UCSF EPI TOOLS WORKSHOP QUANTITATIVE BIAS ANALYSIS

Elizabeth Rose Mayeda, PhD, MPH

UCSF Epi Tools Workshop

Quantitative Bias Analysis

March 4, 2022

ermayeda@ph.ucla.edu

@er_mayeda

Outline

- Part 1: Introduction to quantitative bias analysis (QBA) (1:10-3:00pm)
- Motivation and key concepts
- What is quantitative bias analysis?
- Deterministic bias analysis:
  - Selection bias
  - Unmeasured confounders
  - Misclassification
- Multidimensional bias analysis
- Probabilistic bias analysis

Outline

- Part 2: Monte Carlo simulations for quantifying confounding bias and collider-stratification bias (3:00-4:00pm)
  - Why simulate?
  - Simulation steps
  - Simulation exercise: confounding bias
  - Simulation exercise: collider-stratification bias
  - Briefly: extension to simulating information bias

PART 1: INTRODUCTION TO QUANTITATIVE BIAS ANALYSIS

Elizabeth Rose Mayeda, PhD, MPH

UCSF Epi Tools Workshop

Quantitative Bias Analysis

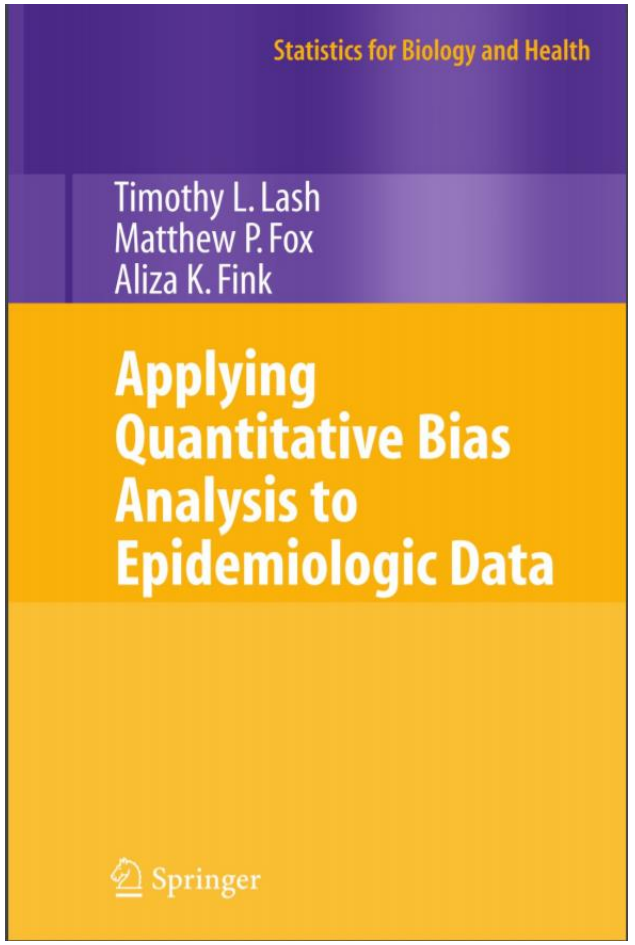
March 4, 2022

ermayeda@ph.ucla.edu

@er_mayeda

Acknowledgements

- Hailey Banack, University of Buffalo
- Maria Glymour, UCSF
- June Chan, UCSF
- Rebecca Graff, UCSF
- Francois Rerolle, UCSF
- Shelley Devost, UCSF
- Alice Guan, UCSF
- Eduardo J. Santiago-Rodríguez, UCSF

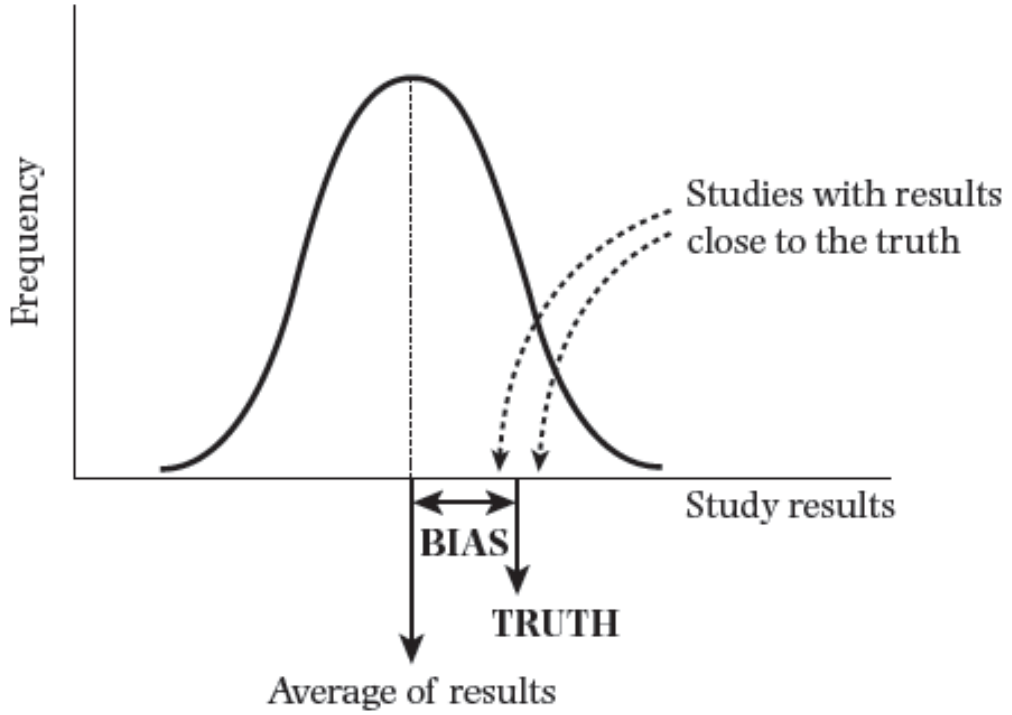


Lash et al., *Applying Quantitative Bias analysis to Epidemiologic Data*, 2009

Key Terms: what is bias?

"Systematic deviation of results or inferences from the truth"

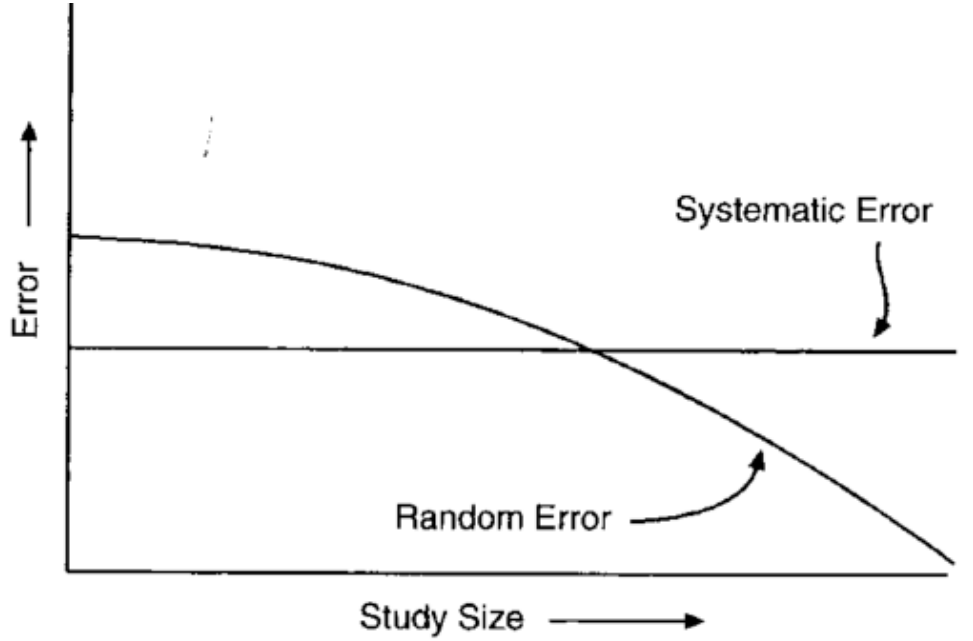
FIGURE 4-1 Hypothetical distribution of results from a biased study design.

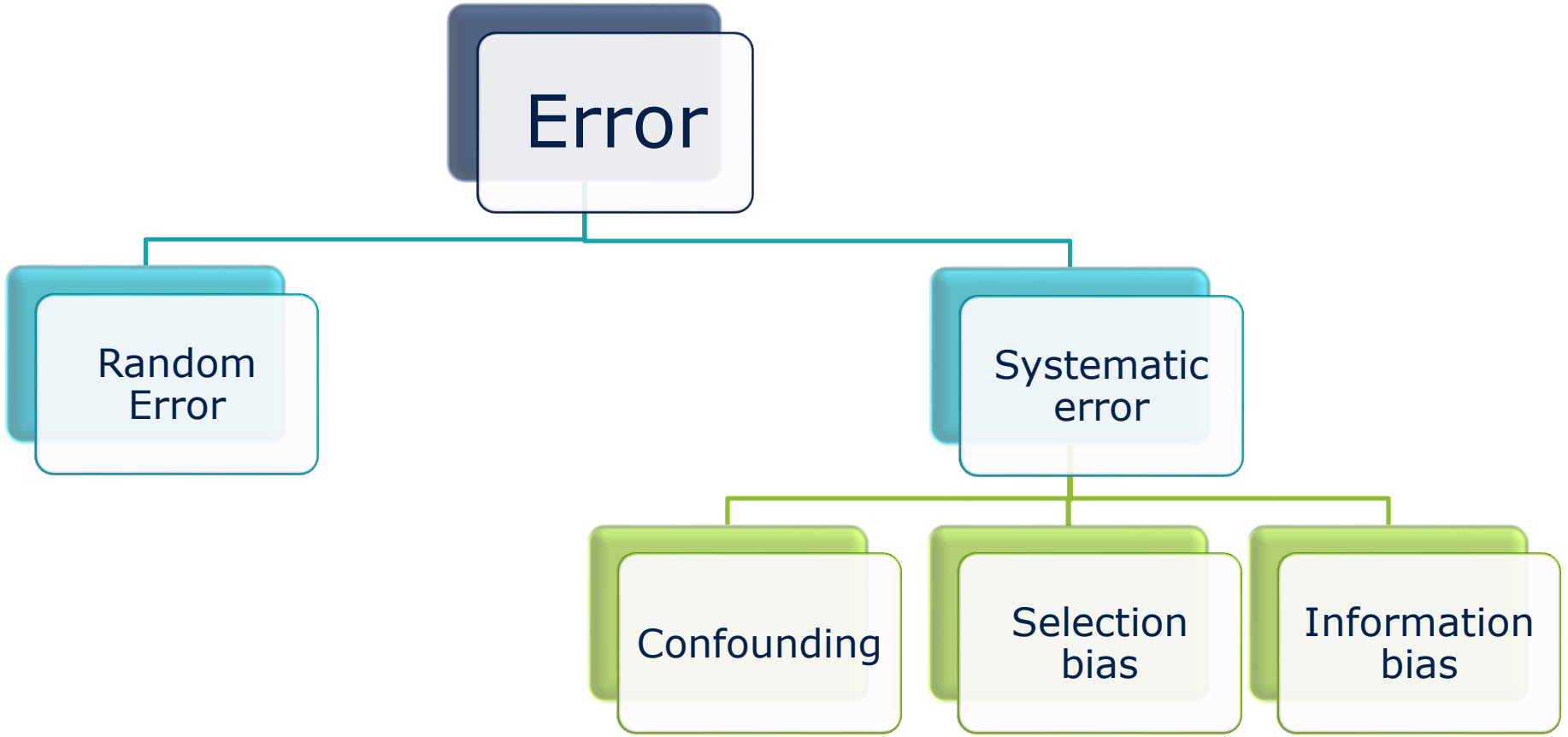


Porta, 2008; Szklo & Nieto,2012

Key Terms: random error vs. systematic error

- Random error = lack of precision
  - Decreases as sample size increases
- Systematic error = bias/lack of validity
  - Present regardless of sample size, even in an infinitely large study
- Validity and precision are both components of accuracy





Precision
Relative lack of random error

Internal Validity
Relative absence of systematic error

American Journal of Epidemiology

DOI: 10.1093/aje/kwy228

Practice of Epidemiology

Target Validity and the Hierarchy of Study Designs

Daniel Westreich*, Jessie K. Edwards, Catherine R. Lesko, Stephen R. Cole, and Elizabeth A. Stuart

* Correspondence to Dr. Daniel Westreich, CB 7435, McGavran-Greenberg Hall, Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599 (e-mail: djw@unc.edu).

Initially submitted November 16, 2017; accepted for publication September 27, 2018.

