

How to estimate the effects of hypothetical interventions

(Hint: First specify the intervention)



Miguel A. Hernán
Harvard School of Public Health
www.hsph.harvard.edu/causal

What was your question again?

The perils of using observational data to answer questions other than the one you might wish to ask



Miguel A. Hernán
Harvard School of Public Health
www.hsph.harvard.edu/causal

Setting:

- We want to make inferences about the causal effect
- of some treatment (or exposure)
 - hormone therapy, lifestyle
- on some outcome
 - CHD risk
- ❖ Observational data available on treatment, outcome, and other variables

Jun-09

Hypothetical interventions

3

Causal inference from **already collected** observational data

- Smart study design is not an option
- We need to either quit or deal with whatever data we have
- A common situation in epidemiologic research
 - Perhaps even more common in social sciences
- How do we do this?

Jun-09

Hypothetical interventions

4

Causal inference from **already collected** observational data

1. Formulate a reasonably well-defined causal question
 2. Propose an answer by combining
 - Available data
 - Untestable assumptions
 - Appropriate analytic method
- Often discussions about causal inference revolve exclusively around stage #2

Jun-09

Hypothetical interventions

5

Stage 1: A reasonably well-defined causal question

- A prerequisite for meaningful causal inference
- Can be expressed in terms of
 - Hypothetical interventions
 - Counterfactual contrasts
- Choice of analytic approach follows naturally from it
 - i.e., formulation of causal question predates discussions about analytic approach

Jun-09

Hypothetical interventions

6

Stage 2: Combination of data, assumptions, and analytic approach

- Data need to be measured with reasonable accuracy and in a population similar to the target population
- Assumptions need to be consistent with expert subject-matter knowledge
- Analytic approach needs to be appropriate to answer the causal question with the above data under the above assumptions

Jun-09

Hypothetical interventions

7

Analytic approach in randomized vs. observational studies

- Randomized experiments (e.g. clinical trials):
 - analysis pre-specified before study is conducted
 - Described in study protocol
- Observational studies
 - Often ad hoc analysis after data have been collected and...
 - explored, massaged, or even tortured
 - Vulnerable to criticism (data dredging?)

Jun-09

Hypothetical interventions

8

Observational studies analyzed like randomized experiments

- Specify the **causal question** of interest
- Design the protocol
 - eligibility criteria, regimes to be compared, period of follow-up, **analytic approach**, ...
- of a hypothetical randomized experiment to answer the causal question of interest
- Try to emulate such experiment with the observational **data + assumptions**

Jun-09

Hypothetical interventions

9

Today I will discuss two examples

1. Relatively well-defined causal question, relatively well-known answer
 - Postmenopausal hormone therapy and risk of coronary heart disease (CHD)
2. Less well-defined causal question, answer unknown or known only qualitatively
 - Lifestyle and risk of CHD

Jun-09

Hypothetical interventions

10

EXAMPLE #1 Hormone therapy and heart disease

- Relatively well-defined causal question
 - Does initiation of postmenopausal hormone therapy (estrogen plus progestin) increase the risk of CHD?
 - Can be expressed in terms of hypothetical interventions or counterfactuals
- Relatively well-known answer
 - A randomized experiment found >20% increased risk of CHD in initiators compared with noninitiators
 - The intervention became non hypothetical

Jun-09

Hypothetical interventions

11

The WHI randomized trial

Manson et al, NEJM 2003

- A large double-blind randomized trial
 - >16,000 U.S. women aged 50-79 yrs
 - Randomly assigned to hormones or placebo
- Women followed approximately every year like in many large observational studies
 - No intervention after baseline
- Analytic approach: Intent-to-treat (ITT) analysis
 - Not all women adhered to their assigned treatment, some guessed their treatment

Jun-09

Hypothetical interventions

12

WHI: ITT effect estimates

Hazard ratio (95% CI) of CHD

- Overall 1.23 (0.99, 1.53)
- Years of follow-up
 - 0-2 1.51 (1.06, 2.14)
 - >2-5 1.31 (0.93, 1.83)
 - >5 0.67 (0.41, 1.09)
- Years since menopause
 - <10 0.89 (0.54, 1.44)
 - 10-20 1.24 (0.86, 1.80)
 - >20 1.65 (1.14, 2.40)

