Recent Development in Epidemiologic Principles

Sept 1975 – March 1976

(20 – 2 hour classes*)

Class Notes

* NB: 1 Semester ≅ 35 hrs

Recent Developments in Epidemiologic Principles – Outlines of Classes – 1975

9/3

9/I Introduction – D & I

A. Course in Minn. – work at Harvard – Miettinen

B. Our objectives

1. communicate (reiterate) course at Minn.
2. place in perspective of our learning experiences
3. document our efforts – notes, etc.
4. apply concepts to our own work
5. evaluate their utility – modification

C. Context – "State of the art"

1. Many ideas well reviewed are unpublished
2. Minn. course only scratched surface – no set of notes would be "complete"
3. There is has been little published criticism of M's work

D. Format / Expectations

1. Background required – memo (little)
2. Format
   a. meet Wed. 2:30–5:00 – w/ break – not 9/10
   b. presentation & discussion of material
      Pres. – develop concepts & analytical principles
      D. – apply concepts to quantitative exercises
   c. readings – reference list – reserve
   d. my handouts – first drafts

E. Vocab. (basic terms (not simple)) building – putative, intuit, obviscate, quantitate

II. General Content

A. What won't be covered

1. Descriptive / Experimental / Eval. epid – I AND I will focus on analytical – ie etiologic Ho's

2. Ho generation – ie, Substantive issues – course is about principles + methodology
3. little emphasis on basic issues of study design – much will be assumed
4. Statistics per se – tho we will make use of it

B. General content
1. first part, mostly theoretical
   second part, mostly computational (applied)
2. outline of topics – handout – review
3. overall object: to address basic epid. issue of identifying + measuring causality (effect) – "obviated" by statisticians
4. a. common nomenclature

III. Measures of Disease Frequency – quantify occurrence of dis. (ie, Epid.)
A. General definitions – other sheet → categorical data of non-disease characteristics (as well as disease)
1. Absolute measures
2. Relative measures
   a. rates
   b. ratios
B. Prevalence Rates – point / period
1. Conceptual quantification
2. limitations in etiologic research
   a. not risk – ie, includes new / existing cases
   b. cause / effect ambiguity
C. Incidence Rates
1. I_c = risk – the concern when applying data to individuals
   a. conceptual quantity
   b. interpretation – referent time period
   c. limitations
      (1) gen'l – dis-specific
      (2) all cause death – (not conditional)
      (3) attack rates – referent time defined by epidemic
      (4) case-fatality – den. is cases

d. Selection into den. – PAR p.6
- exclusion of ... (Cervical Ca.)
- problem – PAR is elusive (eg) Breast cancer

2. I_D – fundamental of dis. occurrence in pop.
a. conceptual – picture "Sea of pop-time"
b. quantify – den. = p-years
c. compare to I_c
- i. limits – I_D may be ≥ 1 ($< \infty$)
- ii. I_D as instantaneous risk

3. I_D (as fundamental measure of dis. occur.) may be used to determine I_c
a. exponential decay model – constant I_D over period
$N = N_0 e^{-Kt}$ $K = I_D$
b. $I_c = 1 - N_t/N_0 = 1 - e^{-I_D \cdot t}$
c. if I_D not constant, subdivide t
$I_c = 1 - e^{-\sum I_{D_j} t_j}$ – t_j = age range
d. interpretation of I_c – over all t_j
e. why compute – I_c? – conceptually simple & easy to communicate

4. I_D as misunderstood concept
a. common criticisms
- I_D not good estimator of I_c ∴ inferior
- which is assumed constant over obser.; but I_D is ↑ if not, it can be stratified into subperiods

5. next time – relate P and I_D

9/17 Outline

I. Review and Elaboration – Rates + Risks + Prevalence

A. handouts – Miettinen's manuscript #4.00
1. Rates & Measures of Effect in Epid. Designs
2. Adj. and Standardization of Rates and Measures of Effect
3. Notations and Symbols
4. Epid. Measures used to Quantify Illness (outline)

B. last time defined freq. measures used to quantify disease occurrence – 3 important ones
1. $P = N_D/N$ – proportion [0–1]
2. $I_c = I/PAR$ – proportion [0–1] – also conditional P_t – time referent
3. $I_D = I/T$ – "force of mortality/morbidity" ("sort of instantaneous risk") [0–∞) to be considered fundamental measure of disease occurrence – ∴ can use I_D to derive I_c or P

C. first change in notations
1. t = point in time
   p = period of time
2. P = prev. rate (gen'l)
   $P_{(D)}$ = " " of disease D

D. we had:

$$I_c = 1 - e^{-\sum I_{Dj} \cdot p_j}$$

where: p_j = hypothetical follow-up period of j-th (age) category – ie, age range of the category

$\Sigma p_j = p$ = total (hypothetical) period of follow-up – total age span

$I_{Dj} = I_D$ for j-th Stratum

(OR) $\boxed{I_c = 1 - \Pi (1 - I_{cj})}$

where: $I_{cj} = 1 - e^{-I_{Dj} \cdot p_j}$

E. for example: I_j = # events observed in j-th category during follow-up

age	I_j	T_j (p-yrs)	I_{Dj}	I_{cj}
30-39	6	2000	.003	.030
40-49	15	3000	.005	.049
50-59	10	1000	.010	.095
Total	31	6000	.0052	$\times \rightarrow$ [NB $I_c' = .144$]

$I_{c(30-59)} = 1 - \Pi (1 - I_{cj}) =$

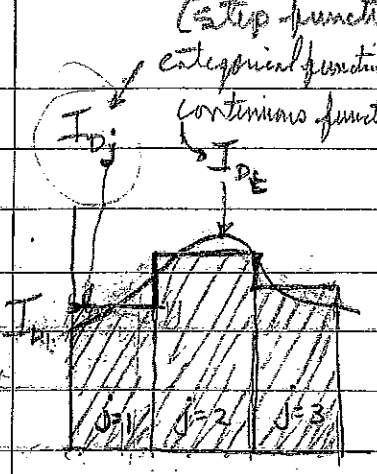
$1 - (1 - .030)(1 - .049)(1 - .095) = \boxed{.165}$

$I_{c(30-59)} = 1 - e^{-\Sigma I_{Dj} p_j} = \boxed{.165}$ also

→ interpretation of I_c: condit. Pr.

* F. 2 very important assumptions:

1. I_{Dj} remains constant throughout each age interval – N.B. this implies a categorical analysis (i.e. categorical assumption)
   a. Since we're assuming that I_D jumps suddenly at ages 40 + 50
   b. I_{Dj} is defined at $t = 0$, so we aren't considering a person of age 39 changing I_{Dj} if he is followed for over a year.



there is no way out of this unless we avoid the assumed constancy of I_D over time periods by:

$$I_t = 1 - e^{-\int_{t_0}^{t_0+P} I_{Dt}\, dt}$$

(I_{Dt} = momentary I_D)

but this requires that we mathematically define the function I_{Dt} — in practice this is rather difficult & I'm not sure whether it is justified — i.e. we categorize into small enough intervals such that I_D remains constant

2. the period-specific experience of my subjects must represent their full longitudinal experience — eg, $I_{D(30-39)}$ won't be same as $I_{D(50-59)}$ group was 20 yrs prior to the study

(or) (hypothetically) if the source pop were stable (ie, distrib. by age remains constant over time), I_{Dj} would also remain constant over time — this is quite analogous to

3. this assumption needed re: annual death rates for a pop. in order to derive an life expectancy at birth (in demographic life table analysis) — X-S experience must represent longitudinal experience

6. this last assumption points out the dual nature of time —
   1. calendar (chronological) time — eg, 1960-1970
   2. age (eg) 30-59 years old

3. this feature allows us to convert an I_D (w/ no period referent) to a I_C (w/ a particular period referent) — this is done by creating a hypothetical follow-up period for the I_C as defined by the age span of the subjects at the onset of the study

H. H_0 I_D is fundamental measure in epid. Why compute I_C at all? 2 reasons
1. I_C is conceptually simpler & therefore easier to communicate
2. this provides an alternative to characterizing the X-year experience of a group w/ a disease when the incidence rate varies considerably w/ age — otherwise we would compute an age-standardized rate which (like any "averaging" process) does not define any identifiable age category but instead suppresses info. Re: the variability of rates among age categories (→ limitation of using the mean) — While we will see the utility of standardization to control for conf. effects; it does not provide any "scientific" info. about the group experience (as a simple SR) ie SR must be compared to other SR's to be of value

II. Relationship between I_D and P (NB. not a rigorous proof)

A. In order to use I_D to derive P, we

1. need to know something about the duration of the disease (ie, from clinical onset to termination) —
   let $\{d_i\}$ = distrib. of durations for a given disease D — d is in years — $d_1 = 1$ etc.
   (categories are 1-yr intervals)
2. let I_i = # new cases per year (same units as d) of a given duration i — we'll assume: I_i remains constant over time (i.e. prob. not rigorous)

B. We can make following relationship (see my paper for graphical illustration)

$N_{(D)i} = I_i d_i$ = # existing cases at 1 pt in time who have a duration i

then: $\bar{d} = \frac{\sum I_i d_i}{\sum I_i}$

$\therefore \boxed{N_{(D)}} = \sum N_{(D)i} = \sum I_i d_i = \bar{d} \sum I_i = \boxed{\bar{d} I}$

Note: relationship of freq. — not rates

C. In order to get relationship of rates:

$I_D = \frac{I}{N - N_{(D)}}$ (hard to intuit)

I_i remains constant
$\therefore I_D$ is constant for each years
(i.e., categorical assumption)

$$P = \frac{N_{(D)}}{N} = \frac{\bar{d} I}{N} = \frac{\bar{d} I_D (N - N_{(D)})}{N} =$$

$$= I_D \bar{d} - \frac{I_D \bar{d} N_{(D)}}{N} = I_D \bar{d} - I_D \bar{d} P$$

$$\therefore \quad P(1 + I_D \bar{d}) = I_D \bar{d}$$

$$\boxed{P = \frac{I_D \bar{d}}{1 + I_D \bar{d}}}$$

$P(1 + I_D \bar{d}) = I_D \bar{d}$

$P + P I_D \bar{d} = I_D \bar{d}$

$I_D(\bar{d} - P\bar{d}) = P$

$I_D = \frac{P}{\bar{d} - P\bar{d}}$

Note: $P \neq I_D \bar{d}$ (as is sometimes assumed; McMahon p66)

D. Miettinen expresses this relationship as an interval (in a more general formulation) & points out the following limits not immediately apparent from the above formula:

$\lim_{\bar{d} \to 0} P = 0$ disease is extremely short in duration

$\lim_{\bar{d} \to \infty} P = I_c$ (from birth to age t...) other extreme: disease is irreversible & non-fatal

III. Adjustment and Standardization of Rates —

A. mentioned before that (SR) is a specific transformation (or averaging) of a stratum-specific rates, to form an overall rate that is used to control for the confounding effects of known factors such as age, race, sex, etc. — & to that extent, forms the heart of epid. methodology

B. Miettinen has actually reconceptualized Standardization to carefully define its utility and limitations

C. But we will begin with the common formulation of 2 separate methods of rate Standardization:

a. Direct Method — DR

category-specific rates observed in the study pop. are applied to (ie, multiplied by) the corresponding category sizes of the "standard" pop (which is an independently observed distribution)

	D	$\bar{D}$	
study	a	b_j	n_j
standard	a_{sj}	b_{sj}	n_{sj}

$R_j = a_j / n_j$

(c) $\bar{R} = \sum a_j / \sum n_j = a/n$

$R_{sj} = a_{sj} / n_{sj}$

(c) $R_s = a_s / n_s$

$$DR = \frac{\sum n_{sj} R_j}{\sum n_{sj}} = \boxed{\frac{\sum n_{sj} R_j}{n_s}}$$

b. Indirect Method — IR

category-specific rates of the "standard pop. are applied to the corresponding subgroup sizes of the study group

$$IR = \frac{\sum n_j R_j \ (R_s)}{\sum n_j R_{sj}} = \frac{a}{E[a]} \ (R_s) =$$

$SMR \cdot R_s$ ← crude rate in Standard

D. Problems w/ common methods
1. concept of "stand. pop" is ambiguous — in DR, it is where the n's come from; in IR, it is where the R's come from
2. it appears that standardy. is actually 2 different methods, that are in one sense opposites — but it will be maintained that a standardy. of (a study) distribution according to a correlate of rate (eg, age) has only a single conceptualization — ∴ there is only one type of standardy. (DR), which we will see has several forms, both for rates and measures of effect (RR)

E. We will first define 2 terms
1. Referent group (or pop) — the series used to compare a given study (test-exposed group — ie, a basis of comparison — eg) E
2. Standard pop. — the series from which we get category-specific "weights" that we will apply to the corresponding rates of the test group, in order to derive an overall (stand.) rate

F. Now define 3 types of rates:

1. CR = no stratification by a potential confounding var. or modifying factor — we can consider num. & den. to be total numbers observed

$$CR = \frac{\sum a_j}{\sum n_j} = \frac{a}{n}$$

this also might be seen as having "latent weights" = the stratum specific sample sizes of the study group. — ie

$$CR = \frac{\sum n_j R_j}{\sum n_j} = \frac{\sum n_j \cdot \frac{a_j}{n_j}}{\sum n_j} = \frac{a}{n}$$

2. AR = includes stratification by potential conf. var. — and incorporates arbitrary weight

$$AR = \frac{\sum w_j R_j}{\sum w_j} \quad \text{(general formula)}$$

eg, $w_j = 1000$

3. SR = Special type of AR in which the weights are derived from the experience of a real group (ie, from the data) — ∴ Standard is an adjustment to a common distrib. (ie, stand. pop.) w/ respect to a potential confounding factor — Same formula as AR

G. So Stand. is a form of weighting rates —
∴ Being more Specific Re: SR, we may define 2 major classes according to the Source of the weights

SR(I) 1. internal Stand. → weights come from Stratum-Specific Sample Sizes of the Study group — $w_j = n_j$
but note! this is equal to the CR, which has n_j as implicit rates

SR(E) 2. external Stand → weights come from a group other than the study group (E
a. Standard = referent
∴ $w_j = n_{Ej}$ (or d_j in $c/\bar{c}$)
b. Standard = another category of E (E'
∴ $w_j = n'_j$ (or d'_j in $c/\bar{c}$)
c. Standard = a general pop. (eg) U.S. (S
∴ $w_j = n_{Sj}$

H. For example: Same data as before + (referent group)

	Study group (E)			referent group ($\bar{E}$)		
	$T_{(E)}$	$I_{D(E)}$	$T_{(E)} I_{D(E)} = I_E$	$T_{(\bar{E})}$	$I_{\bar{E}}$	$I_{D(\bar{E})}$
30-39	2000	.003	6	3000	5	.0017
40-49	3000	.005	15	2000	5	.0025
50-59	1000	.010	10	1000	5	.0050
total	6000	.0052	31	6000	15	.0025

$$DR = \frac{\sum T_{(E)j}\, I_{D(E)j}}{\sum T_{(E)j}} = \frac{9 + 10 + 10}{6000} = \boxed{.0048}$$

NB. • referent group is considered "Standard"!
• weights are similar for I_c as I_D is

$$IR = \frac{\sum T_{(E)j}\, I_{D(E)j}}{\sum T_{(E)j}\, I_{D(\bar{E})j}} \times I_{D(\bar{E})} = SMR \cdot I_{D(\bar{E})}$$

($I_{D(\bar{E})}$: crude rate)

$$= \frac{31}{(3.33 + 7.5 + 5)} (.0025) = \boxed{.0049}$$

* might mention that in IR, it is customary to pool all exposure categories to get what is commonly considered a "standard" pop. — but we see here that the standard is actually the $\bar{E}$ group (where the weights come from) — thus the common pooling technique is actually forming a pooled referent, which represents a rather ambiguous category since it is dependent on the specific distrib. of the pop. — not a known category

$$SR(E) = \frac{\sum T_{(\bar{E})j}\, I_{D(E)j}}{\sum T_{(\bar{E})j}} = DR = \boxed{.0048}$$

(where standard = referent)

$$SR(I) = \frac{\sum T_{(E)j}\, I_{D(E)j}}{\sum T_{(E)j}} = \frac{31}{6000} = \boxed{.0052} = CR$$

$\therefore SR(I) \neq IR$

$SR(\bar{E}) = DR$ (st = neg.)

While it appears that we have simply eliminated the IR from our bag of analytical procedures, we will see next week that using the method of SR(I), IR is highly related to the standardization of effect measures (not of rates)

I. An SR (of any type) is a hypothetical measure and it has no meaning unless compared to other SR's (for other groups) that incorporate the same standard. But first we must discuss the general nature of rate comparisons to see how we measure the association between 2 variables — (E → D)

IV. Measures of Association — 3 general types (see other outline)

A. Non-effect Measures of Assoc.
   1. Correlations (concurrent validity)
   2. Closeness measures (instrumental validity)

B. Measures of Effect
   1. Absolute Measures of Effect
   2. Relative " " Effect

C. Potential Impact Measures
   1. for causal factors (EF)
   2. for protective factors (PF)

I. To be discussed Next time

A. Measures of Assoc. in more detail (esp. B + C)

1. ~~it~~ might be obvious now why "relative" risk" is not used by Miettinen since it does not refer to the ratio of 2 I_D's — Similarly w/ "attributable risk"

2. in C/C study, we estimate RR by calculating an OR — in the past, this approx. was considered valid only if we assumed a rare disease — in fact Miettinen makes this assumption in his manuscript but it is not necessary (→ my paper + his) — based on the distinction between I_D and I_c — measures of assoc. are computed for various study designs in latter portion of paper (I_DR, I_cR, OR, ϕ)

B. Standardization of measures of effect: (RD and RR)

(eg) $\frac{SRR(I)}{SRR(E)} = SMR$ (ie, SMR ≠ SR)

9/24 – 10/1 – 10/7 Outline

I. Introduction
A. discuss class schedule, pace, etc. — 2:30–4:30
B. Debbie — exercises (Latin in class if time)
C. handouts: Quality of Data + Findings (Study Design before 10/8)
D. topics today
1. measures of association – concentrating on 3 measures of effect
2. Standardization of these 3 measures

II. Types of measures of association (ie, indications of criterion validity) (see other outline #4)
A. Non-effect measures of assoc.
1. Correlation (concurrent validity → criterion is another factor, besides the predictor var. at issue; and both are measured at same point in time + independently) – used in prevalence studies –
a. of interest to us: nominal data – generally when analyzing a contingency table, a χ^2 statistic (of indep. or homogeneity is calculated) – but χ^2 only indicates the existence of an assoc., not its strength (or magnitude) –

predictor → criterion

type of correl. coef. depends on kind of data:

in 2×2 table

(genl.) $$\chi^2 = \sum\sum \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$ ↳ $$\chi^2 = \frac{n(ad-bc)^2}{n_D n_{\bar{D}} n_E n_{\bar{E}}}$$

but notice that the statistic is dependent on n; the larger the n, the larger the χ^2 thus the smaller the p-value

b. We can form a correl. coef (ϕ) which measures the strength of the assoc. — $\phi = [-1, +1]$

$$\phi = \sqrt{\frac{\chi^2}{n}} = \frac{ad - bc}{\sqrt{m_D m_{\bar{D}} n_E n_{\bar{E}}}}$$

c. actually this is a Pearson correl. coef. (r) applied to nominal data, assuming any numbers for the categories

d. Significance level is found w/ p-value of the corresponding $\chi^2_{(1)}$

e. for $m \times n$ table: use Contingency coef. (C)

$$C = \sqrt{\frac{\chi^2}{n + \chi^2}} \qquad \text{range: } \left(0 - \sqrt{\frac{k-1}{k}}\right)$$

where: $k = m = n$

with: $(m-1)(n-1)$ degrees of freedom for the χ^2 statistic

C & ϕ: are only stable when the researcher has not arbitrarily fixed any of the marginals:

	D	$\bar{D}$	
E	a	b	n_E
$\bar{E}$	c	d	$n_{\bar{E}}$
	m_D	$m_{\bar{D}}$	n

thus, ϕ can only be used w/ type II Sampling (i.e. X-S Selection of a pop. — ∴ we know the Sampling fraction(s))

2. Correctness of Measurement — (instrumental validity → criterion is other known measure of the same factor, measured at same point in time) — also for X-S Sampling; but generally info. in correl. coef. is Separated into ≥1 measure:

a. Sensitivity
b. Specificity
c. Predictive Value

} Screening tests

B. Measures of effect — (predictive validity → criterion is another factor measured after the predictor var. (in calendar time)

B. 1. Absolute measures of effect
   a. rate difference = RD = $R_E - R_{\bar{E}}$
   b. excess mortality = $a - E[a]$
2. Relative Measures of Effect
   a. relative rate (or rate ratio) = RR = $\frac{R_E}{R_{\bar{E}}}$
   b. excess relative rate ("relative effect")
      = ERR = RR − 1
C. Potential Impact Measures
   1. for causal factors — ARP = % of cases that could have been prevented if persons were unexposed
   2. for protective factors — (PF) = % of potential cases that might have been prevented

III. Measures of of Effect
A. RD
   1. $I_cD = I_{c(E)} - I_{c(\bar{E})}$; when period of follow-up is same for both rates
   2. $I_DD = I_{D(E)} - I_{D(\bar{E})}$; when follow-up is comparable
   3. limits (of both) for causal exposure are:
      $0 \leq I_DD < \infty$
      $0 \leq I_cD \leq 1$

(10/1) B. RR
   1. $I_cR = I_{c(E)} / I_{c(\bar{E})}$
   2. $I_DR = I_{D(E)} / I_{D(\bar{E})}$
   3. limits of both for causal factors:

3. a. $1 \leq I_c R < \infty$

$1 \leq I_D R < \infty$

b. also: $\boxed{\lim_{p \to \infty} I_c R = 1}$ ie, if we could follow everyone for an infinite period of time & they didn't die of any other cause, eventually, everyone would die of the disease at issue — $\therefore 1$ is lower limit of $I_c R$ for causal exposure

c. also: $\boxed{\lim_{p \to 0} I_c R = I_D R}$

ie, assuming I_{D_t} is a continuous function, as p approaches 0, the I_c for each group becomes $I_D \cdot p$ — thus:

$$\lim_{p \to 0} \frac{I_{c(E)}}{I_{c(\bar{E})}} = \frac{I_{D(E)} p}{I_{D(\bar{E})} p} = I_D R$$

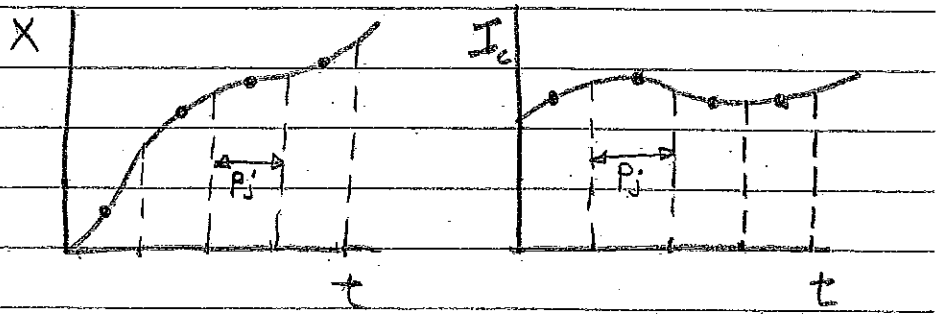
d. putting these 2 limits together:

$$\boxed{1 \leq I_c R \leq I_D R < \infty}$$

if I_{D_j} is constant over time

e. But the constancy of I_D (for any p_j) has important implications to epid. research — we can think of I_{D_t} as analogous to velocity (V_t); and

categorical I_D (ie, I_{Dj}) as analogous to average speed ($\bar{S}$) for time period p (NB! $\Delta t = p$) — distance = X — j = time period



$$\bar{S}_j = \frac{\Delta X_j}{p_j} \qquad I_{D_t} = \lim_{p_j \to 0} I_{Dj}$$

$$V_t = \lim_{p_j \to 0} \bar{S}_j = \text{slope} \qquad I_{D_t} = \lim_{p_j \to 0} \frac{I_{cj}}{p_j}$$

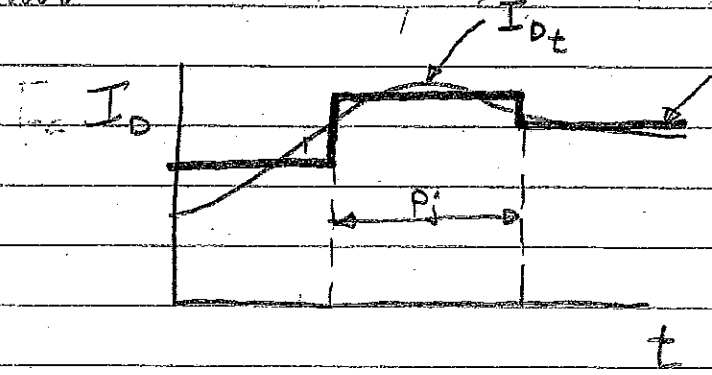
$$\boxed{V_t = \lim_{p_j \to 0} \frac{\Delta X_j}{p_j}} \qquad \boxed{I_{D_t} = \lim_{p_j \to 0} \frac{\Delta I_j / N_j}{p_j}}$$

Despite this analogy, the actual situation in physics is quite different from that of Epid. In physics, we have "laws" (thanks to Newton, etc.) which describe the mathematical relationship between V and other variables — eg,

$$V_t = \sqrt{2gX}$$ (for a free-falling body encountering no friction)

But in epid, we have no such laws of nature describing I_{D_t}. So we approximate it with a categorical measure

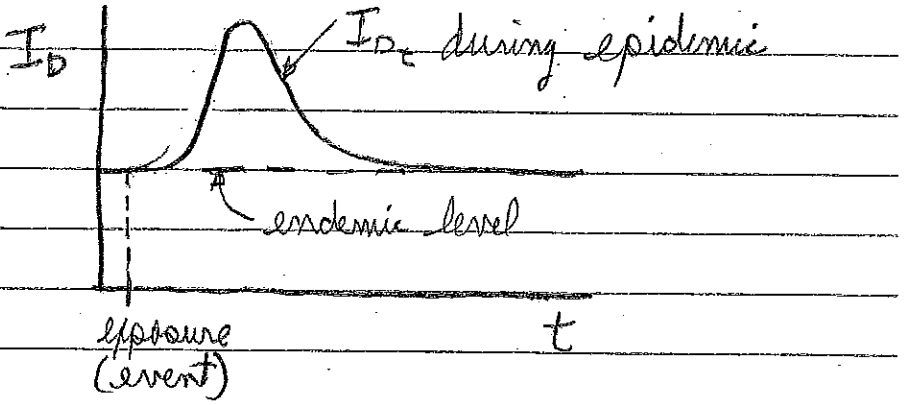
(using person-years of follow-up) — just like we use $\bar{S}$ to characterize an assumed constant velocity over a time period —



where: $I_{C(p)} = 1 - e^{-\int_{t_0}^{t_0+p} I_{D_t}\, dt}$

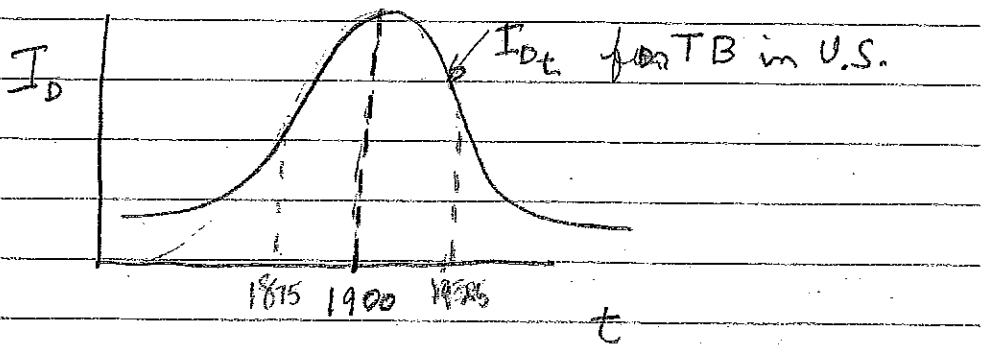
or: $I_{C(p)} = 1 - e^{-\sum I_{D_j} p_j}$

f. Thus Epid. needs a Newton; but is there any evidence that I_D follows any laws of nature. The key to using (categorical) I_D is the assumption of its constancy for p_j and its constancy over time w/in age categories. We have already seen that I_D should not remain ~~const~~ uniform over wide age spans, but is there always reason to believe that I_D remains constant over calendar time. For example, imagine a short term epidemic in which exposure to contaminated food at one meal causes a sudden outbreak.



Thus, there would be a short-term time variation within age bands.

But we know that even chronic diseases follow long-term epidemic trends, that pose a similar phenomenon (very crudely.)



In fact, we even know (according to Dr. C) that this curve in any pop. shows a basic consistency for many diseases: the ratio of F/M and young/old is greater on the rise (when Slope $\pm$ +) than on the decline. It is possible that the curve may be described by sociological/environmental variables in the society. —

One researcher has actually claimed that he can predict (or determine) where a pop. is chronologically in relation to its industrial revolution by the ratio of M/F TB deaths per year. Perhaps there are possibilities for describing some general principles of disease occurrence that would enable us to predict I_{D_t} on a basis independent of observing disease occurrence (post hoc)

4. OR = $O_{(D)} / O_{(\bar{D})}$ (usually computed in C/C study, but may compute anytime

$$Pr[\text{event}] = \frac{[\#\text{ favorable outcomes possible}]}{[\#\text{ total outcomes possible}]} = \frac{1}{2}$$

$$O[\text{event}] = \frac{[\#\text{ fav. outcomes possible}]}{[\#\text{ unfav. " "}]} = \frac{1}{1}$$

for cases: Odds of exposure = $O_{(D)}$

$$O_{(D)} = Pr[E|D] / Pr[\bar{E}|D] = \frac{\frac{a}{a+c}}{\frac{c}{c+d}} =$$

$$= \boxed{e_D / 1 - e_D}$$

for non-c.:

$$O_{(\bar{D})} = Pr[E|\bar{D}] / Pr[\bar{E}|\bar{D}] = \frac{\frac{b}{b+d}}{\frac{d}{b+d}} =$$

$$\boxed{e_{(\bar{D})} / 1 - e_{\bar{D}}}$$

So, $OR = \frac{O_{(D)}}{O_{(\bar{D})}} = \frac{e_D(1-e_{\bar{D}})}{e_{\bar{D}}(1-e_D)} = \boxed{\frac{ad}{bc}}$

(which is a ratio of 2 ratios)

It has also been shown using Bayes theorem that in a f/c study

$$OR \cong IR$$

if $P_r[D]$ is small ($< .20$)

5. In a X-S study (w/ type II sampling)

$$OR_p = \frac{P_{(E)}(1-P_{(\bar{E})})}{P_{(\bar{E})}(1-P_{(E)})} = \frac{e_D(1-e_{\bar{D}})}{e_{\bar{D}}(1-e_D)} = OR$$

if Cases + $\bar{C}$ of both types of studies are from same pop.