In recent years, increasing attention has been paid to problems of external validity, specifically to methodological approaches for both quantitative generalizability and transportability of study results. However, most approaches to these issues have considered external validity separately from internal validity. Here we argue that considering either internal or external validity in isolation may be problematic. Further, we argue that a joint measure of the validity of an effect estimate with respect to a specific population of interest may be more useful: We call this proposed measure *target validity*. In this work, we introduce and formally define target bias as the total difference between the true causal effect in the target population and the estimated causal effect in the study sample, and target validity as target bias = 0. We illustrate this measure with a series of examples and show how this measure may help us to think more clearly about comparisons between experimental and nonexperimental research results. Specifically, we show that even perfect internal validity does not ensure that a causal effect will be unbiased in a specific target population.

causal inference; external validity; generalizability; internal validity; study design; target population; target validity; transportability

https://www.ncbi.nlm.nih.gov/pubmed/30299451

Internal validity, external validity, and target validity

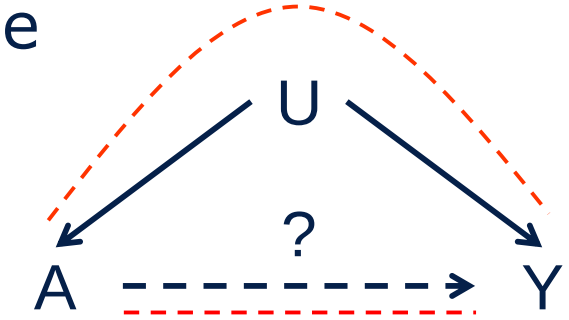
- Internal validity: a result is internally valid when the effect estimated in the study sample is unbiased for the true effect in that sample
- External validity: a result is externally valid if the true effect in the study sample is unbiased for the true effect in the target population
- Target bias = [true causal effect in the target population] – [estimated causal effect in the study sample]
  - Target validity is defined as target bias = 0
- The focus of today's lecture is on internal validity (and QBA for internal validity)

Threats to internal validity

- **Confounding:**
  - Unmeasured confounder (or mismeasured confounder)
  - Unknown confounder
- **Selection Bias (collider-stratification bias):**
  - Baseline
  - Losses to follow-up
- **Information Bias:**
  - Misclassification of categorical variables, measurement error of continuous variables
  - Differential vs. non-differential
  - Dependent vs. independent errors

Confounding (see DAG lecture for more details)

- Confounding: Estimate of effect of exposure on outcome is 'mixed with' the effect of uncontrolled/unmeasured covariate on outcome
  - Unmeasured confounder (or mismeasured confounder)
  - Unknown confounder
- Confounding bias: Bias arising from the existence of a common cause of exposure and outcome



⇨Lack of exchangeability of treated and untreated

Selection bias

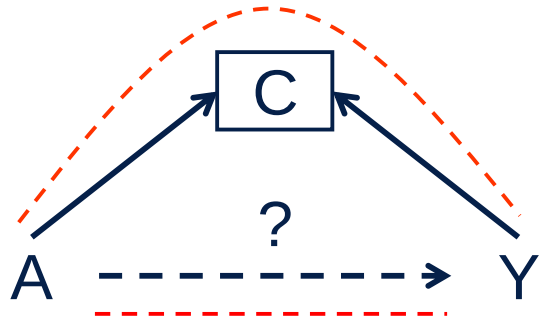
- Occurs as a result of the process by which people are selected into the analysis (e.g. participating in the study, having complete data)
- As a result of the sample selection process, exposure-outcome association in the sample differs from the true causal effect of exposure on outcome in the sample (internal bias) or in the target population (external bias)
- Can refer to threat to internal validity or external validity

Collider-stratification bias (see DAG lecture for more details)

- Bias arising from conditioning on a common effect (collider) of exposure *(or cause of exposure)* and outcome *(or cause of outcome)*
- Conditioning on a common effect of exposure and outcome results in two sources of association between exposure and outcome:

1. Causal effect of A on Y: A→Y
2. Spurious association between A and Y

   induced by conditioning on collider C: A→C ← Y

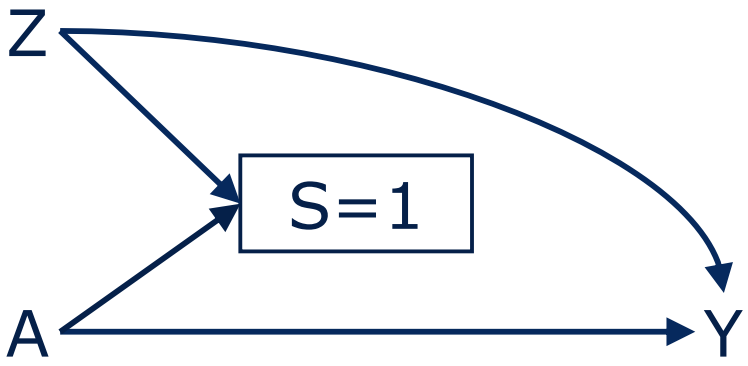


⇨Lack of exchangeability of treated and untreated

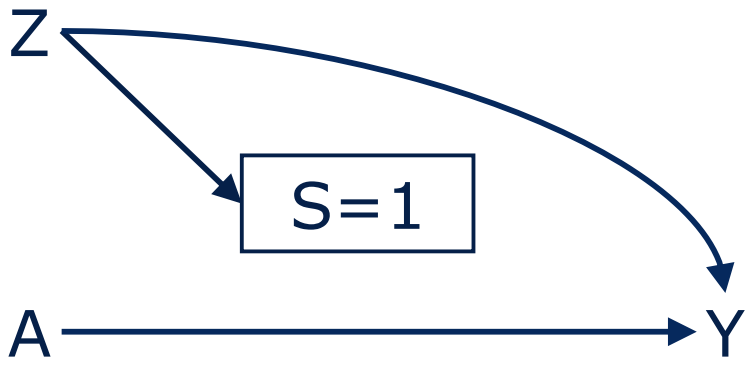
DAGs for selection bias

Example: A = treatment, Y = health outcome, S=1 = selection into the study sample, and Z = sociodemographic and health factors

Exposure and determinants of outcome influence selection (collider-stratification bias): threat to internal validity

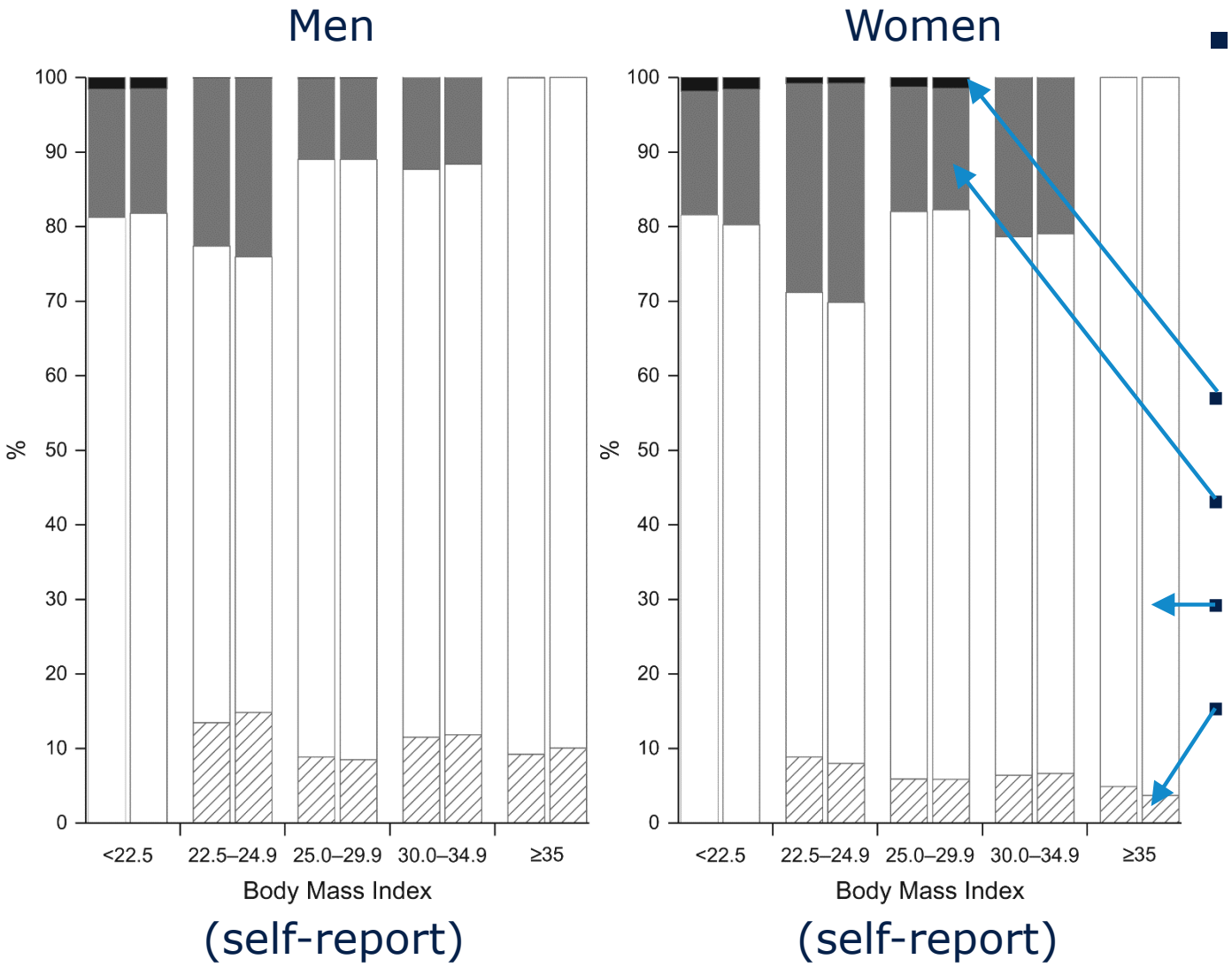


Determinants of the health outcome, but not the exposure, influence selection: threat to external validity



Information Bias (see DAG lecture for more details)

- The effect estimate is distorted by inaccurate measurement or classification of the exposure, outcome, or a covariate
- Misclassification of categorical variables (exposure, outcome, covariate)
- Measurement error (continuous variables)



- Percentage of sample within categories of <u>self-reported</u> BMI that were classified into the same, higher, or lower categories of <u>measured BMI</u> (NHANES)
  - Black: BMI $cat_{measured}$ 2 categories > BMI $cat_{self\text{-}reported}$
  - Gray: BMI $cat_{measured}$ 1 category > BMI $cat_{self\text{-}reported}$
  - White: BMI $cat_{measured}$ = BMI $cat_{self\text{-}reported}$
  - Striped: BMI $cat_{measured}$ < BMI $cat_{self\text{-}reported}$

Limitations of meta-analysis

- If there is a systematic bias affecting study results, a meta analysis will not solve this problem
- More precise, but still biased effect estimate

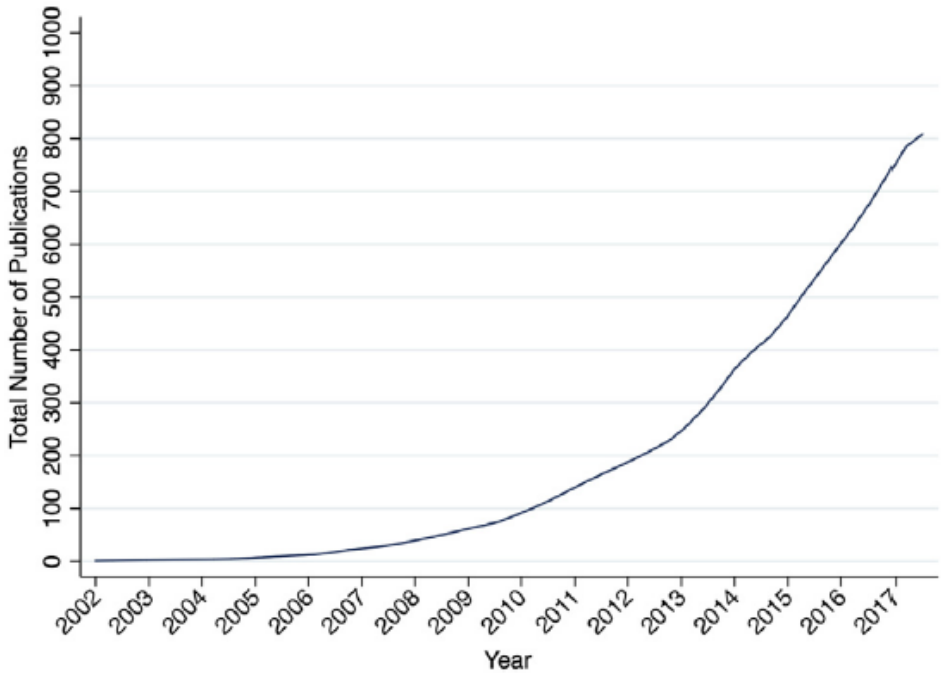


Figure 1. Peer-reviewed publications emphasizing the "obesity paradox." Methods included a search of EMBASE (1947-2017, week 29) for full articles that included the phrase "obesity paradox" in the title, abstract, or keywords. Our search strategy identified 812 articles, with the initial articles appearing in the year 2002.

Stovitz, Banack & Kaufman, 2018

Why QBA?

- As epidemiologists, we all try to design quality studies and reduce the amount of systematic error affecting an effect estimate

- Problem: It's not always possible!

- Incumbent upon investigators to quantify how far an estimate is from the "truth" (an estimate that is free of bias)

When QBA?