Jun-09

Hypothetical interventions

13

Before the WHI

- Several observational studies published in the 1980s and 1990s had apparently found the opposite result
 - lower CHD risk in users of hormone therapy compared with nonusers
- For example, in the Nurses' Health Study the CHD hazard ratio for current versus never use was **0.68** (0.55, 0.83)
 - Grodstein et al (*J Women's Health* 2006)
- Clinical recommendations were based on the estimates from observational studies

Jun-09

Hypothetical interventions

14

Before the WHI

- 1992 American College of Obstetricians and Gynecologists**
"Probable beneficial effect of estrogen on heart disease"
- 1992 American College of Physicians**
"Women who have coronary heart disease or who are at increased risk of coronary heart disease are likely to benefit from hormone therapy"
- 1993 National Cholesterol Education Program**
"Epidemiologic evidence for benefit of estrogen replacement therapy is especially strong for secondary prevention in women with prior CHD"
- 1996 American Heart Association**
"ERT does look promising as a long-term protection against heart attack"

Jun-09

Hypothetical interventions

15

After the WHI: Chain reaction

- There is a clear discrepancy
- Since randomized trials are the gold standard for causal inference...
- Observational studies got it wrong
- Can observational studies ever be trusted again?
 - The end of observational epidemiology?
- *Should we fund observational studies?*

Jun-09

Hypothetical interventions

16

Why did observational studies get it "wrong"?

- Popular answer: the key **assumption** of no unmeasured confounding was violated
 - the randomized-observational discrepancy is due to insufficient adjustment for lifestyle risk factors and socioeconomic indicators
 - Corollary: causal inference from observational data is a hopeless undertaking
- Consider an alternative answer: **causal question** was not explicit
 - Observational studies implicitly asked a question different from the question explicitly asked by the randomized trial

Jun-09

Hypothetical interventions

17

More formally, the different (discrete-time) hazard ratios were

- Randomized: Initiators vs. noninitiators
 - $\Pr[Y_{t+1}=1|Y_t=0, A_0=1] / \Pr[Y_{t+1}=1|Y_t=0, A_0=0]$ for $t \geq 0$
- Observational: Current vs. never users
 - $\Pr[Y_{t+1}=1|Y_t=0, A_t=1] / \Pr[Y_{t+1}=1|Y_t=0, A_m=0]$ for $0 \leq m \leq t$
 - where
 - Y_t : CHD indicator at time t
 - A_t : indicator for use of therapy at time t

Jun-09

Hypothetical interventions

18

Randomized experiment

- First state your question, then decide your analytic approach
 - Explicit causal question: what is the effect of hormone therapy **initiation** on CHD risk?
 - Analytic approach following from that question: Compare risk between women who initiate and do not initiate hormone therapy (ITT)

Jun-09

Hypothetical interventions

19

Observational studies

- First decide your analytic approach, then try to find out the question you are answering?
 - Analytic approach: Compare risk between women who currently use therapy and those who never used it
 - Implicit causal question: what is the effect of hormone therapy **continuation** on CHD risk?

Jun-09

Hypothetical interventions

20

Our strategy

- Use the observational data to answer same question as randomized experiment
 - Re-analyze observational studies to estimate the observational analog of the ITT effect
- Then compare both set of estimates

- For a detailed description see
 - Hernán et al. *Biometrics* 2005
 - Hernán et al. *Epidemiology* 2008
 - Hernán and Robins. *Epidemiology* 2008

Jun-09

Hypothetical interventions

21

The Nurses' Health Study (NHS)

- A large longitudinal observational study
 - >120,000 women recruited in 1976
 - ~80,000 with complete lifestyle data in 1980
- Lifestyle and health information updated by questionnaire every two years
 - Use of hormone therapy
 - Diagnosis of CHD (confirmed by physician)
 - Risk factors for CHD
- Use this observational study to emulate a "trial" of hormone therapy