Note: OR_p (in X-sectional) cannot technically be used as a measure of effect since we don't know the temporal relationship

Note: $OR_p \neq PR = \frac{P_{(D)}}{P_{\bar{D}}}$

PR is often used to estimate RR – eg, Slones manual

insert V p.9 > b. Derive: (see next insert)

$\hat{I}_D R$ in c/$\bar{c}$ study
- a. for existing cases
- b. for new cases (± advantages)

B. General conclusions from my paper #1

c/$\bar{c}$ study
- a. $OR = \hat{I}_D R$ (when d's are equal regardless of $P_{(D)}$ or $P_{(E)}$)
- b. $OR = \hat{I}_{\bar{c}} R$ (only when disease is rare, ie $P_{(D)} \leq .20$)

c. But despite the lack of inherent quantitative flaws in estimating effect in c/$\bar{c}$ study, we must remember the logical flaws that accompany the specific study design – (eg)
- (1) biased recall of exposure
- (2) Selective survival of all cases

d. $\hat{I}_{\bar{c}} R < \hat{I}_D R$ ± this difference widens w/ incr. $P_{(D)}$!

e. OR is quite ineff. in estimating RR in: CH (when: $.10 > P_{(D)} > .20$)
X-S (when: $.25 > P_{(E)} > .50$)
(ie, CI are very wide)
also: OR is unstable (invalid) for CH study (Type I sampling) for $P_{(D)} \geq .20$

f. OR (in genl.) is always ineff. at.
$P_{(E)} < .05$
$P_{(D)} < .10$

g. ϕ is unstable (invalid) for:
C-H: $.10 < \phi < .70$
C/E: $\phi > .10$
but: X-S — fairly invariant (assuming type II sampling)

IV. Potential Impact Measures

A. ARP — % of the disease that could potentially have been prevented if no one had been exposed

1. ARP_E (for exposed cases only)

$$ARP_E = \frac{RR-1}{RR} = \frac{a - E[a]}{a}$$

= % of exposed cases that would have been prevented if no one had been exposed → (potentially)

2. ARP_N (pop. A.R.P) $= e_D ARP_E = \frac{a - E[a]}{a+c}$

= % of all cases that would have been prevented if no one had been exposed

3. ARP requires that cases be representative of the pop. in order for it to apply to the pop. as an "expected prevention rate" (ie) ARP applies to a specific pop. at a period in time (unlike RR + RD which refer to the nature of a disease)

4. ARP pertains only to Observ. Studies — as an expected % of preventable cases, & it is like a PF, not a measure of prevention

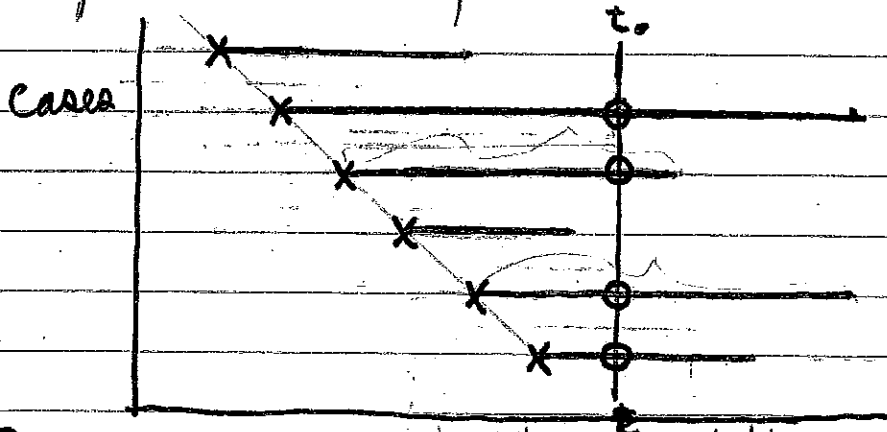
or treatment effect that one would estimate in an experiment — i.e. ARP is a measure of assoc. (like RD or RR)

5. 2 additional assumptions needed to regard ARP as a prevention measure:
   a. E can be Δ'ed in source pop. — ideally all $E \rightarrow \bar{E}$
   b. changing a person from category E to $\bar{E}$ will place him at the same risk (rate) that previously unexposed persons had

V. (next page)
last time discussed Measures of Assoc. moved too quickly over estimation of effect in c/c Study, so I'll go over it in more detail

V. Estimating the $I_D R$ in a c/c study – Review in more detail

a. define incid & prev. cases –



O – prevalent cases at pt. t

X – incident cases identified over time

NB. – duration of disease

(1) prev. cases – identifying existing cases that might have had the disease for any # of years (NB. d is not any duration of cases at time t)

(2) incid. cases – new (fresh) cases identified as they occur over time –

to get $\hat{I_D R}$

(1) $N_{(D)} = I\,\bar{d}$ ... not what $\bar{d}$ is; from onset to termination

$$I_D = \frac{I}{N - N_{(0)}} = \frac{I}{N_E}$$

$$P = \frac{I_D \bar{d}}{1 + I_D \bar{d}}$$

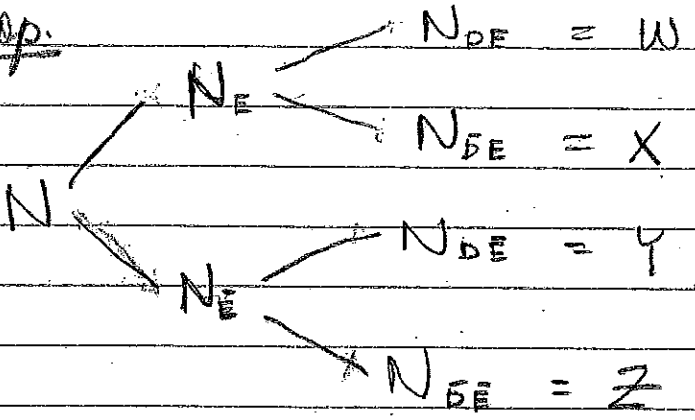
$$I_D = \frac{P}{\bar{d}(1-P)}$$

$$I_D R = \frac{I_{D(E)}}{I_{D(\bar{E})}} = \frac{P_{(E)}(1-P_{(\bar{E})})}{P_{(\bar{E})}(1-P_{(E)})} \cdot \frac{\bar{d}_{(\bar{E})}}{\bar{d}_{(E)}} = OR_P \cdot \frac{\bar{d}_{\bar{E}}}{\bar{d}_E}$$

X-S of pop.

	D	$\bar{D}$
E	W	X
$\bar{E}$	Y	Z

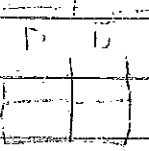
In a pop.



$P_{(E)} = \frac{W}{W+X}$

$P_{(\bar{E})} = \frac{Y}{Y+Z}$

X-S of pop.



$$OR_p = \frac{\frac{W}{W+X}\left(1 - \frac{Y}{Y+Z}\right)}{\frac{Y}{Y+Z}\left(1 - \frac{W}{W+X}\right)} = \frac{\frac{W}{W+X} - \frac{WY}{(W+X)(Y+Z)}}{\frac{Y}{Y+Z} - \frac{YW}{(Y+Z)(W+X)}}$$

$$= \frac{W(Y+Z) - WY}{Y(W+X) - YW} = \frac{WY + WZ - WY}{YW + YX - YW}$$

$$= \boxed{\frac{WZ}{YX}}$$

$$\therefore \quad I_D R = \frac{WZ}{YX} \; \frac{\frac{d_{\bar{E}}}{T}}{\frac{d_E}{T}}$$

E values in C/$\bar{C}$ study: where we have sampled cases + noncases w/ certain sampling fractions

	D	$\bar{D}$
E	$F_D W$	$F_{\bar{D}} X$
$\bar{E}$	$F_D Y$	$F_{\bar{D}} Z$

note both F's are equal in each column if we assume no selective bias

F_D = # cases in Study / # cases in Source pop.

$F_{\bar{D}}$ = # noncases in Study / # noncases in Source pop.

$$OR = \frac{ad}{bc} = \frac{F_D W \, F_{\bar{D}} Z}{F_{\bar{D}} X \, F_D Y} = \frac{WZ}{XY} = OR_P$$

$$\therefore \boxed{I_D R = \left(\frac{ad}{bc}\right) \frac{\bar{d}_{\bar{E}}}{\bar{d}_E}}$$

Thus, (w/ prev. cases) $\hat{OR} = \hat{I_D R}$ if $\bar{d}_{\bar{E}} \approx \bar{d}_E$

(2) If incident cases — don't need to know about durations — but NB: $\bar{d} \neq 0$ (it's the same as before)

All we can say: $\bar{d}_E = \bar{d}_{\bar{E}}$, simply because we have incident cases

but we can say:

	D	$\bar{D}$
E	$F_D I_E$	$F_{\bar{D}} N_{DE}$
$\bar{E}$	$F_D I_{\bar{E}}$	$F_{\bar{D}} N_{D\bar{E}}$

same as before

If we assume I & N_D remain stable over time of Study

from before: $I_D = \frac{I}{N - N_D} = \frac{I}{N_{\bar{D}}}$

for exposed: $I_{D(E)} = \frac{F_D I_{(E)}}{F_D N_{D\bar{E}}}$

for unexposed: $I_{D(\bar{E})} = \frac{F_D I_{\bar{E}}}{F_D N_{D\bar{E}}}$

$\therefore I_D R = \frac{F_D I_E / F_D N_{D\bar{E}}}{F_D I_{\bar{E}} / F_D N_{D\bar{E}}} = \frac{I_E N_{\bar{D}\bar{E}}}{I_{\bar{E}} N_{\bar{D}E}}$

$OR = \frac{ad}{bc} = \frac{F_D I_E F_D N_{D\bar{E}}}{F_D N_{D\bar{E}} F_D I_{\bar{E}}} = \frac{I_E N_{\bar{D}\bar{E}}}{I_{\bar{E}} N_{\bar{D}E}} = I_D R$

$\therefore \boxed{OR = I_D R}$ w/o any assumption about d

So advantages to using incident cases

(1) with - don't need to know duration to estimate $I_D R$

(2) ~~with~~ methodological adv. (inference)

(a) having new cases helps eliminate possibility that D predated E — ie, temporal relationship

(b) also reduce Selection bias since cases have not had much chance to die from the dis. ie, reduce Selective Survival

So what are practical adv. of this:

(1) don't need to make approximation RE: non-diseased — but usually no big deal

(2) important advantage to having new (incident) cases in c/c study

Repeat: c/c study

(1) Incidental math. flaw in estimating I_{DR}

(2) but several important methodological problems that are of particular relevance in c/c studies (differential bias)

(a) biased recall of exposure history

(b) Selective survival of all cases (esp. w/ prev. cases)

(c) often difficult to distinguish cause + effect (esp. w/ prev. cases) & with chronic diseases w/ long induction time

So, practical advantages of this theory:

(1) We don't need to make assumption of rare disease — but in most situations, this isn't an added benefit because most conditions we study are rare (<.20)

(2) Using incidence cases in a C/C study allows us to gain math. + method. advantages

To repeat — how this affects the utility of a C/C study in practice

(1) no inherent math flaw in estimating I.R (esp. w/ incident cases)

(2) but several important methodological problems that are of particular relevance to C/C designs

a. might be biased recall of exposure history by cases + non-cases

b. Selective survival of all cases (esp. w/ prev. cases)

c. often difficult to attribute cause + effect relationship (esp w/ prev. cases and/or chronic diseases that have long induction periods) — this is more of a quantitative problem since the I.P. might vary according to the time that the exposure occurred)

VI. Standardization

A. Review: SR

1. adjusted rate = AR = $\frac{\sum w_j R_j}{\sum w_j}$

2. 2 types of AR

a. general adj. — when w_j are chosen arbitrarily

b. standardized — when wts. are taken from empirical data source — 2 types:

i. $SR(I) = \frac{\sum n_j R_j}{\sum n_j} = \frac{a}{n} = CR$

(ie, wts come from study group)

ii. $SR(E) = \frac{\sum w_{sj} R_j}{\sum w_{sj}}$

in which w_{sj} come from:

(1) referent ⇒ SR(E) = DR
(2) another category(s) of E – ie, E' (besides E + Ē)
(3) general pop. – (actually a special case of SR(E))

3. IR = SMR · R_E ≠ SR
(only 1 way to standardize a rate – the "direct method")

B. Standardization of RR

1. SR is meaningless by itself – i.e. must be compared to another SR – when both SR's have the same standard pop. (ie, a common distribution) we call the ratio of the 2 SR's an SRR
2. At this point, mention difference in terminology between my use of the terms "internal & "external" vs. M.'s – I refer to (I) and (E) as an indication of the source of weights for each SR + for an SRR, the terms internal & external refer to the source of weights relative to the test (exposed) group – M. only uses (I) and (E) to refer to the type of RR standardization (not SR) thus: (M.)
   a. internal (standardized) = both component ratios (R_E & $R_{\bar{E}}$) have been adj. to the same distrib. (in my terms, the defn. of SRR)
   b. external (stand.) = in addition, the Standard is not the Study pop. – i.e., SR_E is not crude

3. Comparison of terminology & defns.

Type	mine	Miettinen's
A.	Internal Stand. (wts. come from E) $= SRR(I) = \boxed{\frac{SR_E(I)}{SR_E(E)}} = \frac{CR_E}{SR_E(E)}$ $= \frac{a}{E[a]}$ (in stand = study) $= SMR$	Internally standardized/ Externally crude = = semi standardized
B	External Stand. (wts. don't come from E) $SRR(E)$	Internally standardized/ Externally standardized
1.	Standard = referent $\boxed{\frac{SR_E(E)}{SR_E(I)}} = \frac{SR_E(E)}{CR_E} = \frac{E[c]}{c}$ (in ref = stand)	= mutually standardized
2.	standard ≠ referent $\boxed{\frac{SR_E(E)}{SR_E(E)}} = \frac{E[a]}{E[c]}$ (in stand.)	
etc.	Standardized excess RR = SERR $= SMR - 1$ (null value)	"crude" v relative effect = "Crude Abs. effect" CR_E v null ratio
	Standardized Rate difference = SRD	crude v absolute effect
	Relative Rate = Rate Ratio = RR	Risk Ratio

(10/15) B. 3. take closer look at: (generally)

a. $$SRR = \frac{SR_E}{SR_{\bar{E}}} = \frac{\sum w_j R_{Ej} / \sum w_j}{\sum w_j R_{\bar{E}j} / \sum w_j} = \boxed{\frac{\sum w_j R_{Ej}}{\sum w_j R_{\bar{E}j}}}$$

recast as:

$$= \boxed{\frac{\sum w'_j \, RR_j}{\sum w'_j}}$$

where: $w'_j = w_j(R_{\bar{E}j})$

$RR_j = R_{Ej} / R_{\bar{E}j}$

	D	den.*
E	a_j	T_{Ej}
$\bar{E}$	c_j	$T_{\bar{E}j}$

* = # persons in I_c study, n
# p-yrs in I_b study

I will compute SRR in a few cases but paper has several others

b. <u>cohort study</u>: computations straightforward

$$SRR(I) = \frac{\sum T_{Ej} R_{Ej}}{\sum T_{\bar{E}j} R_{\bar{E}j}} = \frac{\sum \frac{T_{Ej} a_j}{T_{\bar{E}j}}}{\sum \frac{T_{Ej} c_j}{T_{\bar{E}j}}} =$$

$$= \frac{\sum a_j}{\sum \frac{T_{Ej} c_j}{T_{\bar{E}j}}} = \boxed{\frac{a}{\sum \frac{T_{Ej} c_j}{T_{\bar{E}j}}}} = \frac{a}{E[a]}$$

$$= \boxed{SMR}$$

$$\underset{\text{stand=ref.}}{SRR(E)} = \frac{\sum \frac{T_{\bar{E}j} a_j}{T_{Ej}}}{\sum \frac{T_{\bar{E}j} c_j}{T_{\bar{E}j}}} = \boxed{\frac{\sum \frac{T_{\bar{E}j} a_j}{T_{Ej}}}{c}} = \frac{E[c]}{c}$$

$$= \boxed{SRR}$$

c. Case-Control Study: derivations are more complicated, but results are understood intuitively. — see paper for derivations —

	D	$\bar{D}$
E	a_j	b_j
$\bar{E}$	c_j	d_j

NB. can't compute R's

$$SRR(E) = \frac{a}{\sum \frac{b_j c_j}{d_j}} = \frac{a}{E[a]} = SMR$$

the trick comes, however, in seeing how the denominator is $E[a]$

ie, $E[a]$ = what "a" would be (in the standard) under the H_0 that all strata had RR = 1

$$\therefore OR_j = \hat{F}_D R_j = \frac{E[a]\, d_j}{b_j c_j} = 1$$

$$E[a] = \frac{b_j c_j}{d_j}$$

N.B. wts. (w_j) are implicit in formula & are not n_{Ej} but b_j / F_{Dj} = # exposed controls in the source pop.

$$SRR(E)_{\text{Stand. ref.}} = \frac{\sum \frac{a_j d_j}{b_j}}{c} = \frac{E[c]}{c}$$

→ (over)

* 4. SMR's — 4 common mistakes

a. note correction in Paper #2, p. 19, 3rd line "rate" should be ratio

b. often in epid. research, SMR's for c/c studies are calculated incorrectly

$$\text{"SMR"} = \frac{a}{\sum \left(\frac{b_j}{b_j + d_j}\right)(a_j + c_j)} = \frac{Se_D}{Se_{\bar{D}}} \neq SRR$$

(4') in the general case w/ standard pop. S, which is different from E, $\bar{E}$ or any combination of exposure groups, it is possible for weights to be in # persons and exposure group denominators to be in person-years — in this case: the SRR(E) will not actually be a ratio of observed to expected cases

(eg)

$$SRR(E) = \frac{\sum \frac{N_{Sj} a_j}{T_{Ej}}}{\sum \frac{N_{Sj} c_j}{T_{\bar{E}j}}} \neq \frac{\mathcal{E}[a]}{\mathcal{E}[c]}$$

It should be noted that this does not apply to the situations when the standard is the referent, etc., because in order to compare E and $\bar{E}$ they must have the same kind of incidence rates: either I_c or I_D

c. often SMR's are computed w/ the stated intention of "indirectly" standardizing the Rate — what is called a stand. pop. is actually a ref. — but it is often recommended that we pool the exposure categories to form this so-called "stand." — done for 2 reasons:

(1) get stable rates in the "stand." by creating a group w/ larger #'s

(2) to make the "stand." similar to the indiv. exposure groups w/ respect to ~~std~~ stratum-specific rates — ~~if~~ it is recognized that IR adj. can ~~give~~ lead to weird results if the 'stand. is quite different ~~from~~ any of the compared groups

but forming a pooled referent (not stand) is actually being done — this makes the content of the ref. group variable according to the distrib. of the pop. or the selection procedures — thus any measure of effect has an ambiguous interpretation (what actually is being compared?) — it is thus best to carefully select the referent group — preferable a single category of exposure — N.B. Cole's Bladder Ca. Study uses several ~~pots~~ categories as ref. since they are Hypothesized as unexposed.

d. often SMR's are used as a form of IR adj. (esp. when rates in the exposed group are unstable) - but there is a problem when we attempt to compare 2 or + SMR's (of various exposure categories) even using the same referent (eg) I.e. study

		D	$\bar{D}$	
>1 pk	E1	a_1	b_1	n_{E_1}
<1 pk	E2	a_2	b_2	n_{E_2}
0 cig.	$\bar{E}$	c	d	$n_{\bar{E}}$
		m_D	$m_{\bar{D}}$	n

Suppose we make two comparisons to estimate RR: E1/$\bar{E}$ + E2/$\bar{E}$, controlling for some confounding factors — Since SMR's are actually internally stand. RR's, the weights for each comparison come from the test group: E1 in the first comparison and E2 in the second — thus I.O.M. has maintained that the 2 SMR's would not be comparable - ie, they would give us the wrong idea of the relative effect between E1 and E2; if the variable (E) is ordinal, this means that a comparison of SMR's would not be appropriate in describing a dose-response relationship — thus weighted averages (SMR's) cannot be compared unless the weights (ie, the standard pop.) is identical for all comparisons (eg) — p. 20 #2

apparently common belief: that things compared to equal things can be compared to each other

$\bar{E}$

E1 - - - - E2

but actually it is not the means but (weighted) avg's that are compared.

Ie

S_1

	D	$\bar{D}$	
E1	10	0	10
E2	90	0	90
$\bar{E}$	5	45	50

S_2

	D	$\bar{D}$	
E1	90	0	90
E2	10	0	10
$\bar{E}$	50	0	50

Note that in each stratum of E1 and E2, the rate of D is the same (1) — ie, everyone gets the disease — we should therefore expect that a comparison of effects due to E1 and E2 should be equal — however, the SMR's are not equal; but SRR's are equal because it involves the same stand. ($\bar{E}$) in both comparisons

NB. I would like to question the universality of this argument — ie, that SRR's are always appropriate (& not SMR's) in comparing several groups with the same referent. — 3 questions

(1) What is meant by comparability?

(2) What happens if the "Weights" for both E1 and E2 are equal — could we then use SMR's?

(3) ~~what~~ does it make any difference if the conf. factor also modifies the effect of E on D?

Set up 4 conditions to study — Conclusions

5. Why use SMR's at all? — to determine whether Conf. is present — $RR^* = \frac{CRR}{SMR}$

c. Standardization of RD

1. (General)

$$SRD = SR_E - SR_{\bar{E}} = \frac{\sum w_j R_{Ej}}{\sum w_j} - \frac{\sum w_j R_{\bar{E}j}}{\sum w_j}$$

only done for I rates $= \frac{a - E_0(A)}{n_1}$

$SRR = \frac{a}{E_0(A)}$

2. 2 types:
   a. SRD(I) ⇒ wts. come from E
   b. SRD(Ē) ⇒ " " " Ē
   c. Same principle of comparability as for SRR

3. relationship between SRD + SRR

$$RD = R_E - R_{\bar{E}} = R_{\bar{E}} RR - R_{\bar{E}} = R_{\bar{E}}(RR - 1)$$

$SRD(I) = SR_E(I) - SR_{\bar{E}}(\bar{E})$ ← Stand. = ref.

Since, $R_E = SR_E(I)$ $(= CR_E)$,

$$\therefore SRD(I) = \frac{R_E}{SR_{\bar{E}}(I)}\left[SR_E(I) - SR_{\bar{E}}(\bar{E})\right] =$$

$$= \frac{R_E\, SR_E(I)}{SR_{\bar{E}}(I)} - \frac{R_E\, SR_{\bar{E}}(\bar{E})}{SR_{\bar{E}}(I)} =$$

might use C/E study + other info. to obtain SRD ▷

$$= R_E - \frac{R_E}{SMR} = R_E\left(1 - \frac{1}{SMR}\right) =$$

$$= \left(\frac{SMR - 1}{SMR}\right) R_E$$

(note difference in form from formula above (RD))

D. Standardization of Measures of Potential Impact

1. ARP — causal factors (p. 7)

(S)ARP = EF

a. for exposed pop.: $EF_E = \boxed{\frac{SMR-1}{SMR}} = \frac{a-\mathcal{E}[a]}{a}$

$\frac{SMR-1}{SMR} = \frac{\frac{a}{\mathcal{E}[a]}-1}{\frac{a}{\mathcal{E}[a]}} =$

where: $SMR \geq 1$ ⇒ (EF would be negative)

$\frac{a-\mathcal{E}[a]}{a} = EF_E$

b. for total pop.

$\frac{a}{a+c}\left[\frac{a-\mathcal{E}[a]}{a}\right] =$ ← $EF_N = \boxed{e_D\, EF_E} = \frac{a-\mathcal{E}[a]}{a+c}$

$\frac{a-\mathcal{E}[a]}{a+c}$

c. EF for several categories of exposure var. (i.e. E_i)

$$EF_N = \boxed{1 - \sum \frac{e_{Di}}{SMR_i}}$$

where: $EF_N \geq 0$ (not $SMR_i \geq 1$)

$e_{Di} = \frac{a_i}{\sum a_i}$ = % of cases in the i-th category of E (including ref.)

SMR_i = SMR for i-th category of E

NB. correction ① p. 23 #2, bottom formula: put $\sum$ at beginning of expression after "="

③ #5, p. 11 — cross out "note" under 3.b.

② #3, p. 11, bottom line $\frac{\mathcal{E}[a]}{\mathcal{E}[e]}$ should be $\frac{\mathcal{E}[e]}{\mathcal{E}[a]}$

Harvey's questions — back of p. 18

d. for several exposure vars; E_a, E_b etc.

if effect of E_a + E_b are indep., (not interactive effect)

$$EF_{a,b} = 1-(1-EF_a)(1-EF_b) =$$

$$\boxed{EF_a + EF_b - EF_a\,EF_b}$$

where: EF_a is EF, given non-exposed state of E_b.

EF_b → " "

if $EF_{a,b} > \ldots$ → Synergism

$EF_{a,b} < \ldots$ → antagonism

2. preventive factors (preventive fraction) — analogue to EF but it is not quantitatively an "inverse" of EF — (as was the case for RR)

because → PF is % of potential cases (actual + prevented) that would have been prevented ~~if~~ because of exposure to protective factor

* PF → exposed category is considered the low risk group (eg. heavy physical exercise)

a. for exposed cases: $PF_E = \boxed{1-SMR} =$

$PF_E = 1 - SMR =$

$= 1 - \frac{a}{E[a]} = \frac{E[a]-a}{E[a]}$ where: $SMR \leq 1$

$= \frac{E[a]-a}{E[a]}$

b. for total cases: $PF_N = e_D^* PF_E$

where: e_D^* = % of potential cases that are exposed (ie, in low-risk category) = $\frac{\mathcal{E}[a]}{\mathcal{E}[a]+c}$ (if E is dichotomized) ie, $e_{D0} = 1 - e_D$

(over)

$$PF_N = \frac{e_D(1-SMR)}{e_D + (1+e_D)SMR} = \frac{\mathcal{E}[a]-a}{\mathcal{E}[a]+c}$$ (over)

where: $SMR \leq 1$

c. ~~for~~ Several categories of 1 fp. var. = E_i

$$PF_N = \sum \frac{e_{Di}(1-SMR_i)}{e_{Di} + e_{D0} SMR_i}$$

Sel. notes

where: e_{D0} = % of cases falling in ref. group

SMR_i = doesn't have to be ≤ 1 for all i

$PF_N \geq 0$

d. for Several vars.: E_a, E_b

if effects of $E_a + E_b$ are indep.,

$$PF_{a,b} = 1 - (1-PF_a)(1-PF_b) =$$

$$= PF_a + PF_b - PF_a PF_b$$

where: PF_a: PF for E_a, given non-exposed state (chi-sq) of E_b

If: $PF_a > \ldots$ → Synergism

$PF_b < \ldots$ → antagonism

VI Next time: Study Design

$$e_D^* = \frac{e_D}{e_D + e_{D_0} SMR} = \frac{\frac{a}{a+c}}{\frac{a}{a+c} + \left(1 - \frac{a}{a+c}\right)\frac{a}{\mathcal{E}[a]}}$$

$$= \frac{\frac{a}{a+c}}{\frac{a}{a+c} + \frac{a}{\mathcal{E}[a]} - \frac{a^2}{\mathcal{E}[a](a+c)}} = \frac{a\,\mathcal{E}[a]}{a\,\mathcal{E}[a] + a(a+c)}$$

$$= \boxed{\frac{\mathcal{E}[a]}{\mathcal{E}[a] + c}}$$

$$PF_N = e_D^* \, PF_E = \left(\frac{\mathcal{E}[a]}{\mathcal{E}[a]+c}\right)\left(\frac{\mathcal{E}[a] - a}{\mathcal{E}[a]}\right) = \boxed{\frac{\mathcal{E}[a] - a}{\mathcal{E}[a] + c}}$$

11/5 Outline

next week: end 4:00
next sem: time

I. Intro.
A. Outline for remaining wks. of Semester
1. criteria for evaluating quality of one's study design and data
2. dimensions and types of epid. study designs + methodological implications
3. clarification + elaboration of 3 concepts: conf., EM, and I.
B. today – ~~mostly~~ based on handout #5 which was intended to give some feel for the different perspectives of viewing the issues dealing w/ the quality of one's design: conceptual (outline #1); statistical (#2); and methodological (#3) → focus of today's ~~discussion~~ topic + based on O.M. (ch. 2). – p. 9

II. 2 bases or general criteria on which we can evaluate estim
A. Feasibility of implementation – practical considerations, including economic, social, moral, + political issues – p. 9
B. Informativeness – accuracy in estimating the object of study (ie, measure of assoc.) –
1. in statistical terms, this means studying the variability in one's data, separating 2 components
a. true Var – reflects etiologically based differences in experience in/re: to D
b. Error variability – reflects inaccurate estimation of experience w/ D
i. sampling (usually don't sample the entire source pop.)
ii. non-Sampling (measurement error)

2. from epid. perspective (O.M.), there are 2 components of I.

a. internal validity — the lack of systematic (non-random) error or bias of epid. bias

b. precision — lack of random error; which is the departure of any given estimate from the mean value (if study were repeated, or reliability — internal

III. there are subtle differences between statistical bias (& validity) and epid. bias (& internal validity)

A. both of stat. bias — considers in terms of CI of pt. est.

$$CI_\gamma[\hat{R}] = \hat{R} \pm CV_\gamma \, SE[\hat{R}]$$

- $CI_\gamma[\hat{R}]$ — reflects: error V. (I)
- $\hat{R}$ — validity or lack of stat. bias.
- CV_γ — power of test $(1-\beta) \cong$ Pr of rejecting false H_0 → degrees of freedom
- $SE[\hat{R}]$ — stat. precision — homogeneity of $\hat{R}$ within groups ~~i.e. diff. since of $\hat{R}$ from $E[\hat{R}]$~~

NB. — strat. is efficient only if gain in precision is not outweighted by loss in df.

stat. efficiency

random error — departure of $\hat{R}$ from $E[\hat{R}]$ — reflected in size (width) of CI

see addendum (A)

B — lack of statistical bias ⇒ if study were done an ∞ # times, the est. of the effect (or any parameter) would produce an accurate assessment of the pop. parameter (R) — i.e., "true" value of effect

i.e., $E[\hat{R}] - R = 0$

NB: St.B is dept. on ~~the~~ n

C. — lack of epid bias ⇒ if study were expanded to ∞ size (n=∞) it would provide accurate est. of R

i.e., $\lim_{n \to \infty} \hat{R} = R$

NB. doesn't depend on n — (≡ stat. consistency)

(me) if we sample entire pop, bias must be 0 since every estimate of parameter (w/ n = N) will be same - but this does not address issue of confounding that may be operating - even in total pop. ⇒ want stat. consistency

D ∓ O.M. chooses this defn. to use for internal val. because:
- a. epid. bias is conceptually clearer
- b. its not dept. on n & ∴ reflects the type of bias that is most prevalent & important in epid. studies - he notes that St. bias can often be corrected in analysis by approp. transformations

(III A) ? E. ∓ stat. efficiency is measured by

$$1/V[\hat{R}] = 1/SE[\hat{R}]^2$$

where: $SE[\hat{R}]$ is a function of the "deviation"

$$\hat{R} - E[\hat{R}]$$

F. ∓ epid. precision is basically same thing except, unlike the other, it doesn't isn't defined in the context of unbiasedness — epid. precision is taken measured indep. of validity

IV. Epid. Precision — 2 components

A. Size-efficiency ⇒ I / subject

B. Size — ie, # subjects — gen'l principle: the larger the n (everything else being =), the more precise an estimate is

1. yet in practice, we often think in terms of an optimum size — decide a priori how large an effect is worth detecting & the degree of precision desired ($\alpha + \beta$) — then compute n according to stat. model
2. OM argues this is not optimum size because we don't know the benefit of I that we might get from study — this would involve putting $ value on study results which is impossible
3. ∴ only optimum sizes are 2 extremes:

3. a. $\beta = \alpha = 0$ — if no info is afforded by the study — implies that Val. + Prec. are not to be traded-off for one another — ∴ unwarranted to pursue high Prec. in study that may be quite invalid

b. $\alpha = \infty$ — very high prec. is necessary to "prove a H_0" (i.e. effect = 0) — so w/ any $\alpha < \infty$, all we can do is disprove H_0's, not prove them

V. Internal Validity — Systematic error — 3 components ∴ 3 types of Bias

A. Selection (often undefined) — when response var. of study influences the selection of subjects in such a way as to bias $\hat{R}$ — in:

(specific type of Bias)

1. C-H study — since D cannot influence selection of subjects, selection is not an issue w/ prospective studies (but can in retrospective studies)

2. C/$\bar{C}$ Study — E (response) influences Pr of selection of subjects, but differentially among cases + controls — otherwise no bias — eg, Berkson's fallacy, where differential admission rates into study (via admitting from different diseases) may bias results

(me) except in certain circumstances: we select subjects w/o existing dis. + follow-up for dis. occurrence — Selection bias exists if large % (>20%) of source pop. have dis. (+ thus (unelig.) + dis. prev. is assoc. w/ dis. incidence — ie, sick people are most likely to get sick in future (Lawrence, Public H Reports 63: 1507-21, 1948)

3. X-S study — influence of D on selection of subjects → need not be differential — only that

$Pr[\text{selecting } C] \neq Pr[\text{sel. } \bar{C}]$

$\left(= \frac{\#C \text{ in Study}}{\#C \text{ in Source pop}}\right)$ " "

if $Pr[C] > Pr[\bar{C}] \Rightarrow$ bias toward null — of PR_0 OR_0

" $<$ " $\Rightarrow$ bias away from " "

N.B. 2 Pr's would be equal under conditions of perfect Simple random sampling (w/o strat.)