1. When the effect estimate is reasonably precise
   - Narrow 95% CI reflects a small amount of random error
   - With a wider CI, it is already difficult to make inferences because of the amount of random error, regardless of whether you try to account for systematic error
2. When you believe the effect estimate is affected by a limited number of systematic errors
   - With multiple systematic errors, total error may be too large to reliably quantify
   - QBA may not be productive; perhaps data should be used for generating ideas for future studies
3. When you have a way of realistically estimating bias parameters
   - Required to complete QBA
   - Could come from internal or external validation data or literature review

Bias parameters

- In QBA, equations are used to adjust an effect estimate to see the impact of potential bias
- The parameters in these equations are called “bias parameters”
- Used to determine the direction and magnitude of bias adjustment
- Confounding bias
  - Hypothesized strength of relationship between unmeasured confounder and exposure and unmeasured confounder and outcome
- Selection bias (collider-stratification bias)
  - Sampling fractions: proportion of all eligible subjects who participate in your study
- Misclassification
  - Sensitivity and specificity of exposure/outcome/confounder classification

CONFOUNDING

Identifying confounding

- Confounding refers to a theoretical concept
  - Risk of disease in the reference group (unexposed) equals the risk the index group (exposed) would have had, if they had not been exposed
  - Reference group intended to be a substitute for index group in every way ("exchangeable" groups)
  - When they are not exchangeable, confounding is present

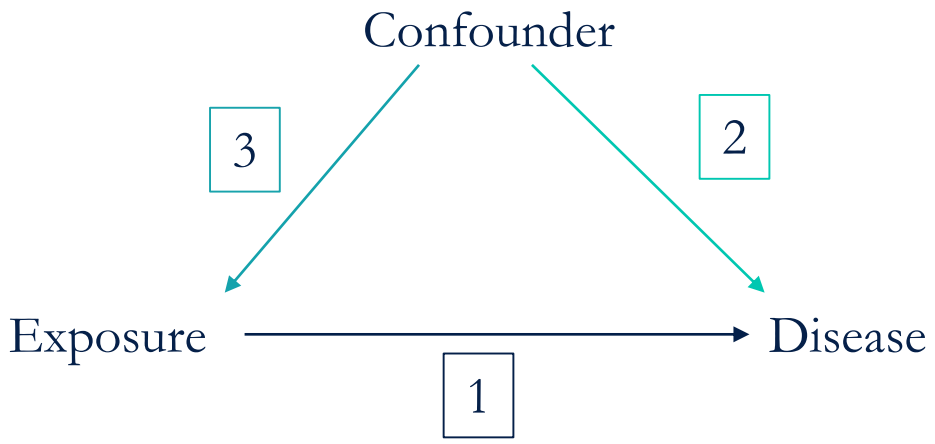
- Methods for accounting for known and measured confounding
  - Stratification, standardization, Mantel-Haenszel (MH) adjustments, regression modeling

Crude vs. adjusted estimates

- In practice, confounding typically identified when crude estimate of association does not equal the adjusted estimate ("change in estimate")
- Some problems:
- Only can adjust for variables measured in your study
  - Can not adjust for:
    - Unknown confounding (variables you can't identify, theoretical concepts)
    - Unmeasured confounding (variables you did not measure)
    - Residual confounding (variables measured poorly/with a lot of error)

Bias analysis for unmeasured/unknown confounding

- Requires:
  1. Crude estimate of the exposure-outcome association
  2. Association between confounder and disease
  3. Association between confounder and exposure (or prevalence of confounder in each exposure group)



Assumption: Unmeasured confounder is independent of measured confounders

Result: Overestimate magnitude of bias

Question of interest

- What estimated effect of EXPOSURE on DISEASE would have been observed in the study if we had collected data on CONFOUNDER and adjusted for it in analyses?
- Formula to adjust for a binary unmeasured confounder, assuming homogeneity of effect of confounder across levels of exposure

$$RR_{adj} = RR_{obs} \frac{RR_{CD}p_0 + (1 - p_0)}{RR_{CD}p_1 + (1 - p_1)}$$

Prevalence of confounder among unexposed (E=0)

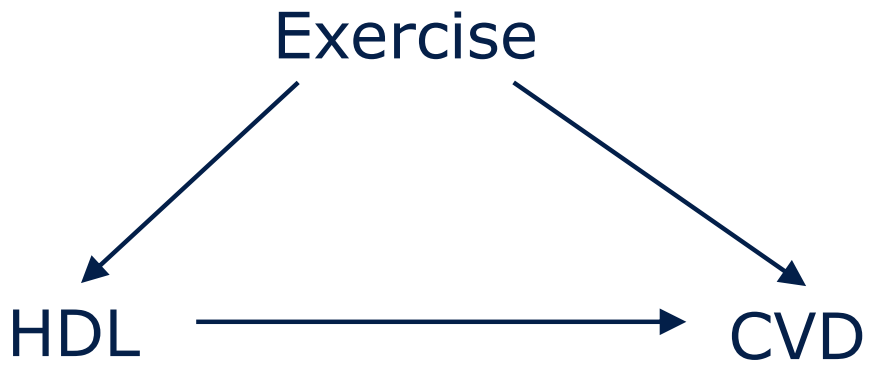
Observed RR without adjustment for confounder

Confounder-disease RR

Prevalence of confounder among exposed (E=1)

Example: HDL cholesterol, exercise & CVD risk

- What estimate effect of HDL on CVD would have been observed in the study had the authors collected data on EXERCISE and adjusted for it in the analysis?

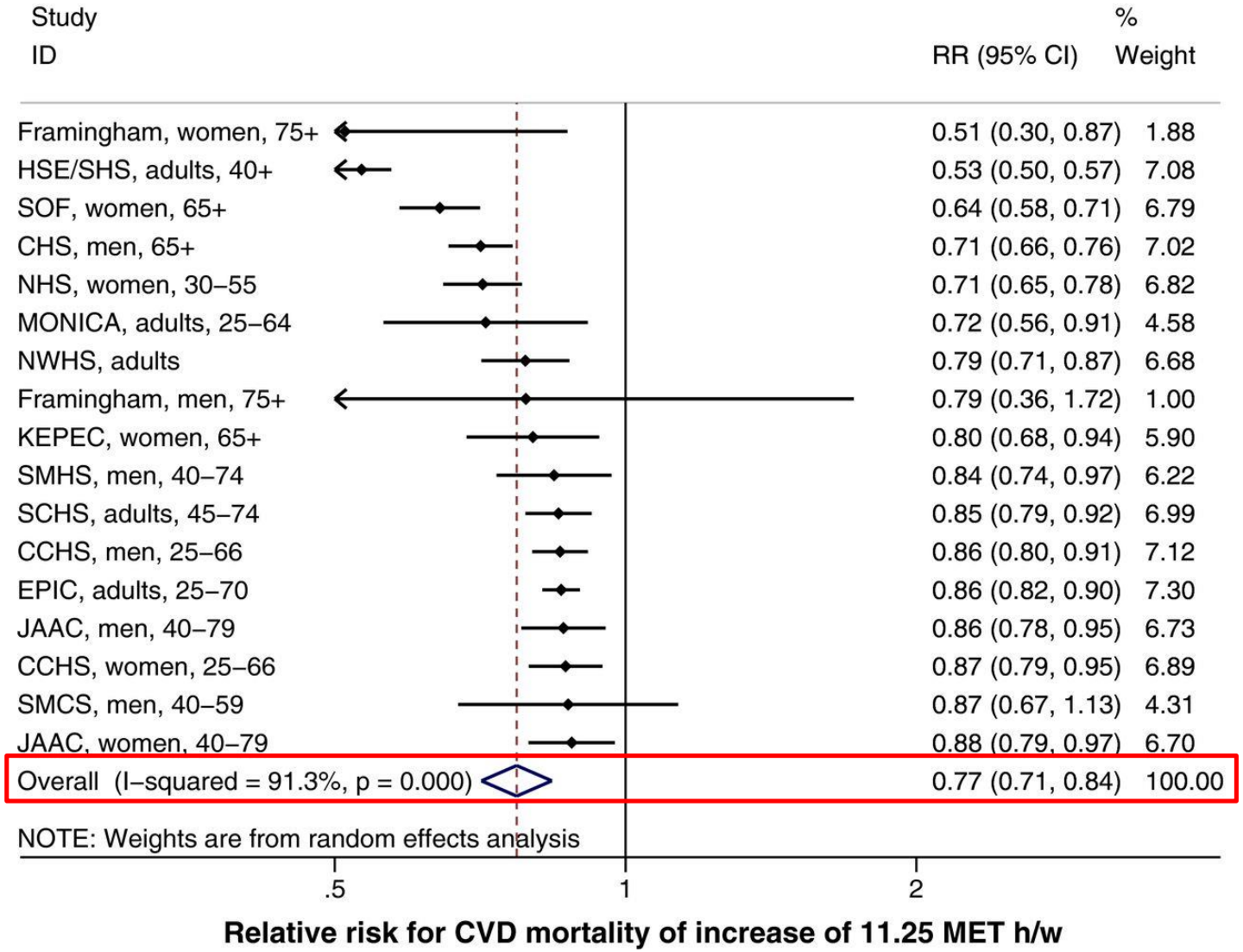


Crude association HDL ⟶ CVD

	Exposed	Unexposed	Total
Cases	515	4357	4872
Noncases	1894	10222	12116
Total	2409	14579	16988
Risk	.2137817	.2988545	.2867907
	Point estimate		[95% Conf. Interval]
Risk difference	-.0850729		-.1030516 -.0670941
Risk ratio	.7153369		.659999 .7753145
Prev. frac. ex.	.2846631		.2246855 .340001
Prev. frac. pop	.0403669		

chi2(1) = 73.15 Pr>chi2 = 0.0000

Exercise $\xrightarrow{?}$ CVD



Bias parameters

Bias Parameter	Description	Value Assigned
RR_{CD}	Association between exercise and CVD	0.77
p_1	Prevalence of exercisers among high HDL	0.80
p_0	Prevalence of exercisers among low HDL	0.05

Observed Data

	Exposed	Unexposed	Total
Cases	515	4357	4872
Noncases	1894	10222	12116
Total	2409	14579	16988

Observed data

Expected data, had the confounder been measured

	Total		C_1		C_0	
	E_1	E_0	E_1	E_0	E_1	E_0
D_+	a	b	A_1	B_1	A_0	B_0
D_-	c	d	C_1	D_1	C_0	D_0
	m	n	M_1	N_1	M_0	N_0

Deterministic bias analysis

Observed data

	Exposed	Unexposed
Cases	515	4357
Noncases	1894	10222
Total	2409	14579

	Total		C_1		C_0	
	E_1	E_0	E_1	E_0	E_1	E_0
D_+	a	b	A_1	B_1	A_0	B_0
D_-	c	d	C_1	D_1	C_0	D_0
	m	n	M_1	N_1	M_0	N_0

1927.2 728.9 481.8 13850.1

1. Solve for M_1 and N_1 using prevalence of confounder in exposed and unexposed:

$M_1 = m*p1 = 2409 * 0.80 = 1927.2$

$N_1 = n*p0 = 14579 * 0.05 = 728.9$

2. Solve for M_0 and N_0 by subtraction:

$M_0 = m - M1 = 2409 - 1927.2 = 481.8$

$N_0 = n - N1 = 14579 - 728.9 = 13850.1$

Deterministic bias analysis

Observed data

	Exposed	Unexposed
Cases	515	4357
Noncases	1894	10222
Total	2409	14579

	Total		C_1		C_0	
	E_1	E_0	E_1	E_0	E_1	E_0
D_+	a	b	A_1	B_1	A_0	B_0
D_-	c	d	C_1	D_1	C_0	D_0
	m	n	M_1	N_1	M_0	N_0

1927.2 728.9 481.8 13850.1

3. Use estimated relationship between exercise (C) and CVD (D) (RR_{CD}=0.77) to fill in the remaining cells:

$$RR_{CD} = \frac{A_1/M_1}{A_0/M_0} \rightarrow A_1 = \frac{RR_{CD} M_1 a}{RR_{CD} M_1 + m - M_1} = \frac{0.77 * 1927.2 * 515}{0.77 * 1927.2 + 2409 - 1927.2} = 388.8$$

$$RR_{CD} = \frac{B_1/N_1}{B_0/N_0} \rightarrow B_1 = \frac{RR_{CD} N_1 b}{RR_{CD} N_1 + n - N_1} = \frac{0.77 * 728.9 * 4357}{0.77 * 728.9 + 14579 - 728.9} = 169.7$$

Deterministic bias analysis

	Total		C_1		C_0	
	E_1	E_0	E_1	E_0	E_1	E_0
D_+	a	b	A_1	B_1	A_0	B_0
D_-	c	d	C_1	D_1	C_0	D_0
	m	n	M_1	N_1	M_0	N_0

4. Fill in remaining cells of the table and calculate effect estimate:

	Total		C1		C0	
	E1	E0	E1	E0	E1	E0
D+	515	4357	388.8	169.7	126.2	4187.3
D-	1894	10222	1538.4	559.3	355.6	9662.7
	2409	14579	1927.2	729.9	481.8	13850.1

Estimated MH RR = 0.82

Stata command -episens-

- Stata command for bias analysis by Nicola Orsini & colleagues (http://www.stata-journal.com/article.html?article=st0138)
  - Add to Stata (type: findit episens)
  - Can be used with data (-episens-) or as an immediate command (-episensi-)

episens *var_case var_exposed* [*var_time*] [*if*] [*in*] [*weight*] [, *options*]

episensi *#a #b #c #d* [, *options*]

- To use -episens- for analysis of unmeasured confounding, need estimates of the bias parameters as previously described:

dpexp: define the prevalence of the confounder among E+

dpunexp: define the prevalence of the confounder among E-

drrcd: define the confounder-disease relative risk

dorce: define the confounder-exposure odds ratio

-episensi-

dpexp: define prevalence of confounder among E+

```
episensi 515 4357 1894 10222, st(cs) dpexp(c.(0.8))
dpunexp(c.(0.05)) drrcd(c(0.77))
```

dpunexp: define prevalence of confounder among E-

drrcd: define confounder-disease risk ratio

```
Pr(c=1|e=1): Constant(.8)
Pr(c=1|e=0): Constant(.05)
RR_cd      : Constant(.77)

Observed Risk Ratio [95% Conf. Interval]= 0.72 [0.66, 0.78]

Deterministic sensitivity analysis for unmeasured
confounding
   External adjusted Risk Ratio = 0.82
   Percent bias = -13%
```

SELECTION BIAS (COLLIDER-STRATIFICATION BIAS)

Selection vs. selection bias

Full sample

hbp	final mortality status: Alive	Deceased	Total
normotensive	6,526	246	6,772
hypertensive	2,656	531	3,187
Total	9,182	777	9,959

OR=5.3

If we select 50% of cases and 20% of controls: no selection bias:

Selected sample

hbp	final mortality status: Alive	Deceased	Total
normotensive	3,263	49	3,312
hypertensive	1,328	106	1,434
Total	4,591	155	4,746

OR=5.3

When does selection bias occur?

