

# 500 Class 2 Slides

[github.com/THOMASELOVE/2019-500](https://github.com/THOMASELOVE/2019-500)

2019-02-07

# Agenda for Class 02

- A Motivating Example: Aspirin and Mortality in Heart Patients
  - How can we avoid being misled?
  - Causal Effects as comparing potential outcomes
- The Hormone Replacement Therapy Story
  - Can Selection Bias Explain the Story?
- Tools for Assessing Causal Effects
  - Subclassification and Cochran's Example
- Homework 1: The DIG Trial and Logistic Regression
- Rosenbaum (2017) Chapters 1-4 (Part I. Randomized Experiments)
  - ① A Randomized Trial
  - ② Structure
  - ③ Causal Inference in Randomized Experiments
  - ④ Irrationality and Polio
- Defining and Motivating the Propensity Score
- Using Matched Sets to Adjust for Overt Bias

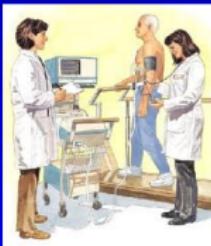
## A Motivating Example (Aspirin and Mortality)

# Aspirin and Mortality in Heart Patients

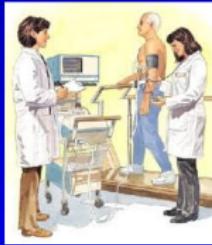
Suppose you want to understand the effect of aspirin (acetylsalicylic acid: ASA) on mortality among patients undergoing stress echocardiography.

- What is the population?
- What is the outcome?
- What are the treatments?

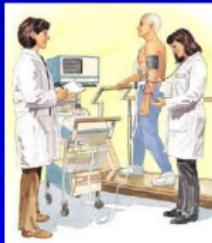
# Potential Outcomes and The Aspirin Study



# Potential Outcomes and The Aspirin Study



# Potential Outcomes and The Aspirin Study



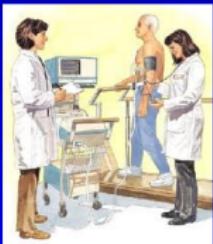
This space  
intentionally  
left blank.

# ASA and Mortality in Heart Patients

Suppose you want to understand aspirin's effect on all-cause five-year mortality among patients undergoing stress echocardiography.

- Comparing ASA to “No ASA”
- What are the potential outcomes here?

# Potential Outcomes and The Aspirin Study

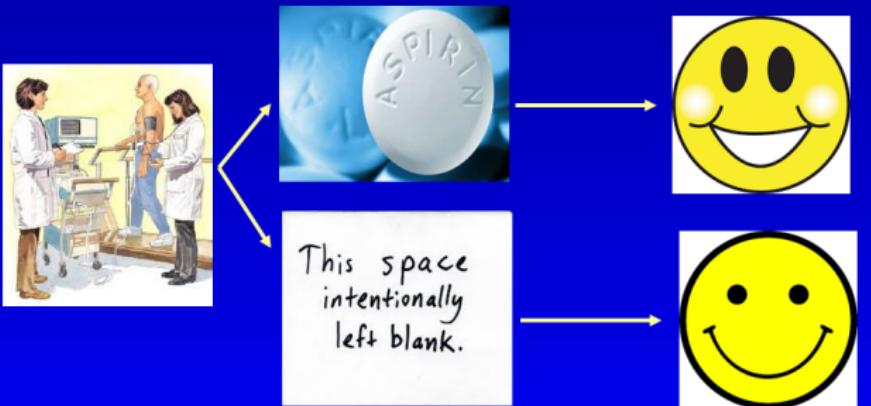


This space  
intentionally  
left blank.

# Potential Outcomes and The Aspirin Study



# Potential Outcomes and The Aspirin Study



Aspirin – “No Aspirin” Effect =

$$\text{Smiley Face} - \text{Smiley Face}$$

# ASA and Mortality in Heart Subjects

- Suppose you want to study the effect of aspirin (acetylsalicylic acid: ASA) on all-cause mortality.
- You identify an interesting group of Subjects as those undergoing stress echocardiography.
  - Your goal is to compare ASA Subjects to “no ASA” Subjects

What would be the **ideal** study?

Step 1. Identify a large group of Subjects from the population at Time 0.

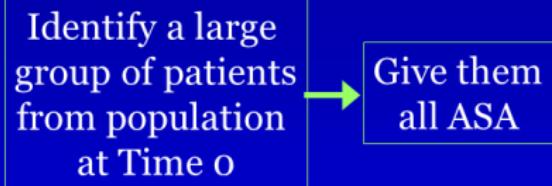
- We want to understand the causal effect of aspirin on all-cause five-year mortality among patients undergoing stress echocardiography.
- Having identified a set of patients, what is the ideal study?

Step 2?