B. Information

1. non-differential bias — error occurs identically in all compared groups — net effect is to bias estimate toward null — not spurious assoc. — (eg) in C/$\bar{C}$, we overestimate % of exposed persons, but equally in C + $\bar{C}$ ∴ $\hat{RR} \to 1^0$

N.B. this occurs whenever: Se ≥ 1 − Sp (almost always)

a. misclassification (of E or D)
b. non-particip. of selected subjects (w/o obs.)
c. loss to follow-up (migration, uncooperativeness, death)

2. differential bias — bias in different between compared groups — results in spurious assoc. — eg over-estimate of E is different in C than controls — eg, lung Ca + Smoking

C. Group Comparison — 2 issues

1. appropriateness of present group (planning stage)
2. confounding factors — (planning + analysis) — conf. is when the effects of 2 D correlate get mixed so that the apparent (crude) effect of one actually includes the other — i.e. assoc. of E + D is spurious owing to the effects of other var's that have not been separated thru statistical analysis —

Crude effect indicty true (unconf.) effect —
if CE is closer to null than TE → "−" conf.
" " " further from " " " → + conf.

note change RCE

We will go into to more detail later + measure the magnitude of conf. effect in order to know when to control.

V. Preferability — combining Val. + Prec. into overall measure of I (both rand + non-random error)

$$I = \frac{1}{B^2 + V[\epsilon]} \qquad \epsilon = \text{random error}$$

but, B + ε are not to be traded-off

VI. Optimality — combining feasibility + I.

A. might assume common units — eg, cost
everything being =, a larger study provides more info
but is costlier —
but optimal design is not identifiable, because we
have no basis for determining whether a more inform.
study is worth the added cost → simply because
we can never place $ figures on the future consequences
of a study that has not been conducted —
but there are some ("rational planners") who devise
complex methods for doing benefit-cost analysis
under these premises ⇒ "Social benefit functions"

B. in practical terms — we should maximize I per cost — i.e.,
consider cost effic → I/$; not trad-off th. 2

VII. External validity — extending empirical results to persons
outside range of source pop. — i.e., generalizability —
moving from particular experience to scientific generalizations
(research to theory) — involves judgmental inferences
about observed results

A. external vs. internal — p. 17
1. types of inference — p. 13-14
2. 3 pts — p. 14-16

B. Again, no trade-off between internal + ext val. —
[illegible] is necessary for the latter.

IX. Next chore is to see how these concepts of E. are considered methodologically — i.e., in context of specific study design — but first we must find some systematic basis for classifying research designs

A. general perspective of epid. pursuits which includes quantification of disease occurrence in groups of individuals — p. 2-3 #6 — 3 types of studies

B. look at analytical (obs.) studies in more detail — division of analyt. studies into 3 types: CH, C/C, X-S, is an over-simplification that might be useful when taking 160 but certainly presents limitations when studying study design in more depth —

C. next time, we'll look at study design in terms of 3 dimensions & combine them to form a typology of study types. — NB: flow diagrams, (#7) — 1 dimension: directionality

11/12 Outline

I. Intro'd.
   A. Handouts:
      1. 2 page replacement for pp 10-11 (#6)
         Types of Study designs
      2. EM & I (#9)
   B. today: dimensions & types of epid. study designs & their methodological implications

II. Underlying task of today's discussion is the classification of design methodologies
   A. broad view of epid.: includes quantification of D occurrence in pop's. — See #6, pp. 2-3 —
      3 types of studies, differ by objectives
      1. Observational
      2. Experiental
      3. Evaluative
   B. focus on Analytical (Obs.) to test etiological H.
      1. traditional division into 3 types (CH, C/C, X-S) may be fine for an intro'd to Epid, but presents limitations when studying design in more depth — eg, what effect measure to use
      2. will try another approach that begins by defining study design (Obs.) into 5 dimensions
         a. directionality — flow diagrams
         b. timing
         c. Subject selection + sampling
         d. temporal arrangement of observations
         e. Units of measurement

See other outline p. 5 (#6)

III. Typology of Study Designs — based on the 5 dimensions — desire categories of study design that may be treated as methodologically ~~separate~~ distinct — See other sheet (just handed out)

~~A~~. Longitudinal
- 1 A. Forward
  - 1. Cohort
  - 2. Panel
- B. Ambidir.
- C. Non-direct.

II. Case-reference
- A. Backward (case-history)
- B. Ambidir.
- C. Non-direct. (X-S)

III. Cross-Sectional — nondirect.

IV. Major methodological considerations for selecting a particular Obs. Study design (based on O&A., Ch. 3) — divide issues into 3 categories
- A. Type of Info. Sought
- B. Internal Validity
- C. Feasibility of Study

(prev.) Somewhat different than the outline already handed out (#6, pp. 13-15)

A Typology of Study Designs in Observational Research Involving Etiologic Hypotheses

I. Longitudinal study - observations of D are made at more than one point in time

A. Forward directionality (follow-up)

1. Cohort study - calculation of incidence rates

a. Type IA (selective) sampling of exposure groups - either prospective, retrospective, or ambispective

i. fixed cohort - I_C (or I_D)

ii. dynamic cohort - I_D

b. Type II (nonselective) sampling of a source population - usually prospective

i. fixed cohort - I_C (or I_D)

ii. dynamic cohort - I_D

2. Panel study - calculation of a change in prevalent status (of the condition - D) - usually prospective

a. Type IA (selective) sampling of exposure groups - (see: I.A.1.a.)

b. Type II (nonselective) sampling of a source population - (see: I.A.1.b.)

B. Ambidirectional study - Type IB (selective) sampling of disease groups within a cohort - retrospective or ambispective

1. Fixed cohort

a. "density type" - subjects are "observed" within the risk period (of D) - for chronic disease - I_D

b. "cumulative type" - subjects are "observed" at the end of the risk period (of D) - for acute diseases, recurrence, complications - I_C or I_D

2. Dynamic cohort - "density type" - I_D

I. C. Non-directionality (momentary)

1. Multiple, cross-sectional study - two or more cross-sectional surveys (see: III) on a dynamic population, over time - calculation of changes in prevalence rates - usually Type II sampling

(case-control)

II. Case-referent study - observations of E and D are made concurrently; with Type IB (selective) sampling of disease groups - retrospective or ambispective

A. Backward directionality ("case-history" study)

1. use of prevalent (i.e., existing) cases

2. use of incident (i.e., new) cases

a. "density type" - subjects are "observed" within the risk period (of D)

b. "cumulative type" - subjects are "observed" at the end of the risk period (of D)

B. Ambidirectionality - (see: I.B.) - i.e., a case-referent study within a cohort that is followed up

C. Non-directionality (momentary)

1. use of prevalent (i.e., existing) cases

2. use of incident (i.e., new) cases - actually a "sequential cross-sectional study" ~ "density type"

? cum. type?

III. Cross-sectional study - observations of E and D are made at one point in time - non-directionality

A. Type I (selective) sampling

1. Type IA sampling of exposure groups

2. Type IB sampling of disease groups - (see: II.C.)

B. Type II (nonselective) sampling of a source population

start

IV. Considerations for Selecting a Study Design in Observational Research

3 criteria: I. Type of Info. sought
II. Internal Validity
III. Feasibility of Study (cost)

I. Type of Info. sought
A. Substantive focus – # (+ type) of primary Var's to be studied (E+D)
1. specific focus – one E and 1 D – i.e. all of the study designs; espec. ambidir. – main interest
2. Semi-specific
a. Several E's, 1 D – all designs, esp. those w/ Type II sampling – also appropriate to use model fitting techniques (fishing expedition) Phil's
b. 1 E and several D's – follow-up study of X-S study w/ Type IA sampling (fishing expedition) – also ambidirectional
3. non-specific – several E's and D's – cohort and X-S w/ Type II sampling – the main idea should be to narrow focus

if we don't have to resample for each new disease by sampling the entire cohort but just noncases

B. Type of Ho unit of measurement
1. • ecologic fallacy – can't attribute to indiv. what we know about groups – also
• individualistic fallacy – can't attribute to groups what we know about indiv. (bias against this concept in epid.) – in paper → paper #6
• conclusion: serve different purposes (ecolg. vs. non-ecol. indiv. analysis → etiology of a dis.
ecol. analysis → programmatic/policy decisions re: the use of preventive action in a given pop.

I. B. 2. O.M (Ch. 3) alludes to this division of objectives by distinguishing:

- Scientific Study Orientation – we attempt to generalize Re: causal determinants of a D. – the findings are based on a source pop., they are not made dept. on the distrib. of the pop. w/ re: other D correlates – (RR)

N.B.: he doesn't define Scientific by the "methods" used – cf. the Scientific method

objectives use Sci. method

- particularistic Study Orient. – aim is to improve the Health status of a particular pop. – ∴ empirical estimates are made dept. on distrib of pop. (eg. by age) – EF (PF) are partic. questions – best to use Type II Sampling (simple or stratified)

NB: refers to external validity, not statistical inference

3. Iye extended this dichotomy (#6, p. 20) b domains of Ho generation – i.e. not only a need to distinguish b domains

b. as PH professionals, we might be more concerned w/ intervention, which requires different Ho's + correspondingly Study designs

views of Renwick, Terris

c. in order to do this, we must put more emphasis on evaluation of public programs and complete record systems

C. Parameter estimation (of assoc.) – (over) by type of Study

C. C. Parameter estimation by type of study

Appropriate Measures of Association by Type of Study Design

I. Longitudinal

A. Forward directionality

1. Cohort study (a. or b.)

a/b. Type IA or II sampling = (forward)

i. fixed cohort — I_eR, I_eD (I_e) especially in acute diseases, complications, recurrences & diseases w/ short induction periods; because I_{Dj} does not remain constant over follow-up (eg, infectious dis. epidemic, congenital disorders, or maternal complications during or following pregnancy) — but often can derive I_DR, I_DD (I_D) $\frac{EF}{\%} \; I_{Dj} = \text{const}$

ii. dynamic "cohort" — I_DR, I_DD (I_D) → can derive I_eR, I_eD — also EF, PF

(1) call it: "dynamic pop."

(2) or "running cohort" (when all subjects who enter are followed to death or some desired event)

Since dyn. cohort provides representative cases; while cases in fixed cohort are not repres. Since the cohort ages (N.B. original cohort is repres.) (~~O.M.~~ O.M. (ch. 3) gives a method for making appropriate adjustments to estimate EF or PF in a fixed cohort

Chronic diseases when I_{Dj} remains constant in follow-up

2. Panel Study — non-categorical procedures (eg, difference in means, model fitting) — can be used to derive certain measures of effect depending on the statistical model — N.B. statistical nature of analysis is different (eg, Mayne's study: ΔBP → CHD; Δsmoking → ΔBP)

* problem (eg) Lung Ca. ~ age-adj. Rates increased by 250% (1950-70)

I. B. Ambidirectional study – type IB sampling

1. fixed cohort

a. density type – I_cR, I_cD (I_c) – if ref. group (R) is a random sample of the cohort – but really want: I_DR, I_DD (I_D), EF, which we can get by adjusting # aggregated persp. (O.M. ch. 3)

b. cumulative type – I_cR, I_cD (I_c) – if ref. series is random sample of cohort – for acute dis.

2. dynamic cohort – I_DR, I_DD (I_D), EF & PF – if ref. series is random sample of cohort (NB: simpler computations than w/ fixed cohort)

show ex. fixed CH. (#1 s

C. Non-directionality

1. Multiple X-S – ΔP (PD) by subgroup of E not very useful etiologically because pop. changes between X-S studies – better for secular trends of D

II. Case-referent study – ("case-control") – See App. C

A/C. Backward ("case-history") or nondirectional (X-S)

	D	$\bar{D}$
E	a	b
$\bar{E}$	c	d

1. prevalent cases –

$$OR \cong \frac{ad}{bc} = OR_p = \frac{P_E(1-P_{\bar{E}})}{P_{\bar{E}}(1-P_E)} \quad \left(OR = \frac{ad}{bc}\right)$$

$$I_DR = OR_p \frac{d_{\bar{E}}}{d_E} = OR \frac{d_{\bar{E}}}{d_E}$$

So: $\hat{I_DR} = OR$ if $d_E = d_{\bar{E}}$

Note: assumption is not good if same factor (E) is a cause of disease & great and prognostic indicator of case-fatality – e.g. alcohol consumption → angina

Cornfield (1951) → ok if D is rare so that $OR_p \cong I_cR \ (\cong I_DR)$

Ambid. Study – Example, Calculation of $I_c R + \hat{I}_c R$

fixed cohort – cum. type

(non-observed quantities in parentheses)

(over – flow diagram)

	D	$\bar{D}$	
E	I_E	$(N_{E\bar{D}})$	(N_E)
$\bar{E}$	$I_{\bar{E}}$	$(N_{\bar{E}\bar{D}})$	$(N_{\bar{E}})$
	I	N_D	$N'_{\bar{D}} = N$

take sample from $N'_{\bar{D}} (= N)$ of size n

estimate $Pr[E] = \frac{N_E}{N}$ from this sample

$\hat{Pr}[E] = n_E / n$ where n_E = # in sample who are exposed

$Pr[\text{being } E+D] = Pr[D|E]\, Pr[E] = Pr[E|D]\, Pr[D]$

$$Pr[D|E] = \frac{Pr[E|D]\, Pr[D]}{\hat{Pr}[E]} = I_{cE}$$

$$I_{cE} = \frac{\left(\frac{I_E}{I} \Big/ \frac{I}{N}\right)}{\frac{n_E}{n}} = \boxed{\frac{n\, I_E}{n_E\, N}}$$

Similarly: $I_{c\bar{E}} = \dfrac{Pr[\bar{E}|D]\, Pr[D]}{\hat{Pr}[\bar{E}]} = \dfrac{\frac{I_{\bar{E}}}{I} \cdot \frac{I}{N}}{\frac{n_{\bar{E}}}{n}} = \boxed{\dfrac{n\, I_{\bar{E}}}{n_{\bar{E}}\, N}}$

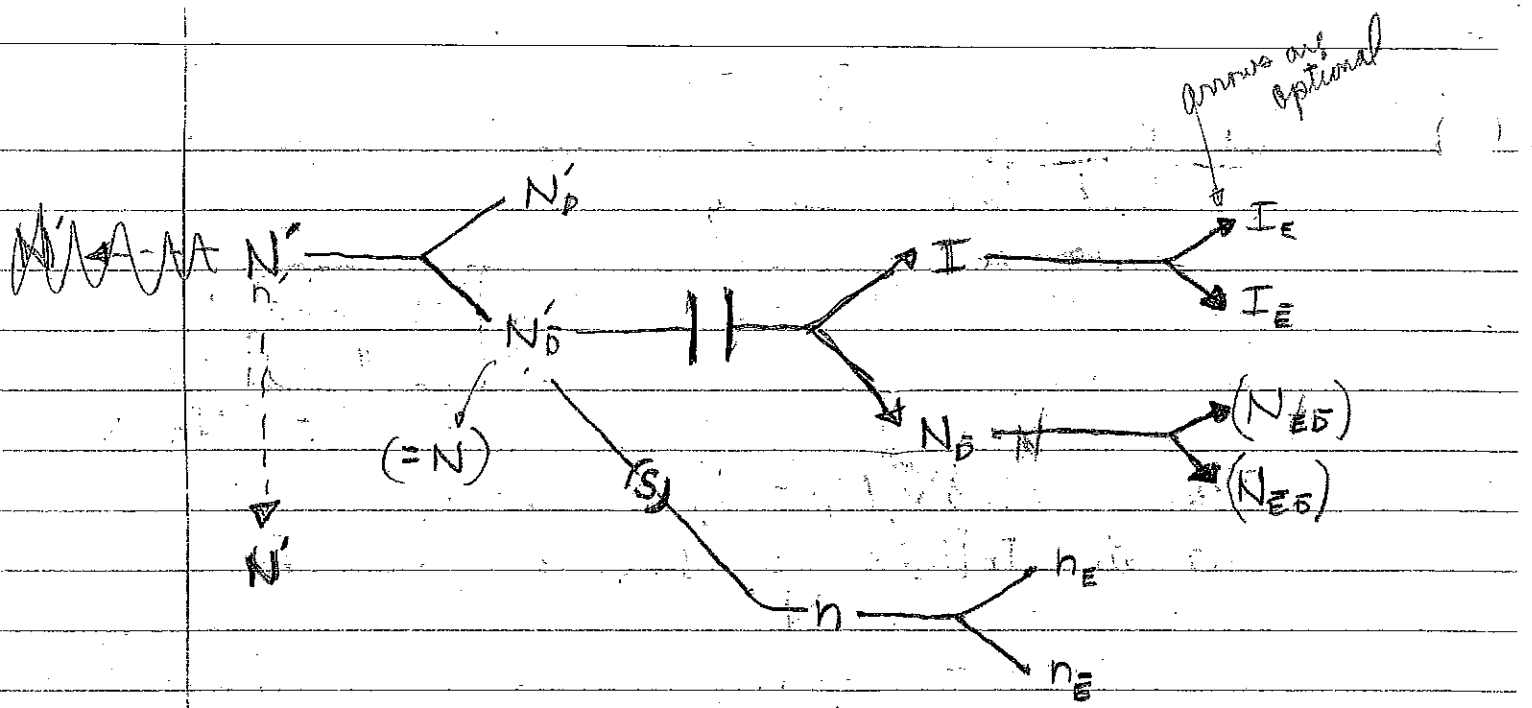
$$\hat{I_c R} = \frac{Pr[E|D]\, \hat{Pr}[E]}{Pr[\bar{E}|D]\, Pr[E]} = \frac{n I_E / n_E N}{n I_{\bar{E}} / n_{\bar{E}} N} = \boxed{\frac{n_{\bar{E}}\, I_E}{n_E\, I_{\bar{E}}}}$$

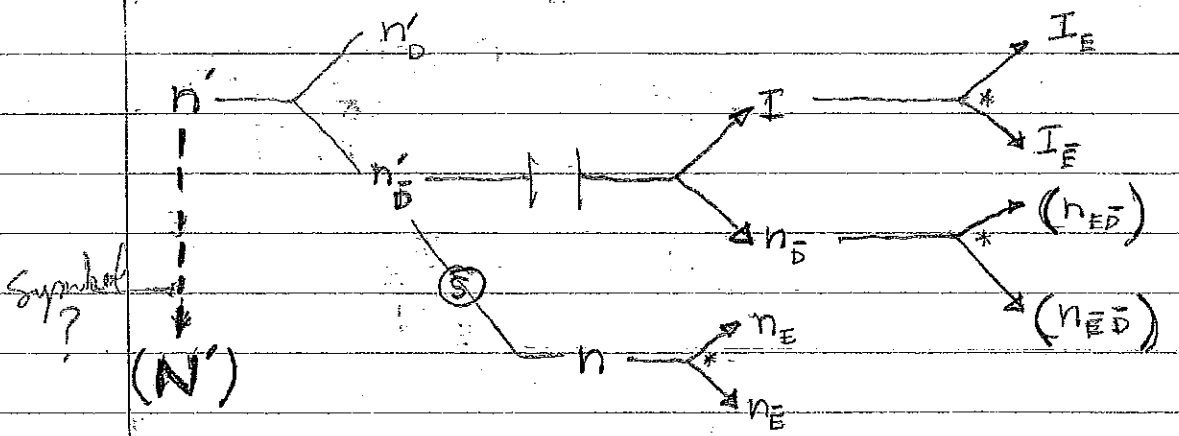
Since $\hat{Pr}[E]$ is an estimate, can get CI of $I_c + \hat{I}_c R$

see Kupper et al., 1975 J ASA (70(351): 524

(also can do stratified analysis)

See: II N/c for dyn. cohort





* – not nec. arrows

( ) – not actually obs.

C-R study

II. A/C. 2. Incident cases

a. density type –

$I_D R = OR_p = OR = \frac{ad}{bc}$ (already shown)

if pop. is dynamically stable

also can get I_D for E and $\bar{E}$ if we have an estimate of I_D (for entire source pop.)

$$\overline{EF}_N \equiv \left(\frac{RR - I_D}{RR}\right) e_D = \frac{I_D - I_{D(\bar{E})}}{I_D}$$

(RR can be standardized)

Can get w/o "I_D"

$$= 1 - \frac{I_{D(\bar{E})}}{I_D}$$

[prove, when I have time]

$$\therefore \frac{I_{D(\bar{E})}}{I_D} = 1 - \overline{EF}_N$$

$$I_{D(\bar{E})} = I_D(1 - EF_N)$$

$$I_{D(\bar{E})}(I_D R) = I_D(I_D R)(1 - EF_N)$$

$$I_{D(\bar{E})} = I_D(I_D R)(1 - EF_N)$$

(OR) $I_{D(\bar{E})} = I_D / (1 - PF_N)$

$$I_{D(E)} = I_D(I_D R)/(1 - PF_N)$$

prev. of D in source pop. (t_1)

	D	$\bar{D}$
E	A	B
$\bar{E}$	C	D

but, sometimes problem in estimating I_D from other source any t_0

because denominator is taken as both persons w/o disease at risk + B prevalent cases

$$I_D^* = \frac{(a+c)}{(A+B+C+D)(t_1 - t_0)} = \frac{\text{total cases}}{\text{p-yrs}}$$

(A+B+C+D) = prev. cases

II. A/c. 2. a. it should be:

$$I_D = \frac{a+c}{(C+D)(t_1 - t_0)}$$

(A susceptible)

So: $I_D = \frac{I_D^*}{1 - \hat{P}_{(N)}}$

prev. rate ($\hat{P}$) may be known from another source or even from ref. pop. & Series P if it was drawn w/o exclusion of prevalent cases

also, non-specific assessment of E (did you ever smoke?) would not good for getting RR (since it will tend toward 1), but since E use also increased, this type of exposure history is good for getting EF (esp. if you divide smokers into groups - by # pack-years)

$\Phi(EF)$

b. Cumulative type: "risk odds ratio"

Incid. in source pop.

	D	$\bar{D}$
E	W	X
$\bar{E}$	Y	Z

$$OR_R = \frac{I_{c|E}(1 - I_{c|\bar{E}})}{I_{c|\bar{E}}(1 - I_{c|E})} = \frac{\frac{W}{W+X}\left(1 - \frac{Y}{Y+Z}\right)}{\frac{Y}{Y+Z}\left(1 - \frac{W}{W+X}\right)} = \frac{WZ}{XY}$$

(can get this)

$$= \frac{W/X}{Y/Z} \left(\approx\right) \frac{a/c}{b/d} = \frac{ad}{bc} = OR$$

(cases ↑, controls ↓)

if D is rare so that $\frac{(1 - I_{c|E})}{(1 - I_{c|\bar{E}})} \approx 1$,

$$OR = \hat{I}_c R \left(= \hat{I}_D R\right)$$

in order to get risks for exposure groups, we need I_c of source pop. from other sources or the study itself — OM (Ch. 3) describes procedures for estimating $I_{c|E}$ and $I_{c|\bar{E}}$

$I_{c|E}$ $I_{c|\bar{E}}$

See addendum for different CR design

III. Cross-sectional study

A. Type I Sampling

1. Type IA (E groups) — PD (or PR - not good)

$OR = \hat{OR}_p = \hat{I_D R}$ if $\bar{d}_E = \bar{d}_{\bar{E}}$

+ distributions are same

but $OR \neq I_D R$, unless d is very short for both groups

2. Type IB (D groups) — eD

Similar to case-referent study (see: II. A/c)

if E status for each person is stable over time

also good for testing validity of a screening detection test — Se, Sp, P_V — in fact, any case-history (backwards) may be used to test validity of screening test based on risk factors

B. Type II Sampling —

correlation coefficients — ϕ w/ categorical data, or

1. or PD or PR or OR_p — (ρ or τ if ordinal)*

Summary

1. Conceptually, methodologically simplest study is density type ambidirectional study — w/ complete ascertainment of cases over a period of time in chronic dis
2. Other studies affording complete estimability of all parameters; follow-up studies — except Type II Sampling tends to be less efficient, the representative of a specific pop.

usually we say:
C-R → can't determine antec./conseq.
CH — can " " "

I. D. Temporality Issues

1. antecedent/consequent determination — theoretically possible from backwards, forward or ambidir. I not from X-S — but even in a follow-up study, determination of temporal relationship may not be established — eg E (as a Δ in status) is measured over the same period of follow-up as when new cases of D are detected → ie, longitudinal nature of ...

eg) ΔSES → CHD if both from 1960-67, 2 problems: (1) temporality ambiguity (2) reduction of effect because of lag time

On the other hand, temporal ambiguity can be largely avoided in a case(-) (history study) (backward direct.) by using new (incident) cases

2. period of follow-up must be selected carefully
eg, if p = ∞ (100 years) — RR → 1 — chronic dis
p ≅ 0 — RR → 1 — for acute dis. or CHD clinical trials (thrombo embol)

eg, if we attempt to observe thromboembolism w/in 6-mo. after taking oral contraceptives, might show no effect because not enough time for D to develop — * but early cases (≈1 mo. after) even in E group do not bias final results only reduce efficiency

1. principle for etiologic (scientific) research: be highly selective of ~~both~~ all comparison groups
2. matching — to be covered later

E. Subject Selection (restriction)

II. Internal Validity — lack of bias in estimation of effect

A. Selection bias — response var. influences selection of subjects in a biased way

1. follow-up study — selection is not an issue if follow-up starts after admissibility criteria are met — NB: may be a problem in retrosp. CH, etc.

3. lapse to follow-up or lack of participation by part of study pop. (exceptions:)
1. remitting/chronic dis. P is assoc. w/ I & P is large — sick people are more likely to get sick in future
2. or (chronic) dis. in which outcome is death (if "some" cases) and cases are included in study — (P is assoc. w/ death)

P's both ... indicate ... case-def. ... (prevalent) ... (positive) ...