Full sample

hbp	final mortality status: Alive	Deceased	Total
normotensive	6,526	246	6,772
hypertensive	2,656	531	3,187
Total	9,182	777	9,959

OR=5.3

If we select 50% of unexposed cases, 80% of exposed cases, 20% of unexposed non-cases, and 10% of exposed non-cases:

Selected sample

hbp	final mortality status: Alive	Deceased	Total
normotensive	3,263	49	3,312
hypertensive	2,125	53	2,178
Total	5,388	102	4,746

OR=1.6

Calculate the cross product of sampling fractions (selection bias odds ratio)

If we select 50% of cases and 20% of non-cases → no selection bias:

	D+	D-
E+	0.5	0.2
E-	0.5	0.2

(0.5*0.2) / (0.5*0.2) = 1.0

If we select 50% of unexposed cases, 80% of exposed cases, 20% of unexposed non-cases, and 10% of exposed non-cases:

	D+	D-
E+	0.8	0.1
E-	0.5	0.2

(0.8*0.2)/(0.5*0.1) = 0.31

Sampling fractions (a.k.a. selection probabilities)

Target

	Exposure+	Exposure-
Disease+	A	B
Disease-	C	D

Selected Sample

	Exposure+	Exposure-
Disease+	$A°$	$B°$
Disease-	$C°$	$D°$

	Exposure+	Exposure-
Disease+	$S_{D+,E+} = A°/A$	$S_{D+,E-} = B°/B$
Disease-	$S_{D-,E+} = C°/C$	$S_{D-,E-} = D°/D$

QBA for Selection Bias

- In the previous exercise with high blood pressure and mortality, we had the "master" 2x2 table for the full sample and applied sampling fractions to get a biased table (and biased effect estimate)

- In general, for bias analysis we need to *work backwards*: we have the biased table (and biased effect estimate), to try to get back to the master 2x2 table

QBA for selection bias with binary exposure & outcome

1. Estimate of selection probabilities

	Exposure+	Exposure-
Disease+	$S_{D+,E+}$	$S_{D+,E-}$
Disease-	$S_{D-,E+}$	$S_{D-,E-}$

2. Use selection probabilities to calculate the selection bias factor

Selection bias factor: $\frac{S_{D+,E+} / S_{D+,E-}}{S_{D-,E+} / S_{D-,E-}}$

3. Bias-corrected estimate: divide effect estimate (e.g. $\widehat{\mathrm{OR}}$) and lower and upper bound of 95% CI by the selection bias factor

$$\mathrm{OR}_{\mathrm{adj}} = \widehat{\mathrm{OR}} / \frac{S_{D+,E+} / S_{D+,E-}}{S_{D-,E+} / S_{D-,E-}}$$

Alternative calculation from Lash et al. (terms rearranged—may be more intuitive)

Table 4.3 Bias parameters required for simple bias analysis to address selection bias from differential initial participation

Selection probabilities	Exposure = 1	Exposure = 0
Cases	$S_{case,1}$	$S_{case,0}$
Noncases	$S_{control,1}$	$S_{control,0}$

Adjust the observed OR by multiplying it by the selection bias OR to account for differential selection:

$$OR_{adj} = \hat{OR} \times \frac{S_{case,0} S_{control,1}}{S_{case,1} S_{control,0}}$$

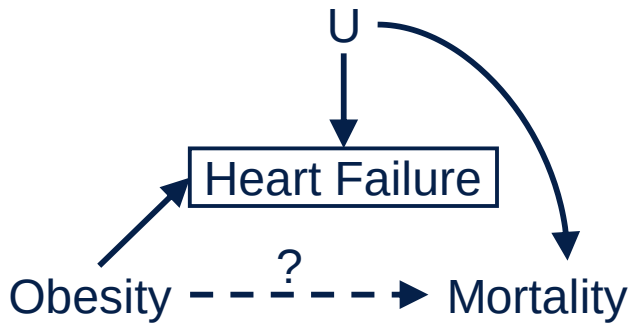
QBA for selection bias

Selection bias factor: $\frac{S_{D+,E+} / S_{D+,E-}}{S_{D-,E+} / S_{D-,E-}}$	
<1.0	$S_{D+,E+} / S_{D+,E-} < S_{D-,E+} / S_{D-,E-}$
1.0 (no bias)	$S_{D+,E+} / S_{D+,E-} = S_{D-,E+} / S_{D-,E-}$
>1.0	$S_{D+,E+} / S_{D+,E-} > S_{D-,E+} / S_{D-,E-}$

Example

- From a study examining the relationship between obesity and mortality in heart failure patients:

```
Proportion
                |   Exposed   Unexposed  |      Total      Exposed
----------------+------------------------+------------------------
           Dead |       998          977 |       1975       0.5053
          Alive |      2086         1606 |       3692       0.5650
----------------+------------------------+------------------------
          Total |      3084         2583 |       5667       0.5442
                |                        |
                |      Point estimate    |    [95% Conf. Interval]
                |------------------------+------------------------
     Odds ratio |          .7864429      |    .7037068    .8789046
```



- You suspect there is collider-stratification bias due to conditioning on a variable (heart failure) affected by exposure that shares common causes with the outcome

Where to get estimates of sampling fractions?

- Need population-representative data to estimate selection probabilities
- Use data from the 1999-2000 and 2001-2002 US National Health and Nutrition Examination Survey (NHANES)

	25.5-29.9	**18.5-24.9**
Dead	258	256
Alive	3270	3096

Total NHANES cohort

	25.5-29.9	**18.5-24.9**
Dead	34	29
Alive	68	23

NHANES Heart Failure = 1

	25.5-29.9	**18.5-24.9**
Dead	?	?
Alive	?	?

Sampling fractions

Calculations

	25.5-29.9	**18.5-24.9**
Dead	258	256
Alive	3270	3096

Total NHANES cohort

	25.5-29.9	**18.5-24.9**
Dead	34	29
Alive	68	23

NHANES Heart Failure = 1

	25.5-29.9	**18.5-24.9**
Dead	34/258 = 0.132	29/256 = 0.113
Alive	68/3270 = 0.0208	23/3096 = 0.00743

Selection probabilities

$$\mathrm{OR}_{\mathrm{adj}} = \widehat{\mathrm{OR}} / \frac{S_{D+,E+} / S_{D+,E-}}{S_{D-,E+} / S_{D-,E-}} = 0.7864429 / \frac{0.132/0.113}{0.0208/0.00743} = 1.88$$

Stata -episens-

- Stata command for bias analysis by Nicola Orsini & colleagues (http://www.stata-journal.com/article.html?article=st0138)
  - Add to Stata (type: findit episens)
  - Can be used with data (-episens-) or as an immediate command (-episensi-)

episens *var_case var_exposed* [*var_time*] [*if*] [*in*] [*weight*] [, *options*]

episensi *#a #b #c #d* [, *options*]

Selection bias options

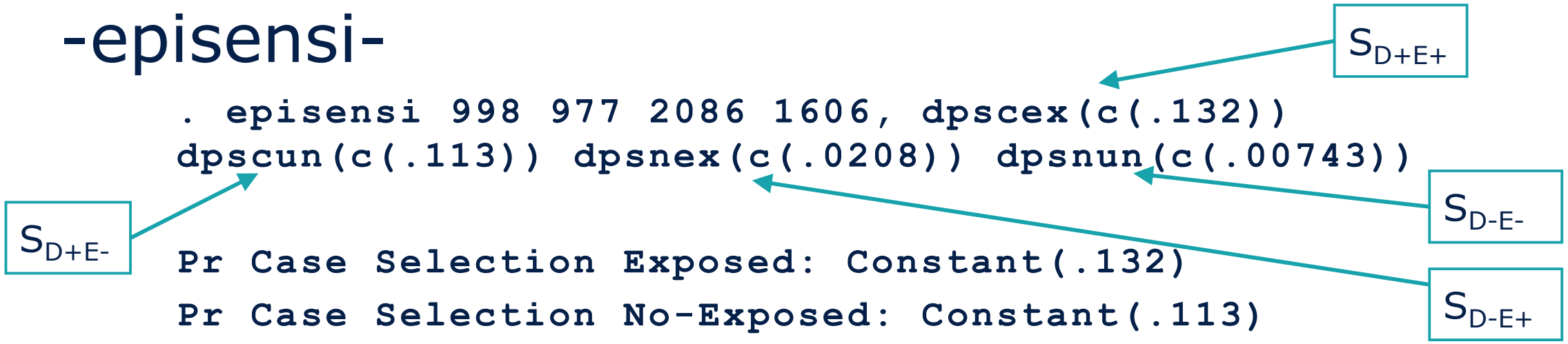
dpscex define the selection probability among cases exposed

dpscun define the selection probability among cases unexposed

dpsnex define the selection probability among noncases exposed

dpsnun define the selection probability among noncases unexposed

-episensi-




```
. episensi 998 977 2086 1606, dpscex(c(.132))
dpscun(c(.113)) dpsnex(c(.0208)) dpsnun(c(.00743))

Pr Case Selection Exposed: Constant(.132)
Pr Case Selection No-Exposed: Constant(.113)
Pr No-Case Selection Exposed: Constant(.0208)
Pr No-Case Selection No-Exposed: Constant(.00743)

Observed Odds Ratio [95% CI]= 0.79 [0.70, 0.88]

Deterministic sensitivity analysis for selection bias

   External adjusted Odds Ratio = 1.88
   Percent bias = -58%
```

MISCLASSIFICATION

Misclassification can arise from:

1. Incorrect measurement and recording of values:
   - Instruments not calibrated correctly
   - Participants not answering questions accurately
   - Incorrect data entry

2. Discordance between definition of variable used in the study and the 'truth'

Result: When dividing participants into exposed/unexposed and diseased/non-diseased some people will be classified into the wrong group (=misclassification)

Terminology

	Truly exposed	Truly unexposed
Classified as exposed	A	B
Classified as unexposed	C	D

$$\text{Sensitivity (SE)} = \frac{A}{A+C} \qquad \text{Positive predictive value (PPV)} = \frac{A}{A+B}$$

$$\text{Specificity (SP)} = \frac{D}{B+D} \qquad \text{Negative predictive value (NPV)} = \frac{D}{C+D}$$

Sensitivity= Probability someone is truly exposed is classified as exposed
Specificity= Probability someone is truly unexposed is classified as unexposed

PPV= Probability of truly being exposed if classified as exposed
NPV= Probability of truly being unexposed if classified as unexposed

Predictive values are highly dependent on the true prevalence of exposure in the population in which the predictive value is measured

Most measurements are imperfect (i.e., measurement is affected by error)

- A and Y = "constructs" of interest
- A* and Y* = measures
- Assumption implicit in many analyses is that the association between A* and Y* approximates the association between A and Y

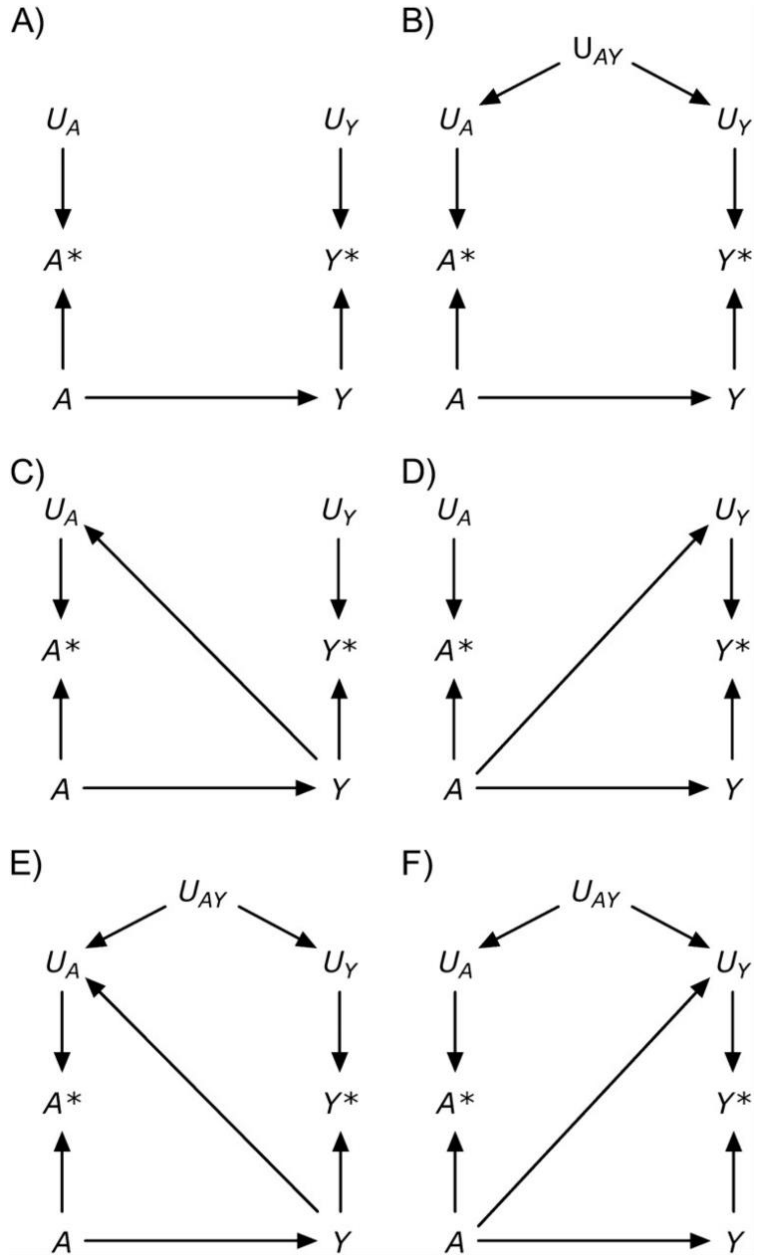
A) Errors for the exposure and the outcome are **independent**

B) Measurement errors for exposure and outcome are **dependent**

C) **Independent but differential measurement error**: true value of the outcome affects the measurement of the exposure (i.e., an arrow from Y to U_A)

D) **Independent but differential measurement error**: true value of the exposure affects the measurement of the outcome (i.e., an arrow from *A* to U_Y)

E and F) Measurement errors that are **both dependent and differential**, resulting from a combination of the settings described above



Misclassification (see DAG lecture for more details)

- Non-differential misclassification: Assignment of exposure is independent of disease status (e.g., cohort study where disease hasn't occurred, blinded ascertainment) *or* assignment of disease is independent of exposure status
- Differential misclassification: Sensitivity and/or specificity for disease classification depends on exposure status *or* sensitivity and/or specificity for exposure classification depends on disease status
- Dependent misclassification: When individuals who are misclassified on one variable (e.g., exposure status) are more likely than individuals correctly classified on that variable to also be misclassified on a second variable (e.g., disease status)
- Independent misclassification: When individuals who are misclassified on one variable are no more likely to be misclassified on a second variable

Where to obtain bias parameters?