Jun-09

Hypothetical interventions

22

Protocol of the NHS "trial" Interventions and Eligibility criteria

- Treatment regimes
 - 1) Initiation of use of oral estrogens plus progesterone at baseline
 - 2) No hormone initiation at baseline

- Similar eligibility criteria as randomized experiment
 - Including washout interval: no hormone use in 2-year period before baseline

Jun-09

Hypothetical interventions

23

Protocol of the NHS "trial" Baseline and Follow-up

- (WHI baseline: randomization time)
- NHS baseline:
 - Initiators: month of initiation in 2-yr period before the 1984 questionnaire
 - Non initiators: average baseline month among initiators
- Follow-up
 - From baseline to CHD diagnosis, death from other causes, loss to follow-up, or June 2000, whichever came first

Jun-09

Hypothetical interventions

24

The NHS "trial" Summary

- The NHS nonrandomized study can be viewed as a nonrandomized, nonblinded trial that mimics the eligibility criteria, definition of start of follow-up, and treatment arms of the WHI randomized trial
- Some differences
 - distribution of baseline characteristics
 - e.g., shorter time since menopause in NHS than in WHI
 - length of follow-up
 - longer in NHS than in WHI

Jun-09

Hypothetical interventions

25

Protocol of the NHS "trial" Intention to treat (ITT) principle

- Compare the risk of CHD between women who initiated and did not initiate hormone therapy at baseline
 - Conditional on potential confounders
- Regardless of future hormone use during the follow-up
- This is the observational analog of the ITT effect

Jun-09

Hypothetical interventions

26

Protocol of the NHS "trial" Analytic approach

- Same as in WHI: Cox proportional hazards model
- Covariates:
 - Indicator for hormone therapy initiation
 - Age, past hormone use, parental history of myocardial infarction before age 60, education, husband's education, ethnicity, age at menopause, calendar time, high cholesterol, high blood pressure, diabetes, angina, stroke, coronary revascularization, osteoporosis, body mass index, cigarette smoking, aspirin use, alcohol intake, physical activity, diet score, multivitamin use, and fruit and vegetable intake

Jun-09

Hypothetical interventions

27

The NHS "trial" Non randomized after all

- To obtain valid effect ITT estimates in a nonrandomized trial, all baseline confounders have to be appropriately measured and adjusted for in the analysis
 - We proceeded as if this condition was at least approximately true in the NHS trial after adding the above covariates to the Cox model
- **Untestable assumption:** combined with **data** and **analytic approach** to answer the **causal question** of interest
 - Unnecessary assumption in truly randomized experiments

Jun-09

Hypothetical interventions

28

The NHS "trials"

- We started the NHS "trial" during the period before the 1994 questionnaire but there is nothing special about the 1984 questionnaire
- We can start our "trial" in the period before the 1986, 1988, ... or 1998 questionnaires
 - Sequence of nested "trials"
- Or we can conduct all possible "trials," pool the data across "trials," and obtain an effect estimate with a narrower confidence interval
 - Need to adjust for within-subject correlation

Jun-09

Hypothetical interventions

29

The NHS "trials"

- We started a separate NHS trial before each questionnaire m
 - $m=0,1,\dots, 8$ representing 1984, 1986,... 1998
- Each woman may participate in a maximum of 8 trials
- For each trial,
 - follow-up started at the trial-specific baseline (as defined above) and ended at diagnosis of a CHD endpoint, death, lost to follow-up, or June 2000, whichever came first
 - Eligibility criteria applied at baseline

Jun-09

Hypothetical interventions

30

Analytic approach (Nested) Cox model

$$\lambda_T[t|G(m) = 1, A(m), \bar{L}(m)] = \lambda_0[t] [\alpha A(m) + \theta' \bar{L}(m)]$$

- Notation
 - T : CHD-free survival time
 - $G(m)$: indicator for eligibility at m
 - $L(m)$: covariates measured before m
- PMLE, robust variance
- Conditional ITT hazard ratio: $\exp(\alpha)$
- Similar results using doubly-robust estimators from nested structural AFT model that incorporates propensity score