# ASA and Mortality: Ideal Study

Identify a large group of patients from population at Time 0

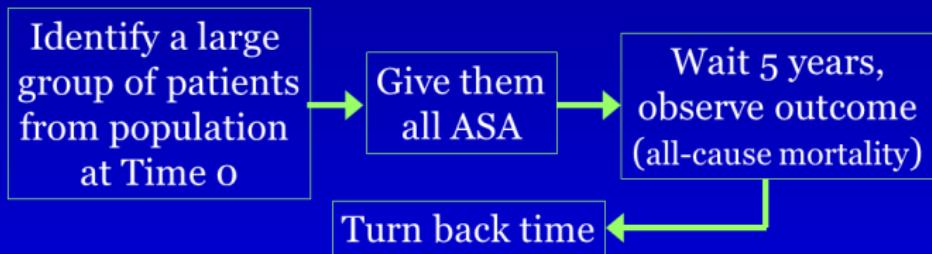
# ASA and Mortality: Ideal Study



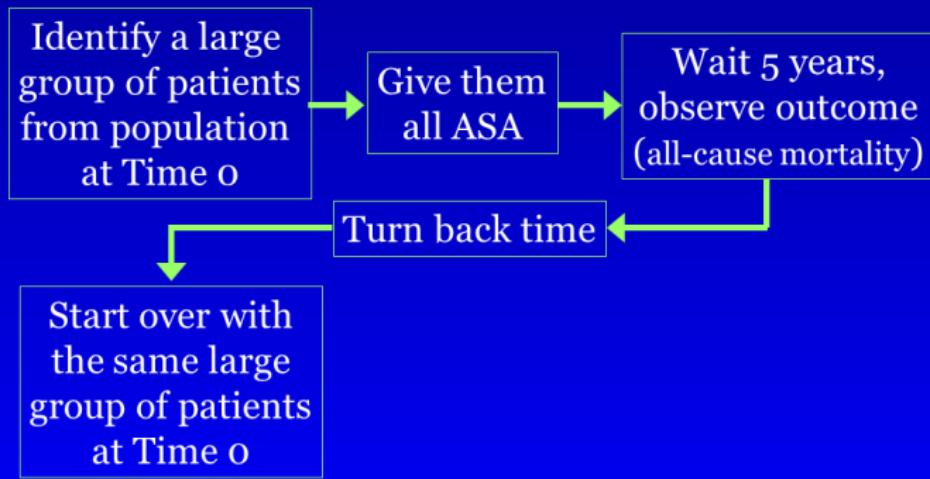
# ASA and Mortality: Ideal Study



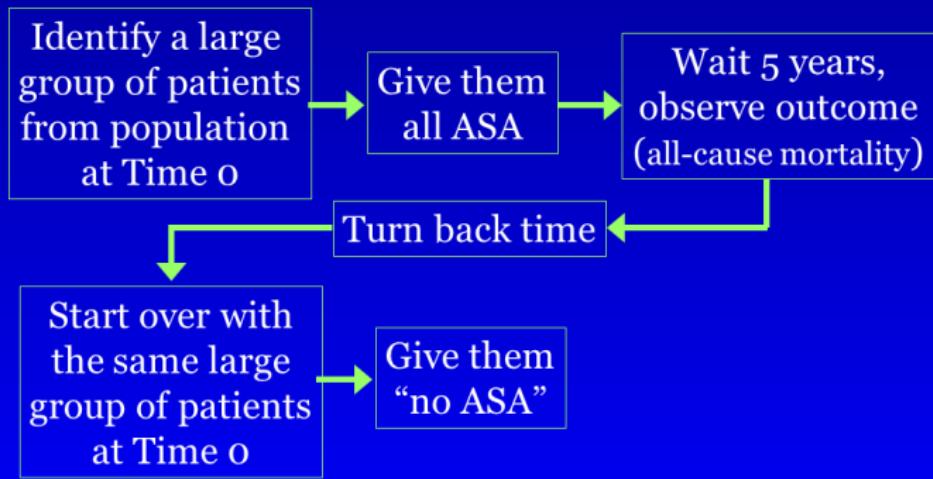
# ASA and Mortality: Ideal Study



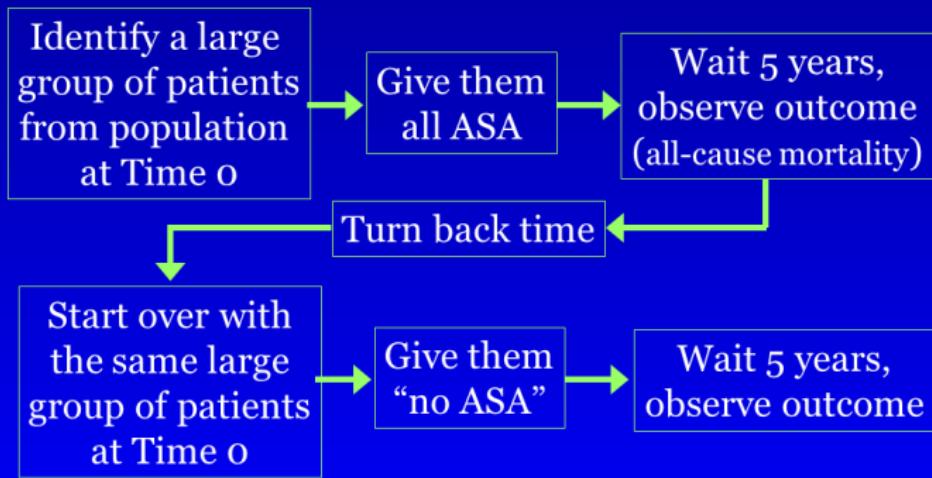
# ASA and Mortality: Ideal Study



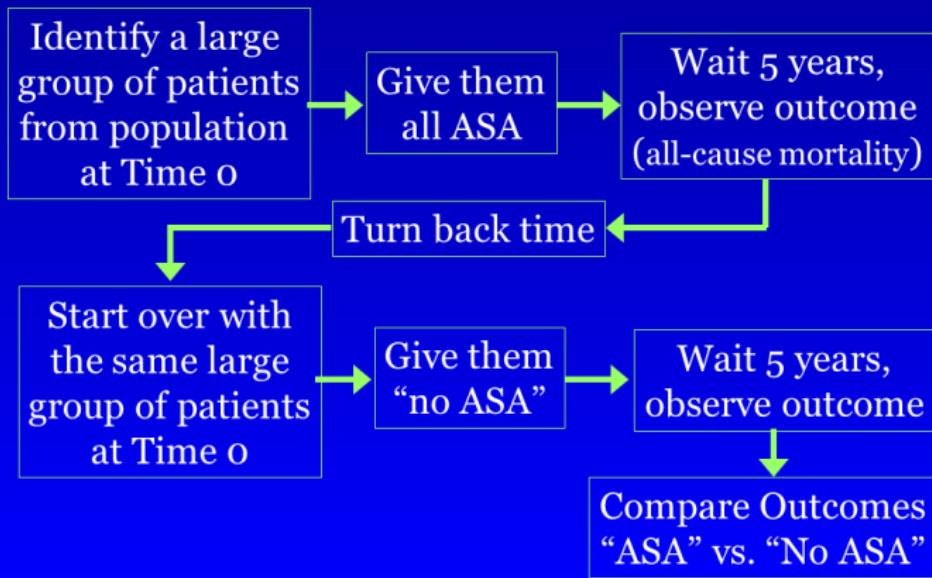
# ASA and Mortality: Ideal Study



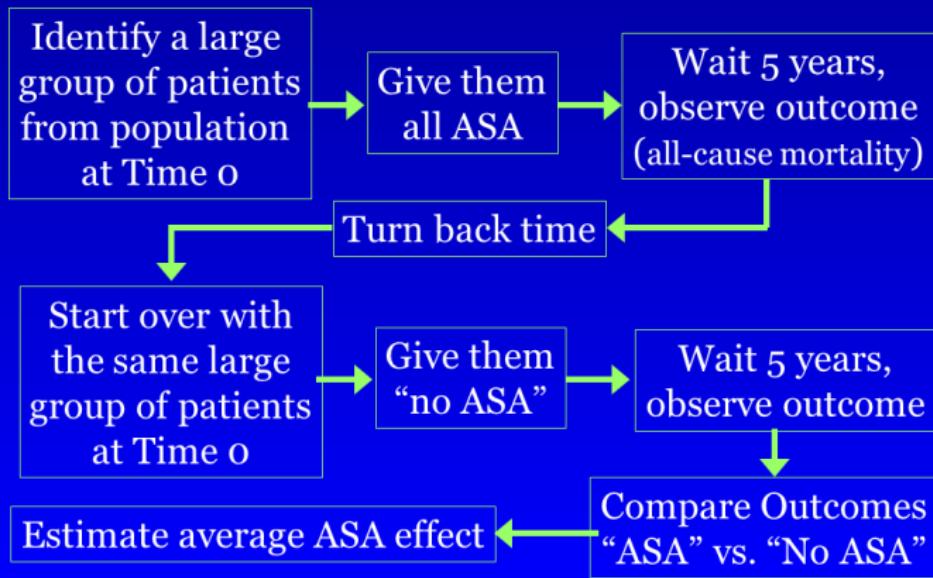
# ASA and Mortality: Ideal Study



# ASA and Mortality: Ideal Study



# ASA and Mortality: Ideal Study



# ASA & Mortality: Next Best Study

Identify a large group of patients from population at Time 0

# ASA & Mortality: Next Best Study

Identify a large group of patients from population at Time 0

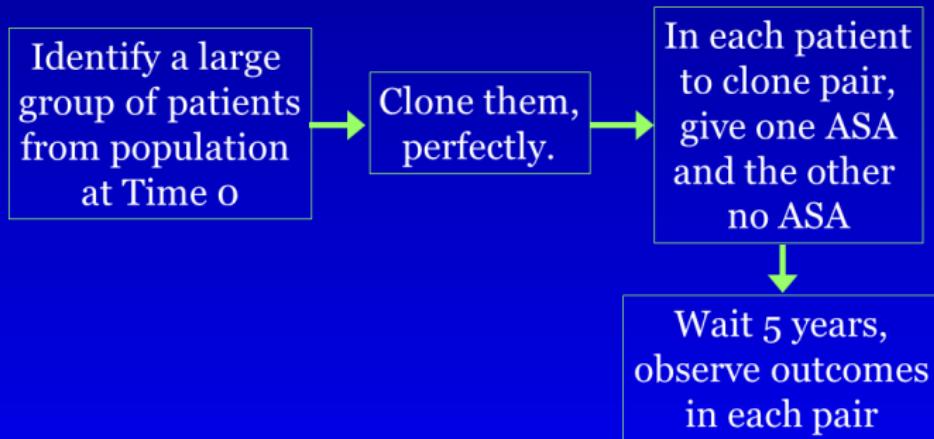


Clone them, perfectly.