- Internal validation study

	Strategy	Parameters
1	Select individuals based on misclassified exposure measurement	PPV, NPV
2	Select individuals based on true exposure status	SE, SP
3	Select a random sample of individuals	PPV, NPV, SE, SP

- External sources (i.e., literature review, expert collaboration, educated guesses)

Example: Self-report vs. measured BMI

Measured BMI

bmi_mcat	Freq.	Percent	Cum.
<18.5	646	2.01	2.01
18.5-24.9	10,140	31.55	33.56
25-29.9	11,136	34.65	68.22
30-34.9	6,345	19.74	87.96
35-39.9	2,758	8.58	96.55
>40	1,110	3.45	100.00
Total	32,135	100.00	

Self-report BMI

bmi_srcat	Freq.	Percent	Cum.
<18.5	579	1.81	1.81
18.5-24.9	11,345	35.50	37.32
25-29.9	11,129	34.83	72.14
30-34.9	5,838	18.27	90.41
35-39.9	2,204	6.90	97.31
>40	859	2.69	100.00
Total	31,954	100.00	

Comparing self-report and measured BMI

bmi_mcat	bmi_srcat					
	<18.5	18.5-24.9	25-29.9	30-34.9	35-39.9	>40
<18.5	346	273	0	0	0	0
18.5-24.9	194	8,758	862	15	0	3
25-29.9	8	1,923	8,206	579	20	2
30-34.9	0	92	1,730	4,047	222	12
35-39.9	0	6	102	988	1,449	93
>40	0	2	7	68	413	524

There does not appear to be a consistent direction to the misclassification, people are just bad at self-reporting their BMI accurately

Table 2. Mortality Hazard Ratios[a] for Measured or Self-Reported Categories of Body Mass Index According to Sex in the Full and Restricted[b] Samples From National Health and Nutrition Examination Survey II (1976–1980), III (1988–1994), and Continuous (1999–2010), United States

Sample and BMI Ascertainment Method	BMI Category							
	Low (<22.5)		Overweight (25–29.9)		Class I Obesity (30–34.9)		Class II–III Obesity (≥35)	
	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI
Men								
Full sample								
Measured[c]	1.28	1.13, 1.45	0.94	0.86, 1.02	1.08	0.95, 1.22	1.70	1.42, 2.03
Self-reported[d]	1.30	1.16, 1.44	0.97	0.89, 1.05	1.07	0.94, 1.22	1.83	1.52, 2.20
Restricted sample								
Measured[c]	0.87	0.56, 1.36	0.85	0.62, 1.17	1.13	0.78, 1.64	1.86	1.12, 3.07
Self-reported[d]	1.26	0.82, 1.94	1.24	0.90, 1.70	1.26	0.84, 1.90	2.88	1.76, 4.72

Abbreviations: BMI, body mass index; CI, confidence interval; HR, hazard ratio.

[a] Hazard ratios from Cox proportional hazards models, with age as the timeline and adjusted for survey, smoking status, racial/ethnic group, and alcohol consumption. The reference category was BMI of 22.5–24.9.

[b] The restricted sample was limited to participants who reported never smoking, who reported no history of heart disease or cancer, and who were examined when younger than age 70 years.

[c] Measured BMI was calculated from measured height and measured weight as weight (kg)/height $(m)^2$.

[d] Self-reported BMI was calculated from self-reported weight and self-reported height as weight (kg)/height $(m)^2$.

Formula using sensitivity and specificity to adjust for exposure misclassification of a binary exposure and outcome assuming non-differential, independent misclassification

	Observed		Corrected data	
	E_1	E_0	E_1	E_0
D_+	a	b	$[a - D_{+\,\text{Total}}(1 - \text{SP}_{D+})]/[\text{SE}_{D+} - (1 - \text{SP}_{D+})]$	$D_{+\,\text{Total}} - A$
D_-	c	d	$[c - D_{-\,\text{Total}}(1 - \text{SP}_{D-})]/[\text{SE}_{D-} - (1 - \text{SP}_{D-})]$	$D_{-\,\text{Total}} - C$
Total	$a + c$	$b + d$	$A + C$	$B + D$

$A + C$ is the corrected total number of exposed individuals and $B + D$ is the corrected total number of unexposed individuals.

Example: Shift work and obesity

- You are conducting a study to explore the relationship between shift work (exposure) and obesity (outcome)
- All that is available in the dataset is self-report BMI

	Exposed	Unexposed	Total
Cases	2670	7957	10627
Noncases	4357	18710	23067
Total	7027	26667	33694
Risk	.379963	.2983838	.3153974

	Point estimate	[95% Conf.	Interval]
Risk ratio	1.273404	1.229504	1.318871

Validation Study

- After taking a QBA workshop, you decide to conduct a validation study to examine potential non-differential misclassification of the outcome in your study (self-report BMI):

ob_sr	ob_m 1	ob_m 0
1	9,279	1,348
0	2,486	20,581

- Calculate sensitivity, specificity, PPV, and NPV from this table

-Diagt- command to calculate sensitivity, specificity, PPV, and NPV

Report summary statistics for diagnostic tests compared to true disease status

---

diagt diagvar testvar [weight] [if exp] [in range] [, prev(#) level(#)

diagti #a #b #c #d [, prev(#) level(#)

diagvar is the variable the contains the *true status* of the patient
testvar is the variable that identifies the *diagnostic test result*

-Diagt- command

diagt diagvar testvar

diagvar = variable for *true status*

testvar = variable for *diagnostic test result*

```
. diagt obm obsr

           |        obsr
       obm |       Pos.        Neg. |     Total
-----------+---------------------------+----------
  Abnormal |      9,279       2,486 |    11,765
    Normal |      1,348      20,581 |    21,929
-----------+---------------------------+----------
     Total |     10,627      23,067 |    33,694

True abnormal diagnosis defined as obm = 1
```

			[95% Confidence Interval]
Prevalence	Pr(A)	35%	34% 35.4%
Sensitivity	Pr(+|A)	78.9%	78.1% 79.6%
Specificity	Pr(-|N)	93.9%	93.5% 94.2%
ROC area	(Sens. + Spec.)/2	.864	.86 .868
Likelihood ratio (+)	Pr(+|A)/Pr(+|N)	12.8	12.2 13.5
Likelihood ratio (-)	Pr(-|A)/Pr(-|N)	.225	.217 .233
Odds ratio	LR(+)/LR(-)	57	53.1 61.2
Positive predictive value	Pr(A|+)	87.3%	86.7% 87.9%
Negative predictive value	Pr(N|-)	89.2%	88.8% 89.6%

-episensi- for misclassification

```
episensi  2670 7957 4357 18710, st(cc) dseca(c(.64))
dspca(c(.83)) dsenc(c(.95)) dspnc(c(.90))

Se|Cases    : Constant(.64)
Sp|Cases    : Constant(.83)
Se|No-Cases: Constant(.95)
Sp|No-Cases: Constant(.9)

Observed Odds Ratio [95% Conf. Interval]= 1.44 [1.36, 1.52]

Deterministic sensitivity analysis for misclassification of the
exposure
   External adjusted Odds Ratio = 1.79
   Percent bias = -19%
```

Formula using PPV and NPV to adjust for exposure misclassification of a binary exposure and outcome assuming non-differential, independent misclassification

- When PPV and NPV data are available, can use those values in place of sensitivity and specificity:

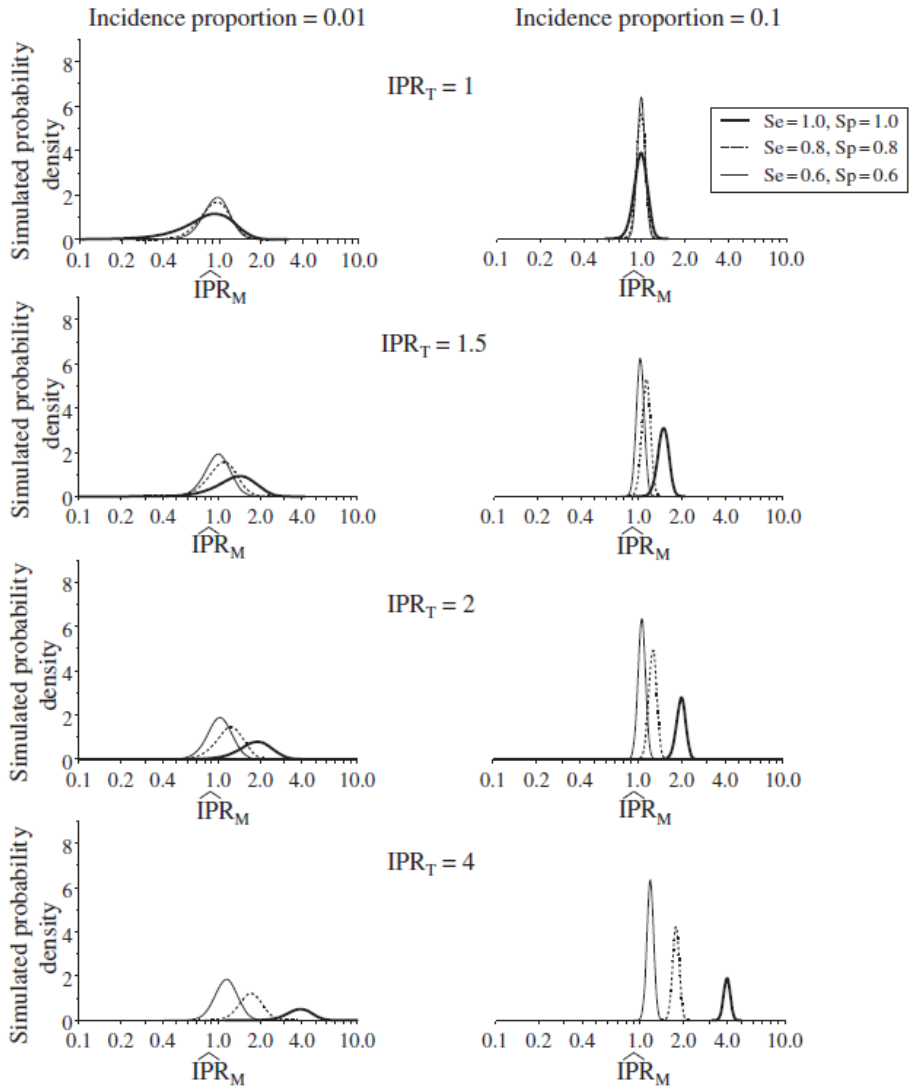
	Observed		Corrected data	
	E_1	E_0	E_1	E_0
D_+	a	b	$a(\text{PPV}_{D+}) + b(1 - \text{NPV}_{D+})$	$D_{+\,\text{Total}} - A$
D_-	c	d	$c(\text{PPV}_{D-}) + d(1 - \text{NPV}_{D-})$	$D_{-\,\text{Total}} - C$
Total	$a + c$	$b + d$	$A + C$	$B + D$

Lash et al., *Applying Quantitative Bias analysis to Epidemiologic Data (Ch. 6)*, 2009

Non-differential misclassification: bias is not *always* towards the null

Non-differential misclassification does not always lead to bias toward the null!

The simulations by Jurek et al. demonstrate that *on average*, results may be biased toward the null, but for some individual iterations, the result was biased away from the null



Jurek et al., 2005

Non-differential misclassification: bias is not *always* towards the null

- When an exposure has more than 2 categories, misclassification of exposure can bias effect estimates in middle categories away from the null

Table 6.14 Association between body mass index and the risk of diabetes, BRFSS 2001 (Mokdad et al., 2003)

	<25	25–30	30+
Diabetics	3,053	5,127	6,491
Nondiabetics	81,416	65,104	33,814
OR	Reference	2.1	5.1

Table 6.15 Corrected data for the association between body mass index and prevalence of diabetes (adapted from Mokdad et al., 2003 and Yun et al., 2006)

	<25	25–30	30+
Diabetics	2,545.9	2,607.5	9,517.6
Nondiabetics	74,977.1	55,776.3	49,580.6
OR	Reference	1.4	5.7

Validation study demonstrated overweight and obesity underestimated in BRFSS

Assuming:

- Specificity 100%
- BMI misclassification independent of diabetes status (nondifferential misclassification)
- Sensitivity for assigning overweight BMI 91%
- Sensitivity for assigning obese BMI 68.2%
- Assume misclassified data only shifted one column (i.e., no obese BMI misclassified as normal weight BMI)

Lash et al., *Applying Quantitative Bias analysis to Epidemiologic Data (Ch. 6)*, 2009

MULTIDIMENSIONAL BIAS ANALYSIS

Concept

- Conceptually, multidimensional bias analysis is no different than the three “simple” types of bias analysis we have already discussed

- One important distinction:
  - Rather than assuming you know the exact value of the bias parameters, you explore several different options for the values of the bias parameters

- Repeat the simple bias analysis several times with different values of the bias parameters

Example

Proportion

	HBP	Normal BP	Total	Exposed
Dead	53	49	102	0.5196
Alive	2125	3263	5388	0.3944
Total	2178	3312	5490	0.3967
	Point estimate		[95% Conf. Interval]	
Odds ratio	1.660879		1.100383	2.510239 (exact)

- You suspect selection bias, but have no way of estimating selection probabilities
- Can take a sensitivity analysis approach and vary the selection bias factor to see how the estimate changes

Selection bias factor

- You suspect selection bias, but have no way of estimating selection probabilities
- Can take a sensitivity analysis approach and vary the selection bias factor to see how the estimate changes

Example

Selection Bias Factor	Corrected RR estimate	Corrected 95% CI
0.4	4.15 (1.660879/0.4)	(2.75, 6.28) (1.100383/0.4, 2.510239/0.4)
0.6	2.77	(1.83, 4.18)
0.8	2.08	(1.38, 3.14)
1.2	1.38	(0.92, 2.09)
1.4	1.19	(0.79, 1.79)
1.6	1.04	(0.69, 1.57)

Divide the biased estimate (and upper and lower bounds of 95% CI) by selection bias factor to get corrected estimate

```
episensi 53 49 2125 3263, st(cc) dsbfactor(c.(0.4))
```

PROBABILISTIC BIAS ANALYSIS

Concept: probabilistic bias analysis

- Extension of simple bias analysis: the analyst assigns *probability distributions to the bias parameters*, rather than single values
- Uses a Monte Carlo sampling procedure to draw values of bias parameters from a probability distribution and use sampled values in the bias correction equations
- Incorporates uncertainty around the bias parameters and produces simulation intervals to quantify magnitude of uncertainty in bias-adjusted effect estimate