2. Case-ref. study — E influence Pr of subject selection but differentially among cases + controls — eg, hospital based study of pill → thromboembolism — it's possible that the use of the pill causes physicians to be more likely to suspect a patient of

II. A. (2.) having thromboembolism (i.e., pill is used as diagnostic evidence of T) → thus "cases" are more likely to have used the pill — also Selective Survival of cases in case-referent study may cause bias if the surviving cases are more (or less) likely to be "exposed" than those who died —

NB: [this kind of selective bias can be sharply reduced by using new cases

possibility of selection bias can often be ruled out judgmentally e.g. Coffee → MI OM: found RR > 1, for selective bias to be inflating effect, MI patients who died must have drank less coffee than cases who were in study → ? not likely

3. X-S study — (a) influence of D on selection of subjects but not necessarily differentially

(see addendum B)

e.g. if Pr [selecting case] ≠ Pr [Selecting $\bar{D}$] there will be bias — also selective survival

(if getting Pr FD)

NB: bias enhanced because P rates are dept. on incidence of dis. as well as duration, which also might vary by E group)

$\bar{d}_E = \bar{d}_{\bar{E}}$ → more likely $Pr[D] \neq Pr[\bar{D}]$

i.e., exposed (or $\bar{E}$) cases are more likely to die + thus not get in study

C-R study (b) influence of E on selection of [subjects] (must be differential) for cases + noncases

$Pr[E] \neq Pr[\bar{E}]$ but different for D + $\bar{D}$

B. Information Bias

1. non-differential

but if getting PR, must be differential

a. mislabelling — can be a source of error in any type of study, esp. retrosp. because of incomplete or inaccurate records — yet

b. lack of participation — e.g. w/ non-retrospective data, esp. when using mailed questionnaires or when compliance with researchers is an inconvenience — can't be more of a problem w/ Type II sampling

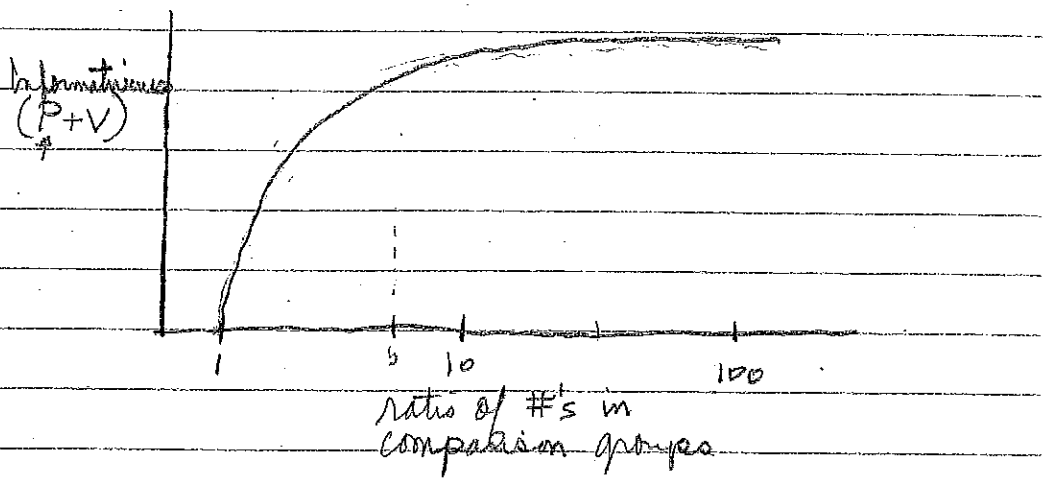
c. loss to follow-up — in any follow-up study, esp. prospective — if too hi (as in b) + reasons are unknown, study is invalid

II. B. 2. differential measurement bias — can be source of error in any study, but esp.
a. Case-history study (backwards) — eg., biased recall of exposure as in Lung Ca. & Smoking.
b. non-retrospective study (prospective, ambispective or X-S) — ie primary data observations of researchers are influenced by what results they expect
C. Comparison bias & confounding
1. choice of referent (Comparison) series
a. in follow-up study w/ rare E, referent group can be Gen'l. pop. — * not likely in case-ref. study in which D is rare unless E was known in pop.
b. generally not good to form a pooled referent unless a logical "unexposed" group does not exist — in which case it is best to assign ref. group a priori, before data is collected
2. confounding factors — may be a problem in any type of study design — esp. in Type II sampling of a source pop. — ∴ better to restrict study pop.

III. Feasibility of the study (cost-efficiency)
A. Freq of E in source pop. — extremely common or rare exposures (Prate) ∴ lend themselves best to follow-up or X-S studies w/ Type IA Sampling —
B. freq of D in study pop. — extremely common or rare dis. lend themselves best to Case-ref. Studies

III c. length of induction period of D — chronic dis. w/ long induction periods ~~are best~~ lend themselves best to retrosp. cohort studies or case-ref. studies (not prospective follow-up studies)

D. cost of getting E data — advantage of ambidirectional study (relative to follow-up) is that we don't need the exposure history on as many subjects (i.e. only a sample of the non-cases in the cohort) — N.B. never gain much efficiency by having groups in ratio of > 4 or 5 (controls on E) to 1 (case on E).



Addendum A $\widehat{OR} \neq I\hat{c}R$

CR: normally done:—

	D	$\bar{D}$	
E_1	a	b	m_1
E_0	c	d	m_0
	n_1	n_0	n

$\widehat{OR} = \frac{ad}{bc} \cong I\hat{c}R$ if D is rare (because of Bayes theorem)

$\widehat{OR} = I\hat{d}R$ if $d_1 \cong d_0$ if prev. cases, no assumpt. if new "

CR: alternative design:—

	D	PAR
E_1	a	n_1
E_0	c	n_0
	m_1	n

— i.e., sample total pop, not non cases

$$\underline{\widehat{OR} = \frac{(a/n_1)}{(c/n_0)} = \frac{a/n_1}{c/n_0} = \underline{I\hat{c}R}}$$

So might use this design when:

① disease is not rare

especially, ② cumulative type study (acute conditions)
i.e., sampling takes place after risk period
So don't want $I\hat{d}R$ but rich. ratio

Otherwise, it would be best to get $I\hat{c}R$ from $I\hat{d}_j$ (j = age categories) using Id (cohort) study or nested study

NB: Sampling PAR may be:

① easier than sampling $\bar{D}$, +

② ∴ might not introduce same kind of selective bias as common in CR study

12/3 Outline

I. Introd. — same time next semester; begin 1/12; same room; change pace
   A. Handouts:
      2. #11, Summary of Quant. Procedures — Conf, EM, I
      1. #10, Elements of Statistical Analysis: χ, CI
         " meat of Minn. Course" — note: Table 1, pp.
   B. today: methodological implications of study designs

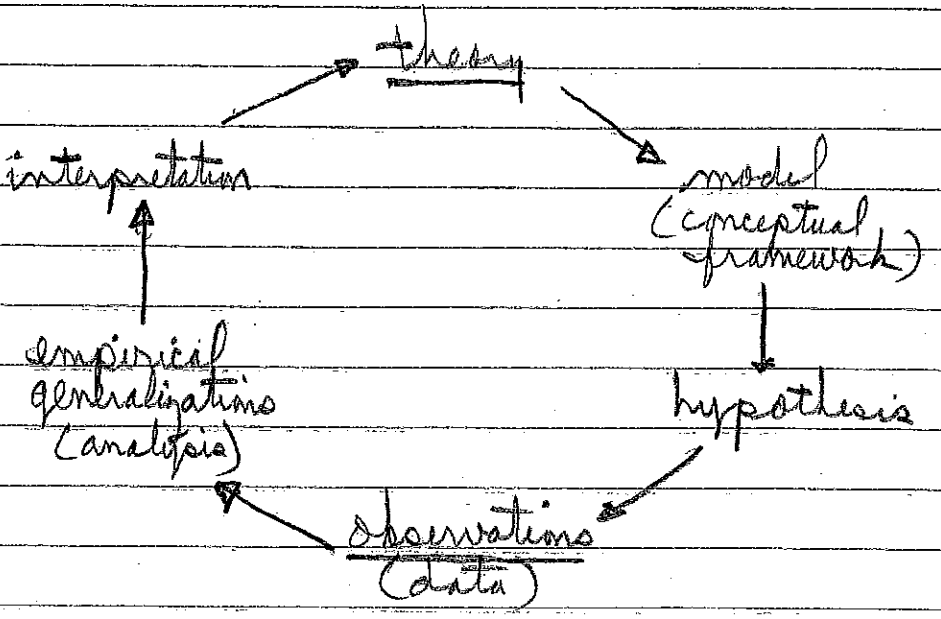
II. Outline (on board) — go to outline: 11/12
   Considerations for Selecting a Study Design in Obs. Research
   I. (3 criteria)
   I. Type/nature of Info. Sought
      A. Substantive focus
      B. Type of Hypothesis
      C. Parameter Estimation
      D. Temporality
      E. Subject Selection (restriction)
   II. Internal Validity (lack of bias)
      A. Selective bias
      B. Information bias
      C. Comparison bias - confounding
   III. Feasibility of Study (cost-efficiency)
      A. Frequency of E in Source pop.
      B. Freq. of D in Study pop.
      C. Lag period of D
      D. Cost of getting E data

1/14/76 Outline + 1/21

I. Intro.
A. would like to finish course (lectures) by March before OM visits (3/24)
B. handouts
#10. Basic Principles of Categorical Analysis: H_0 testing + interval estimation (CI) – heart of Mini. course
#11. A Summary of Quantitative Procedures Used to Identify & Measure Conf., EM, + I
#12. Matching and Matched Analysis in Obs. Studies
#8 (brief revision) Methods of Controlling for conf. var.
C. todays topics
1. principles + vocabulary of causal inference (epid. standpt.)
2. confounding

II. Causal Inference
A. purpose of this discussion is to propose a common language (vocab.) for establishing etiological relationships in epid. research
B. must distinguish 2 basic levels of scientific inquiry
1. theoretical level – involving biological/behavioral mechanism of causation – think in terms of expected changes
2. Empirical level – statistical associations among variables in one's data – think in terms of observed changes (covariation)
C. any scientific process must bridge the gap between these 2 levels – often simplified as the "scientific method" (over)
often causes conceptual + semantic confusion

Scientific Process – overly simplified cyclic conceptualization



Epid. H's are generated on the basis of models or conceptual frameworks which are based on existing knowledge of the subject area as well as "intuitive leaps" or hunches.

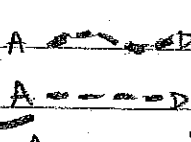
A model is a translation of postulated causal mechanisms; while emp. gen'l. are a translation of analytics of statistical relationships in the data. Thus, study results are analyzed and interpreted, thereby leading to a revised theoretical understanding.

Some confusion exists between modelling & emp. gen'l., largely because of the lack of specificity in denoting theoretical postulates and statistical associations – thus a need to adopt common vocabulary –

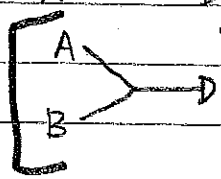
II. D. Begin by defining 4 conceptual levels of assoc. between 2 var. (E x D)
See p. 3 (#9) – hierarchy of assoc.; each level includes evidence for the levels above it. – i.e., graduation from data to theory

1. define 4 levels
2. define 3 types of inference needed to bridge gap between 4 levels
3. define graphical symbols –

A ~ D stat. correlate
A —— D risk (rate) indicator
A ——▸ D risk factor (or broadly, risk ind. w/ temporality)
A ═══ D dept. measurements – same construct



Spurious assoc. resulting from differential measurement errors



A + B are dependent risk indic. or factors

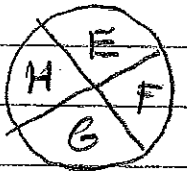
(no good)
From an empirical point of view perhaps theoretically meaningful

4. From analytical viewpoint, philosophical influences are irrelevant. We can only hope to establish that statistical correl. are unbiased (i.e. risk indic.) and build an argument that they are risk factors of D – still probabilistic statements – i.e. stochastic view of causation

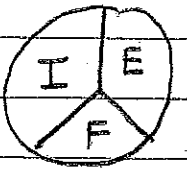
E. One procedure for modelling etiologic theory was presented at Minn. – involving concepts of "sufficient" conditions.

Modelling is the methodologic link for generating
Epid. 11-05

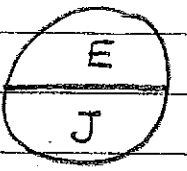
II. (eg.) E.



or



or



→ D

I. II. III.

? any value

3 sets of sufficient conditions
each circle consists of a set of dependent conditions (ie. D results from interplay of all factors)
each set of suff. cond. is indep. of other sets
any factor (eg, E) which is present in all suff. sets, is a necessary condition

F. Crucial point is that we recognize disparity between theory & data; & between models and imp. gen'l. — there is no consistent 1:1 correspondence between biolog./bel. mechanisms of causation & statistical assoc. — made more evident in coming weeks —
(eg.) we must be very specific w/ regards to graphical symbols ⇒ E → D
What does this mean?
Does it refer to a postulated causal mech., or one or more sets of empirical evidence
(N.B.) path analysis
establishing "causal" associations is the essence of analytic epid. & is what distinguishes us from biost.; et al.

* perhaps before talk FM

defn. of indep. risk indic. — E is an indep. risk ind. of D if it is assoc. w/ D in (every) category of C
E is a risk indic. if it is assoc. w/ D in at least 1 category of C — E is a dept. risk indic. if there is statistical interaction

← not same "thing as "main effect?" but a qualitative defn. (categorical analysis)

III. Confounding – most of #8 isn't worth reading except pp. 1-15; revised outline

A. def. – the heart of epid. methodology

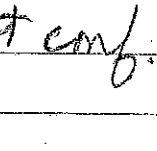
A defn. – conf. is the mixing of two or more effect risk indicators of a given disease – ie, the effect of E on D is confounded by C when the 2 effects have not been separated thru control methods involved in subject selection or statistical analysis – assuming we are primarily interested in the effect of E, we would like to know whether the observed (crude) effect is distorted by C.

Objects:
(1) identify conf. factors
(2) measure their degree of conf.

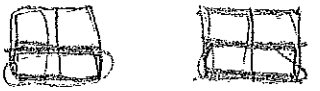
Conceptual B. requirements of a single confounder C:

1. assoc. must exist in the data
   a. C ~ E
   b. C ~ D

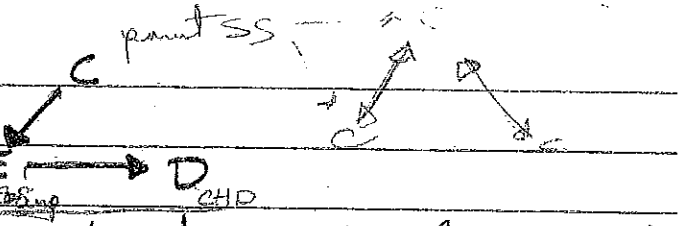
2. + further specifications – (outline on board)
   [1 primary / 2 Secondary]
   a. both assoc. must exist in the data (Study pop.) – ie, conf. is an empirical condition not an inherent property of certain variables
   b. assoc. between C & D cannot be spurious
   C must be a risk indicator of D – ↓ in a prediction sense
   ∴ not conf. if:



   c. the prediction of D by C must occur without mediation by E – ~~if C must be a risk indicator~~ must be assoc. w/ D conditioned on ~~any~~ E
   ∴ not conf. if.



III. B. 2. c. (cont.)



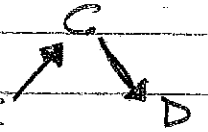
~~thus C must also be indep. (not dep.) risk indicator of D~~

C can be a dep. risk indicator

NB. an indep. risk indicator would mean that C was assoc. w/ D for every category of E (not just $\bar{E}$)

d. C may be conf. even if assoc. between E + C is spurious resulting from selection bias, differential measurement bias, or another conf. factor – also C may not be a true risk factor of D but only an "extensible risk factor" & risk indic. might occur because of differential errors in case detection – actually C might be only assoc. w/ a risk factor; but this distinction is often difficult to distinguish – eg, is age or immunologic competence a risk factor of D?

secondary (e) C must not be intervening link between E + D – ∴ not conf. if:



I think this is still conf. → mostly.

NB: not often possible to tell whether E predicts C since these are generally measured simultaneously – yet 1 piece of info. would suggest this pathway (i.e., C not conf.): no assoc. between E + D conditioned on the (lowest) risk category of C (if C is causal) – also this pathway can often be ruled out a priori – eg, Smoking (E) doesn't predict age (C).

See notes from Blalock, 1964

without [illegible] of any kind

but also true if

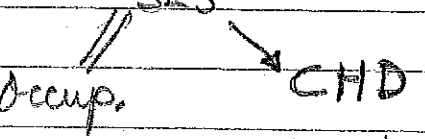
III. B. 2. f. When dealing with several potential conf. ~~they~~ their conf. effects must be considered simultaneously; this viewpoint is often violated in practice by examining pot. conf. sequentially -

→ When looking at the conf. effect of a single variable, it must be examined conditionally on all other conf. - thus the importance of a one conf. is reduced

(one) g. C must not be measuring one aspect or part of E - i.e., the measurement of C should be independent of E (or) C must not be measuring the same underlying theoretical construct as does E - i.e. not conf. if:

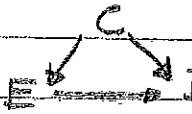
the C would be conf in a math. sense, to control for it would lead to mis-interpretation - eg,

Similar to e.



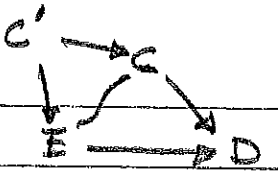
c. how conf. comes about - results from selection of subjects and measurement bias - [might define types of conf. by their causal configurations - Causal Config. may differ

1. "pure" conf - C predicts both E + D - $\emptyset_{EC}$ is not spurious:

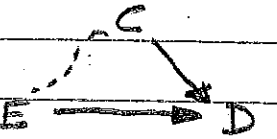


2. "multiple" conf. - spurious assoc. between E + C results from other (exogenous) factors which are themselves confounded - eg.

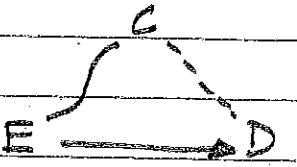
III, C.C. 2.



3. "artificial" conf. – spurious assoc. between C + E results from correlation of their errors (ie, diff. meas. bias)

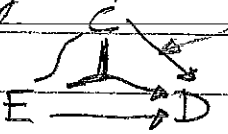


N, in a C-R study,



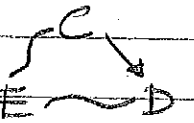
4. "complex" conf. – same factor (C) is both a conf. & dept. risk factor of D (being dept. on the effect of E) –

eg,



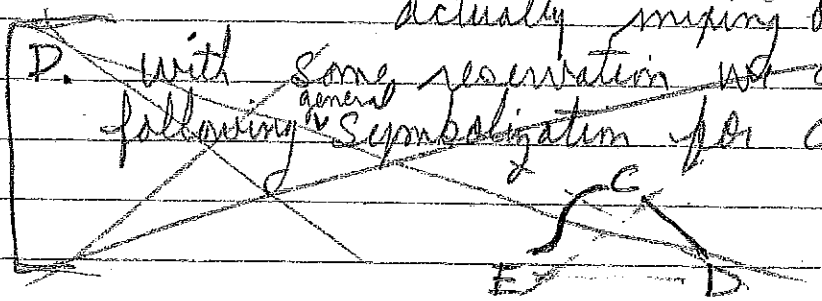
N.B. Strength of arrows may vary

5. "Secondary assoc." (Special case of 1-3) – assoc. between E + D is entirely spurious – eg,



NB: sort of ultimate case of "conf." – yet there is no real assoc. between E + D to conf. – ∴ not ~~spiping~~ actually "miping" of effect

P. with some reservation we can adopt the following general Symbolization for conf.



needed ?

III. E. but at the same time we should not be very careful in identifying conf. factors from statistical correlations – this may be elaborated w/ 3 principles.

1. conf. cannot be identified by observing only crude assoc. – eg, conf. can still occur even if $\phi_{CD} = 0$ – $C \rightarrow D \Rightarrow \phi_{CD \cdot E} \neq 0$

2. pertinent correlations (which reflect conf.) are different for CH + C-R studies – for conf. to exist:

a. CH: $\phi_{EC} \neq 0$, $\phi_{CD \cdot E} \neq 0$ } if either is 0, no conf.

b. C-R: $\phi_{EC \cdot D} \neq 0$, $\phi_{CD \cdot E} \neq 0$ } if either is 0, no conf.

3. degree of conf. (which we have not yet specified how to measure) does not depend on the significance levels of the previous ϕ's because conf. does not involve statistical inferences to a source pop. remember: conf. only characterizes one's data – a large enough n would make even the smallest ϕ correlation significant.

F. to quantify degree of conf. – bids – determine what crude effect (CRR) would be if exposures were to have no effect on the development of D. (ie, under H_0) – this component of CRR (label it RR*) is then the crude effect that

NB: 1 set of symbols inadequate to describe conf. even for CH study (might be dept. effect)

conf. in C-R not spurious

can be spurious only in sense that didn't risk factor → E–C results from selection – we know that if other vars. were controlled, E ⊥ C

III. F. would be expected in the absence of any true (unconf.) effect.

$\therefore RR^* = \mathcal{E}[CRR] = \mathcal{E}[a]$

	D	$\bar{D}$	
E	a	b	n_E
$\bar{E}$	c	d	$n_{\bar{E}}$

1. CH Study:

$$RR^* = \mathcal{E}[CRR] = \frac{\mathcal{E}[a]/n_E}{c/n_{\bar{E}}}$$

$$SMR = \frac{a}{\mathcal{E}[a]}$$

$$\mathcal{E}[a] = a/SMR$$

$$\therefore RR^* = \frac{\frac{a}{SMR}/n_E}{c/n_{\bar{E}}} = \frac{a/n_E}{c/n_{\bar{E}}}\left(\frac{1}{SMR}\right) =$$

$$\boxed{RR^* = \frac{CRR}{SMR}}$$

2. C-R study:

$$RR^* = \mathcal{E}[CRR] = \frac{\mathcal{E}[a]\,d}{bc}$$

$$= \frac{ad/SMR}{bc} = \frac{CRR}{SMR} \quad (\text{Same})$$

$\therefore$ CRR has 2 components: RR^* & SMR

i.e. if $RR^* \neq 1$, then is confounding

G. but in order to measure the amount of confounding, it is necessary to compare the conf. effect $(RR^* - 1)$ with the unconf. effect $(SMR - 1)$ — i.e. in adj

III. C. Thus a useful index of conf. may be derived in order to decide whether stratification is necessary in analysis

$$RCE = \frac{RR^* - 1}{SMR - 1}$$

if RCE > 0 ⇒ "positive conf."
RCE < 0 ⇒ "negative conf."
RCE = 0 ⇒ no conf.

pos. conf. ⇒ CRR is further from null value (1) than is SMR + both are > or < 1

neg. conf. ⇒ CRR is closer to null value than is SMR and both are > or < 1 (or) CRR > 1 / SMR < 1 (or) CRR < 1 / SMR > 1

still no single cutoff for substantial conf. w/ RCE exp. for any single factor

H. this procedure implies that conf. is a quant. (not qual.) phenomenon - it takes the first 4 specifications into consideration but not the 5th (I believe) - i.e., just because C is math. conf. does not mean that it should be controlled in the interest of internal validity.

IV. Illustrations and general principles of conf.
A. computations in a CH study.

C1

	D	$\bar{D}$	
E	6	14	20
$\bar{E}$	9	91	100
	15	105	120

C2

	D	$\bar{D}$	
E	8	52	60
$\bar{E}$	2	18	20
	10	70	80

IV. A. $CRR = \frac{14/80}{11/120} = 1.909$

$$SMR = \frac{6+8}{\frac{9}{100}(20) + \frac{2}{20}(60)} = 1.795$$

$$RR^* = \frac{1.909}{1.795} = 1.064$$

$$RCE = \frac{1.064 - 1}{1.795 - 1} = .081$$ ~~pos. conf.~~

would statistically only conf. but don't know w/ more

Note: there was conf. even tho $\phi_{CD} = 0$

	D	$\bar{D}$	
C1	15	105	120
C2	10	70	80
	25	175	200

$$\phi_{CD} = \frac{15(70) - 105(10)}{\sqrt{120(80)(25)(175)}} = 0$$

but C was still a risk indicator of D because its assoc. w/ D conditional on $\bar{E}$

B. computations in a C*R study

C1

	D	$\bar{D}$	
E	40	10	50
$\bar{E}$	10	40	50
	50	50	100

C2

	D	$\bar{D}$	
E	25	25	50
$\bar{E}$	25	25	50
	50	50	100

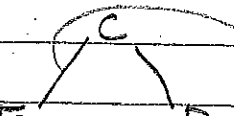
$$CRR = \frac{65(65)}{35(35)} = 3.449$$

$$SMR = \frac{40+25}{\frac{10(10)}{40} + \frac{25(25)}{25}} = 2.364$$

$$RR^* = \frac{3.449}{2.364} = 1.459$$

IV. B. (cont.) $RCE = \frac{1,459-1}{2,364-1} = .337 \Rightarrow$ poss. conf.

NB. there is conf. even tho $\phi_{EC} = \phi_{CD} = 0$
(ie, crude assoc are 0)
note: that criteria for conf. in C-R study:



* can be "spurious" only in the sense that C doesn't cause E, but if we controlled for other relev vars, E ~ C —(p.9)

C. "procedure often used in epid. study to eliminate possibility that C (eg, age) was conf. after finding sig. assoc. between E & D (complications)

1. C-R study: compare cases & noncases with respect to age distrib. $\Rightarrow \phi_{CD} \neq 0$
no good since criterion is $\phi_{CD \cdot \bar{E}} = 0$

eg, age	D (#'s)	$\bar{D}$
Y	—	—
O	—	—
⋮	—	—

(Kolmogorov-Smirnoff test)

2. CH study: compare $E + \bar{E}$ w/ respect to age distrib. $\Rightarrow \phi_{EC} = 0$
this is ok since it is a criterion for conf.

CH study

IV. D' –

C

	D	$\bar{D}$	
E	10	10	20
$\bar{E}$	6	24	30
	16	34	50

$\bar{C}$

	D	$\bar{D}$	
E	30	70	100
$\bar{E}$	10	40	50
	40	110	150

$$CRR = \frac{40/120}{16/80} = 1.667$$

$$SMR = \frac{40}{4+20} = 1.667$$

because: $\phi_{CD,\bar{E}} = 0$

$RR^* = 1 \quad RCE = 0 \Rightarrow$ no conf. due to C

E

	D	$\bar{D}$	
C	10	10	20
$\bar{C}$	30	70	100
	40	80	120

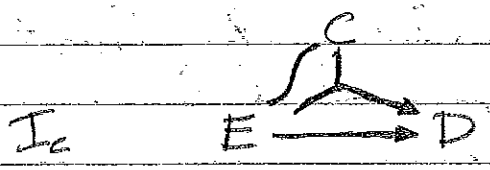
$\bar{E}$

	D	$\bar{D}$	
C	6	24	30
$\bar{C}$	10	40	50
	16	64	80

$$CRR_c = \frac{16/50}{40/150} = 1.2$$

$$SMR_c = \frac{16}{6+6} = 1.333$$

$RR^* = .900 \quad RCE = -.300 \Rightarrow$ neg. conf. due to E



both C + E are risk indicators but C is not indep. risk indicator

	C	$\bar{C}$
E	.5	.5
$\bar{E}$	.3	.1

∴ better example because:

SMR_c, SMR_E $\neq 0$

NB: Same results if analysed as C-R study

$RCE_c = 0 \qquad RCE_E = -.303$

on D'

IV. B. Sometimes interested in causal effect of 2 vars. (E+C) conf. by C (of E-D assoc.) is not necessarily = to conf. by E (of C-D assoc.) —

eg. consider IV. A. — RCE = .081 (pos conf.)

Imagine E as a potential conf. of C-D assoc.

E:

	D	$\bar{D}$	
C1	6	14	20
C2	8	52	60

$\bar{E}$

	D	$\bar{D}$	
C1	9	2	11
C2	91	18	109

$$CRR_c = \frac{15/31}{99/169} = .826$$

$$SMR_c = \frac{6+9}{\frac{8}{60}(20) + \frac{91}{109}(11)} = 1.266$$

$$RR_c^* = \frac{.826}{1.266} = .652$$

$$RCE_c = \frac{.652-1}{1.266-1} = -1.307 \quad \text{neg. conf.}$$

Thus magnitude + sign of conf. is different for E + C.

"Conf. is not symmetrical" in (CHD C-R study)

IV. E. control tables should not be used to identify & measure conf. – tho they often are, under the mistaken view that conf. is a qualitative phen. – i.e., seeing conf. as the presence of secondary bases – eg, CH study

C1

	D	$\bar{D}$	
E	30	70	100
$\bar{E}$	4	16	20
	34	86	120

C2

	D	$\bar{D}$	
E	4	16	20
$\bar{E}$	6	54	60
	10	70	80

Control table: I_c

	C1	C2
E	.30	.20
$\bar{E}$	.20	.10

⇒ could say still E effect after controlling for C, but is C conf.?

$$CRR = \frac{34/120}{10/80} = 2.267$$

$$SMR = \frac{30+4}{\frac{24}{100}(100) + \frac{6}{60}(20)} = 1.545$$

$$RR^* = \frac{2.267}{1.545} = 1.467$$

$$RCE = \frac{1.467-1}{1.545-1} = .856 \quad \text{pro. conf.}$$

conceptual (p. 15)

3 reasons why we need to know if conf. exists & how much

1. know whether simple analysis (more eff) can be used
2. when considering several pot. conf. simult. we would like to select the combin. which accounts for the most bias (most eff.)
3. generally, conf. is quantitative phen. & thus we should like to know the magn. of unconf. effect (RR^*)

but Suppose data were:

	D	$\bar{D}$	
E	15	35	50
$\bar{E}$	10	40	50
	25	75	100

	D	$\bar{D}$	
E	10	40	50
$\bar{E}$	5	45	50
	15	85	100

IV. E. (cont.) Same control table,

$$CRR = \frac{25/100}{15/100} = 1.667$$

$$SMR = \frac{15+10}{\frac{10}{50}(50)+\frac{5}{50}(50)} = 1.667$$

$RR^* = 1$ $RCE = 0$ no conf.