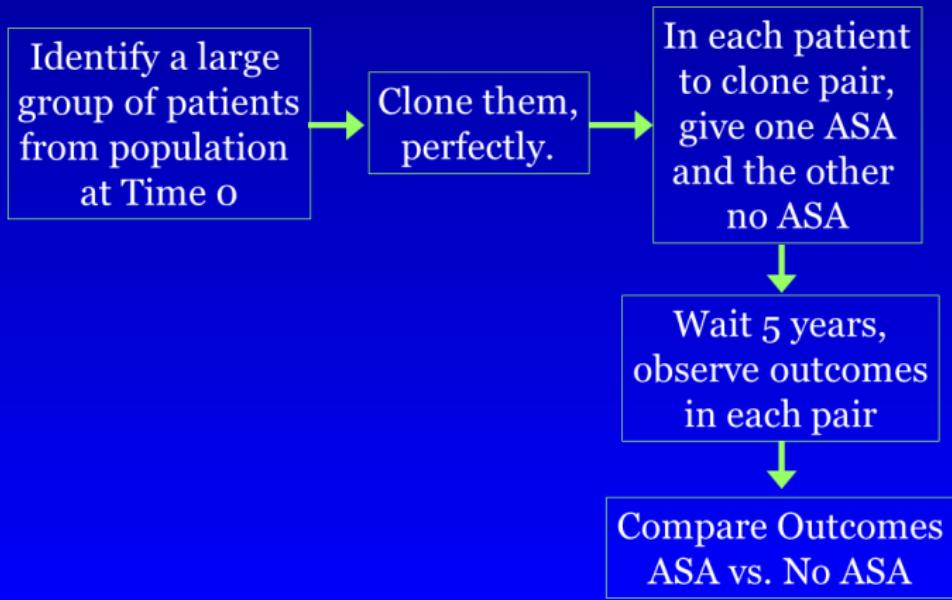
# ASA & Mortality: Next Best Study



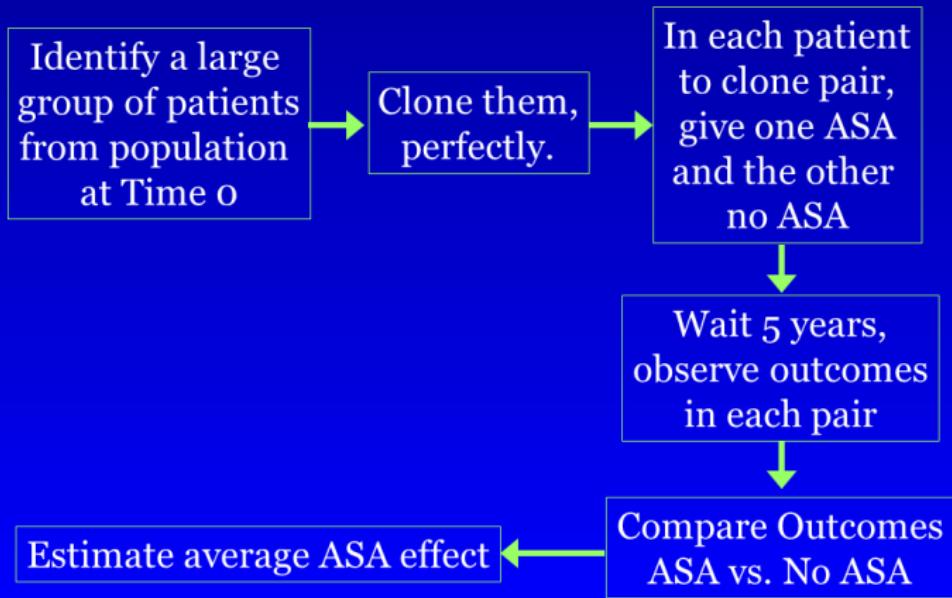
# ASA & Mortality: Next Best Study



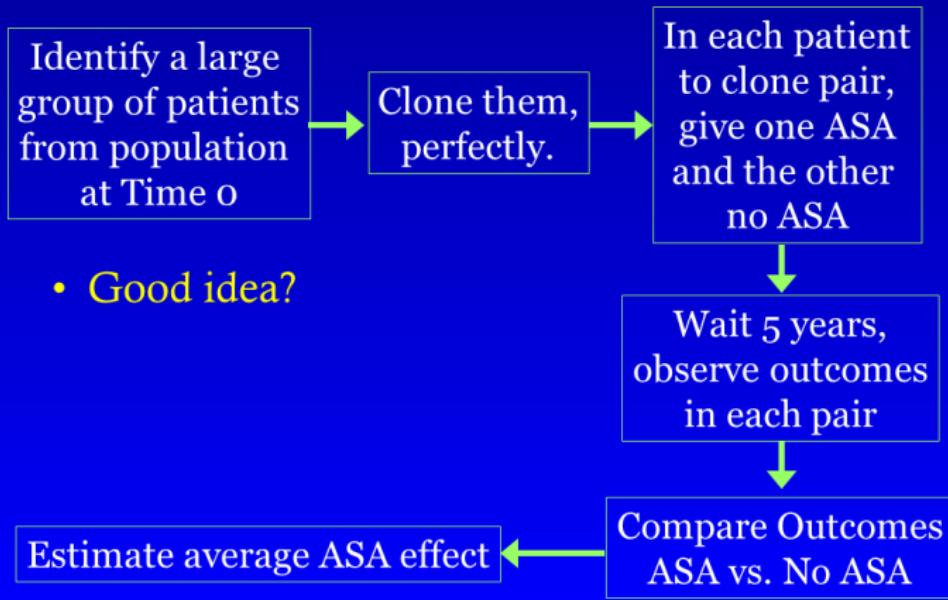
# ASA & Mortality: Next Best Study



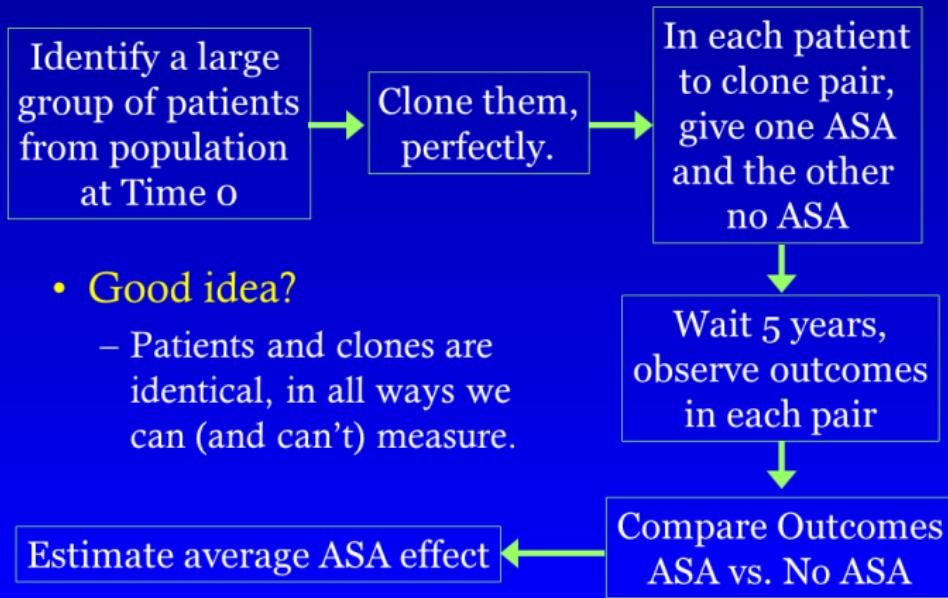
# ASA & Mortality: Next Best Study



# ASA & Mortality: Next Best Study



# ASA & Mortality: Next Best Study



# ASA and Mortality in Heart Patients

- Designing the Study

We want to understand aspirin's effect on all-cause five-year mortality among patients undergoing stress echocardiography.

- OK.
- What's the best **practical** study?