Probabilistic bias analysis

1. Estimate bias parameters

2. Specify a probability distribution for each bias parameter

3. Use Monte Carlo sampling to select bias parameter values from the probability distributions

4. Use the sampled bias parameter values to correct the naive effect estimate

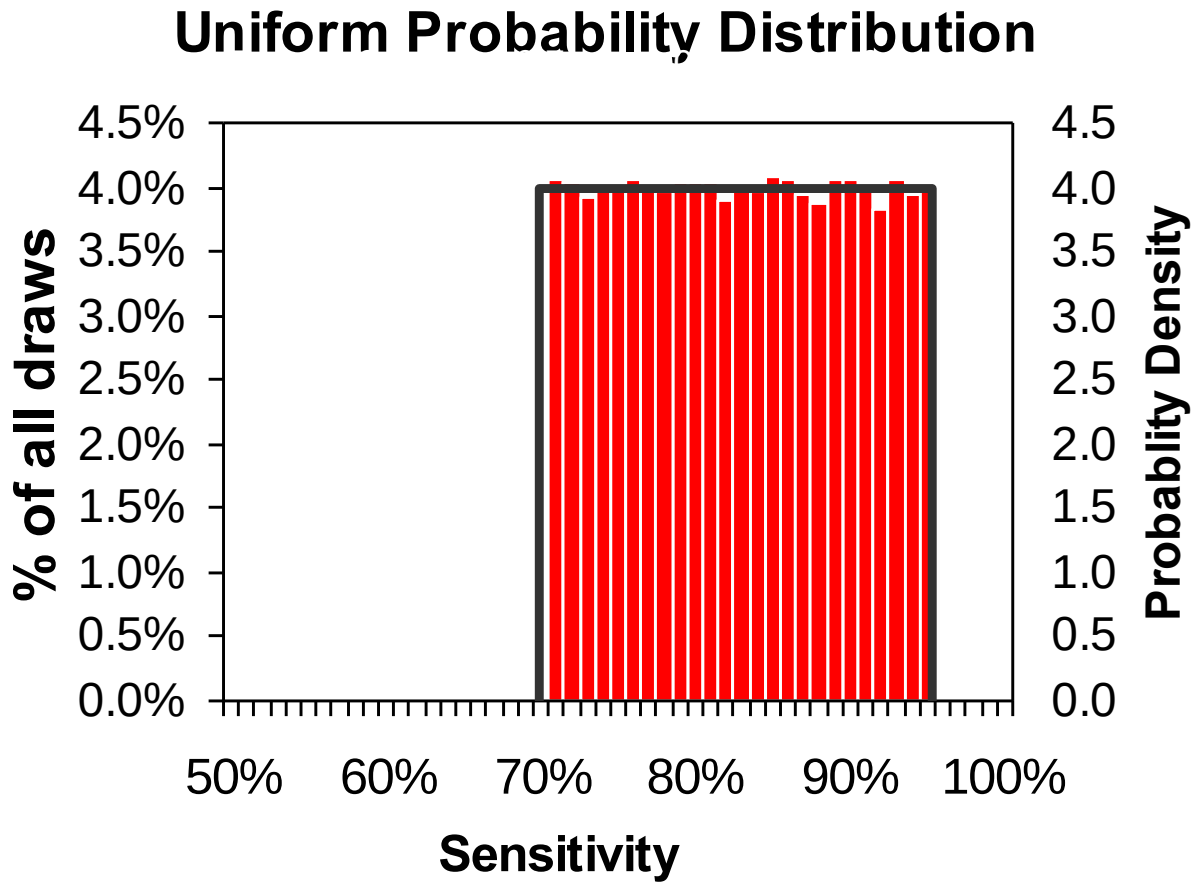
5. Repeat steps 3 and 4 (*many* times)

6. Generate a frequency distribution of corrected effect estimates

- 50^{th} percentile of this distribution = bias-corrected effect estimate
- 2.5 and 97.5 percentiles = 95% simulation interval that accounts for random and systematic sources of error

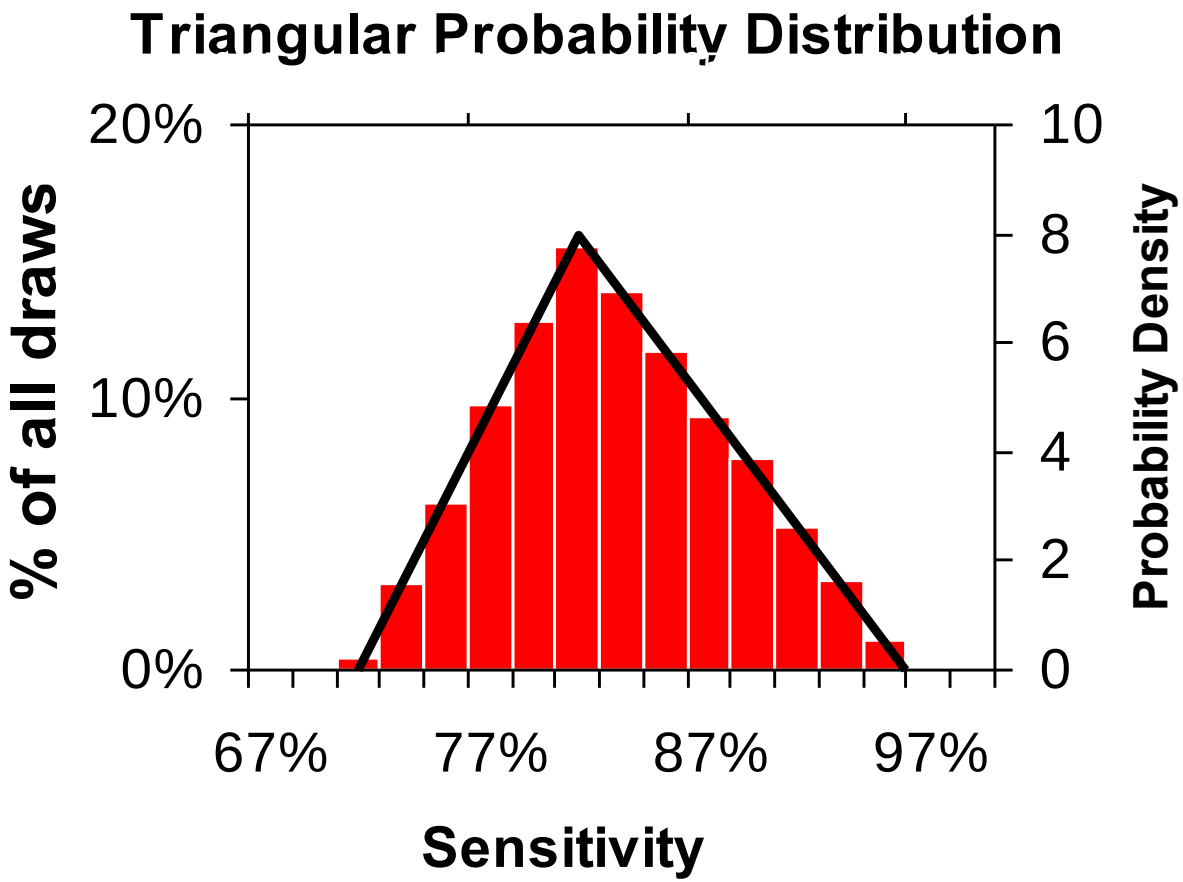
Probability Distributions: Uniform

- **Uniform**
  - Min and max
  - Not very realistic (sharp drop on sides plummeting to 0)



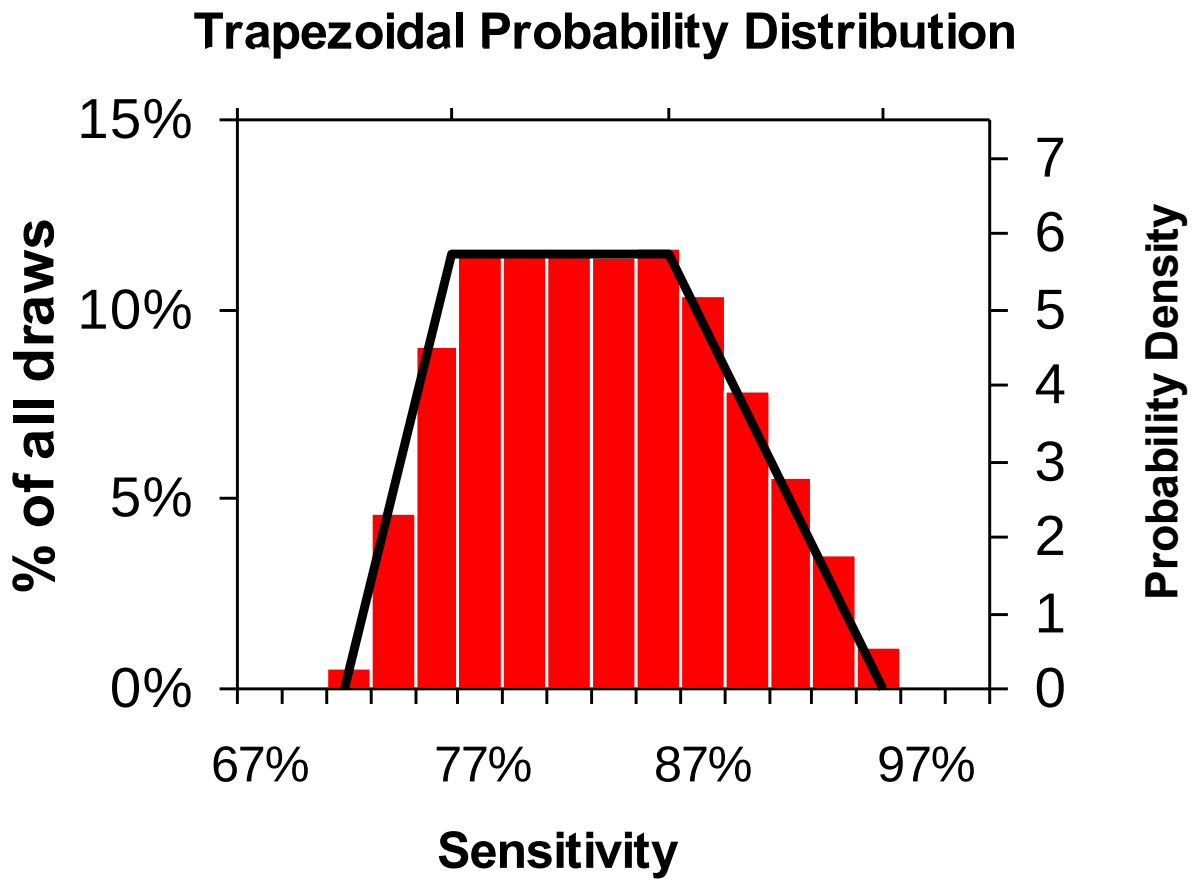
Probability Distributions: Triangular

- Triangular
  - Min, mode, max
  - More realistic, but rarely is one point known to be superior to others so strongly



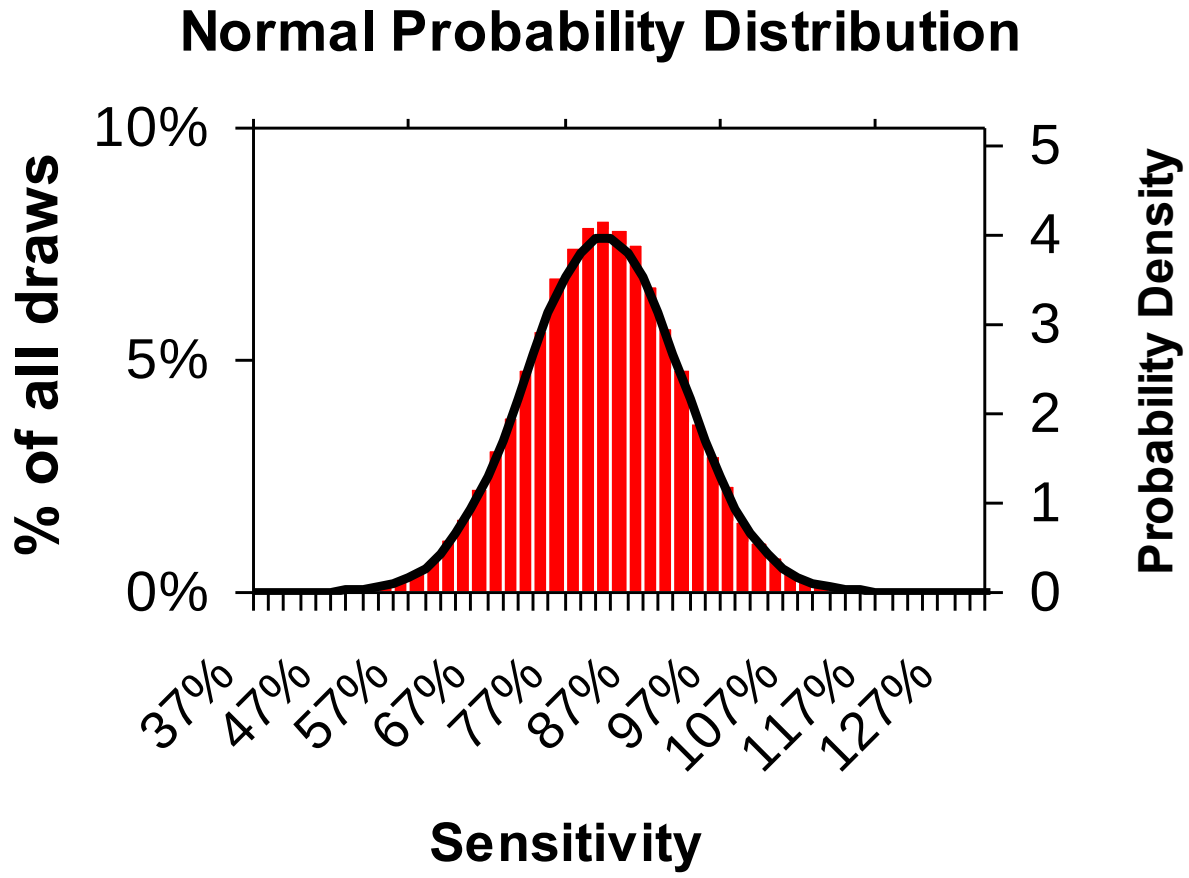
Probability Distributions: Trapezoidal

- Trapezoidal
  - Min, lower mode, upper mode, max
  - More realistic, has a zone of confidence (plateau at the top)



