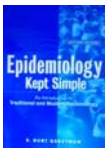
Epidemiology Kept Simple



Chapter 12

Error in Epidemiologic Research

11/4/2009 1

§12.1 Introduction

- Analytic studies are exercises in **measurement**
- All measurements are **prone to error**
- Understanding **measurement error** begins by distinguishing between parameters and estimates

Train wreck

11/4/2009 2

Parameters and Estimates

Parameters	**Statistical estimates**
– Error-free	– Error-prone
– Hypothetical	– Calculated
– Causal effect	– Association

Epidemiologic statistics are merely approximations of the parameters they wish to estimate.

11/4/2009 3

Target Analogy

Let ϕ represent the *RR* parameter.

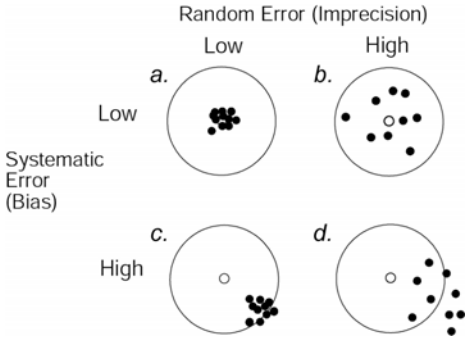
Replicate the study many times

Let $\hat{\phi}_i$ represent the RR estimate from study *i*

How closely do the estimates reflect the parameter? How accurate is a given estimate?

4

Target Analogy



5

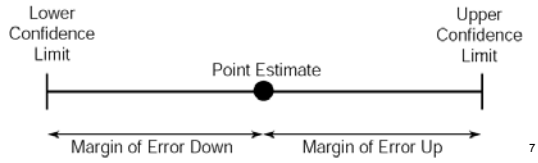
Random Error (Imprecision)

- Balanced "scatter"
- Use statistical methods such as confidence intervals and hypothesis tests of significance to address random error

11/4/2009 6

Confidence Intervals (CI)

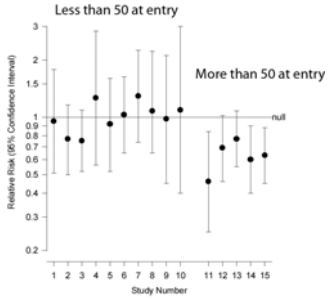
- Surround point estimate with **margin of error**
- **Locates** parameter with given level of confidence (e.g., 95% confidence)
- **CI width** quantifies precision of the estimate (narrow CI → precise estimate)



7

Mammagraphic Screening and Breast Cancer Mortality

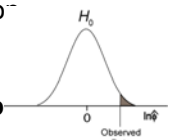
- E = mammographic screening
- D = Breast CA death
- First series: women in their 40s
- Second series: women in their 50s



11/4/2009

P-values In a Nutshell

- Start with this null assumptio
  H_0: RR = 1 (no association)
- Small *P*-value provide **evidence *against* H_0** (do **no** use P-value cutoffs -- surely, G-d loves *P* = .06")
- By **convention**
  - *P* ≤ .10 is considered marginally significant evidence against H_0
  - *P* ≤ .01 is considered very significant evidence against H_0



R.A. Fisher

11/4/2009 9

Childhood SES and Stroke

Factor	*RR*	*P*-value
Crowding (persons/room)	< 1.5 = 0.4 1.5 – 2.49 = 1.0 (ref.) 2.5 – 3.49 = 0.6 ≥ 3.5 = 1.0	trend *P* = 0.53
Tap water	0.73	*P* = 0.53
Toilet type	flush/not shared = 1.3 flush/shared = 1.0 (referent) no flush = 1.0	trend *P* = 0.67
Ventilation	good = 1.0 (ref.) fair = 1.7 poor = 1.7	trend *P* = 0.08
Cleanliness	good= 1.1 fair = 1.0 (ref.) poor = 0.5	trend *P* = 0.07

Nonsignif. evidence against H_0

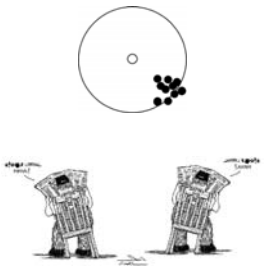
Marginally significant evidence against H_0

Source: Galobardes et al., *Epidemiologic Reviews*, 2004, p. 14

10

Systematic Error (Bias)