# ASA & Mortality: RCT

Identify a large group of patients from population at Time 0

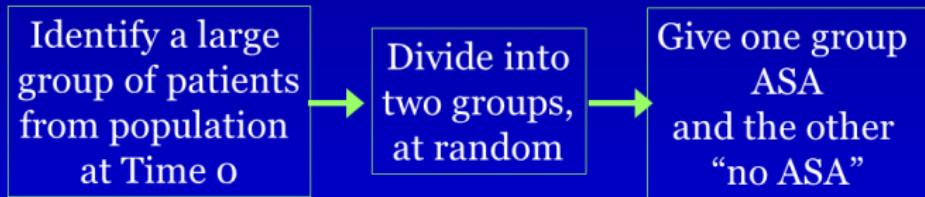
## ASA & Mortality: RCT

Identify a large group of patients from population at Time 0

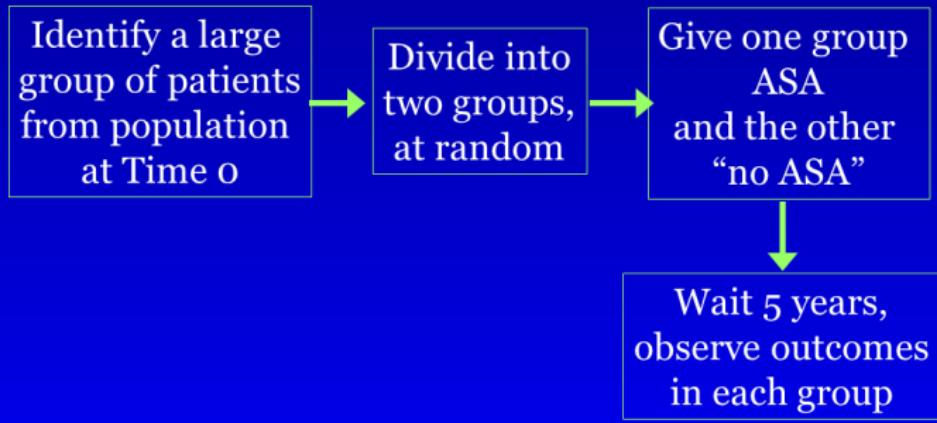


Divide into two groups, at random

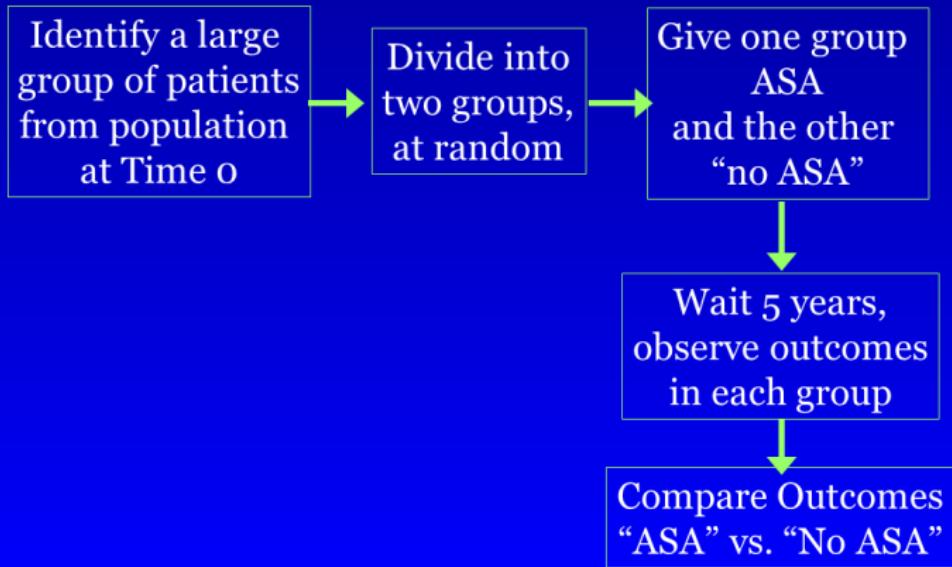
# ASA & Mortality: RCT



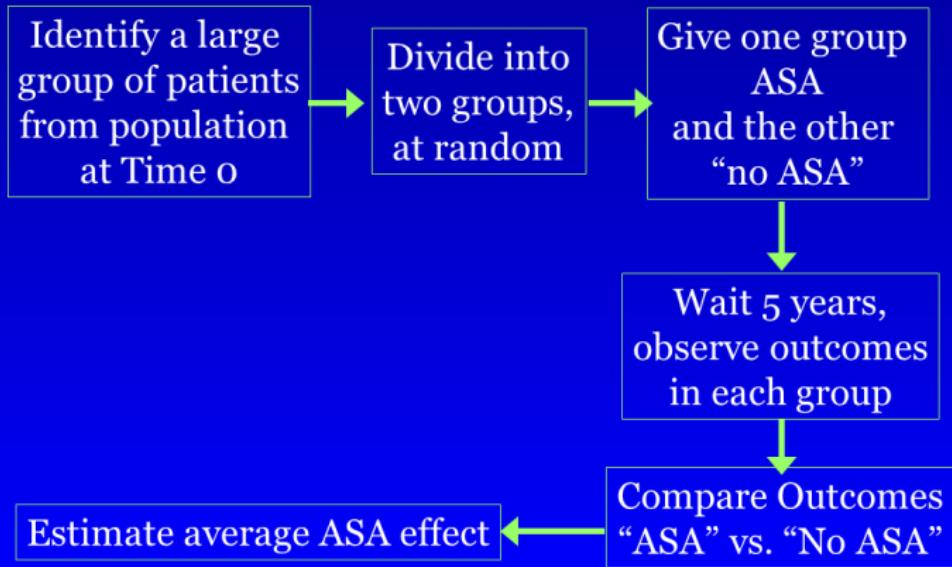
# ASA & Mortality: RCT



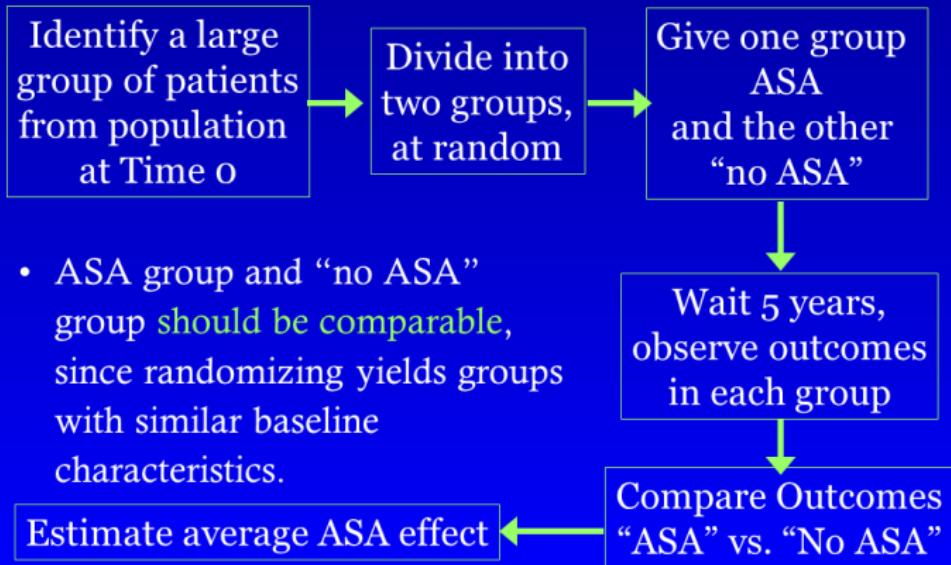
# ASA & Mortality: RCT



# ASA & Mortality: RCT



# ASA & Mortality: RCT



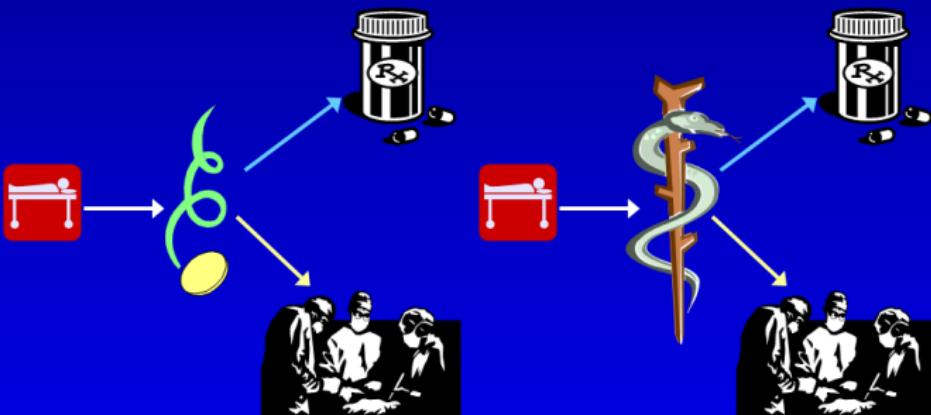
# ASA and Mortality in Heart Patients

- Designing the Study

We want to understand aspirin's effect on all-cause five-year mortality among patients undergoing stress echocardiography.

- But what if we **cannot** do an RCT?

## Randomized vs. Observational Studies



Randomization ensures  
that subjects receiving  
different treatments  
are comparable.

In observational studies,  
the researcher  
does not randomly  
allocate the treatments.