Jun-09

Hypothetical interventions

31

More formally, the discrete-time hazard ratios are

- Initiators vs. noninitiators
 - $\Pr[Y_{t+1}=1|Y_t=0, A_t=1, A_m=0 \text{ for all } 0 \leq m < t] /$
 $\Pr[Y_{t+1}=1|Y_t=0, A_t=0, A_m=0 \text{ for all } 0 \leq m < t]$
- rather than Current vs. never users
 - $\Pr[Y_{t+1}=1|Y_t=0, A_t=1] /$
 $\Pr[Y_{t+1}=1|Y_t=0, A_m=0 \text{ for all } 0 \leq m \leq t]$
- where
 - Y_t : CHD indicator at time t
 - A_t : use of therapy at time t

Jun-09

Hypothetical interventions

32

Results Women eligible for NHS trials

- 34,472 women contributed to trials
 - 1,021 CHD cases
- Pooling over "trials"
 - On average, each woman participated in 4.4 trials
 - 152,479 participants
 - 6,602 initiators
 - 3,597 CHD cases

Jun-09

Hypothetical interventions

33

ITT effect estimates WHI NHS

	WHI	NHS
□ Overall	1.23 (0.99, 1.53)	1.05 (0.82, 1.34)
□ Years of follow-up		
■ 0-2	1.51 (1.06, 2.14)	1.43 (0.92, 2.23)
■ >2	1.07 (0.81, 1.41)	0.91 (0.72, 1.16)
□ Years since menopause		
■ <10	0.89 (0.54, 1.44)	0.88 (0.63, 1.21)
■ 10-20	1.24 (0.86, 1.80)	1.13 (0.85, 1.49)
■ >20	1.65 (1.14, 2.40)	--

Jun-09

Hormone therapy and CHD

34

Example #1 Conclusions

- No shocking discrepancy between observational and randomized studies when the same **causal question** is asked
 - though wide confidence intervals in both the WHI and the NHS
- What about the popular response? Is there any unmeasured confounding?
 - Probably, but insufficient to explain the original discrepancy

Jun-09

Hypothetical interventions

35

Example #1 Conclusions

- The choice of the **analytic method** was guided by the **question**
 - Our sequence of nested "trials" is a particular case of g-estimation of nested structural models
- Not a vindication of ITT analyses!!
 - We used the observational data to estimate ITT effect only to facilitate comparison with the trial
- ITT analyses are problematic in the presence of substantial noncompliance
 - Better to estimate adherence-adjusted effects like the effect of continuous hormone use

Jun-09

Hypothetical interventions

36

What is the effect of continuous hormone therapy versus no use?

- A hypothetical intervention over time
 - rather than only at baseline
- We estimated this effect in both the WHI and the NHS via inverse probability weighting
 - weaker assumptions than the “current vs. never” comparison
- Again, little discrepancy
 - Hernán et al. *Epidemiology* 2008
 - Toh et al. (under review)

Jun-09

Hypothetical interventions

37

Effect estimates for continuous use

	WHI	NHS
<input type="checkbox"/> Overall	1.60 (0.90, 2.84)	1.30 (0.76, 2.21)
<input type="checkbox"/> Years of follow-up		
■ 0-2	2.45 (1.64, 3.67)	1.71 (1.03, 2.83)
■ >2	1.25 (0.81, 1.95)	1.07 (0.44, 2.63)
<input type="checkbox"/> Years since menopause		
■ <10	0.75 (0.23, 2.47)	0.68 (0.24, 1.91)
■ ≥10	1.96 (1.02, 3.76)	1.57 (0.86, 2.85)

Jun-09

Hormone therapy and CHD

38

Example #1

Conclusions

- A clear specification of the question of interest helps
 - state our assumptions precisely
 - compare apples with apples, or oranges with oranges
 - Randomized vs. observational ITT estimates, randomized vs observational adherence-adjusted estimates
 - We can then discuss whether we like apples or oranges better

Jun-09

Hypothetical interventions

39

EXAMPLE #2

Lifestyle and heart disease

- Less well-defined causal questions
 - What is the causal effect of smoking, alcohol, physical activity, diet, and body mass index (BMI) on the risk of CHD?
 - Need to express this question in terms of hypothetical interventions or counterfactual contrasts
- Unknown answer
 - No large, long-term randomized trials of lifestyle with full adherence to the relevant interventions
 - The interventions remain mostly hypothetical