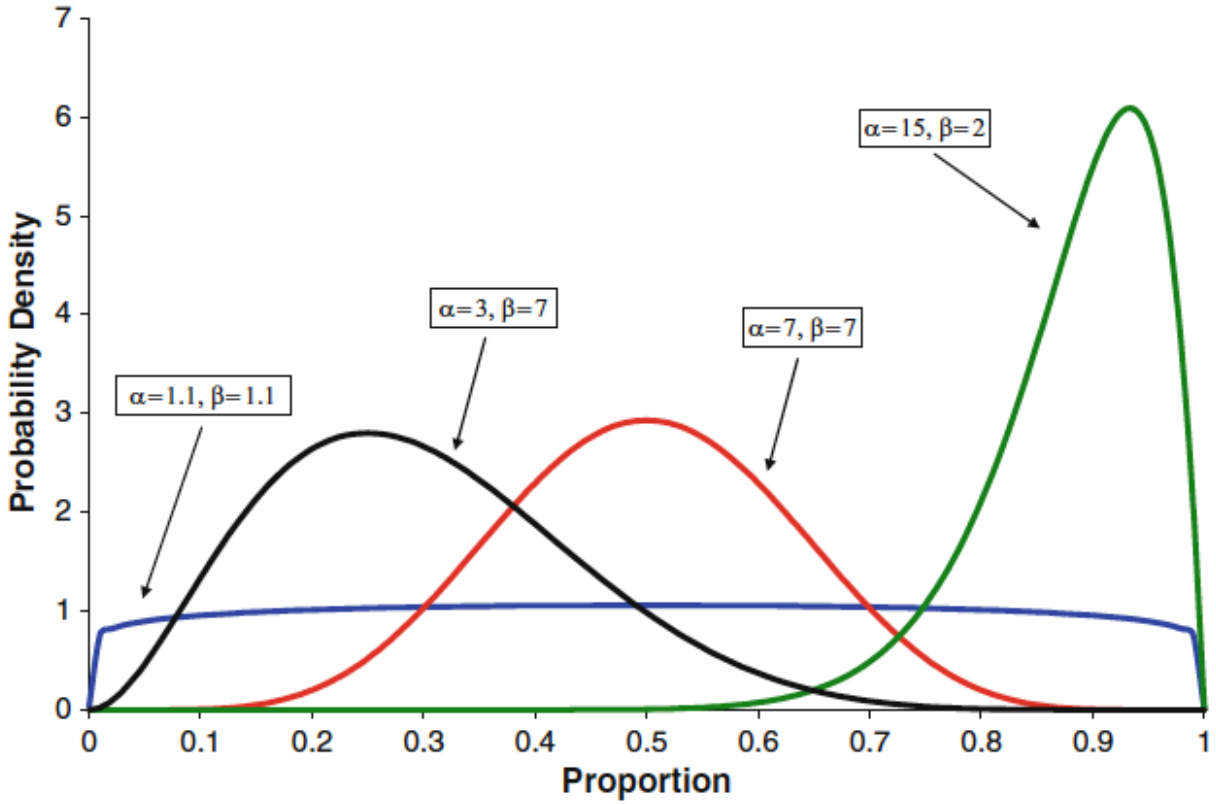
Probability Distributions: Normal

- Truncated normal
  - Mean, std
  - Truncated so that cannot draw outside range 0-1
  - Can allow for empirical truncation



Probability Distributions: Beta

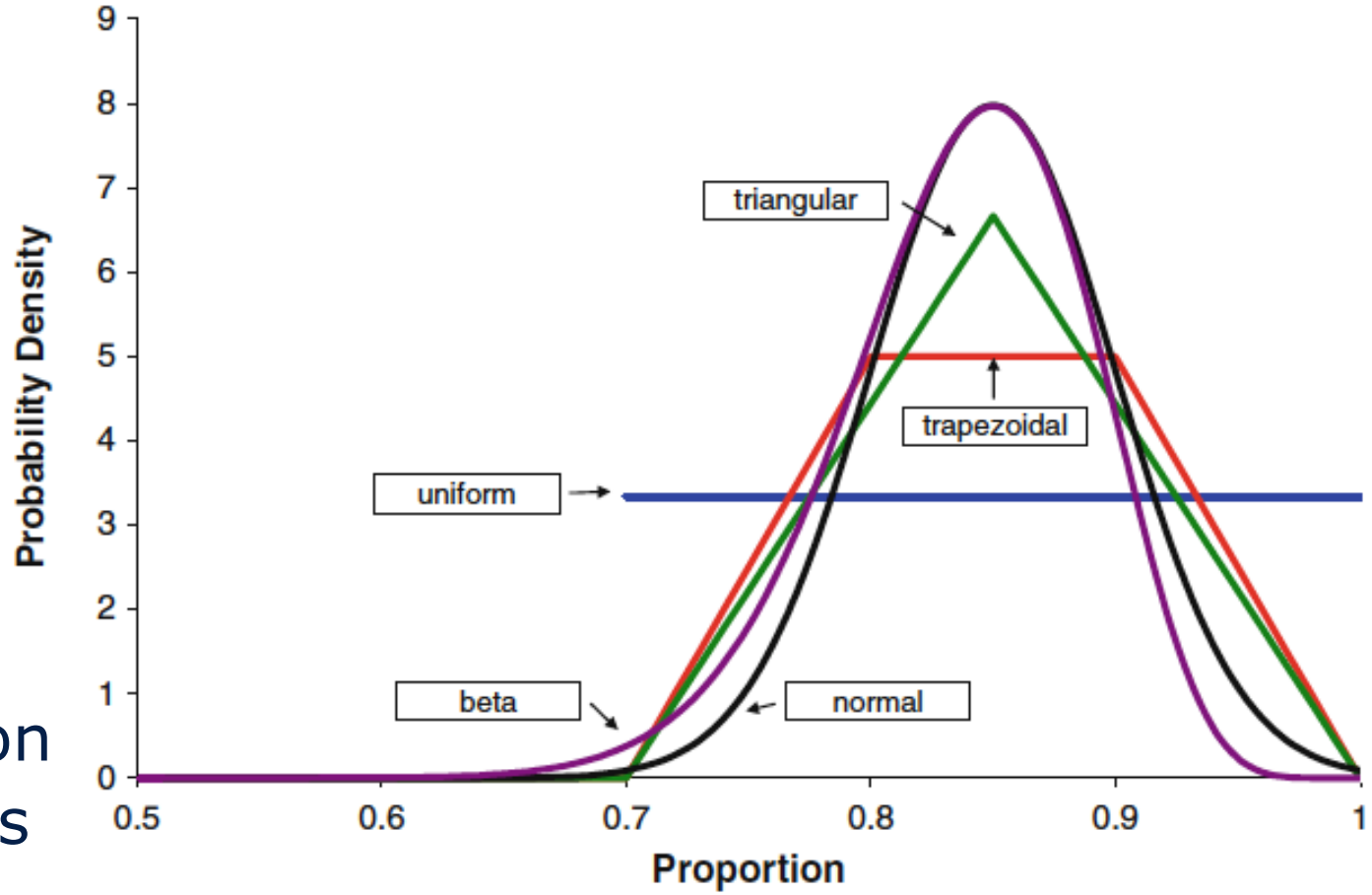
- Beta
  - Alpha and beta shape parameters
  - Works well for proportions and validation data (and a lot of the data we use are proportions)
  - For sensitivity, among true positives, set
    - Alpha = N test positives + 1
    - Beta = N test negatives + 1



Comparing distributions

- All of the distributions here are centered around 0.85
- Majority of the density lies between 0.7 and 0.1

⇨Choice of distribution not likely to have a major effect on the results of the bias analysis



Software options

- Stata:
  - rbeta
  - runiform
  - episens has several probability density functions built-in

- SAS:

- standard uniform: ranuni(seed)
- quantile(‘uniform’,u,{l,r})
- rannor(seed) ← standard normal
- quantile(‘normal’,u, μ,σ)
- quantile(‘beta’,u,α,β,l,r)

-episens- for probabilistic bias analysis (exposure misclassification example)

- Case Control study: Exposure to resins & lung cancer cases (Greenland et al., 1994)

	Exposed	Unexposed	Total	Proportion Exposed	
Cases	45	94	139	0.3237	
Controls	257	945	1202	0.2138	
Total	302	1039	1341	0.2252	
	Point estimate		[95% Conf.	Interval]	
Odds ratio	1.760286		1.202457	2.576898	(Woolf)
Attr. frac. ex.	.4319106		.1683693	.6119365	(Woolf)
Attr. frac. pop	.1398272				

chi2(1) = 8.63 Pr>chi2 = 0.0033

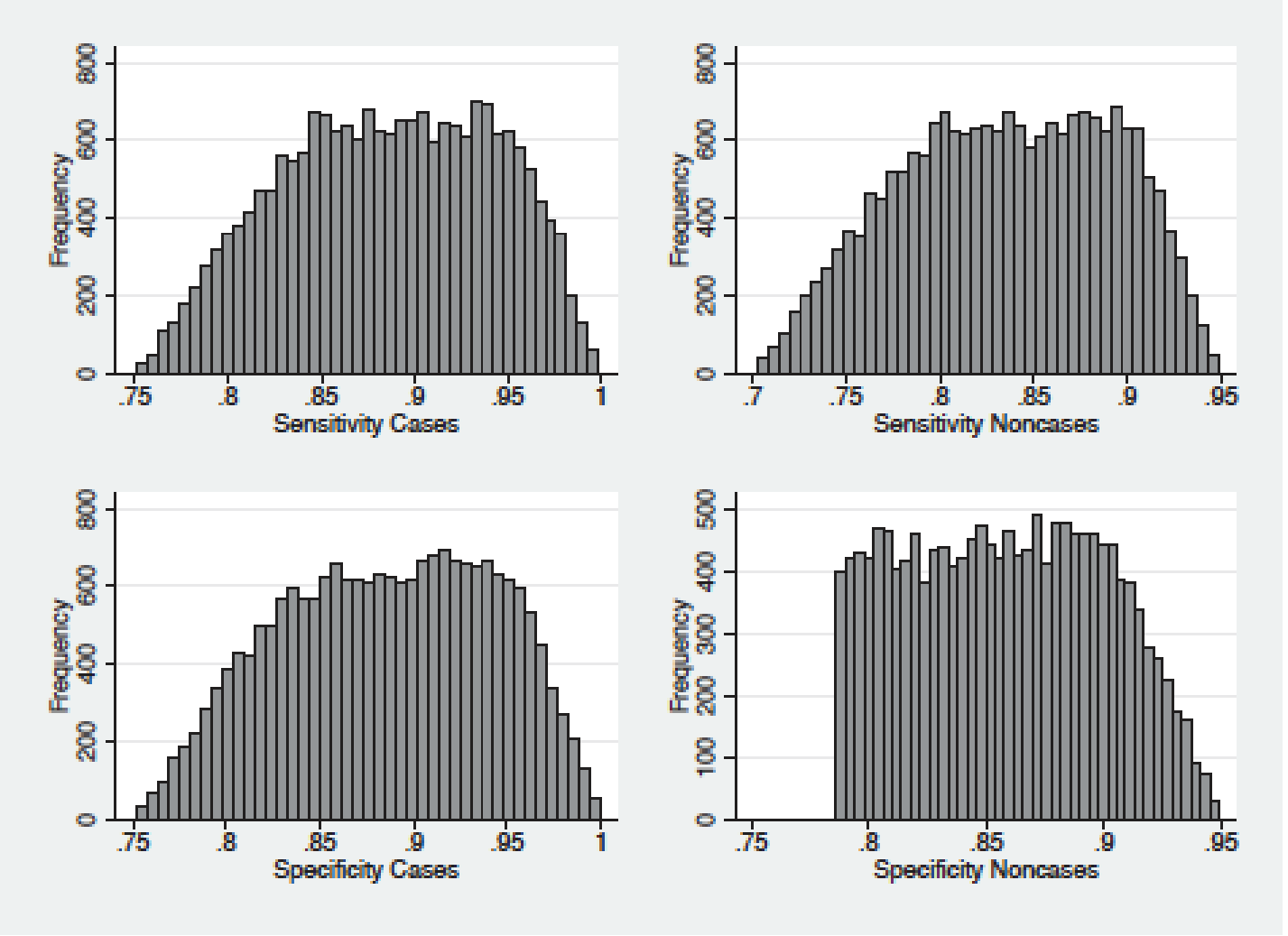
Specify bias parameter *probability distributions* trapezoidal(a b c d) → trapezoidal with min=a, lower mode=b, upper mode=c, max=d

```
. episensi 45 94 257 945, st(cc) reps(20000) nodots dseca(trap(.75 .85 .95 1))
> dspca(trap(.75 .85 .95 1)) dsenc(trap(.7 .8 .9 .95))
> dspnc(trap(.7 .8 .9 .95)) corrsens(.8) corrspec(.8) seed(123) grprior
Se|Cases    : Trapezoidal(.75,.85,.95,1)
Sp|Cases    : Trapezoidal(.75,.85,.95,1)
Se|No-Cases: Trapezoidal(.7,.8,.9,.95)
Sp|No-Cases: Trapezoidal(.7,.8,.9,.95)
Corr Se|Cases and Se|No-Cases : .8
Corr Sp|Cases and Sp|No-Cases : .8
Probabilistic sensitivity analysis for misclassification of the exposure
                                    Percentiles          Ratio
                          2.5       50        97.5       97.5/2.5
                          ----------------------------------------
Conventional              1.20      1.76      2.58       2.14
Systematic error          1.81      3.48      48.19      26.57
Systematic and random error 1.61    3.60      48.92      30.47
```

"The 95% simulation limits including systematic and random error were 1.6 and 49, with a median estimate of 3.6"

Exposure misclassification histograms

- Histograms of 20,000 draws from trapezoidal prior distributions for the sensitivity and specificity among cases and non-cases