# ASA & Mortality: Observational Study

Identify a large group of patients from population at Time 0

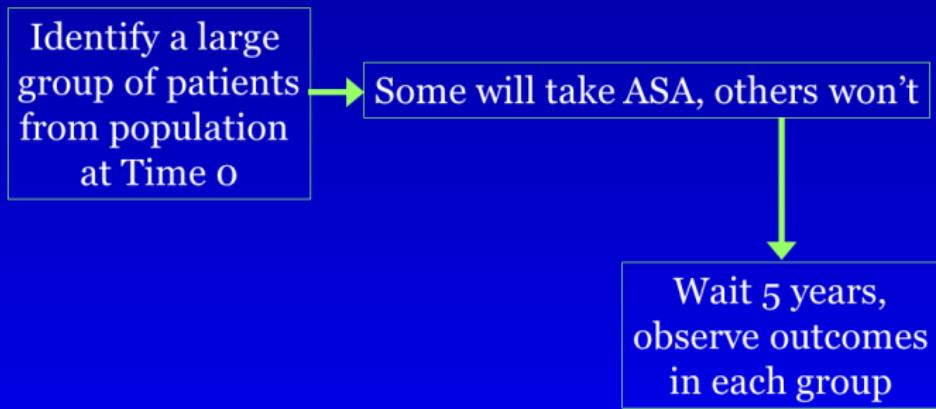
## ASA & Mortality: Observational Study

Identify a large group of patients from population at Time 0

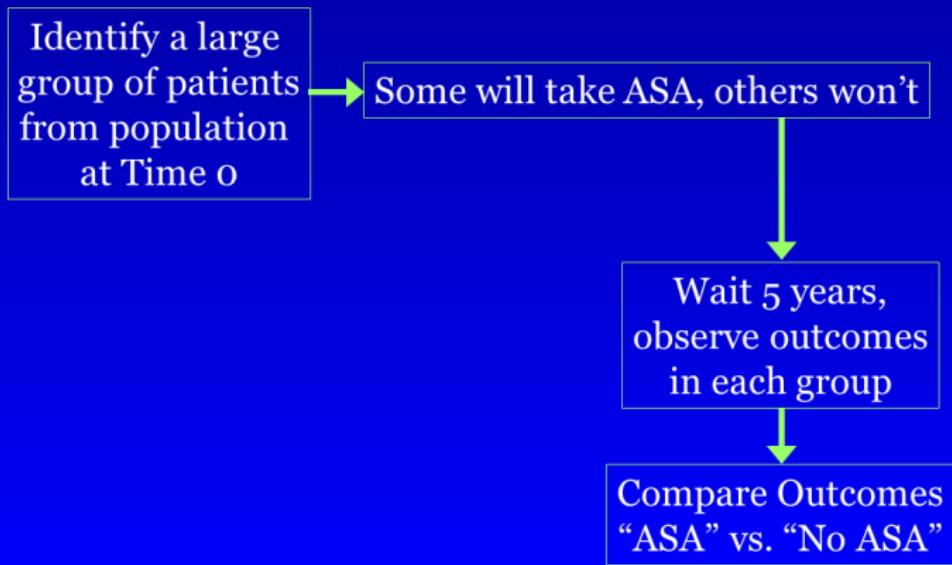


Some will take ASA, others won't

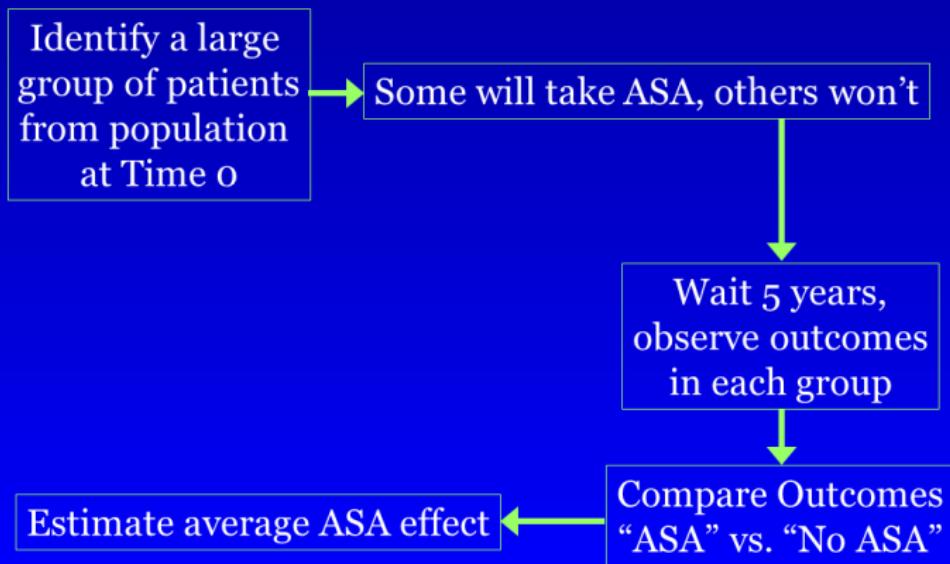
## ASA & Mortality: Observational Study



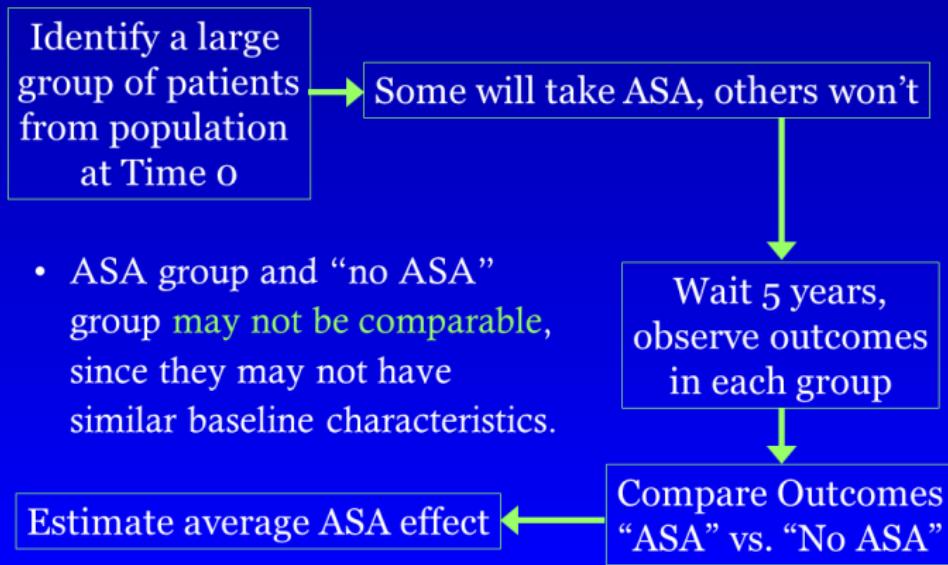
## ASA & Mortality: Observational Study



## ASA & Mortality: Observational Study



## ASA & Mortality: Observational Study



# How Do We Avoid Being Misled by Observational Studies?

- What differentiates an observational study from a randomized controlled trial?
  - One key element: potential for selection bias.
- What is selection bias and what can we do about it?
  - Baseline characteristics of comparison groups are different in ways that affect the outcome.

We will often distinguish between overt and hidden bias.

- Overt Bias (seen in data - propensity scores can help)
- Hidden Bias (required data not collected - requires sensitivity analyses)

## Aspirin Use and Mortality - The Real Study

6174 consecutive adults at CCF undergoing stress echocardiography for evaluation of known or suspected coronary disease<sup>1</sup>.

- 2310 (37%) were taking aspirin (treatment).
- Main Outcome: all-cause mortality
  - Median follow-up: 3.1 years
- Univariate Analysis: 4.5% of aspirin patients died, and 4.5% of non-aspirin patients died.
  - Unadjusted Hazard Ratio: 1.08 (0.85, 1.39)

More on this study to come.

---

<sup>1</sup>Gum PA et al. 2001

# The Hormone Replacement Therapy Story

# Testing out Cause and Effect: Comparing Potential Outcomes

- The causal effect of a treatment is based on a comparison of two potential outcomes.
  - Outcome patient would have if treated.
  - Outcome patient would have if untreated.
- Causal effect = Treated - Untreated difference (or ratio, or whatever)

The key problem is that we only get to observe **one** of these outcomes.

# Hormone Replacement Therapy and Dementia

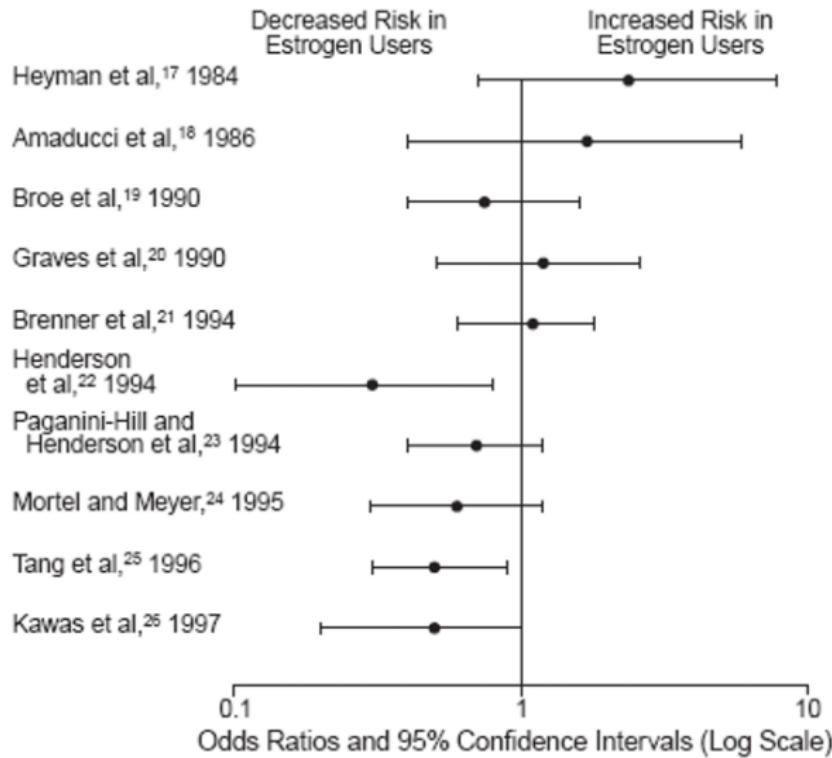
1998 Meta-Analysis<sup>2</sup>

- Estrogen associated with a 29% decreased risk of dementia
- Promising results for Alzheimer's disease (see next slide)

---

<sup>2</sup>Yaffe et al 1998

# 1998 Meta-Analysis of Estrogen & Alzheimer's Disease



# Hormone Replacement Therapy and Dementia

- 2001: “Estrogen and HRT users have . . . a 20% to 60% reduction in the risk of Alzheimer’s disease.”

# Hormone Replacement Therapy and Dementia

BUT

- These studies discussed in the 1990s are, for the most part,
  - small in size
  - short in duration
  - non-randomized,
  - and uncontrolled.
- The largest and most methodologically sound observational study (Barret-Connor et al. 1993 JAMA) of the effect of estrogen use on *cognition* in non-demented women showed no benefit.

# Hormone Replacement Therapy and Dementia

- 2001: “Estrogen and HRT users have . . . a 20% to 60% reduction in the risk of Alzheimer’s disease.”<sup>3</sup>
- 2005: “Estrogen with or without progestin, given to women 65 years and older . . . substantially increases the risk of dementia of any cause and cognitive decline.”<sup>4</sup>

---

<sup>3</sup>Burkman et al. August 2001 Am J Obstet Gynecol

<sup>4</sup>Craig et al. March 2005 Lancet Neurol

# Cache County Memory Study

Prospective study of incident dementia<sup>5</sup>

- This was a prospective study of incident dementia among 1357 men and 1889 women residing in a single county in Utah. Patients were first assessed in 1995-97, with follow-up 3 years later.
- Adjustments in models included terms for age and age squared, years of education, and presence of 1 or 2 APOE  $\epsilon 4$  alleles, and interactions.

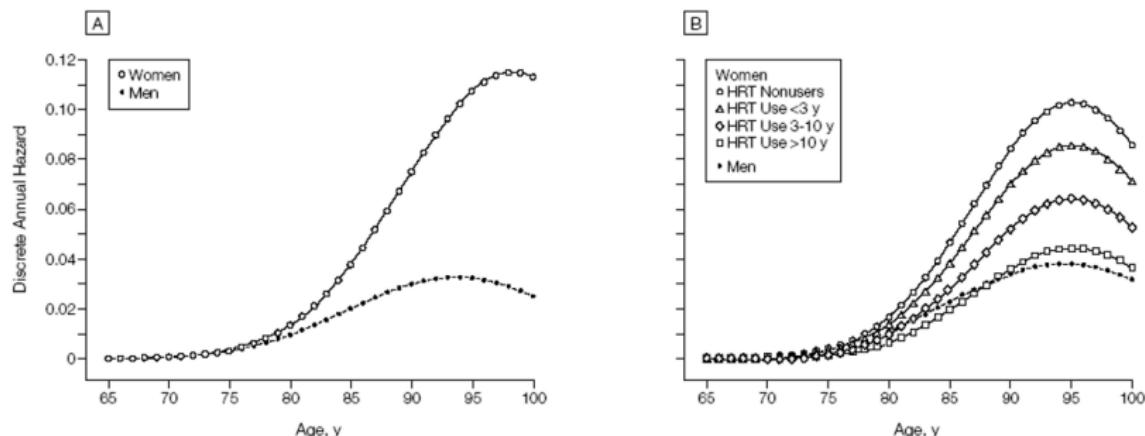
Figure on next slide...

---

<sup>5</sup>Zandi et al. 2002

# Cache County Memory Study

**Figure 2.** Estimated Discrete Annual Hazard of Alzheimer Disease for Men and Women by Age, and by Duration of Hormone Replacement Therapy Use for Women



Both figures indicate risks estimated for an individual with the mean value of 13 years of education and no  $\epsilon 4$  alleles at APOE. A, The curves depict the annual hazards predicted by fitting the base model including an age-by-sex interaction term. The annual hazard for Alzheimer disease (AD) appears similar for men and women before 80 years of age but diverges rapidly afterward with an excess risk found in women. B, The curves depict the annual hazards predicted by fitting model 7 of Table 3 to the women with available hormone replacement therapy (HRT) exposure information and, in filled circles, the corresponding annual hazards for men after omitting the terms for HRT. There were 35 instances of incident AD among 1357 men. Ordinate values for women differ slightly from those in panel A due to omission of women lacking HRT exposure information, several of whom experienced incident dementia.

## Conclusions from Cache County

- Women using HRT had a reduced risk of AD compared with non-HRT users (adjusted HR is 0.59).
- Risk varied with duration of HRT use, so that a woman's sex-specific increase in risk disappeared entirely with more than 10 years of treatment.
- Conclusions: Prior HRT use is associated with reduced risk of AD, but there is no apparent benefit with current HRT use unless such use has exceeded 10 years.