3 reasons p. 14

However, control tables do allow us to identify secondary associations, tho they seldom exist in practice – (i.e.) SMR = 1 ; but doesn't tell whether conf. exists which is the criterion for using simple or stratified analysis (simple analysis is more efficient) – eg, I_c

	C1	C2	
E	.3	.3	⇒ no E effect ∴ SMR = 1
$\bar{E}$	.1	.1	but what is CRR ?

* NB: (1) can't be used in C-R studies (2) but do allow us to look for EM + I

(note: can do as C-R study)

F. Consider 2 potential conf. – Sex + race CH

	WM D	WM $\bar{D}$		BM D	BM $\bar{D}$		WF D	WF $\bar{D}$		BF D	BF $\bar{D}$	
E	56	104	160	2	18	20	1	19	20	46	69	11
$\bar{E}$	14	26	40	6	54	60	3	57	60	6	9	15
	70	130	200	8	72	80	4	76	80	52	78	13
	RR = 1.00			1.00			.105 1.00			.40 1.00		

Note: $I_c R_j = 1$ for all 4 strata $= RR_j$

$$CRR = \frac{105/315}{29/175} = 2.01 \qquad n = 490$$

D	$\bar{D}$
134	356
.27	

IV. F. (cont.): control for race only

W

	D	$\bar{D}$	
E	57	123	180
$\bar{E}$	17	83	100
	74	206	280

B

	D	$\bar{D}$	
E	48	87	135
$\bar{E}$	12	63	75
	60	150	210

$$SMR = \frac{57+48}{\frac{17}{100}(180)+\frac{12}{75}(135)} = 2.01$$

$RR^* = 1$ $RCE = 0$ no conf.

control for Sex only

M

	D	$\bar{D}$	
E	58	122	180
$\bar{E}$	20	80	100
	78	202	280

F

	D	$\bar{D}$	
E	47	88	135
$\bar{E}$	9	66	75
	56	154	210

$$SMR = \frac{58+47}{\frac{20}{100}(180)+\frac{9}{75}(135)} = 2.01$$

$RR^* = 1$ $RCE = 0$ no conf.

control for Sex + race Simultaneously

$$SMR = \frac{105}{\frac{14}{40}(160)+\frac{6}{60}(20)+\frac{3}{60}(20)+\frac{6}{15}(115)} = 1$$

$RR^* = 2.01$ $RCE = \frac{2.01-1}{1-1} = ?$ ∞ pos conf.

IV. F. (cont.) Thus when controlling for both simult., there is no E effect – yet this effect is not discovered by controlling for Sex + race sequentially!

Also Control table analysis of each var. fails to show no E effect.

I_c

	M	F
E	.322	.348
$\bar{E}$	.200	.120

	W	B
E	.317	.356
$\bar{E}$	.170	.160

In order to determine whether one var. (eg, Sex) is conf., it must be viewed conditionally on the other – thus,

for whites

$$CRR = \frac{57/180}{17/100} = .863$$

$$SMR = \frac{56+1}{\frac{14}{40}(160)+\frac{3}{60}(20)} = 1$$

$RR^* = 1.863$ $\quad RCE = \frac{.863}{0} = ?$ ∞ pos. conf

for Blacks

$$CRR = \frac{48/135}{12/75} = 2.222$$

$$SMR = \frac{2+46}{\frac{6}{60}(20)+\frac{6}{15}(115)} = 1$$

$RR^* = 2.222$ $\quad RCE = \frac{1.222}{0} = ?$ ∞ pos. conf

IV. G. It is conceivable that when considering more than one pot. conf. variable more bias will be discovered when stratification is done by combinations of factors — eg, C×C' as an interactive term — this was done by Kleinbaum et al. in Evans county when they derived a Multiple logistic risk function predicting CHD. e.g, age × chol. ; age × BP — only problem might come if we insist on interpreting the central variables; what do they mean — yet with several pot. conf. & a limited N, considering interactive terms systematically can make the estimation more precise — OM, however has formulated an alternative method that we will discuss later in the term; it is designed to deal with several conf. factors by combining them into a single risk function then multiv. analysis — it appears that such a methods makes the consideration of interactive conf. terms a secondary issue

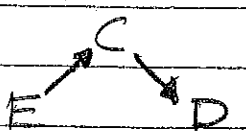
must needed if not want to control for conf.

V — methods of control — see revised outline 2 basic approaches (#8)

H. (p. 19)

addendum: also relates to Interaction

IV. H. recall spec. "c" : math. conf. but not to be controlled if.



eg, CH study

C

	D	$\bar{D}$	
E	30	120	150
$\bar{E}$	5	45	50
	35	165	200

$I_{c(c)} = .175$

$\hat{RR}_c = 2$

$\bar{C}$ (low risk)

	D	$\bar{D}$	
E	2	48	50
$\bar{E}$	16	144	150
	8	192	200

$I_{c(\bar{c})} = .04$

$\hat{RR}_{\bar{c}} = 1$

$$C\hat{R}R = \frac{32}{11} = 2.909$$

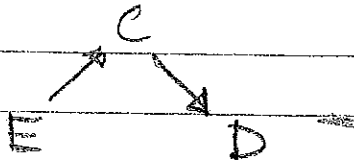
$$S\hat{M}R = \frac{32}{15+2} = 1.882$$

$RR^* = 1.545$ $\quad RCE = .618$ (pro. conf.)

but $\hat{RR}_{\bar{c}} = 1$ (no effect)

I_c

	C	$\bar{C}$
E	.2	.04
$\bar{E}$	.1	.04



NB @ might be C = E

ⓑ interaction?

? what about if $\hat{RR}_c = 1$? (over)

IV. H' C+H study — $\hat{RR}_E = 1$ $\hat{RR}_E > 1$

	D	$\bar{D}$	
E	15	135	150
$\bar{E}$	5	45	50
	20	180	200

D	$\bar{D}$	
5	45	50
5	145	150
10	190	200

I_c

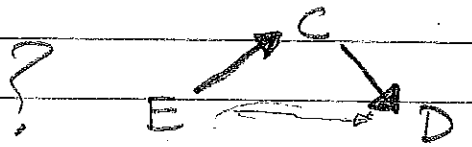
	C	$\bar{C}$
E	.1	.1
$\bar{E}$	.1	.0333

$CRR = 20/10 = 2$

$SMR = \frac{20}{15 + 5/3} = 1.2$

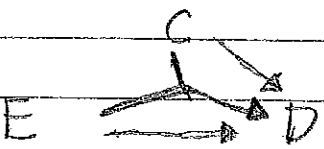
$RR^* = 1.667$ $RCE = 3.333 \Rightarrow$ (pos. conf.

$\hat{RR}_E = 1$ (hi risk cat of C)



not as much
no? because "C" subjects (ie, hi risk) may have gotten that way because of other causes of C besides E

* Can this config. also manifest in the data as statistical (and/or "causal") interaction? yes



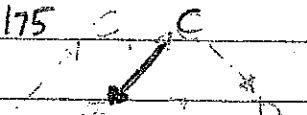
∴ underlies difference in theoretical & empirical levels: (eg)

smok
age → CHD

→ Same problem w/ $\hat{RR}_E = 1$ (p. 19)

Notes on Causal Inference

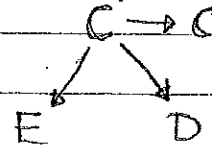
Blalock (1968)

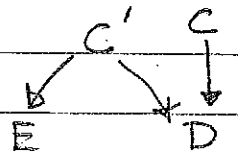
1. p.175  (Blalock, 1964) [E as background variable]

observe: r_{ED} vs $r_{ED.C}$ (partial correl.)
[C is not a conf.] b_{ED} vs. $b_{ED.C}$ (partial regression

if E is intervening (as drawn), controlling for C will reduce r (even if $C \to D$), but will not reduce E[b] → (similar w/ $RR_{E.C} = SMR_E$)
the stronger r_{EC}, the more reduction in $r_{ED.C}$

 ⇒ both r and b are reduced by controlling for C (see # 4)

2. p.177  ⇒ $r_{ED.C} < r_{ED}$

 ⇒ $r_{ED.C} \gtrless r_{ED}$ $(r_{C'C} \neq 0)$

3. p.190

$$r_{XY} = b_{XY}\left[\frac{S_X}{S_Y}\right]$$

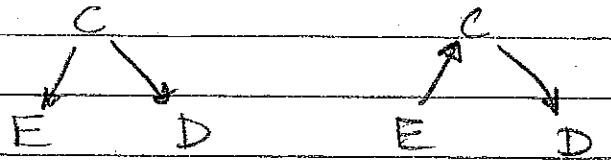
[in brackets] → peculiar to pop.
b's → peculiar to variable

$$r_{XY.Z} = b_{YX.Z}\left[\frac{\sqrt{1-r_{XZ}^2}}{\sqrt{1-r_{YZ}^2}} \cdot \frac{S_X}{S_Y}\right] = \frac{r_{XY} - r_{XZ}\,r_{YZ}}{\sqrt{1-r_{XZ}^2}\sqrt{1-r_{YZ}^2}}$$

(me) (4)  $RR_{E.C}$ (=SMR) ≠ CRR_E the C should not be controlled for in interest of interpretation

p.175

4.



both cases: $r_{ED.C} \cong 0$

$r_{ED.C} < r_{ED}$

Same w/ b

1/21/76 Outline 1/28 2/4

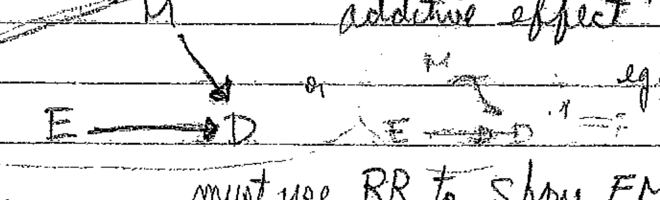
I. Introd.
  A. handouts — revised outline of Epid. Measures — #4
  B. today — finish conf.
  EM + I

(change to 3/4 levels of EM.)

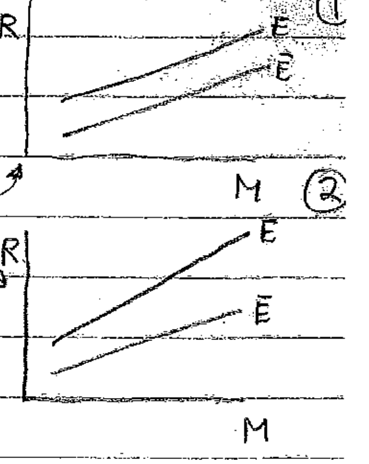
II. Confounding (cont.) see: 11/14/76 p. 11 RR* RCE

III. Effect Modif. — defn.
  A. multiple risk indicators (factors) assoc. w/ dis. in all dis.
    1. Stat. correl. of dis. = factor is assoc. w/ rate of di[s]
  (EM) 2. correlate of effect = factor (M) is assoc. with the effect of E on D ⇒ M is an effect modifie[r] if the effect (RR or RD) between E + D is not uniform among all categories of M.
  B. EM requires that 2 (or +) variables (E + M) are risk indic. of the of dis. — doesn't
  (risk indicators ? Stat. correlates)
  C. EM ∴ doesn't represent a singular biosocial mechanism of causation, but may result from oth[er]
  [4 primary classes of conceptual configurations — ie, 4 types of EM
    1. E + M are (indep.) risk factors (indicators) "additive effect"

(no good)



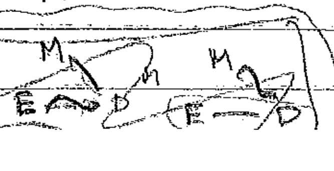
eg,



must use RR to show EM
" " RD " " "

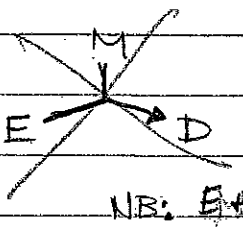
✱ ie, E is assoc. w/ D for every category of M + vice-versa

(needn't be causal) (1A)

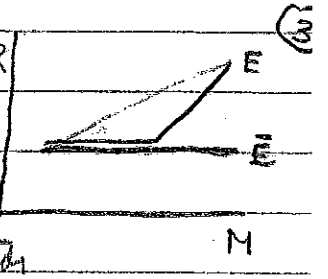


but not additive effect

II C. 2. M + E are dept. risk factors



"interactive effect" (when curves are non-parallel)



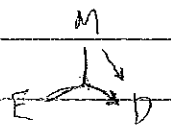
(3)

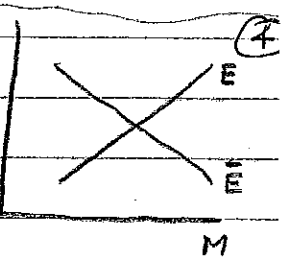
no good

NB: E + M are not indep. risk factors

see (2) – using RD, no EM, yet interaction

yet sometimes I does necessitate EM →



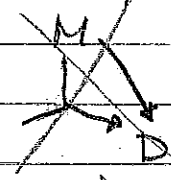


(4)

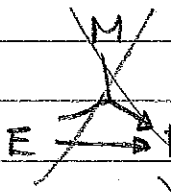
3. indep + dept. risk factors aren't mutually exclusive categories – a var. can be both

3. M + E as combination

(eg),



n

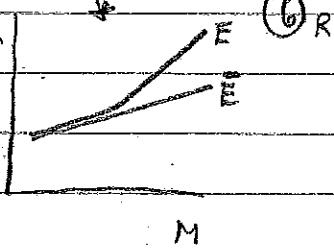


n

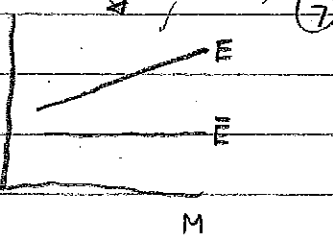
N.B. from a statistical stdpt. just because a var. (E or M) is not an indep. risk factor doesn't mean it does not have a significant "main effect" on D (in qualit. sense)

(eg),

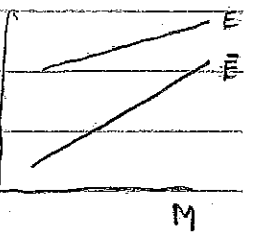
(5)



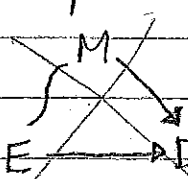
(6)

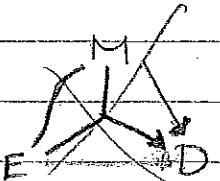


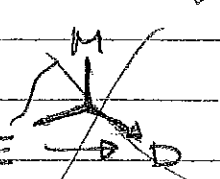
(7)



4. Complex conf. – M is both EM + confounder







D. confusing pt. C.5 does not mean/imply that conf. automatically suggests EM – altho the diagram seems to indicate such – will be illustrated later

III. E. Note difference in terminology

* mention gen statistical patterns by dept/subdisp

me	OM.	Fleiss
EM	"trends in effect"	"homogeneity of assoc."
I.	"EM"	—

begin → IV. Comparison of Conf., EM, + I as separate issues – p. 14
A. 1. Conf. & purpose in identifying – how internal validity is threatened
Conf. is a f. of (determinants)
pooling estimates no stratification p. 5 — B. 2. EM – " " how to control for conf
C. 3. I – " " (no modifying variables outside data) causal dept.
(D) note disagreement w/ Rothman – Conf. is f(data); I = f(causal assoc. but with empirical relationships)

V. Identifying EM + I – control table analysis (CH study)
A. EM – I_c

	M	$\bar{M}$
E	R_{11}	R_{01}
$\bar{E}$	R_{10}	R_{00}

ref.

RR

	M	$\bar{M}$
E	RR_{11}	RR_{01}
$\bar{E}$	1	1

RD

	M	$\bar{M}$
E	RD_{11}	RD_{01}
$\bar{E}$	0	0

$RR_{11} = RR_{01} \Rightarrow$ no EM
$RD_{11} = RD_{01} \Rightarrow$ " "

B. I. control table I_c

	M	$\bar{M}$
E	R_{11}	R_{01}
$\bar{E}$	R_{10}	R_{00}

RD

	M	$\bar{M}$
E	RD_{11}	RD_{01}
$\bar{E}$	RD_{10}	0

$RD_{11} = RD_{01} + RD_{10} \Rightarrow$ no I. (ie, additive effects)
" $>$ " " $\Rightarrow$ I – synergistic
" $<$ " " $\Rightarrow$ I – antagonistic
* poor words because they suggest biolog. processes, not statistical

ERR = RR-1

IV. B. also, can't use control table via C-R study

$ERR_{11} = ERR_{01} + ERR_{10} \Rightarrow$ no I

	M	$\bar{M}$
E	ERR_{11}	ERR_{01}
$\bar{E}$	ERR_{10}	0

C. Examples.

1. CH study:

M	D	$\bar{D}$	
E	60	90	150
$\bar{E}$	10	40	50
	70	130	200

$\bar{M}$	D	$\bar{D}$	
E	10	40	50
$\bar{E}$	15	135	150
	25	175	200

I_c	M	$\bar{M}$
E	.4	.2
$\bar{E}$	.2	.1

$RR_M = RR_{\bar{M}} = 2 \Rightarrow$ no EM (RR)

NB: $RD_M = .2$, $RD_{\bar{M}} = .1$ $\Rightarrow$ EM (RD)

ERR

	M	$\bar{M}$
E	3	1
$\bar{E}$	1	0

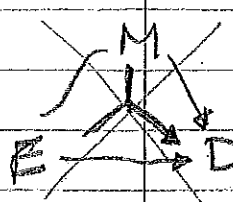
$3 > 1+1 \Rightarrow$ I (Synergy)

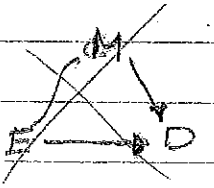
EM + I can be deduced from initial table

$$CRR = \frac{70/200}{25/200} = 2.8$$

$$SMR = \frac{10+60}{\frac{10}{50}(150) + \frac{15}{150}(50)} = 2.0$$

$$RR^* = \frac{2.8}{2} = 1.4 \qquad RCE = \frac{1.4-1}{2-1} = .4 \Rightarrow \text{pos. conf.}$$



∴ conf.  doesn't imply EM (RR)

(don't know I_c)

I. c. Ex. - C-R study - no control table, but similar principle

adj.

M	D	$\bar{D}$	
E	50	10	60
$\bar{E}$	30	10	40
	80	20	100

$\bar{M}$	D	$\bar{D}$	
E	60	20	80
$\bar{E}$	20	20	40
	80	40	120

n/b.

$$OR_{11} = \hat{I}_D \hat{R}_M = \frac{50(10)}{10(30)} = 1.6$$

$$OR_{01} = I_D R_{\bar{M}} = \frac{60(20)}{20(20)} = 3.0$$

stat. test. $\Rightarrow$ EM (OR)

$\therefore$ overall $\hat{RR}$ derived from pooling strata of M would be misleading - tho, recall we can standardize to control for conf.

* 2 basic ways to form weighted avg of strata spec. estimate

$$SRR = \frac{\sum w_j R_{Ej}}{\sum w_j R_{\bar{E}j}}$$ (RR$_0$)

① adjustment

$$A_0 = \frac{\sum W_j A_j}{\sum W_j}$$

A_0 = overall M of assoc. (e.g. RR)

A_j = measure of Assoc.

w_j = based on crude/stand.

W_j = based on $1/Var[A_j]$

② gen'l. stratification

ERR

M	$\bar{M}$	
E	(4)	(2)
$\bar{E}$	(2)	0

$3-1 = 2$

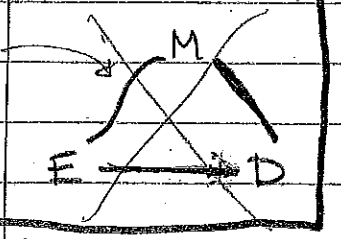
$\frac{30(20)}{10(20)} - 1 = 2$

$\frac{50(20)}{10(20)} - 1 = 4$

$4 \neq 2+2 \Rightarrow$ no I.

I. C. (cont.) $\hat{CRR} = \frac{110(30)}{130(50)} = 2.2$

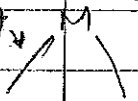
NB: E~M but not conf. in CR study



$\hat{SMR} = \frac{50+60}{\frac{10(30)}{10} + \frac{20(20)}{20}} = 2.2$

$RR^* = 1 \quad RCE = 0 \Rightarrow$ no conf.

not M



i.e. I. does not necessarily imply conf. actually conf, EM, + I are separate issues

VI. Quantifying EM + I – statistical procedures

A. EM – qualitatively – ie, is EM stat. sign?

1. partitioning of $\chi^2 = f[A]$ q strata of r

$$\chi^2_{u(q-1)} = \sum \chi^2_{j(q)} - \chi^2_{0(1)}$$

χ^2_j = usual χ^2 for 1 strata – several ways to compute, most common = $\sum\sum_{\text{all cells}} \frac{(O-E)^2}{E}$

will be discussed in more detail later (#10)

χ^2_0 = test of overall assoc. between E and D controlling for M (if it is conf.)

$= \frac{(A_0 \sum w_j)^2}{\sum w_j}$ A_0 = overall M of assoc.

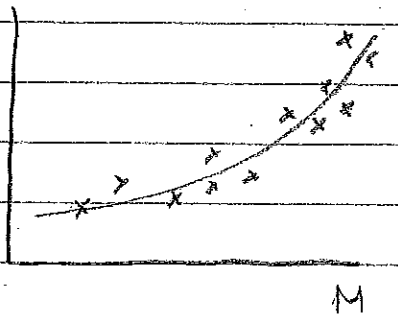
$w_j = 1/V[A_j]$

will also be covered later

χ^2_u = test of uniformity of effect (EM); ie, if p^0 is signif ($\leq .10$) there is sign. EM

∴ A_0 is not good estimate of overall effect

VI. A. 2. χ^2 test is not very sensitive test since no order is assumed in the categories of M – ∴ one highly weighted stratum could distort the overall picture of EM –
a more sensitive test might be made if M is at least ordinal – we could fit the data to a specific statistical model



B. I – normally measured qualitatively thru statistical models ("multivariate") – eg,

linear: $\hat{R} = \beta_0 + \beta_E E + \beta_M M + \beta_{ME} ME = \beta_i X_i$
not good fit for $R < .2$ or $> .8$

log linear: $\ln \hat{R} = \beta_i X_i$
$\hat{R} = \exp[\beta_i X_i]$ often gives E values < 0 or > 1

logistic model: $\ln\left[\frac{\hat{R}}{1-\hat{R}}\right] = \beta_i X_i$

$$\hat{R} = \frac{1}{1 + \exp[-\beta_i X_i]}$$

2 or 3 disadvantages of model-fitting:

(1) disadv: can't interpret since mixture (trends) of interaction remain hidden

(2) don't get degree of interaction (except β)

(3) there is also another disadvantage in using model fitting which we'll discuss later (ie, when we'd like to control for conf. effects + study I, simultaneously)

while this model is fine (esp. for many var's) there is another easier, more direct way to not only detect interaction but to measure it quantitatively (based on contingency tables)

VI. B. Rothman's index of interaction (S) = RIE = relative interaction effect

= ratio of observed joint effect from indep. vars. (E+M) to the effect expected if each var. were acting indep.

$$RIE = \frac{ERR_{ME}}{\hat{E}[ERR]_{indep.}}$$ how to compute:

	M	$\bar{M}$
E	ERR_{11}	ERR_{01}
$\bar{E}$	ERR_{10}	0

ERR

could have more categories

$$RIE = \frac{ERR_{11}}{ERR_{01} + ERR_{10}}$$

"+" pos. interaction

RIE = 1 ⇒ no I

RIE > 1 ⇒ "synergy"

RIE < 1 ⇒ "antagonism"

Rothman also gives SE & CI of RIE

Criticisms of RIE:

NB. stat. interaction does not have to imply some biologic interaction (of any factors)

1. if there are many categories of E and/or M, the effect for each cat. of E (or M) controlling for the other must be increasing (or ↓) monotonically — if not, computation of RIE is ambiguous & misleading

(eg) ERR

	M1	M2	M3
E1	(4)	3	(2)
E2	3	6	3
E3	(2)	3	0

RIE = ?

interaction is an artifact

in both cells to same extent/magnitude

2. when RIE < 1, interpretation of antagonism may be deceptive, esp. if M or E is an abstract var. eg, SES — if M = E (measuring same thing), RIE would be < 1, i.e. this measure of I may not relate to biological or soc. processes — ie, the statistical assoc. may have no etiologic significance

not really a criticism of RIE but how it is interpreted (theoretically)

(8A)

~~A. there is no~~

A. stat. interaction does not imply there is any bio-logical/behavioral interaction (of any sort) because might have:

? not sp. (1) M + E are measuring same thing ⇒ RIE < 1
spurious (2) differential measuring error (eg, correlated errors)
spurious (3) M$^{E}_{E}$ is assoc. w/ lag period of Dis.
not sp. (4) non-diff. meas. error ⇒ RIE → 1 ?

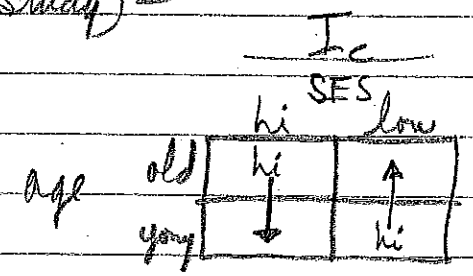
~~A~~ types of ~~artifacts~~ bias

(5) selection — i.e., 3 factors that prod. bias between E + D
(6) conf. — 2 factors above (1 + 3)

B. ~~there~~ in general, there is no 1:1 correspondence between stat. ~~to~~ and bio-beh. mechanisms →

VI. c. this underscores the lack of direct correspondence between biosocial mechanisms of causation (theoretical level) and statistical assoc. (empirical)

consider the example of infectious hepatitis (CH study) –



opposite trends for each stratum of SES

∴ Age + SES interact. But so what?
does it indicate "synergy" in a biomedical sense? no

the hypothesis put forth + later validated was that SES was actually reflecting the difference in hygienic practices ~~&~~ young persons of low SES families w/ poor hygiene have a hi risk of developing dis (the symptoms are minimal); as they get older, immunity is developed + fewer low SES persons are susceptible – w/ high SES, young children are protected from virus until they get older, at which time the risk of D increases

VII

thus: relationship between statistical interaction + biosocial interaction is not necessarily direct nor is the connection always obvious, esp. when presented w/ statistical info. + attempting to infer biosocial mechanisms – yet biological or bio. interaction is extremely important to pathogenesis of chronic disease. acknowledgment the importance of (eg) Env + genetic factors

or Env stress and social supports

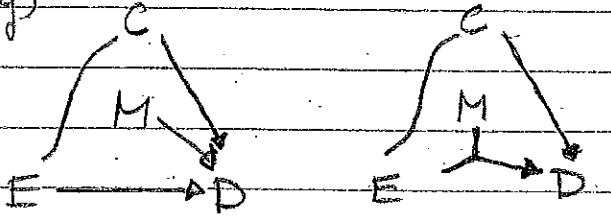
also control for conf.

VII. EM and I — Illustrations + principles

A. Conf + EM+I

Consider the study of EM or I, when having to control for pot. conf. factors simultaneously

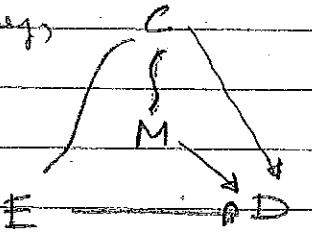
eg)



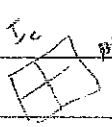
etc.

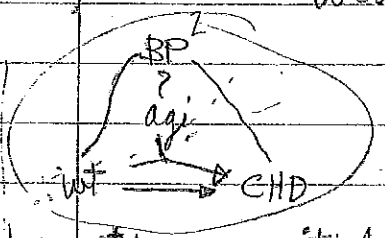
usual procedure of studying interaction of E+M & controlling for C is thru multiv. analysis Specifically covariance — OM has pointed out that multivariate would be deceptive if C were highly assoc. w/ M (which they usually are (even spuriously))

eg,



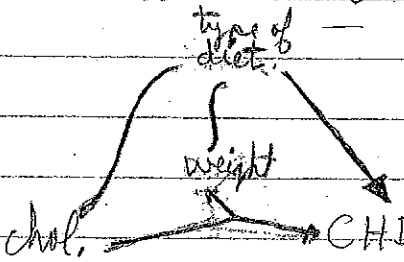
the model controls for C by keeping it "fixed" i.e. any correlate of C (eg, M) will be made statistically meaningless with regards to its interactive effect on D





Ho: wt. is more critical for young than old

eg)



if we control for dieting, wt will be reduced a hollow concept (to a certain extent)

OM's solutions: involve correct choice of effect measure

1. study EM — use SMR (internal stand.) for each category of M & compare — N.B. that a different "standard" is used for each stratum of M, but not a problem of

begin → (2/4)
review 2. disadvantages of model-fitting to study interaction
@ interp. b. conf. p.7

VII. A. 1. (cont.) comparability because each stand. measure is computed within 1 stratum of M, comparing E and $\bar{E}$ (ie, only 2 categories of E)
∴ the importance C as a conf. is within each stratum of M

2. Study I — can find RIE by externally standardizing for C within each stratum of M — ∴

$$RIE = \frac{SERR(E)_{\bar{ME}}}{\mathcal{E}[SERR]_{indep.}}$$

where: SERR = SRR − 1 (ext. stand.)
must be ext. stand. since stand. for each comparison must be identical.

stand: 1. ref. categories (ext) 2. all 3. other genl. pop.

eg. #9 p.42 Oral Ca.