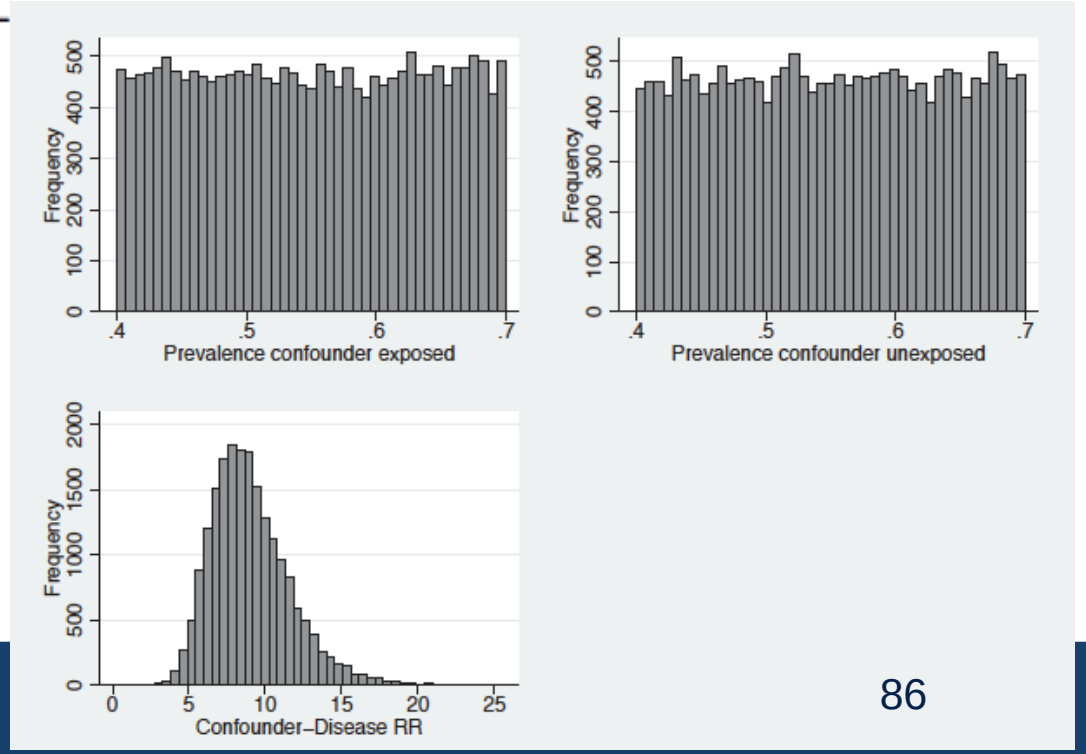


-episens- for probabilistic bias analysis

```
. episensi 45 94 257 945, st(cc) reps(20000) nodots dpexp(uni(.4 .7))
> dpunexp(uni(.4 .7)) drrcd(log-n(2.159 .280)) seed(123)
> grarrtot grprior

Pr(c=1|e=1): Uniform(.4,.7)
Pr(c=1|e=0): Uniform(.4,.7)
RR_cd      : Log-Normal(2.16,0.28)

Probabilistic sensitivity analysis for unmeasured confounding

                                    Percentiles          Ratio
                          2.5       50          97.5     97.5/2.5
                          ------------------------------------
Conventional              1.20      1.76        2.58     2.14
Systematic error          1.25      1.76        2.49     2.00
Systematic and random error 1.05    1.76        2.96     2.83
```

Unmeasured confounder (smoking)

"From the 20,000 draws for each bias parameter, the median smoking-adjusted resins lung-cancer OR is 1.76 with 2.5th and 97.5th percentiles of 1.05 and 2.96"



Collider-stratification bias example: the obesity paradox

Association of Weight Status With Mortality in Adults With Incident Diabetes

Mercedes R. Carnethon, PhD
Peter John D. De Chavez, MS
Mary L. Biggs, PhD
Cora E. Lewis, MD
James S. Pankow, PhD
Alain G. Bertoni, MD, MS
Sherita H. Golden, MD, MS
Kiang Liu, PhD
Kenneth J. Mukamal, MD, MPH
Brenda Campbell-Jenkins, PhD
Alan R. Dyer, PhD

Objective To test the association of weight status with mortality in adults with new-onset diabetes in order to minimize the influence of diabetes duration and voluntary weight loss on mortality.

Design, Setting, and Participants Pooled analysis of 5 longitudinal cohort studies: Atherosclerosis Risk in Communities study, 1990-2006; Cardiovascular Health Study, 1992-2008; Coronary Artery Risk Development in Young Adults, 1987-2011; Framingham Offspring Study, 1979-2007; and Multi-Ethnic Study of Atherosclerosis, 2002-2011. A total of 2625 participants with incident diabetes contributed 27 125 person-years of follow-up. Included were men and women (age >40 years) who developed incident diabetes based on fasting glucose 126 mg/dL or greater or newly initiated diabetes medication and who had concurrent measurements of BMI. Participants were classified as normal weight if their BMI was 18.5 to 24.99 or overweight/obese if BMI was 25 or greater.

Conclusion Adults who were normal weight at the time of incident diabetes had higher mortality than adults who are overweight or obese.

JAMA. 2012;308(6):581-590 www.jama.com

Analytic cohort comprised of 2,625 men and women with incident diabetes (based on fasting glucose or newly initiated medication for diabetes)

Table 2. Association Between Weight Status at the Time of Diabetes Incidence and Mortality[a]

	Total Mortality		Cardiovascular Mortality	
	Overweight/Obese	Normal Weight	Overweight/Obese	Normal Weight
	Full Sample			
No. of participants	2332	293	2086	265
No. of events	371	78	149	24
Event rate (per 10 000 person-years)	152.1	284.8	67.8	99.8
Pooled analysis[b]				
Unadjusted HR (95% CI)	1 [Reference]	1.70 (1.33-2.18)	1 [Reference]	1.31 (0.85-2.02)
Multivariable model 1[c]	1 [Reference]	1.49 (1.15-1.93)	1 [Reference]	1.04 (0.65-1.66)
Multivariable model 2[d]	1 [Reference]	2.08 (1.52-2.85)	1 [Reference]	1.52 (0.89-2.58)

	Normal weight	Overweight/obese
Dead	78	371
Alive	215	1961

Carnethon paper

	Normal weight	Overweight/obese
Dead	1835	2735
Alive	5705	7812

NHANES III Complete cohort

	Normal weight	Overweight/obese
Dead	203	612
Alive	99	521

NHANES III Diabetes

	Normal weight	Overweight/obese
Dead	0.11	0.22
Alive	0.02	0.07

Selection probabilities

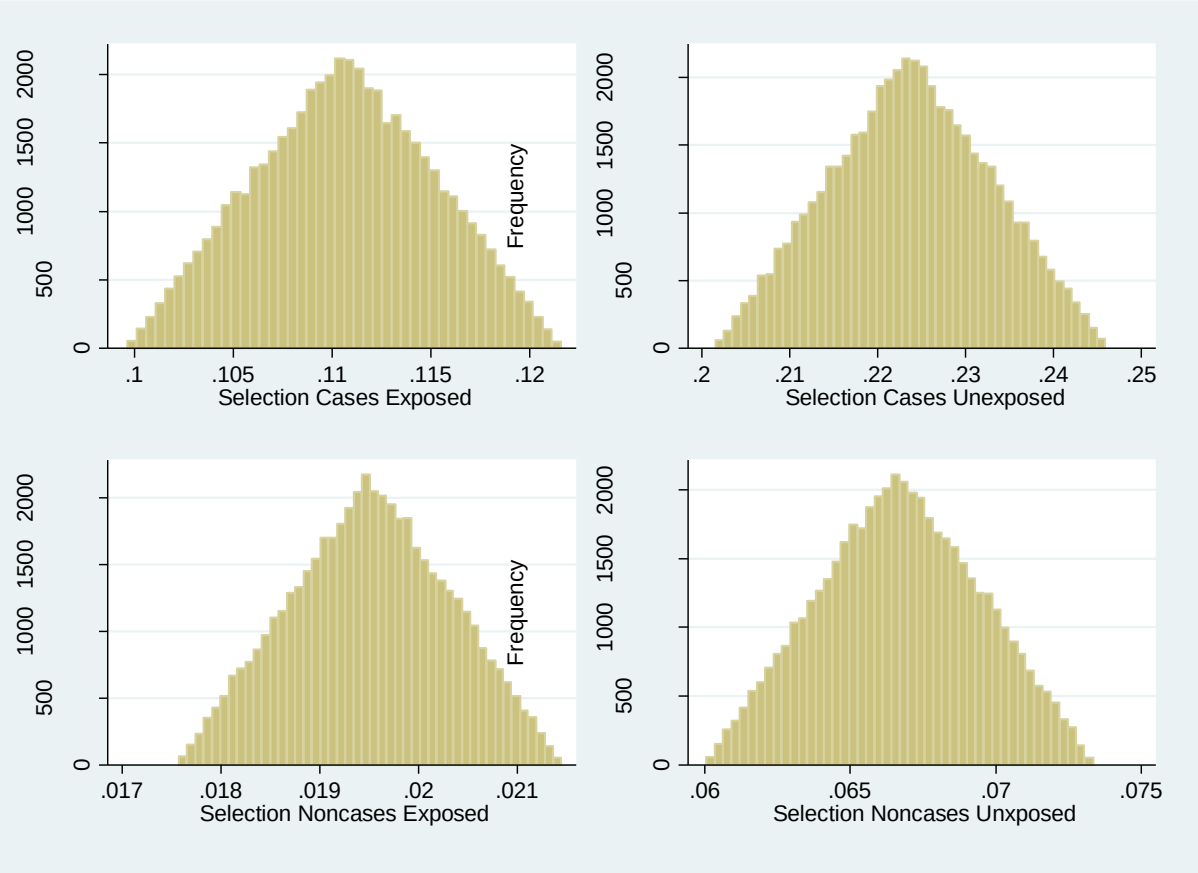
Analytic cohort comprised of 2,625 men and women with incident diabetes (based on fasting glucose or newly initiated medication for diabetes)

Table 2. Association Between Weight Status at the Time of Diabetes Incidence and Mortality[a]

	Total Mortality		Cardiovascular Mortality	
	Overweight/Obese	Normal Weight	Overweight/Obese	Normal Weight
	Full Sample			
No. of participants	2332	293	2086	265
No. of events	371	78	149	24
Event rate (per 10 000 person-years)	152.1	284.8	67.8	99.8
Pooled analysis[b]				
Unadjusted HR (95% CI)	1 [Reference]	1.70 (1.33-2.18)	1 [Reference]	1.31 (0.85-2.02)
Multivariable model 1[c]	1 [Reference]	1.49 (1.15-1.93)	1 [Reference]	1.04 (0.65-1.66)
Multivariable model 2[d]	1 [Reference]	2.08 (1.52-2.85)	1 [Reference]	1.52 (0.89-2.58)

	Normal weight	Overweight/obese
Dead	78	371
Alive	215	1961
	2332	293

$$\mathrm{OR}_{\mathrm{crude}} = \frac{78 * 1961}{371 * 215} = 1.92$$

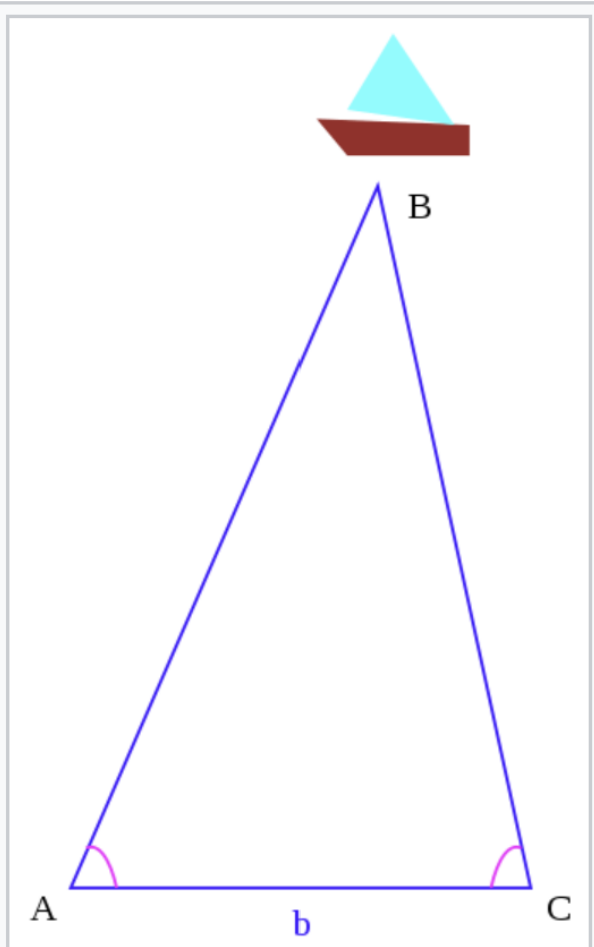


	HR Estimate from article	Crude OR Estimate	Selection bias corrected OR estimate
Normal weight vs. overweight-obese BMI	1.70 (1.33, 2.18)	1.92 (1.45, 2.54)	1.13 (0.82, 1.57)

WRAP UP

Triangulation of evidence

- All studies have potential sources of bias
  - We should do our best to remediate and quantify potential bias
- We can increase confidence in findings through "triangulation" of evidence" from studies different strengths and weaknesses



Finding the position of a distant object B with the angles observed from points A and C and the baseline b between them

Review of what we covered today:

- Key concepts about bias and different types of bias
  - Unmeasured confounding
  - Selection bias
  - Misclassification
  - Deterministic and probabilistic QBA: hand calculations, Stata commands
- Did not cover: Bayesian bias analysis, multiple bias analysis, more complex scenarios (e.g., polytomous unmeasured confounders, effect measure modification with unmeasured confounding)

To download the Part 2 simulation tutorials from GitHub…

1. Navigate to github.com/ermayeda

2. Click on the link to the Monte_Carlo_simulation_tutorials repo

3. Click clone or download and choose Download ZIP

The whole repo is now in your computer's downloads folder