# WHIMS (Women's Health Initiative Memory Study)

Randomized, controlled trial, reported in 2003<sup>6</sup>

- 4,352 post-menopausal women age 65 or more
- Estrogen + Progestin HRT
  - increased risk (hazard ratio 2.05) for probable dementia
  - treating 434 women age 65+ with combination HRT would cause one new dementia case.
  - NS impact on mild cognitive impairment

---

<sup>6</sup>Shumaker et al 2003, Shumaker et al 2004, Espeland et al 2004

# WHIMS Baseline Comparisons

No significant baseline differences between the two arms of the trial in

- Age, Education,
- Smoking, Diabetes, Prior HRT or Aspirin use,
- or 3MSE score.

Significant differences (comparing E & P to placebo) in

- History of Stroke (1.0% vs. 1.9),
- Statin use (12.0 vs. 9.8), and
- Adherence (E & P less than Placebo)

# HRT and Cardiovascular Disease

- Stampfer et al 1985 [Nurses' Health Study] . . . "estrogen reduces the risk of severe CHD."

# HRT and Cardiovascular Disease

- Stampfer et al 1985 [Nurses' Health Study] . . . "estrogen reduces the risk of severe CHD."
- Col et al 1997 (JAMA) . . . "HRT should increase life expectancy for nearly all postmenopausal women"

# HRT and Cardiovascular Disease

- Stampfer et al 1985 [Nurses' Health Study] . . . "estrogen reduces the risk of severe CHD."
- Col et al 1997 (JAMA) . . . "HRT should increase life expectancy for nearly all postmenopausal women"
- WHI trial 2002 (JAMA) . . . "(HRT) should not be initiated or continued for primary prevention of coronary heart disease."

# Selection Bias?

Comparing NHS (OS) to WHI (RCT)

- Healthy User Effect
  - Women with healthy behaviors may select to use postmenopausal hormones. (prevention bias)
- In the NHS,
- HRT users tended to have better CV risk profiles
- HRT users were generally better educated
- Perhaps women taking HRT / ERT were “compliant” and such people have lower CHD risk.
- HRT users have more contact with physicians, and are perhaps more health conscious, generally.

# So... How Can We Avoid Being Misled?

- ① What differentiates an observational study from a randomized controlled trial?
  - One key element: potential for selection bias.
- ② What is selection bias, and why should I care about it?
  - Baseline characteristics of comparison groups are different in ways that affect the outcome.
  - We often split this into **overt** bias we observe in our measures
  - As compared to **hidden** bias across measures we didn't think to observe.
- ③ What can be done to deal with selection bias in observational studies?
  - **Propensity score** methods for overt bias.
  - *Sensitivity analyses* to deal with hidden bias.

# Tools for Assessing Causal Effects

# Assessing the Causal Effect of an Exposure on an Outcome

Objective: Draw causal inferences between [use of exposure vs. non-use] and outcome

- Standard Approach: Risk Adjustment
- Problem: Selection Bias (exposed people are different from unexposed people at baseline, in ways that affect the outcome)
- Idea: Compare exposed to unexposed subjects that looked similar (had similar propensity for exposure) prior to the exposure decision

## Overt, but no Hidden Bias Model

Two units with the same value of the covariates  $\mathbf{x}$  have the same probability  $\pi$  of receiving the exposure.

- An observational study is **free of hidden bias** if the unknown  $\pi_j$ s are known to depend only on the observed covariates  $x_j$ .
- Sometimes this is referred to as “randomization based on covariates”

# How can we adjust for overt bias?

Simplest approach: stratify on the covariates  $x$

- Exact stratification - two units are in the same stratum only if they have the same value of  $x$ .
- If there is no hidden bias and we stratify exactly, then all units in the same stratum have the same probability of treatment, so we can use methods appropriate for a randomized experiment.

# A Simple Survival Comparison

Status at 30 days	Alive	Dead	Pr(Alive)
Without Exposure	80	120	0.40
With Exposure	130	70	0.65

- Without Exposure (perhaps as estimated by historical records) only 40% of subjects survived.
- With Exposure, we see a “clinically meaningful” improvement (65% of subjects survived.)
- $p$  value from Fisher’s exact test is  $< 0.001$ .

But was this a randomized trial, or an observational study?

## Simple Survival Comparison

Suppose in addition to

- our **outcome** (Alive or Dead at 30 days)
- and **exposure** status,

we also had a **covariate**, say, sex, available for each subject. Suppose 200 of the subjects in the study are Male, and 200 are Female.

Suppose also that sex might be related to the outcome.

- Can we adjust for sex's effect in assessing the impact of our exposure on that same outcome? How?

# Stratification in our Survival Comparison

ALL PATIENTS	Alive	Dead	Pr(Alive)
Without Exposure	80	120	0.40
With Exposure	130	70	0.65

Now, 200 of these subjects are Male, and 200 are Female.

# Survival Comparison among Male Subjects

MALE	Alive	Dead	Pr(Alive)
Without Exposure	40	60	0.40
With Exposure	40	60	0.40

No difference between the exposed and unexposed group in terms of survival, among males. Is that also the story for our female subjects?

## Survival Comparison among Female Subjects

MALE	Alive	Dead	Pr(Alive)
Without Exposure	40	60	0.40
With Exposure	40	60	0.40

FEMALE	Alive	Dead	Pr(Alive)
Without Exposure	40	60	0.40
With Exposure	90	10	0.90

Stratification allows comparison adjusting for sex.

# Cochran's Smoking Example

## Cochran's Smoking Example

Outcome: mortality rates of US male [1] cigarette smokers, [2] cigar/pipe smokers and [3] non-smokers

US Death Rates per 1,000 person-years

Smoking Group	Unadjusted Death Rate
Non-Smokers	20.2
Cigarettes only	20.5
Cigars, pipes	35.5

## Cochran's Smoking Example

Outcome: mortality rates of US male [1] cigarette smokers, [2] cigar/pipe smokers and [3] non-smokers

Let's look at an important covariate - any suggestions?

## Cochran: US Death Rates per 1000 person-years

Smoking Group	Mean Age in Years	Unadjusted Death Rate
Non-Smokers	54.9	20.2
Cigarettes only	50.5	20.5
Cigars, pipes	65.9	35.5

Now, how could we adjust for the impact of age on our estimates of the death rate?

## Subclassification on Age (3 subclasses)

Divide the subjects into 3 subclasses on the basis of age (low, middle, high)

- ① Calculate “low age” mortality rate in each smoking group
- ② Then calculate “middle age” and finally “high age” mortality rate in each smoking group
- ③ Within the “non-smokers” combine the “low”, “middle” and “high” mortality rate estimates by weighting according to the population proportions of “low”, “middle” and “high” age.
- ④ Repeat to obtain estimates for “cigarettes only” and “cigars, pipes”

## Cochran: US Death Rates per 1000 person-years

Smoking Group	Mean Age	Unadjusted Death Rate	Adjusted for Age
Non-Smokers	54.9	20.2	20.3
Cigarettes only	50.5	20.5	28.3
Cigars, pipes	65.9	35.5	21.2

## Cochran's Smoking Example

Outcome: mortality rates of US male [1] cigarette smokers, [2] cigar/pipe smokers and [3] non-smokers

- Subclassification on the covariate **age**
- Key Finding: Five subclasses are often sufficient to remove over 90% of the bias due to the subclassifying variable or covariate<sup>7</sup>.
  - Even as few as 2 or 3 subclasses can have a big impact.

---

<sup>7</sup>Cochran WG 1968 \*Biometrics\* 24, 205-213.

## Why can't we always just do direct adjustment like this for covariates?

- Because we don't (typically) have only one covariate.
- As the number of covariates increases, the number of subclasses grows exponentially
  - 2 categories for each of  $p$  covariates yields  $2^P$  subclasses.
- Also, if  $p$  is large, some subclasses will contain no units, or will contain only exposed or unexposed units but not both.

A solution? Propensity scores.

# Homework 1

# Homework 1

## Task 2: The DIG trial

- ① What comparison do you want to make? (And what comparison did the DIG trial want to make?)
- ② Why is this of interest? What (direction of) effect is hypothesized?
- ③ What are the key measures to help address your question of interest?
  - The exposure/treatment (and how will you confirm receipt)
  - The primary outcome (and what type of variable)
  - The important covariates (related to exposure or to outcome, measured prior to exposure)

## Task 3: Fitting a Logistic Regression Model

See posted answer sketch. There will be more opportunities to demonstrate this set of skills.

# Rosenbaum, Chapters 1-4

# Rosenbaum, Part I

- ① A Randomized Trial
  - ② Structure
  - ③ Causal Inference in Randomized Experiments
  - ④ Irrationality and Polio
- 
- What was the most **important** thing you learned from reading these chapters?
  - What was the **muddiest**, least clear thing that arose in your reading?
  - What questions are at the front of your mind now?

## Aspirin Use and Mortality (Gum 2001)

6174 consecutive adults at CCF undergoing stress echocardiography for evaluation of known or suspected coronary disease<sup>8</sup>.

- 2310 (37%) were taking aspirin (treatment).
- Main Outcome: all-cause mortality
  - Median follow-up: 3.1 years
- Univariate Analysis: 4.5% of aspirin patients died, and 4.5% of non-aspirin patients died.
  - Unadjusted Hazard Ratio: 1.08 (0.85, 1.39)

---

<sup>8</sup>Gum PA et al. 2001

# Matching on the Covariates, X

- We can create a **matched sample**, where we match treated subjects to control subjects, on the basis of their covariates.
  - Simplest is exact matching - but this can pose problems unless we have few covariates to deal with, with very limited possible values.
  - Often exact stratification or matching is impossible, but when it is, things go smoothly.