3. do a problem * numerically — CH study look for interaction of E and race (M) controlling for sex (C)

	WM D	WM $\bar{D}$		WF D	WF $\bar{D}$		BM D	BM $\bar{D}$		BF D	BF $\bar{D}$	
E	25	75	100	15	35	50	10	40	50	50	50	100
$\bar{E}$	5	45	50	15	85	100	5	95	100	5	45	50
	30	120	150	30	120	150	15	135	150	55	95	150

$\hat{I_cR}$ = 2.5 —— 2.0 —— 4.0 —— 5.0

a. Suppose we ignored M (race) & determined if C (sex) were conf.

$$CRR = (I_cR) = \frac{100/300}{30/300} = 3.33$$

$$SMR = \frac{100}{\frac{10}{150}(150) + \frac{20}{150}(150)} = 3.33$$

$RR^* = 1$ RCE = 0 ⇒ no conf. due to sex (but wrong)

VII, A. 3. (Cont.) — we might then proceed to calculate the interactive effect between E+M (race), ignoring C (sex); if we did,

WM

	D	$\bar{D}$	
E	40	110	150
$\bar{E}$	20	130	150
	60	240	300

BL

	D	$\bar{D}$	
E	60	90	150
$\bar{E}$	10	140	150
	70	230	300

I_c

	W	B
E	.267	.400
$\bar{E}$	.133	.067

→ ERR

	W	B
E	3	5
$\bar{E}$	1	0

Since $3 < 5+1 \Rightarrow$ I ("antagonism")

$RIE = \frac{3}{1+5} = \boxed{.5}$

b. Suppose we now consider C (sex) in the above analysis thru ~~Stratif.~~ Standardiz.

SERR(E) SRR-1

	W	B
E	3.00	3.50
$\bar{E}$	.75	0

$SERR = \frac{\frac{10}{50}(100) + \frac{50}{100}(50)}{10} - 1 = 3.5$

$SERR = \frac{\frac{5}{50}(100) + \frac{15}{100}(50)}{10} - 1 = .75$

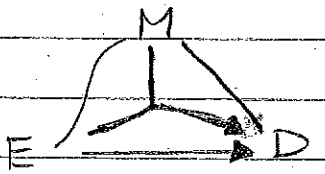
$SERR = \frac{\frac{25}{100}(100) + \frac{15}{50}(50)}{10} - 1 = 3.0$

$RIE = \frac{3}{3.5+.75} = \boxed{.71}$ (antag.)

* but: $\boxed{.71 \neq .50}$ — why?
because Sex was conf. — RR* (above) was calculated incorrectly

VII. / A. 3. RR* was incorrect because it ignored race (M) which is itself a conf. — when considering the conf. effect of a variable it must be viewed conditionally on all other confounders — thus in getting RIE controlling for sex, we must standardize separately for each comparison made; and stand. must be done externally to insure that the same stand. is used for all comparisons

4. another possibility is that C=M, ie, "complex conf" Since C is same as M, it is impossible to consider it simult. as both a conf. + E mod. —



perhaps we could control for M within each category of M?

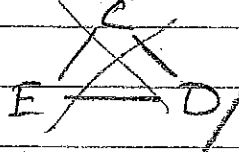
can't

in the analysis M as a E mod. becomes irrelevant since we must control for conf. in the interest of internal validity — use SMR w/ 2 categories of E or SRR w/ > 2 categories if looking for dose-response gradient

5. planning a study on the basis of expected conf. in the data (or lack of it) can then be complicated by the presence of EM or I. — eg, C-K study:

ABO Blood group —?→ Cervical Ca.
(E) (D)

VII. A. 5. can we have a control series of male non-cases, altho this ~~paradoxical~~ design seems paradoxical? we might defend it on the basis that Sex is not assoc. w/ ABO Bl group in the source pop. & thus conf. won't be expected in the data (if no selection is introduced) — but recall that $\Phi_{CE} \neq 0$ is not a condition for conf. — instead $\Phi_{CE,D} \neq 0$; ie age can't be assoc. w/ ABO Bl. group, conditional on being a non-case:



this condition will not be met, since the risk of D is greater for unexposed females than unexp. Males (when it is 0) — this in fact suggests an interactive effect between Sex & ABO Bl groups — to illustrate, consider a hypoth. C-R study of all females (assume all other pos. have been controlled for)

ABO	E	D	
E	20	10	30
Ē	30	40	70
	50	50	100

$\hat{I}_{DR} = \frac{20(40)}{10(30)} = 2.67$

if done w/ F cases & M controls — would expect following e's

$e_D(F) = \frac{20}{50} = .4$ (from F above)

$e_D(M) = \frac{30}{100} = .3$ (from total)

ie, $\Phi_{Sex,ABO} = 0$

VII. A. 5. ∴ if did study w/ 50 F cases + 50 M controls — ?

	D	$\bar{D}$	
E	(20)	(15)	35
$\bar{E}$	30	35	65
	50	50	100

.4 × 50 = 20

.3 × 50 = 15

$\hat{I}_D R = \frac{20(35)}{15(30)} =$

1.56 < 2.67

∴ bias

in order to use M controls, it would have been necessary for either:

$\Phi_{\text{sex ABO, D}} = 0 \Rightarrow$ no conf. (difficult to empirically assume no selection)

or, $\Phi_{\text{ABO D, E}} = 0$ which is impossible for env. Cs. because of interaction

we may conclude that there was conf. (indirectly) but notice that it can't be computed because an SMR is undefinable

F

	D	$\bar{D}$	
E	20	0	20
$\bar{E}$	30	0	30
	50	0	50

M

	D	$\bar{D}$	
E	0	15	15
$\bar{E}$	0	35	35
	0	50	50

$$SMR = \frac{20+0}{\frac{0(30)}{0} + \frac{15(0)}{35}} = ?$$

8.D

VII. B. EM and I – the relationship

1. as noted before, EM depends on the meas. of A. chosen as well as the type of design selected (ie, I_D vs I_C) – See: p. 48 for illustration –

* there are 4 ways of determining EM – I_DD, I_CD, I_DR, I_CR – all give different results

2. there exists an inherent relationship between EM + I such that computing EM, using I_DD, amounts to a test for interaction – the proof of this is given in #9, pp. 47-53 – eg,

I_D

	M	$\bar{M}$
E	.04	.02
$\bar{E}$	.02	.01

$(.04 - .02) \neq (.02 - .01) \Rightarrow$ interaction

the other way:

ERR

	M	$\bar{M}$
E	3	1
$\bar{E}$	1	1

$RIE = \frac{3}{2} = 1.5 \Rightarrow$ interaction

Can use control table (I_D) to look for interaction

∴ observing trends in I_DD is tantamount to identifying an interactive effect

NB. – this also implies that I_D (not I_C) should be used in computing RIE – something we have not done thus far

VII. B. 3. These principles only apply, however, to causal factors — when analyzing the interaction of protective factors, the appropriate measure is PF (instead of $I_{\&}D$) = % of all pat. cases that have been prev. because of exp.

$$PF_{11} = PF_{a,b,\ldots} \overset{?}{=} 1-(1-PF_a)(1-PF_b)\ldots$$

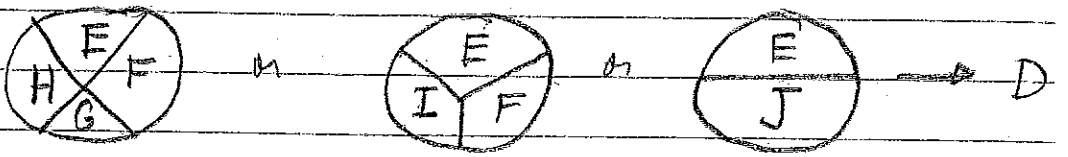
$>$ ⇒ Synergism (+ interaction)

$<$ ⇒ antagonism (− int.)

4. Similar relationship applies to causal factors

$$EF_{a,b,\ldots} = 1-(1-EF_a)(1-EF_b)\ldots$$

Suppose we had computed EF's for the following model.



I. EF = .20 II. EF = .30 III. EF = .40

? ∴ EF for (H) = .20 (i.e., b/c indep. effect of H were 0)

$EF_{12} \overset{?}{=} 1-(1-.2)(1-.3) = .44 = EF_F$

∴ 44% of all cases could have been prevented if ✱ factor F had been absent in Source pop. (i.e., no one had been in hi risk category of F)

d F had no indep. effect

$EF_E = 1-(1-.2)(1-.3)(1-.4) = .664$

VII, B. 4. (cont.) thus all measured factors (A-J) did not account for 1 - .664 = 33.6% of all cases of disease D — new model or factors?

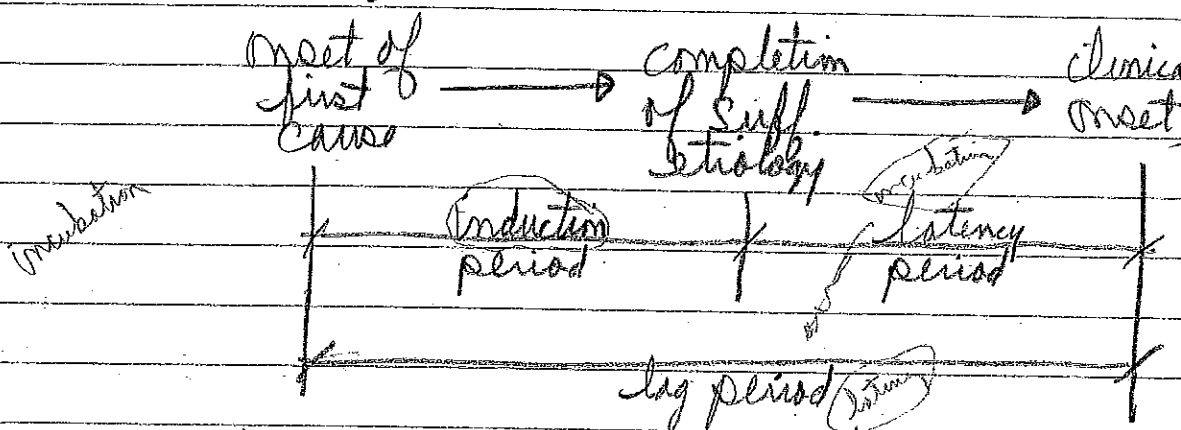
* Biased estimate of statistical I.: 3 factors (Selection, meas., conf.) 2 " (M = E, lag period)

VIII, Lag period —

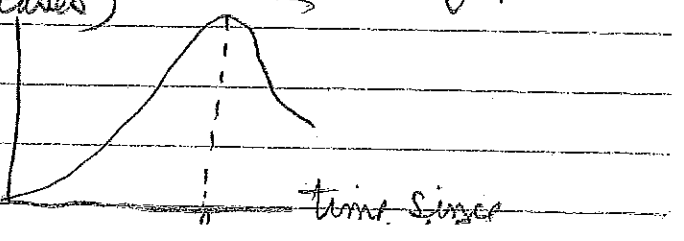
A. ~~Since~~ it is unlikely that any single factor is a "sufficient" cause of chronic V.D., since it is bound to statistically interact w/ a no. of other factors

stat. interaction need not imply any bio-beh. interaction

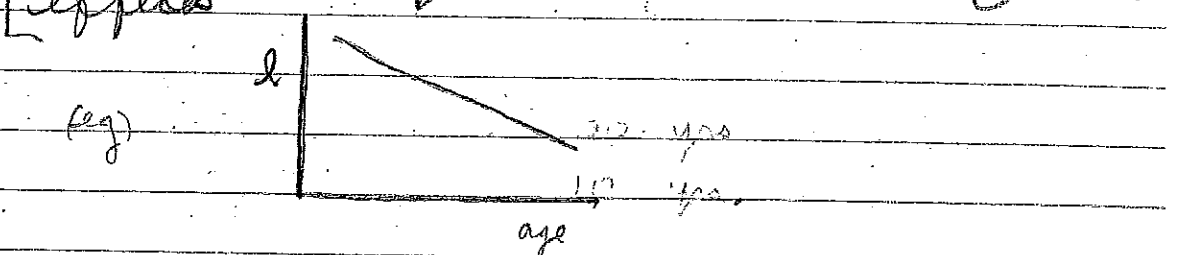
B. in studying diseases w/ multiple risk factors, the lag time between onset of first cause, & clinical onset of dis. may create complex analytical problems — consider the following conceptualization



C. but from practical standpt. it is only possible to estimate the "lag" period — if ~~the~~ E's begins at a certain age & remain stable thereafter — if this is possible we could estimate lag period empirically (new cases)



VIII. D. (cont'd) the fact that the lag period is not constant but, in fact, has Variance > 0, ~~[illegible]~~ complicates the estimation of EM + I − thus: even without a true EM effect, there may appear to be EM (Spurious) − ie, the effect of E on D may be observed to vary along categories of M, even if M is not a risk indicator (i.e. not true E mod), any vars. which is assoc. with the lag time in study pop (eg age) may appear to modify the effect of E or interact w/ it, thereby distorting the true nature of effects



modifying effect of age would be distorted (tho not entirely spurious)

E. #11 −

2/4/76 Outline 2/11

I. Introd.
- A. topic - principles of stat. analysis of categorical data - today: Simple (unstratified) analysis
- B. analysis of etiologic research question involving assoc. of E → D — 2 components
  - (is there an assoc.?) 1. hypothesis testing - qualitative
  - (how much?) 2. estimation of effect - quantitative (more important)
    - a. pt. estimation - eg RR̂
    - b. interval estimation - eg. CI[RR̂]

II. H. testing: nominal (categorical data) — simple analysis
- (A) common procedure: χ^2 test of homogeneity (c-R or CH indep. (2-S)

	D	D̄	
E	a	b	n_E
Ē	c	d	$n_{\bar{E}}$
	m_D	$m_{\bar{D}}$	n

in gen'l. $\chi^2 = \sum_h \sum_i \frac{(O_{ih} - \mathcal{E}_{ih})^2}{\mathcal{E}_{ih}} = \frac{n(ad-bc)^2}{n_E n_{\bar{E}} m_D m_{\bar{D}}}$ (2×2 table)

- (B) $Z = \sqrt{\chi^2} = \chi$ ("chi-statistic") (more exact tables)

Suggesting that a test of homog. is also testing the difference between proportions — same test

c. actually χ^2 or χ can be derived from any measure of assoc. (A) that is N distrib.

A as a variable

$$\chi = Z = \frac{A - \bar{A}}{\sqrt{V[A]}} = \frac{A - \mathcal{E}[A]}{SE[A]} \cong \frac{A}{\mathcal{E}[A]}$$

("diff in prop.")

if $A \equiv RD\ (I_D, \mathcal{E}_D, P_D) \Rightarrow \chi_{RD} = \frac{A}{SE[A]_{\mathcal{E}}}$

$= \sum\sum \frac{(O-\mathcal{E})^2}{\mathcal{E}}$ as above

Fcn of C-R
hypg model

	D	$\bar{D}$	
E_1	a	b	n_1
E_0	c	d	n_0
	m_1	m_0	n

So we sample n_1 exposed from pop. of n; — # of hyp cases (a) is distrib. as hypg. distrib.

$N = n$

$n = n_1 \quad p = \frac{m_1}{n}$

$D = m_1$

$$E[a] = \frac{nD}{N} = np = \boxed{\frac{n_1 m_1}{n}}$$

$$V[a] = \frac{n D (N-D)(N-n)}{N^2(N-1)} = np(1-p)\left(\frac{N-n}{N-1}\right) = \boxed{\frac{n_1 m_1 m_0 n_0}{n^2(n-1)}}$$

II. C. (cont.) but could use any A that is N distrib., tho the value of χ might be slightly different

begin → D. consider: RR — logN distrib ∴ $\ln RR \sim N$

$$\therefore \chi = \frac{\ln RR - \mathcal{E}[\ln RR]}{SE[\ln RR]} = \frac{\ln RR}{SE[\ln RR]}$$

could then estimate SE (lge. sample approx.) eg. $SE[\ln \hat{OR}] \cong \sqrt{\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}}$

E. another A: we will use for remaining discussion on analysis — $A \equiv \hat{MH}$

$$A \equiv \boxed{\hat{MH}} = \frac{n(n-1)(RD)}{m_D m_{\bar{D}}} = \frac{n(n-1)(ad-bc)}{n_E n_{\bar{E}} m_D m_{\bar{D}}}$$

→ not usually computed directly because hard to interpret — NB. null value = 0

but we can use to get χ → M-H test

$$\chi_{MH} = \frac{\hat{MH} - \mathcal{E}[\hat{MH}]}{SE[\hat{MH}]} \qquad \hat{MH} \sim N$$

in I_c study: $\mathcal{E}[\hat{MH}] = 0$

(M+H) $SE[\hat{MH}] = \sqrt{\frac{n^2(n-1)}{n_E n_{\bar{E}} m_D m_{\bar{D}}}}$

$$\therefore \chi_{MH} = \frac{n(n-1)(ad-bc)}{n_E n_{\bar{E}} m_D m_{\bar{D}}} \Bigg/ \sqrt{\frac{n^2(n-1)}{n_E n_{\bar{E}} m_D m_{\bar{D}}}} =$$

calculating formula / conceptual formula

$$= \frac{a - \left(\frac{m_D n_E}{n}\right)}{\sqrt{\frac{n_E n_{\bar{E}} m_D m_{\bar{D}}}{n^2(n-1)}}} = \frac{a - \mathcal{E}[a]}{\sqrt{V[a]}}$$

consider distrib of "a"

(n-1) (n is known)

gen'l. form.

N: using hypergeometric model: $\mathcal{E}[a] + \sqrt{V[a]}$

sample w/replacement

II. E. (cont.) NB. not even necessary to compute $\hat{M}_{MH}$

F. χ_{MH} is same for C-R study $= E[a] + V[a]$

	D	D̄	p-w.
E	a	T_E	
Ē	c	$T_{\bar{E}}$	
	m_D	T	

but somewhat diff. for I_D study: use binomial model because n is unknown

$$E[a] = \frac{m_D T_E}{T} = np \qquad V[a] = \frac{m_D T_E T_{\bar{E}}}{T^2} = np(1-p)$$

E (np)

$n = m_D$

$p = T_E/T$

G. Sample: I_c study

	D	D̄	
E	40	50	90
Ē	30	80	110
	70	130	200

Sample m_D cases from pop. w/ $p = T_E/T$; What is # cases, a?

$$\chi_{MH} = \frac{a - E[a]}{\sqrt{V[a]}} = \frac{40 - \frac{70(90)}{200}}{\sqrt{\frac{90(110)(70)(130)}{200^2(199)}}} = 2.527$$

(2-p = .012)

we will see ~~next week~~ that this technique is easily adapted to stratified analysis, but first ⊥ CI

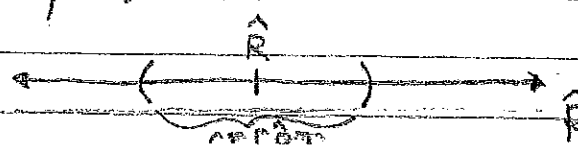
III. Estimation of Effect

A. Point Est. – covered previously: $\hat{RR}$, $\hat{RD}$ (crude) w/o conf.

B. Interval Est. – CI

1. CI = empirically derived range of values (or interval) expressing the P_r that the interval (constructed around pt. est.) includes the true pop. measure

eg, R, $\hat{R}$ CI$[\hat{R}]$



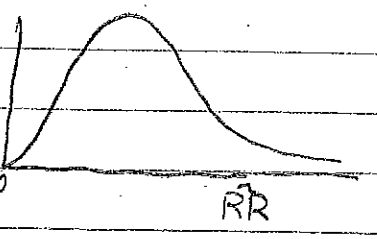
III. B. 2. We express CI as $CI_\gamma[\hat{R}]$
where: γ = conf. level & represents designated precision of the CI

3. in genl., if $\hat{R}$ is N distrib.

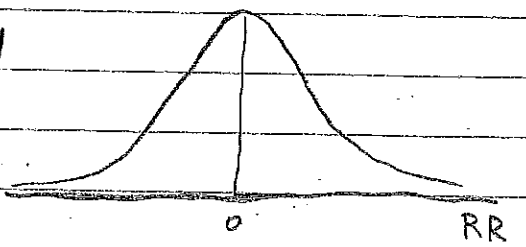
$$CI_\gamma[\hat{R}] = \hat{R} \pm CV_\gamma \, SE[\hat{R}]$$

where: CV_γ = crit. value, from known distrib. Z or $t_{(n-1)}$ if σ^2 is not known

C. CI around ratio, eg $\hat{RR}$ ~ not N distrib
but $\hat{RR} \sim \ln N$ distrib



so, $\ln \hat{RR} \sim N$



Thus,

$$CI_\gamma[\ln \hat{RR}] = \ln \hat{RR} \pm Z_\gamma \, SE[\ln \hat{RR}]$$

$$CI_\gamma[\hat{RR}] = \exp\left[\ln \hat{RR} \pm Z_\gamma \cdot SE[\ln RR]\right]$$

but how to estimate SE?
if RR = OR $\Rightarrow$ $SE[\ln OR] \overset{\text{large samples}}{\approx} \sqrt{\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}}$

could also use approx. from Taylor Series (genl)

$$SE[\ln \hat{RR}] \approx \frac{SE[\hat{RR}]}{\hat{RR}}$$ ← but what is this

CI[$\hat{EF}$] – proof:

If: $(1-\hat{EF}) \sim lnN$

$ln(1-\hat{EF}) \sim N$

$\therefore CI[ln(1-\hat{EF})] = ln(1-\hat{EF}) \pm Z_\alpha \, SE[ln(1-\hat{EF})]$

$SE[ln(1-\hat{EF})] = ln(1-\hat{EF})/\chi$

$\therefore CI[1-\hat{EF}] = exp\left[ln(1-\hat{EF}) \pm Z_\alpha \cdot \frac{ln(1-\hat{EF})}{\chi}\right]$

$= exp\left[ln(1-\hat{EF})\left(1 \pm \frac{Z_\alpha}{\chi}\right)\right]$

$= (1-\hat{EF})^{(1 \pm Z_\alpha/\chi)}$

$$\boxed{CI[\hat{EF}] = 1-(1-\hat{EF})^{(1 \pm Z_\alpha/\chi)}}$$

... but is this good for the study?

If: $\hat{EF} \sim N$ (or $(1-\hat{EF}) \sim N$)

$\therefore CI[\hat{EF}] = \hat{EF}(1 \pm Z_\alpha/\chi)$

CI[$\hat{Id}$] – CI[$\hat{Ic}$] – #12 on ref. list

III. D. new method for computing CI – method of "test-based intervals" – avoids direct computation of SE

recall: $\chi = \dfrac{A - \mathcal{E}[A]}{\sqrt{V[A]}} = \dfrac{\ln \hat{RR}}{SE[\ln \hat{RR}]}$

$$SE[\ln \hat{RR}] = \boxed{\frac{\ln \hat{RR}}{\chi}}$$ ← as computed before: $\frac{a - \mathcal{E}[a]}{\sqrt{V[a]}}$

an exact expression

Subst. into formula for CI:

$$CI_\gamma[\hat{RR}] = \exp\left[\ln \hat{RR} \pm Z_\gamma \cdot \frac{\ln \hat{RR}}{\chi}\right] =$$

$$= \exp\left[\ln \hat{RR}\left(1 \pm \frac{Z_\gamma}{\chi}\right)\right] =$$

antilog of log $= \boxed{\hat{RR}^{\left(1 \pm \frac{Z_\gamma}{\chi}\right)}}$ where: χ

E. can also use this method w/ RD ~ N distrib.

$$\chi_{RD} = \frac{\hat{RD}}{SE[\hat{RD}]}$$

$$SE[\hat{RD}] = \frac{\hat{RD}}{\chi}$$

$$\therefore\ CI_\gamma[\hat{RD}] = \hat{RD} \pm Z_\gamma \frac{\hat{RD}}{\chi}$$

$$= \boxed{\hat{RD}\left(1 \pm \frac{Z_\gamma}{\chi}\right)}$$

F. + for EF (mPF)

$(1 - \hat{EF}) \sim \ln N \Rightarrow$ $\boxed{CI_\gamma[\hat{EF}] = 1 - \left(1 \pm \hat{EF}_N\right)^{\left(1 \pm \frac{Z_\gamma}{\chi}\right)}}$ (w/ EF ≥ 0)

(note CI, must be ≥ 0 & ≤ 1)

proof ↪

new

III F'. $CI_\gamma[\hat{I}d_i] = \exp\left[\ln \hat{I}d_i \pm Z_\gamma SE_i[\ln \hat{I}d_i]\right]$ (i = (ref, or unref.))

$= \hat{I}d_i \, e^{\pm Z_\gamma SE[\ln \hat{I}d]}$

could remove i's to get $CI[\hat{I}d]$ ⇒ only need

$e^{\ln Id}$ $e^{\pm Z\,SE}$

where: $SE_i^2[\ln Id_i] = V_i[\ln Id_i] =$

$V_i[\ln \hat{I}d] + \dfrac{(\ln \hat{I}d_i - \ln \hat{I}d)^2}{\chi^2}$

where: $V[\ln \hat{I}d] = \dfrac{1}{a+c}$ inverse of total # cases in the estimate

F''. $CI_\gamma[\hat{I}_{c_i}]$ = similar to above except

$V[\ln \hat{I}_c] = \dfrac{1}{\hat{I}_c^2} \sum \left[\dfrac{Id_j^2}{(a_j+c_j)}\right](p_j^{\,2})$ (duration)

NB. $\hat{I}d_j$ is for both E+Ē

When $\hat{I}_c$ is computed from Id_j s

$\hat{I}_{c_{(p)}} = 1 - e^{-\sum Id_j p_j}$

but what about in Tc study?
I guess, you can consider $\hat{I}_c \sim N$?

III. G. NB: "$\pm$" indicates calculation of CI_L + CI_u

implicitly suggests:

2-sided $\begin{cases} H_0: & \hat{RR} = 1 \quad \text{(no effect)} \\ H_1: & \hat{RR} \neq 1 \end{cases}$

yet in epid there may be no biologic or soc. basis (plausibility) for hypothesizing E as protective (or causal) ∴ can increase power of test by employing 1-sided test

$H_0: \hat{RR} = 1$

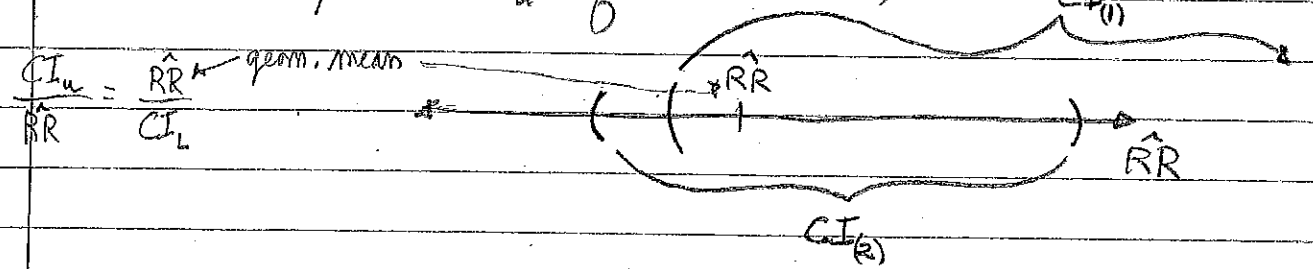
$H_1: \hat{RR} > 1$ or $\hat{RR} < 1$

∴ p-value will be cut in half (given same γ)

$Z_{1-\alpha/2} \rightarrow Z_{1-\alpha}$

OM recommends that even a CI reflect the one-sidedness of most epid H's by computing a tighter 1-sided CI_L w/ an implied CI_u of ∞ (same γ)

$\frac{CI_u}{\hat{RR}} = \frac{\hat{RR}}{CI_L}$ ← geom. mean



∴ increasing ability of rejecting the H_0

$CI_\gamma[\hat{RR}] = [CI_L, \infty] = \left(RR^{\left(1 \pm \frac{Z_\gamma}{\chi}\right)}\right)$

H. Example - C-R study

	D	$\bar{D}$	
E	40	23	63
$\bar{E}$	60	77	137
	100	100	200

III. H. let: $\gamma = .95$

$\therefore \alpha = .05$ (1-sided test)

$Z_{.95} = 1.645$

$\hat{I}_D R = OR = \frac{ad}{bc} = \frac{40(77)}{23(60)} = 2.232$

$$\chi_{MH} = \frac{40 - \frac{63(100)}{200}}{\sqrt{\frac{63(137)(100)(100)}{200^2(199)}}} = 2.581$$

(use χ^2 table, $p = .005$ (1-sided))

$CI_L[\hat{RR}] = 2.232^{\left(1 - \frac{1.645}{2.581}\right)} = 1.338$

$(CI_u = \infty)$

If we had used traditional method for calculating CI by est. SE, CI_L would have been somewhat less (closer to 1

I. advantages of test-based method - p. 48

J. Summary of χ calculations for different study designs - pp. 28-32, #10

next: Stratified analysis

Addendum

Ic on CR

Hypergeometric

	D	$\bar{D}$	
E_1	a	b	n_1
E_0	c	d	n_0
	m_1	m_0	n

$a \sim$ Hyp

$N = n$

$n = n_1$

$D = m_1$

$\left(\frac{m_1}{n} = p\right)$

$$\mathcal{E}[a] = \frac{nD}{N} = \frac{n_1 m_1}{n}$$

correction factor

$$np(1-p)\left(\frac{N-n}{N-1}\right) = V[a] = \frac{nD(N-D)(N-n)}{N^2(N-1)} =$$

$$V[a] = \frac{n_1 m_1 m_0 n_0}{n^2(n-1)}$$

$m_1 \sim$ Bin or $a_1 \sim$ Bin

$p = m_1/n$

$m_1 \dot{\sim}$ Nor if n is large

$$\frac{m_1 - np}{\sqrt{np(1-p)}} = \frac{\frac{m_1}{n} - p}{\sqrt{\frac{p(1-p)}{n}}} = \frac{\hat{p} - p}{\sqrt{V[\hat{p}]}}$$

Can p be $Id = \frac{m_1}{T}$?

