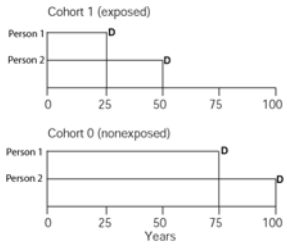
- *Detection bias*: different diagnostic practices ⇒ artificial differences
- *Reverse-causality bias*: "Disease" causes "exposure" (e.g., psychosis causes low SES)
- *Prevalence-incidence bias*: Difference in prevalence but not incidence (e.g., high SES have more persistent diagnoses)
- Limitations in cross-sectional studies stimulated development of better quality studies

§11.4 Cohort Studies

- Classify individual (exposed / nonexposed)
- Follow individual experience over time



Cohort: Simple Example

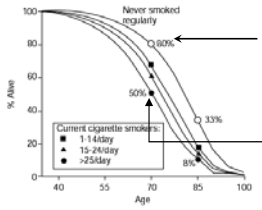


$$R_1 = \frac{2 \text{ deaths}}{75 \text{ person - years}} = 0.0267 \text{ deaths per p - yr}$$

$$R_0 = \frac{2 \text{ deaths}}{175 \text{ person - years}} = 0.0114 \text{ deaths per p - yr}$$

$$RR = \frac{R_1}{R_0} = \frac{.0267}{.0114} = 2.33$$

British Doctors Cohort



80% of nonsmoker survived to age 70

50% heavy smokers survived to 70

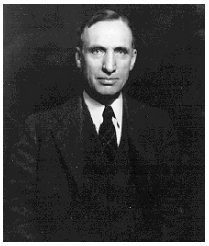
Source: Doll, R., Peto, R., Wheatley, K., Gray, R., & Sutherland, I. (1994). Mortality in relation to smoking: 40 years' observations on male British doctors. *British Medical Journal, 309*(6959), 901-911.

British Doctors Cohort

Cause of Death	No. of Deaths	Death Rates of Men Smoking a Daily Average of: Non-smokers	1-14 gms.	15-24 gms.	25+ gms.
Lung cancer	36	0.00	0.48	0.67	1.14
Other cancers	92	2.32	1.41	1.50	1.91
Respiratory diseases other than cancers	54	0.86	0.88	1.01	0.77
Coronary thrombosis	235	3.89	3.91	4.71	5.15
Other cardiovascular disease	126	2.23	2.07	1.58	2.78
Other diseases	247	4.27	4.67	3.91	4.52
All causes	789	13.61	13.42	13.38	16.30

Wade Hampton Frost

- First Professor of epidemiology in U.S.
- First Dean of US School of Public Health
- Bridged gap between infectious disease epi and chronic disease epi with TB studies (infectious disease with long latency)



1880 – 1938

Frost's TB Studies

TB morality per 100,000 p-yrs

Age	1880	1890	1900	1910	1920	1930
Males						
0-4	760	578	309	209	108	41
5-9	43	49	31	21	24	11
10-19	126	115	90	36	49	21
20-29	444	361	288	207	149	71
30-39	378	368	296	253	164	115
40-49	364	336	253	253	175	118
50-59	366	325	267	252	171	127
60-69	475	346	304	246	172	95
70+	672	396	343	163	127	95

Columns ⇒ cross-sectional rates by age

(NO follow-up of *individuals*)

Frost's TB Studies

TB mortality per 100,000 p-yrs

Age	1880	1890	1900	1910	1920	1930
Males						
0-4	760	578	309	209	108	41
5-9	43	49	31	21	24	11
10-19	126	115	90	36	49	21
20-29	444	361	288	207	149	71
30-39	378	368	296	253	164	115
40-49	364	336	253	253	175	118
50-59	366	325	267	252	171	127
60-69	475	346	304	246	172	95
70+	672	396	343	163	127	95

Rows ⇒ cross-sectional rates by year

(NO follow-up of *individuals*)

Frost's TB Studies

TB morality per 100,000 p-yrs

Age	1880	1890	1900	1910	1920	1930
Males						
0-4	**760**	578	309	209	108	41
5-9	**43**	49	31	21	24	11
10-19	126	**115**	90	36	49	21
20-29	444	361	**288**	207	149	71
30-39	378	368	296	**253**	164	115
40-49	364	336	253	253	**175**	118
50-59	366	325	267	252	171	**127**
60-69	475	346	304	246	172	95
70+	672	396	343	163	127	95

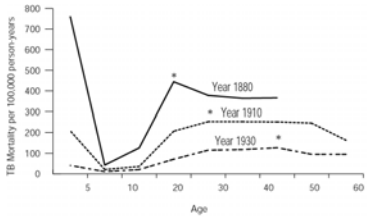
Diagonals ⇒ ***mimic*** experience of **birth cohort** (1870 birth cohort shaded)

TB Study

Cross-Sectional Rates

Note shifting peak occurrence (marked by *) with age

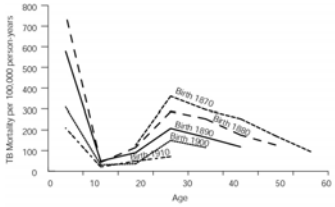
Not a true cohort: *False* impression of cohort experience.



(c) Gerstman Ch 11: Observational Designs 20

Frost's Birth Cohorts

True cohort perspective. Note consistent peak at in late 20s (corresponds childbearing age)



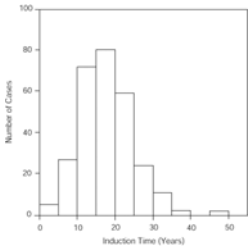
Comstock, G. W. (2001). Cohort analysis: W.H. Frost's contributions to the epidemiology of tuberculosis and chronic disease. *Social and Preventive Medicine, 46*(1), 7-12.

(c) Gerstman Ch 11: Observational Designs 21

Retrospective Cohort

Bladder cancer example

- Historical info on exposure to aniline dyes (from work records) were used to compile cohorts
- Retrospective data on cause of death was recorded from death certificates
- Figure: incubation periods
- SMR of bladder cancer was approx 100 (i.e., 100-fold increase in occurrence)



"Incubation curve"

(c) Gerstman Ch 11: Observational Designs 22