## What's the difference between Aspirin Users and the other patients?

Variable	Aspirin Users	No Aspirin
Patients	2,310	3,864
Age, Mean (SD)	62 (11)	56 (12)
Male, %	77.0	56.1

Might it be reasonable to match up patients who are the same gender and similar in age? Or to stratify into groups by age and gender?

# What's the difference between Aspirin Users and the other patients?

Variable	Aspirin Users	No Aspirin
Patients	2,310	3,864
Age, Mean (SD)	62 (11)	56 (12)
Male, %	77.0	56.1
Prior CAD, %	69.7	20.1
Beta Blocker, %	35.1	14.2

But now what do we do?

- How can we match on Age **and** Gender **and** history of CAD **and** beta-blocker prescription?
- Or (if that's not hard enough) how about the complete set of 31 covariates?

# Using Matched Sets or Strata to Adjust for Overt Selection Bias

- Observe a set of  $p$  covariates, collected in  $\mathbf{X}$
- Even if each covariate is binary, there are  $2^p$  possible values of  $\mathbf{X}$ 
  - Many subjects are likely to have unique values of  $\mathbf{X}$ .
- Realistic Goal: compare treated and control groups with similar distributions of  $\mathbf{X}$ , even if matched individuals have differing values of  $\mathbf{X}$

Key tool for doing this well: propensity score

# What Do We Want to Know about a Clinical or Health System Intervention<sup>9</sup>?

- Response: Can we estimate the impact of the intervention? Can we estimate costs and benefits?
- Predictors: Can we “mine” for attributes that help predict response to the intervention?
- Evaluation: Can we fairly estimate the average health impact of our intervention?
- Target Evaluation: Can we identify likely responders? Subgroup analyses?

---

<sup>9</sup>from a marketing list at [www.anabus.com](http://www.anabus.com)

# The Data You Wish You Had

Subject	Health if exposed	if unexposed
A	12	8
B	7	4
C	7	3
D	12	9

**ALL** potential outcomes available!

# The Data You Wish You Had

Subject	Health if exposed	if unexposed	Exposure Effect
A	12	8	4
B	7	4	3
C	7	3	4
D	12	9	3

Wouldn't this be great!

# Grim Reality

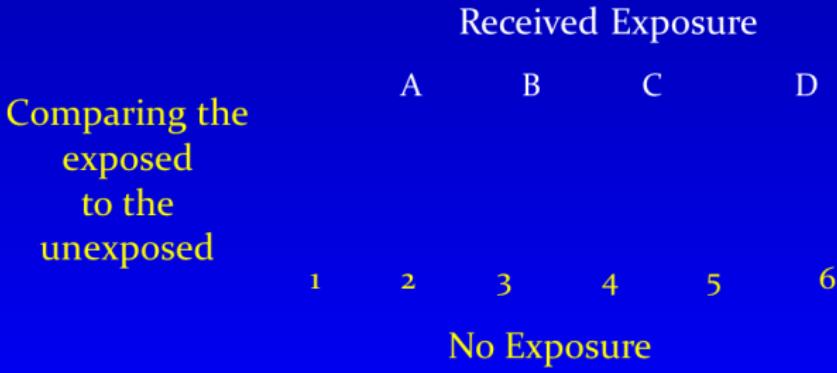
Subject	Health if exposed	if unexposed	Exposure Effect
A	12	?	?
B	7	?	?
C	?	3	?
D	?	9	?

Causal inference is a **missing data** problem.

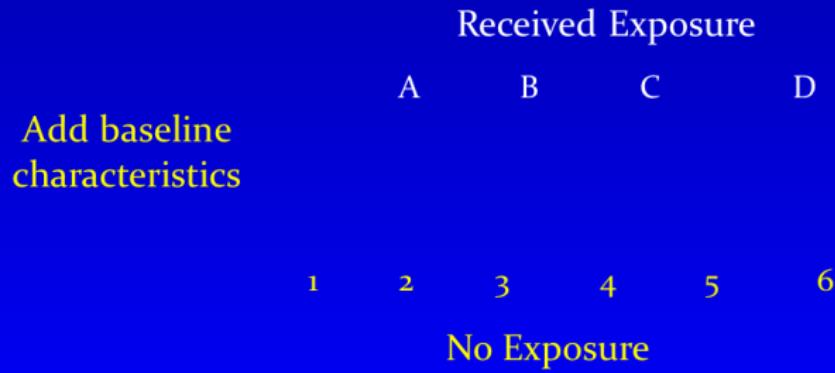
How should we fill in those question marks?

# Matching and Causal Effects

# Causal Analysis



# Causal Analysis



# Causal Analysis

Received Exposure



A

B

C

D

Add baseline  
characteristics

1

2

3

4

5

6

No Exposure

# Causal Analysis

Received Exposure



A

B

C

D

Add baseline  
characteristics

1

2

3

4

5

6

No Exposure



# Causal Analysis

Received Exposure



Add baseline  
characteristics

1

2

3

4

5

6

No Exposure



# Causal Analysis

Received Exposure



A

B

C

D

Add baseline  
characteristics

1

2

3

4

5

6

No Exposure



# Causal Analysis

Received Exposure



A

B

C

D

Which pairs  
should we  
compare?

1

2

3

4

5

6

No Exposure



# Causal Analysis

Received Exposure



Which pairs  
should we  
compare?



B

C

D

1

2

3

4

5

6

No Exposure



# Causal Analysis

## “Comparing Apples to Apples”

Received Exposure



Which pairs  
should we  
compare?



1

2

3

4

5

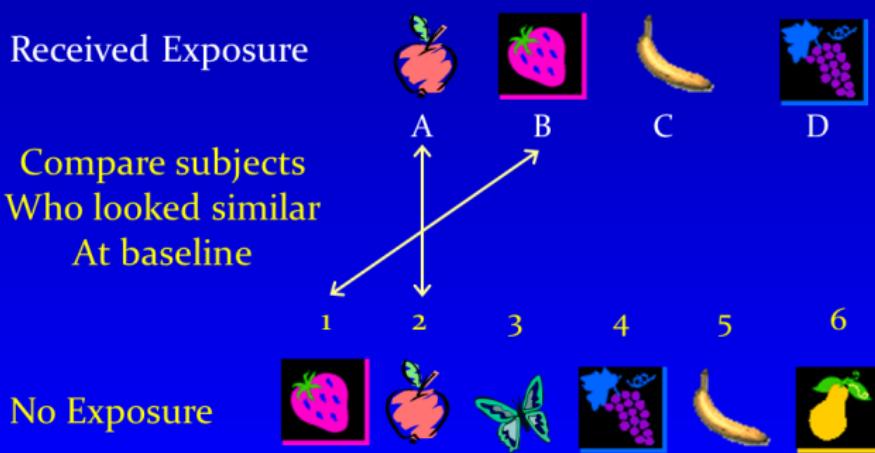
6

No Exposure



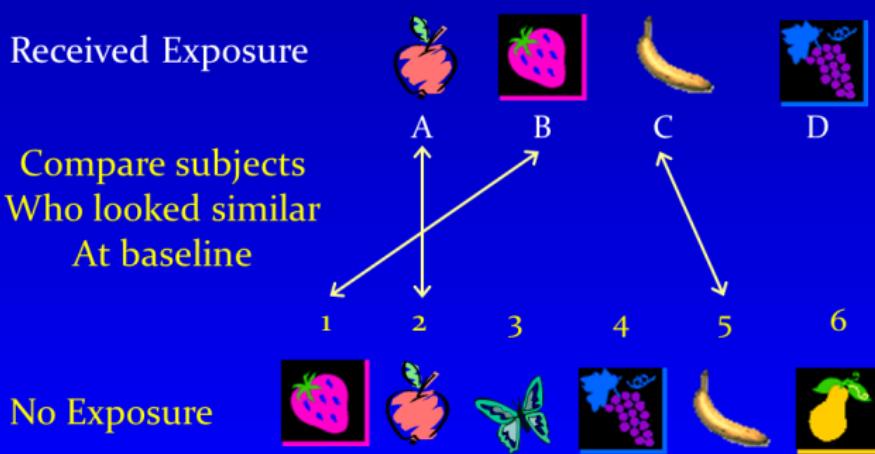
# Causal Analysis

## “Comparing Apples to Apples”



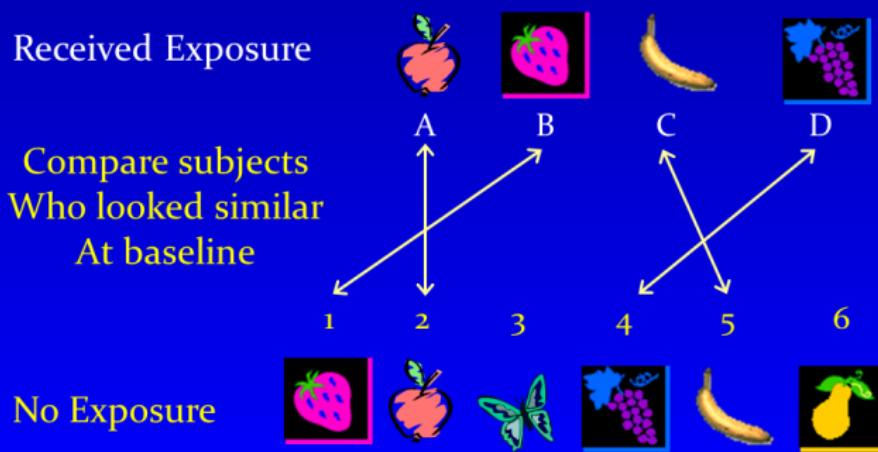
# Causal Analysis

## “Comparing Apples to Apples”



# Causal Analysis

## “Comparing Apples to Apples”



# The Propensity Score

# The Propensity Score

Definition: The conditional probability that a subject receives an exposure given the values of their vector of covariates.