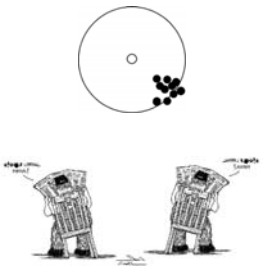
- ***Bias*** = systematic error in inference (*not* an imputation of prejudice)
- **Direction** of bias
- **Toward the null** (underestimates risks or benefits)
- **Away from null** (overestimates risks or benefits)



11/4/2009 11

Categories of Bias

- **Selection bias:** participants selected in a such a way as to favor a certain outcome
- **Information bias:** misinformation favoring a particular outcome
- **Confounding:** extraneous factors unbalanced in groups, favoring a certain outcome



11/4/2009 12

Examples of Selection Bias

- Berkson's bias
- Prevalence-incidence bias
- Publicity bias
- Healthy worker effect
- Convenience sample bias
- Read about it: pp. 229 – 231

11/4/2009 13

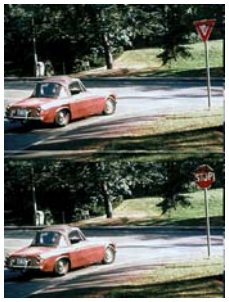
Examples of Information Bias

- Recall bias
- Diagnostic suspicion bias
- Obsequiousness bias
- Clever Hans effect
- Read about it: p. 231

11/4/2009 15

The Misinformation Effect

- Memory is constructed rather than played back
- Social contagion of memory
- The misinformation effect: Elizabeth Loftus (http://en.wikipedia.org/wiki/Misinformation_effect)



11/4/2009 16

Recall Bias in Case-Control Studies

	Cases	Controls
Exposure reported	a ↑	b
Exposure not reported	c	d ↓

Cases recall past exposure better than controls & controls tend to underreport their past exposures: ↑ a or ↓ d in case-control studies → ↑ OR

11/4/2009 17

Differential & Nondifferential Misclassification

- **Non-differential misclassification:** groups equally misclassified ⇒ biases results toward null
- **Differential misclassification:** groups unequally classified ⇒ biases in either direction

11/4/2009 18

Table 17.6 Examples of nondifferential and differential misclassification, and their effect on relative risk estimates.

Scenario A: Nondifferential misclassification: 10% of the cases in the exposed and nonexposed groups are misclassified as noncases.

	Accurate data +	Accurate data −	Total		Misclassified data +	Misclassified data −	Total
Exposed	100	900	1000	→	90	910	1000
Nonexposed	50	450	500		45	455	500

$\widehat{RR} = \frac{100/1000}{50/500} = 1.00$ $\widehat{RR} = \frac{90/1000}{45/500} = 1.00$

Scenario B: Differential misclassification: none (0%) of the cases are misclassified in the exposed group, while 10% are misclassified in the nonexposed group.

	Accurate data +	Accurate data −	Total		Misclassified data +	Misclassified data −	Total
Exposed	100	900	1000	→	100	900	1000
Nonexposed	50	450	500		45	455	500

$\widehat{RR} = \frac{100/1000}{50/500} = 1.00$ $\widehat{RR} = \frac{100/1000}{45/500} = 1.11$

11/4/2009 19

Confounding

- **Confounding** ("to mix together") ≡ a distortion brought about by extraneous variables
- Properties of confounder:
  - Confounder is associated with the exposure
  - Confounder is an independent risk factor
  - Confounder is *not* in causal pathway

Accident Evacuation
Crude Analysis

	Died	Survived	Total
Helicopter	64	136	200
Road	260	840	1100

- R_1 = 64 / 200 = .3200
- R_0 = 260 / 1100 = .2364
- *RR* = .3200 / .2364 = 1.35
- 35% increase in risk of death? This makes no sense!

Accident Evacuation
Serious Accidents (Stratum 1)

	Died	Survived	Total
Helicopter	48	52	100
Road	60	40	100

- R_1 = 48 / 100 = .4800
- R_0 = 60 / 100 = .6000
- RR_1 = .48 / .60 = 0.80
- Helicopter evacuation cut down on risk of death by 20% among serious accidents

Accident Evacuation
Minor Accidents (Stratum 2)

	Died	Survived	Total
Helicopter	16	84	100
Road	200	800	1000

- R_1 = 16 / 100 = .16
- R_0 = 200 / 1000 = .20
- RR_2 =.16 / .20 = 0.80
- Helicopter evacuation cut down on risk of death by 20% among minor accidents

Accident Evacuation
Properties of Confounding

- Seriousness of accident (C) associated with method of evacuation (E)
- Seriousness of accident (C) is an independent risk factor for the outcome (D)
- Seriousness of accident (C) *not* in the causal pathway (the helicopter evaluation does not cause the accident to become more serious)