Id

Binomial

	D	$p-n$
E_1	a	T_1
E_0	c	T_0
	m_1	T

$a \sim$ Bin

$n = m_1$

$p = T_1/T_0$

$$\mathcal{E}[a] = np = \frac{m_1 T_1}{T}$$

$$V[a] = np(1-p) = \frac{m_1 T_1 T_0}{T^2}$$

$$\frac{\frac{a}{T_1} - \frac{m_1}{T}}{\sqrt{\frac{m_1 T_0}{T_1 T^2}}} \quad Id_1 - Id_0$$

$$\hat{I}_d = \frac{10 \text{ cases}}{1000 \text{ p-yrs.}} = .01$$

$$CI[\hat{I}_d] = \ln(.01) \cdot e^{\pm 1.96 \cdot SE[\ln .01]}$$

$$SE[\ln .01] = \sqrt{V[\ln .01]} = \sqrt{\frac{1}{10} + }$$

$$CI = .01\, e^{\pm \frac{1.96}{\sqrt{10}}} = [.005 - .019]$$

$$CI = .01 \pm 1.96\sqrt{\frac{.01(.99)}{1000}} \quad [.004 - .016]$$

	D	p-yrs.
E_1	15	500
E_0	10	500
	25	1000

$$\hat{I}_{dR} = 1.5$$

$$CI[\hat{I}_{dR}] = 1.5^{\left(1 \pm \frac{1.96}{z}\right)} = 1.5^{1 \pm 1.96}$$

$[.68 - 3.32]$

$$z = \frac{15 - \frac{25(500)}{1000}}{\sqrt{\frac{25(500)(500)}{1000^2}}} = 1$$

$$\hat{I}_{dD} = .01$$

$$CI[I_{dD}] = .01(1 \pm 1.96) = [.0096 - .0296]$$

$$CI[I_{dD}] = .01 \pm 1.96\sqrt{\frac{.03(.97)}{500} + \frac{.02(.98)}{500}} = [.0093 - .0293]$$

2/11/76 Outline + 2/18

I. Introd.

A. extend principles of simple analysis to stratified analysis

II. Hypothesis testing – MH test

A. consider q strata of C (pot. conf. factors) – $j = 1 \ldots q$

$$\chi_{MH_j} = \frac{a_j - E[a_j]}{\sqrt{V[a_j]}} = \frac{\hat{MH}_j}{SE[\hat{MH}_j]} \qquad \hat{MH}_j \sim N, \; E = 0$$

$$\sum_{j=1}^{q} \chi_{j} = \sum \frac{\hat{MH}_j}{SE[\hat{MH}_j]} = \text{"total chi"} \quad (\text{or } \Sigma \chi^2_{j(q)})$$

H_0: no assoc. in any strata ∴ no good because strata may distort results

B. what is needed is an overall test statistic:

χ_0, which is a function of an overall measure of assoc. where A_0 combines strata to form a pooled est.

1. 2 ways of getting A_0
   a. ML est. – esp. when #'s are small (per cell)
   b. weighted average of A_j – requires large #'s in each cell

2. (1.b.) → $A_0 = \frac{\Sigma w_j A_j}{\Sigma w_j}$ where: $w_j = \frac{1}{V[A_j]} = \frac{1}{SE^2[A_j]}$

ie, wts. maximize efficiency of the estimate (unlike SE)

$$\therefore SE[A_j] = \frac{1}{\sqrt{w_j}}$$

Since A_0 is a wted. sum of A_j's, $SE[A_0] = \frac{1}{\sqrt{\Sigma w_j}}$

II. B. 3. using $A_j = \hat{M}H_j$, w/o proof (in paper)

$$\chi_{o(1)} = \frac{a - \mathcal{E}[a]}{\sqrt{V[a]}} = \boxed{\frac{\sum a_j - \sum \mathcal{E}[a_j]}{\sqrt{\sum V[a_j]}}}$$

[NB: #'s in cells needn't be large]

∴ same form as in simple analysis

where: $\sum \mathcal{E}[a_j] = \sum \frac{m_{Dj}\, n_{Ej}}{n_j}$ in I_c or CR

$= \sum \frac{m_{Dj}\, T_{Ej}}{T_j}$ in I_D

$\sqrt{\sum V[a_j]} = \sqrt{\sum \frac{n_{E\cdot} n_{\bar{E}j}\, m_{Dj}\, m_{\bar{D}j}}{n_j^2 (n_j - 1)}}$ in I_c or C-R (hypgeom. V)

$= \sqrt{\sum \frac{m_{Dj}\, T_{Ej}\, T_{\bar{E}j}}{T_j^2}}$ I_D (bin V.)

III. Test for uniformity

A. χ_o ($\sim \chi_o^2$) is applicable to any study design, where C is conf. —
but, if C is also an effect mod., A_o (or $\hat{M}H_o$) is a deceptive measure because different estimates are pooled —
so, you don't get A_o unless there is no sign. EM — ie, all strata-sp. est. need not be exactly equal, but the data must be consistent with an assumption of uniformity — (ie, not sign. diff. from uniformity)

III. B. the test: $\chi^2_{u(q-1)} = \sum \chi^2_{j(q)} - \chi^2_{o(1)}$

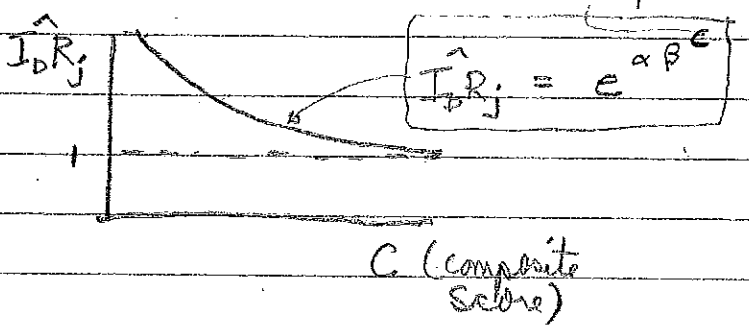
computed as described

if p is small ⇒ reject H_o of uniformity ; don't pool strata to get A_o – tho can still use χ_o – can get SMR, SRR

if p is large ⇒ no EM ∴ use A_o as pt. est.

* if $A_o = \hat{M}H_o$ (ie, $\chi_j = \chi_{MH}$), then χ^2_u is only good in a C-R study – I guess because $\hat{M}H$ isn't a stable est. in a CH study (like OR)

C. recall that χ^2_u isn't a very sensitive test because it only requires nominal (unordered) data we might fit the data to a model instead – eg, (OM et al, 1976) – Cig. Sm. –?→ acute MI (non-fatal).



testing: ① $\alpha > 0$ ⇒ assoc. between cig Sm + AMI
② $\beta < 1$ ⇒ $\hat{RR}_j$ approaches unity w/ ↑C (ref. risk)
③ fit of model

IV. Pt. est. of overall effect (A_j) – don't usually get $\hat{M}H_o$ because it has little meaning (not RD nor RR) must have large #'s in each strata, or use ML est.

↑ who Kleinbaum computes it in 287

IV. A. I_c study —

$$\hat{RD}_0 = A_0 = \hat{I_c D_0} = \frac{\sum w_j \hat{I_c D_j}}{\sum w_j} \qquad \hat{RD_j} \sim N$$

$$w_j = \frac{n_{Ej}\, n_{\bar{E}j}\, n_j}{m_{Dj}\, m_{\bar{D}j}} \simeq \frac{1}{V}$$

$$\hat{RR}_0 = A_0 = \hat{I_c R_0} \qquad \hat{RR_j} \sim \ln N$$

$$\therefore \ln \hat{I_c R_0} = \frac{\sum w_j (\ln \hat{I_c R_j})}{\sum w_j}$$

$$\hat{I_c R_0} = \exp\left[\frac{\sum w_j (\ln \hat{I_c R_j})}{\sum w_j}\right]$$

$$w_j = \frac{m_{Dj}\, n_{Ej}\, n_{\bar{E}j}}{m_{\bar{D}j}\, n_j} \simeq \frac{1}{V}$$

B. I_D study

$$\hat{RD}_0 = A_0 = \hat{I_D D_0} = \frac{\sum w_j \hat{I_D D_j}}{\sum w_j}$$

$$w_j = \frac{T_{Ej}\, T_{\bar{E}j}}{m_{Dj}}$$

$$\hat{RR}_0 = A_0 = \hat{I_D R_0} = \exp\left[\frac{\sum w_j (\ln \hat{I_D R_j})}{\sum w_j}\right]$$

$$w_j = \frac{m_D\, T_{Ej}\, T_{\bar{E}j}}{T_j^2}$$

IV. C. C-R study

$\hat{RR}_o$ — could use same technique as above, w/ $OR_j \equiv \hat{I}_b R_j$; but there is a simpler way:

M-H Procedure for est. RR in a C-R study

$$\hat{RR}_o = \hat{I}_b \hat{R}_o = \frac{\sum \frac{a_j d_j}{n_j}}{\sum \frac{b_j c_j}{n_j}} = \frac{\sum w_j \cdot OR_j}{\sum w_j} \qquad w_j = \frac{b_j c_j}{n_j}$$

can still use χ_o (χ_{MH})

D. Table 1, #10, pp. 28-32 — Summary of χ_{MH} + A_o for different study designs + analysis including Matched analysis, to be covered next week

V. Interval Estimation — CI

A. for overall meas. of assoc. — A_o
use test-based intervals exactly as in Simple anal.

eg, $CI_{1-\gamma}[\hat{RR}_o] = \hat{RR}_o^{\left(1 \pm \frac{z_\gamma}{\chi}\right)}$

where: $\hat{RR}_o$ is one of the previous weighted formulas or a ML est.

this method (CI) is only good when the A has maximized effic. — which occurs when there is uniformity of effect over Strata (of c)

Similar to SMR

CH $\quad CI_{\gamma}[sRD] = \hat{sRD} \pm Z_{\gamma}\sqrt{\hat{V}[sRD]}$

where: $\hat{V}[sRD] \cong \dfrac{\sum_j n_{1j} Ic_{1j}(1-Ic_{1j})}{\left(\sum_j n_{1j} Ic_{0j}\right)^2} + \dfrac{\sum_j n_{0j} Ic_{0j}(1-Ic_{0j})}{\left(\sum n_{0j} Ic_{0j}\right)^2}$

optional CH only

$$CI_\gamma[SRD] = SRD \pm Z_\gamma \sqrt{\hat{V}[SRD]} \qquad \left(\frac{s_j}{S} = \frac{n_{0j}}{n_j}\right)$$

$$\hat{V}[SRD] \cong \sum_j \left(\frac{n_{sj}}{n_s}\right)^2 \left[\frac{I_{c_1 j}(1-I_{c_1 j})}{n_{1j}} + \frac{I_{c_2 j}(1-I_{c_2 j})}{n_{2j}}\right]$$

V B. for standardized measures - SMR, SRR, which haven't max. effic. - must estimate SE

assume: E_j ~ Bin'l indep. for each category

eg, $CI_\gamma[\hat{SMR}] = \hat{SMR} \pm Z_\gamma \cdot SE[\hat{SMR}]$ $\quad \sqrt{\hat{V}[SMR]}$

(tho SMR not Bin'l)

$$\hat{V}[\hat{SMR}] \cong \frac{\sum n_{Ej}\, \hat{R}_{Ej}(1-\hat{R}_{Ej})}{\left(\sum n_{Ej}\, \hat{R}_{Ej}\right)^2}$$

n_i

? SRR ? by changing

but can't compute SE (or $\hat{R}$) in C-R studs so I don't know if SE has been estimated for C-R study

VI. Example - C-R study

C_1

	D	$\bar{D}$	
E	35	55	90
$\bar{E}$	20	90	110
	55	145	200

C_2

	D	$\bar{D}$	
E	25	35	60
$\bar{E}$	10	30	40
	35	65	100

A. Is there conf.?

$$CRR = \frac{60(120)}{90(30)} = 2.667$$

$$SMR = \frac{60}{\frac{55(20)}{90} + \frac{35(10)}{30}} = 2.512$$

$RR^* = 1.062$ $\quad RCE = .041$ (Some pos. conf.)

B. EM? $\quad OR_1 = 2.86$, $OR_2 = 2.14$ } is this diff. stat. sign?

II. c. But first, is there an overall assoc. –

Impact V

$$\chi_{o(1)} = \frac{a - E[a]}{\sqrt{V[a]}} = \frac{(35+25) - \left(\frac{55(90)}{200} + \frac{35(60)}{100}\right)}{\sqrt{\frac{90(110)(55)(145)}{200^2(199)} + \frac{60(40)(35)(65)}{100^2(99)}}}$$

$$\approx 3.627 \qquad p = .00014 \text{ (1-sided)}$$

$$\chi^2_o = 13.157$$

D. check to see whether EM occurred by chance (to determine the type of pt. est. to use)

$$\chi^2_{u(g-1)} = \sum \chi^2_{j(g)} - \chi^2_{o(1)}$$

$$\sum \chi^2_j = \chi^2_1 + \chi^2_2 = \frac{\left(35 - \frac{55(90)}{200}\right)^2}{\frac{90(110)(55)(145)}{200^2(199)}} + \frac{\left(25 - \frac{35(60)}{100}\right)^2}{\frac{60(40)(35)(65)}{100^2(99)}}$$

$$= 13.494$$

$$\therefore \chi^2_u = 13.494 - 13.157 = .337$$

$$\chi_u = \sqrt{.337} = .587 \qquad p = .561 \text{ (2-sided)}$$

$\therefore$ no EM

E. $\Rightarrow$ can get $\hat{RR}_o$ using MH procedure

$$\hat{RR}_o = \frac{\sum \frac{a_j d_j}{n_j}}{\sum \frac{b_j c_j}{n_j}} = \frac{\frac{35(90)}{200} + \frac{25(30)}{100}}{\frac{55(20)}{200} + \frac{35(10)}{100}} = 2.583$$

VI. F. CI_γ — $\gamma = .95$

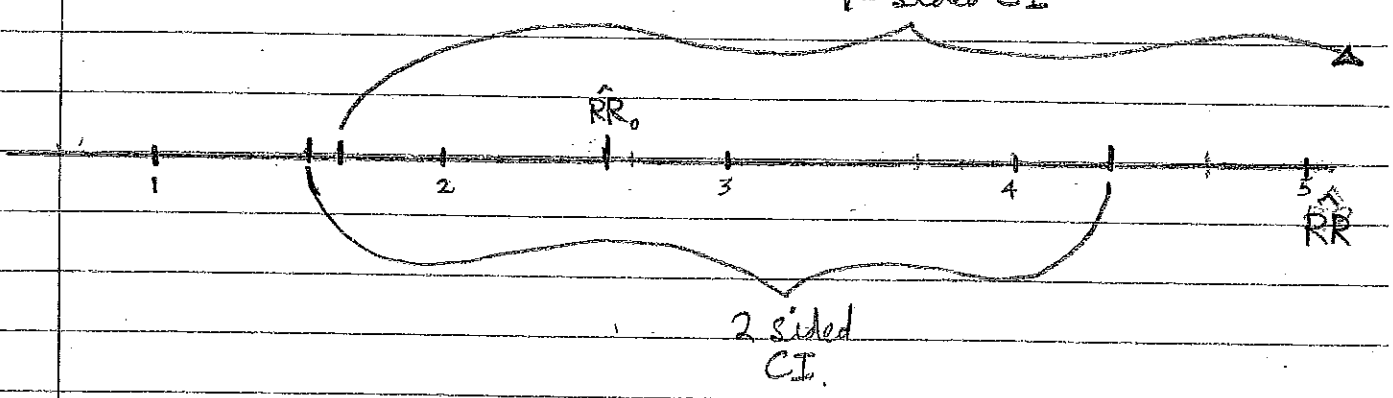
$\alpha = .05$ (1-Sided) $Z_\gamma = 1.645$

$95\%\ CI_L = \hat{RR}_0^{\left(1 - \frac{Z_\gamma}{\chi}\right)} = 2.583^{\left(1 - \frac{1.645}{3.627}\right)} =$

$= \underline{1.68}$ ∴ 95% confid. that "true RR" is > 1.68

2-Sided 95% CI $= 2.583^{(1 \pm 1.96/3.627)}$

$= [\underline{1.55} ; \underline{4.31}]$



2/18/76 Outline

I. Intro. → Ch. 8 (OM) –

A. $CI_0[\hat{EF}]$
$CI_0[\hat{I_0}] + CI[\hat{I_c}]$ } other notes

B. today:
1. finish stratified analysis
2. Matching + M. analysis (#12)

II. Defn's

A. design: predictor + outcome (response) var's
predictor var → test + referent groups

B. M – one way to control for conf. factors in selection of subjects

C. M (in selection of subjects in obs. study) ≡ restricting the selection of referent series by making it similar to the test group (selected indep.) w/ re: to C (pot. conf.)

1. indiv. M.: each test subject is grouped with one or more referent subjects according to their value on C. → making ref. series comparable to test series w/ re: to C – each set of 1 test + 1+ ref's ≡ stratum (strata)

a. fixed-ratio (r) M – all sets of M. subjects contain same ratio of ref.: test subj.

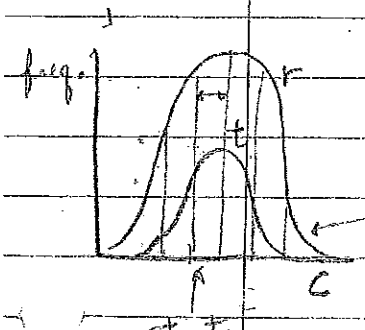
i. pairwise M – r = 1

ii. non-pairwise – r > 1

b. variable-ratio M – r is not uniform across strata

2. freq. M – forces comparability of distrib.'s of test + ref. series w/ re: to C –

OM: freq M is like other side of continuum from pairw. M – Strata defined over C, as in non-M
∴ must have lg #'s w/in each category of C (stratum) – analyzed like gen. stratified



D. procedure: all M factors are divided into a few (2–4) classes (ht) combining factors produces M. categories = Πk_i
eg. sex + race → 4 M cats.

II. E. Freq. M has less power than Ind. M. because ~~the~~ freq. M. has fewer df (= # strata - 1) where strata are larger - than Ind. M (df = # strata - 1) but strata are smaller (as little as 2)
∴ in planning a study there is little reason to use freq. M unless Ind. M. is too cumbersom say because of too many M categ. (>12) — on the other hand, we will see that such extensive M. may not be called for
— ∴ only consider Ind. M.

III. Analysis
A. Cohort - pairwise
1. n subjects - n/2 pairs - each E is M'ed w/1 $\bar{E}$
∴ n/2 strata of 4 types

	D	$\bar{D}$	
E	1	0	1
$\bar{E}$	1	0	1
	2	0	2

concordant

	D	$\bar{D}$	
E	1	0	1
$\bar{E}$	0	1	1
	1	1	2

	D	$\bar{D}$	
E	0	1	1
$\bar{E}$	1	0	1
	1	1	2

	D	$\bar{D}$	
E	0	1	1
$\bar{E}$	0	1	1
	0	2	2

concordant

Since all pairs of similar type give same info. (validity + precision) eg w/re: $\hat{RR}$ — they can be combined, which amounts to stratif. in analysis — summarized by 2×2 table

		E: D	E: $\bar{D}$	
E	D	P	Q	n
	$\bar{D}$	R	S	

n/2 (pairs)

$$\hat{ML}\ \&\ RR = \frac{P+Q}{P+R}$$

2. if Stratif. had been ignored

	D	$\bar{D}$	
E	P+Q [a]	R+S [b]	n/2
$\bar{E}$	P+R [c]	Q+R [d]	n/2
			n (subj.)

$$\hat{IcR} = \frac{\frac{P+Q}{n/2}}{\frac{P+R}{n/2}} = \frac{P+Q}{P+R} = \hat{ML}\ \text{w/ stratif.}$$

so failure to stratify in M. of study doesn't result in bias

III. A. 3. H_0 test — χ_{MH}

$$\chi_{MH} = \frac{a - \mathcal{E}[a]}{\sqrt{V[a]}} = \chi_0$$

where: $a = \sum a_j = P+Q$

$$\mathcal{E}[a] = P + \frac{Q+R}{2}$$

$$V[a_j] = \frac{n_{Ej}\, n_{\bar{E}j}\, m_{Dj}\, m_{\bar{D}j}}{n_j^2 (n_j - 1)} \quad \text{hypergeom. V}$$

$$= \frac{0}{2^2(1)} = 0 \text{ for concordant pairs}$$

$$= \frac{1(1)(1)(1)}{2^2(1)} = \frac{1}{4} \text{ for disc. pairs}$$

$$\therefore V[a] = \sum V[a_j] = \frac{Q+R}{4}$$

$$\chi_{MH} = \chi_0 = \frac{(P+Q) - \left(P + \frac{Q+R}{2}\right)}{\sqrt{\frac{Q+R}{4}}} = \boxed{\frac{Q+R}{\sqrt{Q+R}}} \quad \text{"McNemar's test"}$$

can use this χ to compute CI [$\hat{RR}$] as described

4. effect of fixed k M.: ck — no assoc. between E + C, but can still study assoc. between C + D (C as risk after having M. ind.)

$$\phi_{CD} = \frac{PS - QR}{\sqrt{(P+Q)(R+S)(P+R)(Q+S)}} \qquad 0 \le \phi \le 1$$

III. B. C-R – pairwise

1. n Subjects – n/2 pairs or strata – each D is M'd w/ 1 $\bar{D}$ –
Stratified analysis is represented as:

$$\hat{IdR} = \frac{X}{Y}$$

		$\bar{D}$	
		$\bar{E}$	E
D	E	W	X
	$\bar{E}$	Y	Z

n/2 (pairs)

2. Suppose: no stratif.

	D	$\bar{D}$	
E	$W+X$	$W+Y$	
$\bar{E}$	$Y+Z$	$X+Z$	
	n/2	n/2	n(subj)

$$\hat{IdR} = OR = \frac{(W+X)(X+Z)}{(W+Y)(Y+Z)} \neq \frac{X}{Y}$$

∴ failure to stratify in analysis can bias $\hat{RR}$.
often done

3. H_0 test – Similar derivation of χ_{MH}

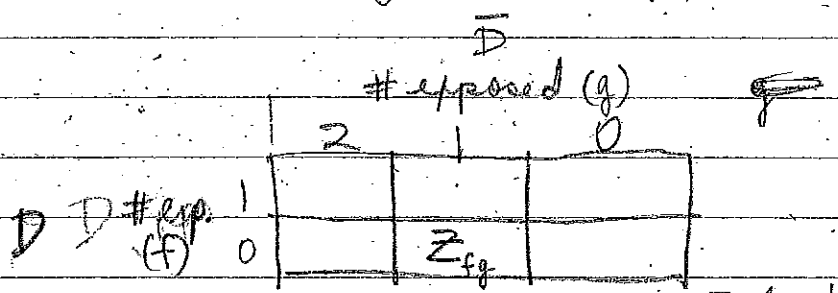
$$\chi_{MH} = \frac{X-Y}{\sqrt{X+Y}}$$ "McNemar test"

can use to compute test-based interval (CI)

4. effect fixed b/t M. – no assoc. between C+D so can't study C as risk indic of D after matching but can get assoc between C+E

$$\phi_{EC} = \frac{WZ - XY}{\sqrt{(W+X)(Y+Z)(W+Y)(X+Z)}}$$

MH est. of $\hat{RR}$ in fixed-r c-c study



z (sets of r controls + 1 case)

$f = 0, 1$

$g = 0, 1 \ldots, r-1$

$$(\text{not ML})\ \hat{IdR} = \frac{\sum_f \sum_g z_{fg}(r-g)f}{\sum_f \sum_g z_{fg}(1-f)g}$$

Last time: defined M. – restricting selection of ref. Series
analysis of pairwise M. in CH + C-R – $\hat{RR}$, χ_0, CI
study in analysis – CH / C-R

III. C. Fixed-ratio M – ($r > 1$)

1. can still extend basic MH test to $r>1$ for either C-R or CH design – the calculations are a little tricky (see: my paper)

$$\chi_{MH} \quad \frac{a - E[a]}{\sqrt{V[a]}}$$

2. NB: Stratif. analysis involves >4 cells

(eg) M- C-R study $r=2$

		$\bar{D}$: 2E	E/Ē	2Ē
D	E	U	V	W
	Ē	X	Y	Z

how many opp.?
$a+b = U+V+W + 2(U+X) + V+Y = 3U+2V+2X+W+Y$

so each stratum has 3 subjects (strata); 1D + 2D̄

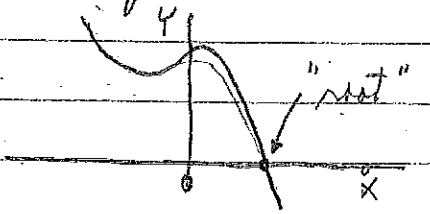
3. $\hat{RR}$ – most simple algebraic Soln.
but requires iterative soln. to get $\hat{ML}$ $r>2$

Mantel-Haenszel do have simple algebraic Soln.: but not $\hat{ML}$.

$$\hat{RR} = \frac{2W+V}{2X+Y}$$

⇒ Suppose we set up an equation to solve for an unknown (X or $\hat{RR}$), in which there is no alg. Soln. – eg,

$$X^3 + X - 1 = 0$$



gen'l approach: get successive approx.'s that narrow in on the exact root – how?

w/ recursive function of form: $X_{n+1} = f(X_n)$
(keep replacing X_n w/ X_{n+1}) X_0

$$X(X^2+1) = 1$$

$$X = \frac{1}{X^2+1} \Rightarrow X_{n+1} = \frac{1}{X_n^2+1}$$

X_n X_{n+1}

eg, start $X_0 = 0$ ∴ $X_1 = 1$

$X_1 = 1$ ∴ $X_2 = .5$

$X_2 = .5$ & $X_3 = .8$ ...

III. C. 3. in the case of $\hat{RR}$, OM, has shown how to get $\hat{ML}$ for M'ed C-R study - in order to derive the proper recursive function, need diff. calculus — use: N-R technique

D. Var-r M. — still use basic χ_{MH} (r_u) also can use iterative soln. to get $\hat{RR}$, treating each type of strata separately — r = 1, 2, ...

when needed: eg. ① M. sibs; ② limited # of ref. Subjects in certain M. categories

still use CI

E. optimum r: (similar Q. in unmatched designs)

1. w/ r > 4 or 5, little gain in size-eff. & thus informativeness — NB: usually limited in # cases
2. mostly a function of cost-eff. — ie, the relative costs of getting data on test and ref. subjects

$$r_0 \approx \sqrt{\frac{\$_t}{\$_r}} \quad \text{(one rule of thumb)}$$

∴ use r < 4 only when: $4 > \sqrt{\frac{\$_t}{\$_r}}$

$4 > \sqrt{16}$

∴ use r < 4 when ref. subj. costs > $\frac{1}{16}^{th}$ of test subj

ie: when ref. subj. are not readily available

~~if using secondary data, use r < 4 when there are fewer than 16~~

IV. When & on what vars. should one M?

A. M. effects accuracy & inform. of study, according to (write on board)

1. Assoc's in source pop
   a. C ~ E
   b. C ~ D
   c. among C's
2. Methodology (most important)
   a. type of study design — CH - CR
   b. " " M. — indiv (r) or freq.
   c. selection of subj. — add bias? (do assoc. from source pop. Δ
   d. type of analysis — stratified?

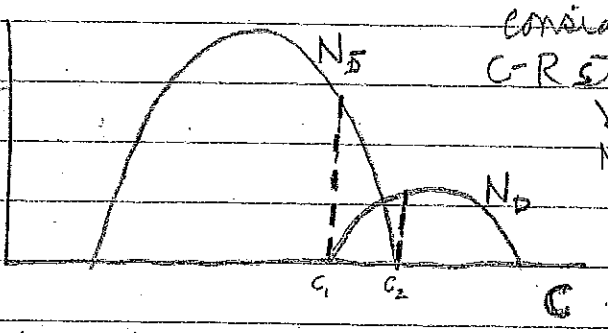
IV. B. 2 criteria for evaluating effects of M: 2 components of inaccuracy

1. Validity (internal) ↓ Systematic error → ↑ accuracy of pt est. → lack of bias
2. Size-eff. – lack of random error, reflected in range of CI for a given n.

C. See: Table 1, pp 30 (#12) for Summary – explain

1. NB: (1): ↓ indicates "overmatching"
   ↑ M was more effective than not M.
   (1): only E_s (constr val.) is effected by; this is not to say that M. is not done to control for conf. factors – only that bias may also be eliminated thru methods in analysis (eg, standardiz.)

explain how M. can improve E_s



in C-R study (or CH), info. is achieved by comparing D w/ $\bar{D}$; Stratific. works by holding C constant + making comparisons w/in each stratum – now suppose, we took a simple random sample of D + $\bar{D}$ in our C-R study: most noncases lie to left of c_1 and most cases to right of c_2 ∴ thus no info. would be gained from these subjects since they cannot be compared in their respect strata – the effect of M. is to ensure that all subjects lie w/in "overlap" region thus achieving

(*) not definitive → C. Statements of M. but based on 1 crucial assumption: no additional bias introduced in selection of study pop. from the source pop. decided (as assoc.) in table 1 – eg, Suppose age is risk factor of D in source pop. but restrict subj. to all 50–55 yr. olds at entry; age could not be predictive of dis. in study nor conf.

∴ table 1 is a guide, designed to illustrate principles of M not a system for making decisions

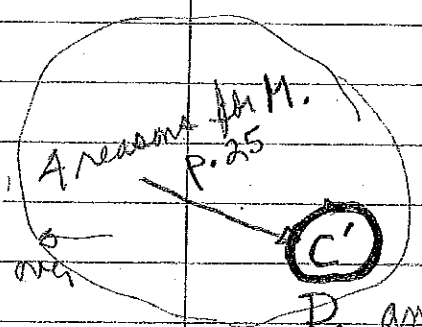
∴ table 1 represents expected effect of M. given the type of design, assoc. in source pop. + typ

IV. C. Reasons for M

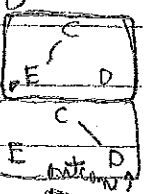
* 1. gain eff. (E_s) – to prevent empty cells in stratified analysis – esp. important when C is meas. on nominal scale w/ many categories (can't combine) eg "neighborhood"
3. cost of M. can be great (primary), but cost-eff. of M can be improved if there is an inexpensive supply of nf. Subjects (secondary) – assuming benefit in E_s exists
4. in primary data – M obviates need to collect info on C ⇒ ↓ compute processing – can't study modifying effect of C
5. M. simplifies analysis conceptually, making it intuitively appealing as a means of controlling for conf. – easy for non-stat. to see how M. removes bias (vs. multiv. analysis)
* 2. In hybrid or C-R study (with new cases), there is a practical advantage in matching since the comparison series (noncases) should ideally be comparable to cases w/ respect to time of diagnosis. If cases are thus identified over time, the most feasible way to control for time of diagnosis is to individual match. Since if all noncases are selected at one time (say end of case detection), the time of diagnosis can't be controlled for thru stratification.

relates to (1)

IV. C. 1. (cont.) more info./subject-ie, improving E_s
* ∴ M. reduces the loss in E_s that must result from Strat'g.
* N.B. even if distrib. are overlapping in their range, same benefit accrues
2. this gain in E_s is greatest when outcome var. is dichotomized
3. dividing contin. conf. var. into > 3 or 4 categories produces little gain in E_s
(Battersby empirical (not theor. results))
4. analogy: M in obs. studies is like blocking (B) in exp.
in exp. — control for conf. factors by randomization
Blocking (prior to R.) makes randomiz. more eff. by creating homogeneous groups w/ lots of between-B. variance i.e. that the (true) effect being meas. is not a part of experimental error



D. another obs. from table is that effects of M. differ for C-R & CH studies
(1) CR: M → ↓ Size-eff. (tho not always)
CH: M → ↑ " " (usually)



M factor ? → outcome

* why this paradox? — in CH (as in exp) Subj. develop. D over follow-up period & so D (outcome) is not regulated by M. procedure — so if E is perfectly correlated w/ D, E var. has a better chance of showing its effect — but,
in C-R, E is often not but is predetermined at onset of study ∴ E variability is influenced by M. — if perfect cond. exists between C and E, & M. only limits variability of E which then has less of chance of showing its effect on D.