- $PS = \Pr(\text{ exposed} \mid \text{covariates})$

Reduces the baseline information to a single, composite summary of the covariates, between 0 and 1.

- Of course, we know whether a subject in fact either receives or doesn't receive the exposure.
- But we will estimate this probability for each subject as a convenient way of expressing the impact of covariate information on the exposure assignment decision, as a scalar value between 0 and 1.

# Estimating the Propensity Score (most common approach)

Estimate a Logistic Regression Model:

- $Y$  = Exposure Group
  - 1 = exposed, 0 = unexposed
- Predictors are the observed covariates

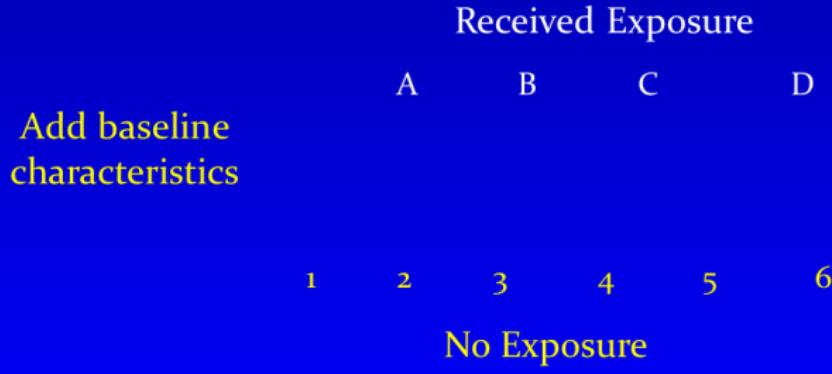
Use anything related to exposure decisions that can be collected prior to exposure assignment.

Propensity Scores = Predicted  $\text{Pr}(\text{exposure})$  for each subject, i.e. the **fitted values**

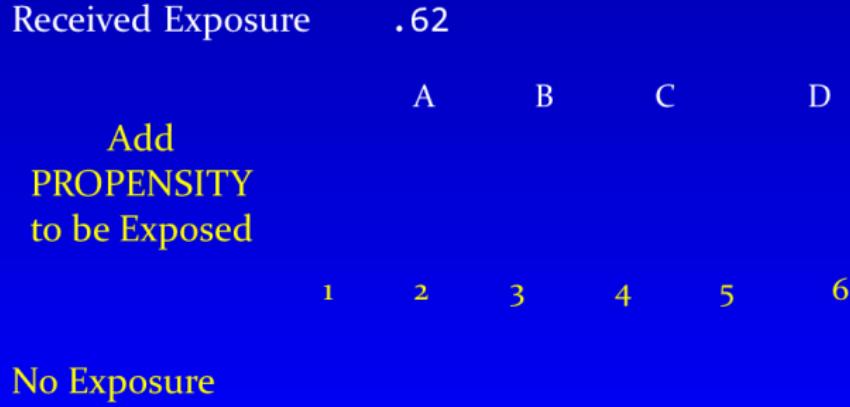
# Why Estimate the Probability that a subject was “exposed”?

- Using  $\Pr(\text{subject would have been exposed})$ , we create a quasi-randomized experiment.
- If we have two subjects, one treated and one control, with the same propensity score, we can imagine that these two subjects were randomly assigned to each group - just as if we were doing an experiment!
- Except that here we can't assume that we control for anything that we didn't measure.

# Causal Analysis



# Causal Analysis



# Causal Analysis

Received Exposure	.62	.74	.59	.81
	A	B	C	D

Add  
PROPENSITY  
to be Exposed

1	2	3	4	5	6
---	---	---	---	---	---

No Exposure

# Causal Analysis

Received Exposure						
	.62	.74	.59	.81		
	A	B	C	D		
Add PROPENSITY to be Exposed						
	1	2	3	4	5	6
	.74	.62	.36	.80	.58	.23
No Exposure						

# Causal Analysis

Received Exposure

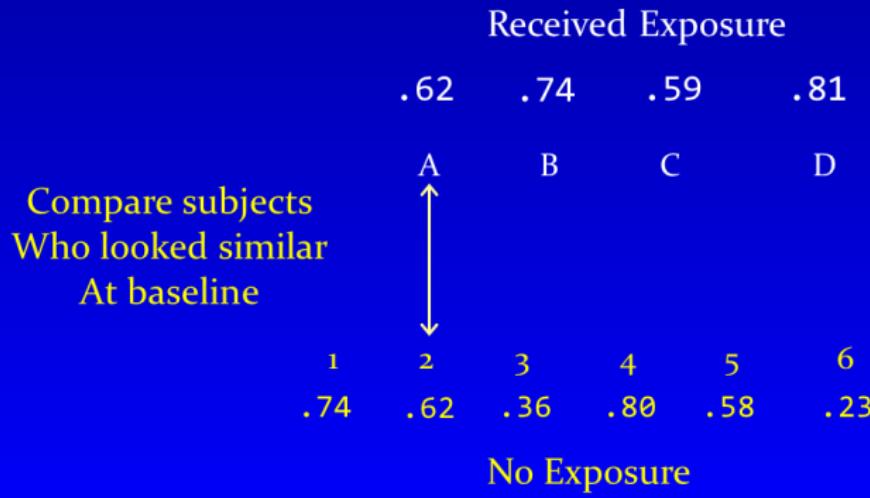
.62	.74	.59	.81
A	B	C	D

Which pairs  
should we  
compare?

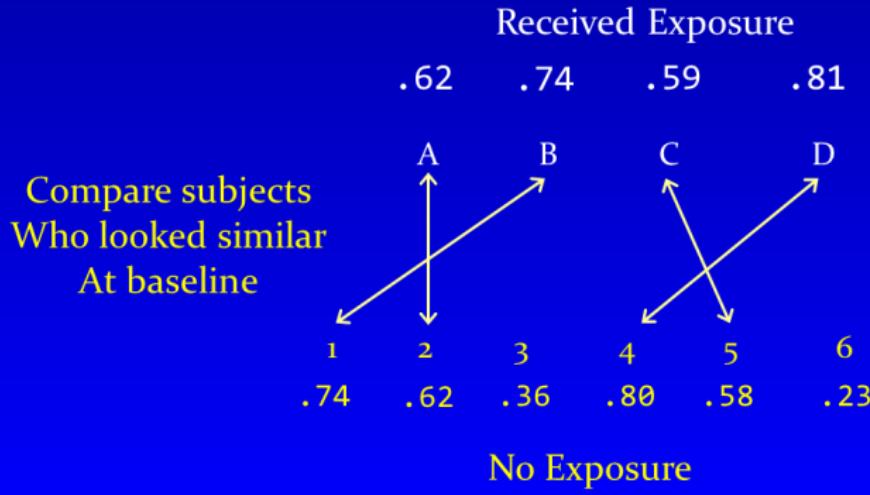
1	2	3	4	5	6
.74	.62	.36	.80	.58	.23

No Exposure

# Causal Analysis



# Causal Analysis



# Grim Reality

Subject	Health if exposed	Health if unexposed	Exposure Effect
A	12	?	?
B	7	?	?
C	?	3	?
D	?	9	?

# Improving Grim Reality

Subject	Propensity for Exposure	Health if exposed	if unexposed
A	0.80	12	?
B	0.50	7	?
C	0.51	?	3
D	0.79	?	9

- Can we use the propensity score to guide our matching approach?
- Can we plug in resulting estimates for our question marks?

## Propensity Score Matching yields a new Database

Subject	PS	Health if exposed	if unexposed	Exposure Effect
A	0.80	12	[9]	[3]
B	0.50	7	[3]	[4]
C	0.51	[7]	3	[4]
D	0.79	[12]	9	[3]

Now, we can estimate the **impact of the exposure** on each matched patient.

# How Do We Use the Propensity Score?

- ➊ Start with a sample where the exposed subjects don't look very much like the unexposed subjects.
  - ➋ Adjust the sample (in some manner) to make the distributions of exposed and unexposed subjects look more similar prior to exposure.
  - ➌ This will let us attribute the differences we see in outcomes between these adjusted samples more easily to the exposure's causal effect, and not so much to the original differences between the groups.
- 
- To do this, we estimate the propensity score: the probability of receiving the exposure for each subject given their covariate values.
  - Then, we use the propensity score in one of the ways listed on the next slide to fuel our estimates of causal effects.

# Methods for Using Propensity Scores

- Subclassification / Stratification on the Propensity Score
- Direct (Regression) Adjustment using the Propensity Score
- Matching using the Propensity Score
- Weighting using the Propensity Score
- Combining Approaches for More Robust Estimation

All of these are found in the toy example in the Data and Code section of our web site. We'll discuss this in detail in Class 04.

## Next Time (Class 03) ...

- Rosenbaum Chapters 5 and 6
  - ⑤ Between Observational Studies and Experiments
  - ⑥ Natural Experiments
- Matching and Causal Effects
- Estimating the Propensity Score
- Using the Propensity Score to account for observed selection bias