Similar misconceptions:
M — max. comparability not M — representativeness of controls

* selection of controls in C-R study

* w/ CR: ↑ in E_s when E ~ C is string points out a common fallacy of both M+ unM. C-R studies — D+E should not be alike on all variables except E, ⇒ ↓ E_s because of less var. in E

IV. D. (3) effects of stratif in analysis — (in M. ~~area~~ design)

C+I: (a) no loss in validity - ie. Selection bias can't be introduced by M.

(b) generally only gain in E_s thru stratif.

C-R: (a) can lose val. by not stratif. in analysis as we already saw because when θ_{EC} is large enough

(b) ↑ or ↓ in E_s

E. Table doesn't include X-S studies (of 1 pop) — M. is most costly ✱ is seldom warranted unless we have a fixed n — otherwise its cheaper to increase n and not match

F. if Outcome var. is meas. on interval scale: M. is generally less eff. than no M. & the use of Cov. analysis

✱ G. 2 rules of thumb for selecting C's (on which to M.

1. C's. should be strong predictors of dis. (esp. to avoid overM'g in C-R)

2. if several pot. C's (as there usually are), only M. on those that are not highly assoc. w/ each other. — even tho the cost of M on 1 var. isn't much more than 2, their intercorrel. will result in ↓ E_s because of M. on redundant factors. —

note example, p.10 Same IdR

3. can use selective M — but stratify (by categories of C) on all variables (including M var.) in the analysis

V. After M. — stratify or not

A. Can't tell if conf. after having M'ed, but can determine one assoc. — θ_{CE} in C-R; θ_{CD} in C+I —

B. see Table 2;

function of both rel & eff

C. must do both unless M. is E.

VI. Illustrative Ex. — C-R; hypoth. source pop. — p. 36 (#12)

A. do study in 4 ways (each way w/ n = 140)

1. pairwise matched – strat. (by pairs)
2. " – unstrat. + strat. (by categories of C)
3. unmatched – strat (since C is conf.)

4. matching – strat. by categories of C results in A1

B. Note: Id R for M. strat ≅ SMR unmatched strat
any diff. is due to rounding error, but –
they would have been different if ~~though~~ C
had also been an effect modifier — in that
case, fixing the ratio of test to ref. subj. (by M.)
implicitly alters the distrib. of C in the study pop
∴ creating a different weighting of ~~strata~~ category
Specified effects (of course applies to unM. analysis as well)

* Thus. M influences the Id R in a ~~matched~~ C-R
in which C is also an effect modifier ∴
we might want to make separate
analyses of the E → D assoc. for diff.
categories of C, but this requires
enough subjects per category of C →
seriously reducing the power of each test
∴ M in C-R not only prevents us from
studying C as ~~a~~ risk indicator, but also
makes it more difficult to study EM.

NB. M-strat analysis is more precise

* C. M. – than forgetting pairs, but stratifying in
the analysis by categories of C – get same R̂R
as strat. by pairs, + even slightly better p-
value → perhaps because of rounding errors

2/25 Outline 3/3/76

I. Intro.

A. today: new method for considering several pot. conf. OM calls: "Confounder Summarization"

B. the problem: we would like to determine assoc. between E + D controlling for several pot. conf.

C. Solutions (already discussed)

1. sequential strat. – no good because we saw that C's must be considered simultaneously

2. gen'l. stratified analysis – but when we stratify C's, we wind up w/ lots of strata (or cells) some w/ very few or no subjects – also strata in which there is either mostly test or ref. subjects ∴ loss of E_s – these cells w/ very few subjects or w/ mostly only one group provide little or no info.

3. model fitting (mathematically) – eg, covariance – problem
   (a) hard to do, esp. for non-statisticians
   (b) more difficult to interpret – since specific trends get buried in coef's etc. – eg, might want to know how interaction occurs, not just if is sign. – to draw inferences
   (c) emphasis is usually on sign. levels, not estimation of effect – tho, certain stat. models allow for easy computation of certain measures of Assoc. – ch 6 ff

   i. linear – $R = \alpha + \beta X_i$ $RD = \beta(X_E - X_{\bar{E}})$
   ii. ln linear – $R = \exp[\alpha + \beta X]$ $RR = \exp[\beta(X_E - X_{\bar{E}})]$
   iii. logistic – $R = 1/(1 + \exp[-(\alpha + \beta X)])$ $OR = \exp[\beta(X_E - X_{\bar{E}})]$
   (true even w/ X_i if no interaction terms)

(d) w/ step-wise model, showing sign. contrib. of a variable may be reduced by additional confounding variables (which would be partially but not fully) – just like stratification problem of too many strata

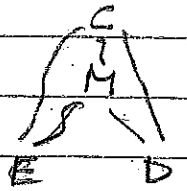
note: (a) stepwise regression looks for parsimonious model – ie, model w/ least terms explaining most variance – ie, eliminate redundancy which may have causal significance

(c) added assumptions may be needed – eg, in linear relationships no partial correl. or nonnormality = variance, or no interaction (in covariance)

⇒ prediction technique

if also conf.

I. C. 3.* (8). determining interaction creates a problem — if modifier (eg, age) is poor. w/ C, it is kept "fixed" w/ C in the analysis ∴ M rendered meaningless as EM



→ NB: could do partial correl. or regression (interpretation is a "worse problem"

II. New Method — (I'll call "composite stratification")

A. idea is to summarize all conf.'s into a single multiv. function + use this F to rank all subj.'s into a small (5) # of strata — then use genl strat. analysis as if we had 1 conf.

(one alternative — might also use binary variable multiple reg. Feldstein)

B. Multiv. statistics (model fitting) is used but it not interpreted as a predictive model, but only to control for conf. in one's data

III. Summarizing all conf. into 1 score (S) amounts to finding a way to pool the many strata that would result from genl strat. — i.e., the problem of too many strata will be dealt w/ by pooling them according to some criterion

A. this criterion: When can 2 strata be combined w/o introducing conf. in a study

1. $\frac{D}{E}$ for any type of study is requirement for no conf.
2. so can combine any strata in which the expected rate (or %) of D is identical for condition of $\bar{E}$ — eg, if rate (P, I, Id) of D is same for old fem. + young males, cond. on $\bar{E}$, can combine these 2 strata

B. So, the Conf. Summary Score (S) will be based on some multiv. function (F) conditional on $\bar{E}$ — i.e., S will be assuming $\bar{E}$ in order to rank subj.

III. C. Its not necessary that F actually estimate null risk ($Pr[D|\bar{C}\bar{E}]$), but only need allow ranking according to the null risk for Π in the study pop — & in fact this Pr is not even a risk in most C-R studies, only Pr[being a case | C, Ē]

IV. Obtaining the Conf.-Summarization Scores (S)

A. Use stat. model to fit data — F

1. type of model (dep't var)
   a. D is dichot. → linear discrim. function
   b. D is interval → multiple regression
2. indep. var's
   a. interval — use the value, eg, age
   b. categorical — create k−1 dummy var's — for k categories of X — eg, race: W, B, R — i.e. race = factor

$$X_1 = \begin{pmatrix} B=1 & \bar{B}=0 \end{pmatrix} \Big\} \text{—— = variables}$$
$$X_2 = \begin{pmatrix} R=1 & \bar{R}=0 \end{pmatrix}$$

   c. let E_i = exposure (may be > 1 var. if categorical) (E's dichotomous)
   d. let C_j = pot. conf. variables — including all dummy vars — or some transform of var. — eg, ln age

how to select initial C_j — from RR* one at a time — tho only a crude guide

B. create F (assume 2 categ.)

$$\hat{D} = a + bE + \sum c_j C_j = F$$ (a, b, c_j are coef. from model)

1. NB: Since object will only be to rank Subj, assumption (eg. ~N) need not be met
2. Size of c_j (coef.) doesn't indicate how conf. C_j is —
3. Since object is only to rank subjects, a is not needed

IV. C. Since any stratif. analysis reduces E_s, we would like as few C_j in model + still control for conf. to minimize loss in E_s — ie, reduction of ~~model~~ F

1. Should it be done by step-wise procedure or any technique based on sign. testing — Because object is to control for conf., which isn't a matter of making inferences to a source pop, but refers only to study pop.

2. to do: criterion — take out each C_j (one at a time) and obtain new F w/ new coef.: if b (for E) does not Δ much, ≠ that C was not very conf., so remove from model

Is there any order in removing C_j's — 2 ways:
→ (1) try model without each of the C_j's & drop any that don't Δ b much
(2) Sequential: drop one C_j then get new b for which we would evaluate next C_j ... (2')

$$F' = a' + b'E + \sum_{j=1}^{r-1} c_j' C_j$$ (if $b \cong b'$, remove C_j from model)

D. then must convert F into individual Score (S) on which we will stratify — done by setting E to null value ($\bar{E} = 0$) —

$$S = F - E = a + \sum c_j C_j$$ (same coef. — reduced F)

reduced (denote w/ (same) symbol F?

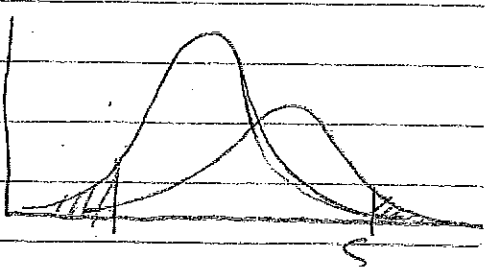
ie, each subj. is given score summary of C_j's (reduced F)

V. Stratify subjects into ≈5 categories along S — trickiest part of this whole procedure method

A. Prelim. step: eliminate subjects from which we get little or no info. — ie, delete fringe areas of S where either D or $\bar{D}$ predominate

this is equivalent to subj restriction in selection

so bounded range of S



V. B. make initial stab at stratif. by simply dividing all D (or $\bar{D}$ if they are fewer) into quintiles & along S & let the $\bar{D}$ fall as they may

C. we would like to create reasonably homogeneous strata w/in range of S just defined; because we would like ~~intra-stratum~~ conf. to occur between strata, not w/in them — 2 ways of maximizing the conf. between strata & thereby min. variability of S w/in strata

(to pool) 1. uniformity of $e_{\bar{D}}$ over strata, or

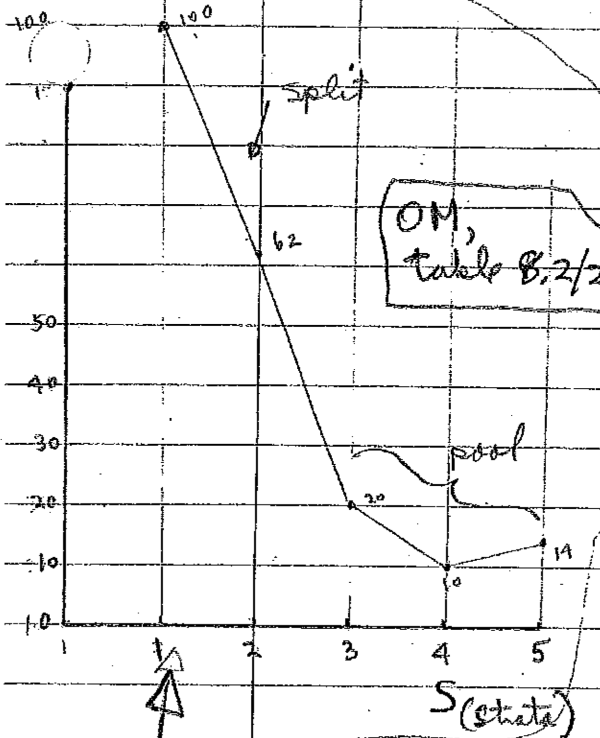
(2.) uniformity of $\bar{D}/D$ ratio over strata, conditional on $\bar{E}$ (or E) — NB. must be cond. on E, so we don't confuse conf. w/ E

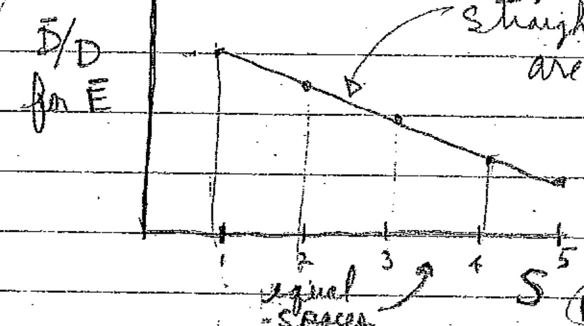
= my procedure: ratio ($\bar{D}/D$) | $\bar{E}$

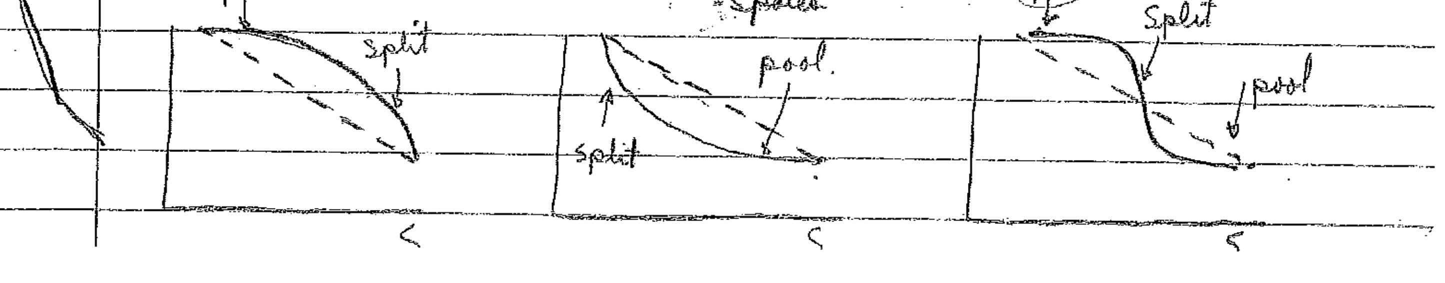
(eg) ($\bar{E}$ only)

S	D	$\bar{D}$	$\bar{D}/D$	~~cum $\bar{D}/D$~~
1	10	30	3.00	~~3.00~~
2	10	25	2.50	~~2.75~~
3	10	20	2.00	~~2.50~~
4	10	15	1.50	~~2.25~~
5	10	10	1.00	~~2.00~~

assume evenly spread out







V. D. but ~~this~~ while this process max. conf. between strata, it does not insure that there is only a minimal amount of conf. w/in strata (absolutely) because F itself may not have been adequate
- So we ~~determine~~ check homogen. w/in strata by comparing of $\hat{D}^+$ and $\hat{\bar{D}}$ as to their distrib. by each of the indiv. C_j, ~~separ~~ for only $\bar{E}$ subjects (& so as not to confuse EM w/ conf.) — we do this by comparing the %'s of $\hat{D}$ and $\hat{\bar{D}}$ that are ~~dichot~~ at "high risk" of each indiv C_j, within each Stratum, in $\bar{E}$ subjects

(% of subj. in hi risk cat. / stratum)

C_j	S=1		2		3		4		5	
	$\hat{D}$	$\hat{\bar{D}}$	$\hat{D}$	$\hat{\bar{D}}$	$\hat{D}$	$\hat{\bar{D}}$	$\hat{D}$	$\hat{\bar{D}}$	$\hat{D}$	$\hat{\bar{D}}$
C_1 = eg, age (hi-lo)	20	20	30	30	40	40	50	50	60	60
C_2 = sex = M	10	10	15	30						
⋮										

% of exposed cases in S_1 who are M

conf

see: table 8.2/1 (OM) for all Subj.

Dich at: MI — of age — not, but may be EM

E. Note that getting strata involves a lot of judgment + ∴ takes practice

VI. Analysis

A. Proceed as in Gen'l. strat. — as if w/ one C = S
1. H_0 testing — χ_0
2. Est. of Effect
   a. pt. est. — eg, $\hat{RR}_0$ or $\hat{SMR}$, $\hat{SRR}$, $\hat{EF}$
   b. CI

Confusion / mistake / M! parody → F_r → S → Stratify (judgmental) → gen'l strat. analysis

begin:

VI. B. F (or S) does not necessarily represent a risk function – ie, if a C_j (eg, age) is not included in the reduced F, it doesn't mean that age is not a risk factor of D – S is only a classification score used to rank subjects in the study

1. CH study: 2 ways to see S ≠ risk
   a. we might select subj. or restrict their selection in such a way as to reduce the variability of a C_j (eg, age) in the study pop, so that age cannot predict the D in the particular data.
   b. also, age might be important predictor of D, but not be assoc. w/ E in the study ∴ age would not be conf. – ie it would be dropped from F (⇒ Δ b)
2. C-R study:
   a. never really considering risk ($Pr[\text{getting } D | \bar{C}_j]$ in most C-R studies, but $Pr[\text{being } D | \bar{C}_j]$
   b. we do not even get this Pr if the ref. series is unrep. of the source pop. Of the cases – might occur in 3 ways
      i. Matching – ∴ (eg) if age-M, age won't appear in F because it can't discrim. D from $\bar{D}$
      ii (even w/o M.), selection might operate so that source of ref. subj. is unrep. of the source of cases – eg, dilute role of age ~~in~~ in F by selecting D + $\bar{D}$ of similar ages distributions – e.g. from same hospital (Sort of like partial Matching or freq. M.)

VI. C. Suppose: E has >2 categories ∴ requiring 2 or + dummy variables, eg, race

E_1 = Blackness — 1 (if B); 0 (if not B)

E_2 = Redness — 1 (if R); 0 (if not R)

W = ref.

$F = a + b_1 E_1 + b_2 E_2 + \sum c_j C_j$ (unreduced)

1. one Way: 2 separate analyses: one comparing B w/ W and R w/ W

F_1' — depends on Δb_1 (treating E_2 as conf.)

F_2' — " " Δb_2 (" " E_1 " ")

∴ $S_1 = F_1' - b_1 E_1$ $\quad S_2 = F_2' - b_2 E_2$

problem: if interested in gradient of effect (over E_i), Standardiz. requires comparable set of weight for each comparison - i.e., same standard → SRR — but multiple analyses changes the strata + thus the weights, so...

2. F' — remains unreduced + live w/ some inefficiency — or perhaps use conf. of previous analysis (assuming E is dich.)

$S = F' - b_1 E_1 - b_2 E_2$ (one analysis)

then, standardize externally ("mutually stand.") to get comparable estimates of effect over categories of E_i — O.M. used in S_m → M.H.

3. can also assume E is ordinal and assign integers (eg, 0, 1, 2...) but this adds extra assumption — not recommended unless E is nearly interval

(recall that main problem w/ partial correl, stepwise regression, or any model fitting is inability of looking for effect mod. of var. that are assoc. w/ conf.)

VI. D. Looking for E.M. (in broad sense) - uniformity of eff

1. S as EM — not to draw specific inferences about associations but to determine whether strata of S may be pooled to create $\hat{RR}_o$ (than MH or weighting by inverses of Variances of $\hat{RR}_j$)

just calculate $\hat{RR}_j$ — test for uniformity

a. $\chi^2_u = \sum \chi^2_j - \chi^2_o$ or,

b. a more specific model — eg, using weighted regression

(?) 2. a Single component factor (of S) as the effect modifier(M) — more concerned w/ specific assoc's — 2 approaches — which is best?

~~a. Separate analyses for each category of M — w/ diff. E, S, Strata, etc.~~

~~Eg, Sm → CHD do analysis for:~~

~~i. all subjects in Study~~

~~ii. hypertensive~~

~~iii borderline hypert.~~

~~iv. Normotensive~~

~~i.e., we control for confounding separately within each category of M, calculate SMR's and compare~~

recommended

b. single analysis (one S w/all subj) — calculate SMR's w/in each category of M — may have been what OM used in Sm → MI —

(me) it seems to me that this approach may only be justified on the basis of convenience

?

VI. E. How to consider interaction of ~~Variables~~ factors

1. Interaction of C_j's – in order to account for more conf. in the data – Since the importance of any single conf. is reduced when there are several, interaction of C_j may not be particularly useful (unlike a prediction model), but under certain conditions, it might be desirable – eg, w/ drugs which are known to interact pharmacologically – eg, $\underline{C_1 C_1 C_2}$ → treat as another conf. or $\underline{\text{age} \times \text{X}}$

2. Interaction of E and C_j – not interested in studying interaction per se, but improving model (F) – ie, getting more variance in D

(eg) E = Smoking
C_1 = age

$$F = a + bE + b_1 E \cdot C_1 + c_1 C_1 + \sum_{j=2} c_j C_j$$

$$S = F - bE - b_1 E \cdot C_1 \quad (\text{when } E=0)$$

NB: can't reduce F in normal way because both E + $C_1 \cdot E$ become 0 when $E \neq 0$, so would have to observe Δ in 2 b's –
so, might get C_j first w/o interactive term then do ranking w/ it – tho might have to sacrifice some effic.

3. Interaction of 2 E's – ie, when we are interested in studying interaction per se – eg, E_1, E_2

$E_1 = 1 (Sm); 0 (\overline{Sm})$
$E_2 = 1 (dr); 0 (\overline{dr})$

$$F = a + b_1 E_1 + b_2 E_2 + b_3 E_1 E_2 + \sum c_j C_j$$

$$S = F - b_1 E_1 - b_2 E_2 - b_3 E_1 E_2$$

VI. E. 3. (cont.) – again, F cannot be reduced in normal way, so might have to sacrifice some E_j – then analyze as if S were 1 var., which we would like to control in looking for interaction

SERR = SRR-1 (sept. stand.)

	S_m	$\bar{S}_m$	
D_i			
$\bar{D}_i$		0 ← ref.	need large n

ie, control for conf. w/in each comparison

VII. "Conf.-Summarization" vs. alternative methods:

A. Comparison points: some say

1. it is new method
2. it is same thing that is done by researchers who stratify by a MLF (eg. Evans Cty.)
3. (me) actually, the 2 procedures are related, but do not have same utility.

B. differences between conf.-sum. and MLF

1. major diff.: MLF is created as a predictor of D (ie risk function) – while S is a classification score for ranking subj. in order to control for conf.
2. often B's of MLF are applied to another study pop. – not done w/ S; requires different ranking for each pop. – different selection factors in each study
3. MLF is chosen as the model because it creates a function that remains between 0 & 1 (like a Pr) – but not necessary w/ S; the might want to take log transformation – eg. ln Age
4. MLF only reflects assoc of C_j w/ D, not C_j w/ E, which is also a prereq. for conf. ∴ MLF as a stratification device is liable to be very inefficient

VII. B, 5. an attempt is made to reduce or limit # terms (C_j) in MLF by statistical procedure, eg, step-wise procedure, not done in S.C.F.

*6. MLF is not generally converted into a score conditional on E — so strong predictive MLF may not be very Conf. — ie, θ_{CD} is not criterion for conf.

So, using MLF (as is) to control for conf. may be
1. very inefficient (cost)
2. possibly invalid, in the sense that $\hat{V}^{max}_{conf.}$ is not being controlled, between

(VIII.) Selective Matching — eg. M. on age, race, sex
Sm → CHD
also would like to control for obesity, BP, chol.

So: use M. only as restrictive device
(ignore matched 'sets' in the analysis but
stratify on 'all' 6 var's — using
Composite Stratification

3/17 Outline

I. Intro'd.
A. O.M on 3/24 – will meet w/ him
B. today:
1. finish composite stratification
2. Class exercise (Syme et al., 1964)

II. Criticisms of Syme et al. (1964)
A. Study design and parameters
1. type of design – had all the ingredients of an ambidir. study but treated as C-R
2. could have computed Id, Id_E, ID_E, IdR, IdD, EF, Ic; – $\hat{Id} = \frac{328}{20,000+1}$ + CI
but did not actually compute any of these

note!
① dynamic pop.
② Should be age-spec. Id

B. Matching
1. analyzed data without stratifying for age (the M. var.) – in C-R analysis this could cause both invalidity & loss of sig-effic.
2. what should have been done (as suggested by authors)

D \ $\bar{E}$	2E	E/E	2E
E	U	V	W
E	X	Y	Z

ML of IdR requires iterative soln. because r=2 – the

$\hat{M\text{-}H}$ (not ML) $= \frac{2W+V}{2X+Y}$ shortcut

203 (sets)

O/E ratio ⇒ $a = U+V+W$

$E[a] = \sum \frac{a_j + b_j}{r+1} = U\left(\frac{2}{3}\right) + (X+V)\left(\frac{2}{3}\right) + (Y+W)\left(\frac{1}{3}\right)$

$= \frac{3U+2X+2V+Y+W}{3}$

$a/E[a] = \frac{3U+3V+3W}{3U+2X+2V+Y+W}$

II. B. 3. What authors did.

	D	$\bar{D}$
E	$u+V+W$	$2u+2x+V+Y$
$\bar{E}$	$x+Y+Z$	$V+Y+2W+2Z$

203 406 609 (subj.)

O/E ratio → $a = u+V+W$ (same)

$$E[a] = \left(\frac{2u+2x+V+Y}{406}\right)203 =$$

$$= \frac{u+x+\frac{V+Y}{2}}{1} = \frac{2u+2x+V+Y}{2}$$

$$a/E[a] = \frac{2u+2V+2W}{2u+2x+V+Y} \quad (\neq \text{ as above})$$

4. Clearly the authors got a different R than if they had stratified in the analysis — But suppose they had simply computed the OR — instead of these R's (ignoring age ...)

	D	$\bar{D}$
E	a	b
$\bar{E}$	c	d

$OR = \frac{ad}{bc}$ but authors:

$$R' = \frac{R_E}{R_{\bar{E}}} = \frac{a/E[a]}{c/E[c]} =$$

authors never do this formally & give it a name

$$= \frac{\frac{a}{\frac{b}{(b+d)}(a+c)}}{\frac{c}{\frac{d}{(b+d)}(a+c)}} = \frac{ad}{bc} = OR$$

So, $R' = R_E/R_{\bar{E}} = OR$

II, B, 5. So forgetting the basic mistake of not stratifying by age in analysis, the authors actually got the IdR, which they "correctly" interpreted – see: p. 3 –
but this estimate actually involved 3 mistakes w/re: to what they said they were doing.

a. They included the missing data in the totals (for $D + \bar{D}$) – thus, they maintained a constant ratio of cases to controls

b. the expected values were computed incorrectly – for a C-R study

$$E[a] = \frac{bc}{d} \neq \frac{b}{(b+d)}(a+c)$$

NB: $SMR = \frac{a}{\sum \frac{b_j c_j}{d_j}} = \frac{ad}{bc} = OR$ (for only 1 stratum)

if IR had been computed correctly it would have been an appropriate meas. of assoc. itself + might have even been used to standardize for pot confounders – but were forced to compare

c. 2 R_s'

e. each R_i treats the other E category as the referent – so,

$R_1 \rightarrow R_2$ is referent

$R_2 \rightarrow R_1$ " "

it wasn't until authors computed R_s' (R') that they arrived at a meas. of assoc.

II, B, 5, (c) cont.

this might result from a fundamental misunderstanding that analytical epid. requires a comparison of groups – the authors treat this idea as secondary & seem to believe that a single category (of a var.) may be observed to be assoc. with the disease

c. Confounding –

1. never actually determined conf., just whether there was still an assoc. w/in categories of pot. conf. (c) – i.e., as if they were studying Eff. Modif. and searching for "secondary assoc." in fact what they did resembled control table analysis but instead of using Ia rates they used "R."
2. more important – did not consider all pot. conf. simultaneously – eg, table 11, p.5 – conf. effect of BP should be seen conditionally on all other conf.
3. also table 11, $\chi^2_{(5)}$ is meaningless – both SC mob. + Hypert. are treated as one factor in getting a p-value → i.e. they looked for the combined effect of both variables, not for the SC effect, controlling for BP – they could have used χ^2_4
4. It would be important to see whether certain confounders might represent intervening mechanisms – eg., SC → diet → CHD
this might show diet as meth. conf., but does not detract from obvious importance of SC

II. C. 4. (cont.)
might have an indication of this intervening mech
$RR_{(\text{of hi SC mob.} \mid \text{low fat diet})} \cong 1$

but $RR_{(\text{of hi SC mob})} > 1$

5. When comparing > 2 categories of E (eg, 3 classes of SC mob.) would like to identify "dose-response" gradient – table 11, p. 5
requires SRR, not SMR to have comparable standards for each comparison
SC$_1$ = ref.
$CRR_3 = 2.28$
$SRR_3 = 2.64$ ∴ neg. conf.
$CRR_2 = 1.34$
$SRR_2 = 1.33$ ∴ no conf.

	SC stable	SC mod mobile	SC highly mob.	
SRR	1.00	1.33	2.64	⇒ dose-response

D. Effect Modif. & Interaction
could have used SRR
1. authors seem to confuse conf. + EM.
2. looked explicitly for SC interaction (table 15, p. 6) but failed to actually determine whether effects were nearly additive and failed to control for other potential conf. –
could have used SRR to analyze the interact between any 2 indep var. (eg, SC + BP)

SC	Norm	Hyp
(1)		
(2)		
(3)		

SERR = SRR – 1