Jun-09

Hypothetical interventions

40

The Nurses' Health Study

Stampfer et al, *NEJM* 2000

- “We defined subjects as **low risk** as those who
 1. were not currently smoking
 2. had a BMI under 25
 3. consumed an average of at least half a drink of an alcoholic beverage per day
 4. engaged in moderate-to-vigorous physical activity for at least half an hour per day
 5. scored in the highest 40 percent of the cohort for consumption of a diet high in cereal fiber, marine n-3 fatty acids, and folate, with a high ratio of polyunsaturated to saturated fat, and low in trans fat and glycemic load”

Jun-09

Hypothetical interventions

41

The causal question of interest

Stampfer et al, *NEJM* 2000

- Not explicitly defined in the article
- The hormone therapy example showed that the absence of an explicit definition of causal question can lead to confusion
- We tried to infer the causal question of interest from the authors' interpretation of the results

Jun-09

Hypothetical interventions

42

Results and interpretation

Stampfer et al, *NEJM* 2000

- “Women in the low-risk category for all five factors considered together, as compared with all other women, had a relative risk of 0.17 (95% CI, 0.07 to 0.41)”
- “The population attributable risk was 82% (95% CI, 58 to 93), suggesting that 82% of the coronary events in this cohort might have been prevented if all women had been in the low-risk group”
- “Closer adherence to a more healthful lifestyle might reduce the risk of coronary heart disease still further”

Jun-09

Hypothetical interventions

43

The causal question of interest

Stampfer et al, *NEJM* 2000

- Appears to involve some unspecified joint (simultaneous) intervention on the 5 risk factors over some unspecified period
 - **TABLE 2.** Risk Of Coronary Events In Low-risk Groups Defined According To Different Constellations Of Modifiable Risk Factors For Coronary Disease In The Nurses' Health Study, 1980 To 1994

Jun-09

Hypothetical interventions

44

Analytic approach

Stampfer et al, *NEJM* 2000

- Hazard ratio of CHD for current “low-risk” vs. no “low-risk” lifestyle
 - $\Pr\{Y_{t+1}=1|Y_t=0, A_t=1, L_t\} / \Pr\{Y_{t+1}=1|Y_t=0, A_t=0, L_t\}$where
 - Y_t : CHD indicator at time t
 - A_t : indicator for “low-risk” lifestyle at t or through t (cumulative average)
 - L_t : potential confounders at time t
 - t is measured in 2-year units

Jun-09

Hypothetical interventions

45

Analytic approach

Stampfer et al, *NEJM* 2000

- Explicitly defined in the article
 - Described in detailed in the Methods section
- Is this analytic approach appropriate to answer a causal question about lifestyle changes during the follow-up?
 - e.g., can the conditional association between current BMI and 2-year CHD risk be interpreted as the causal effect of intervening on BMI during the follow-up?
 - It depends on the causal question and the investigators' assumptions

Jun-09

Hypothetical interventions

46

The assumptions

Stampfer et al, *NEJM* 2000

- Appear to involve exchangeability of low-risk and no low-risk women in each period conditional on certain variables
 - age, calendar time, parental history of myocardial infarction before the age 60, menopausal status and use of postmenopausal hormones, history of hypertension, history of high cholesterol levels
 - But not others like baseline lifestyle
- Also stated as the assumption that “there was a causal relation between the risk factors and coronary heart disease”

Jun-09

Hypothetical interventions

47

Our strategy

1. Define the causal question(s) explicitly
 - Specify the hypothetical interventions as explicitly as possible
 2. Propose an answer by using updated NHS data under explicit assumptions and using a consistent analytic method
 - e.g., the parametric g-formula
- For a detailed description, see
 - Taubman et al. *Int J Epidemiol* 2009
 - Taubman et al. *JSM Proceedings* 2008

Jun-09

Hypothetical interventions

48

Consider 9 hypothetical interventions

1. Avoid smoking
2. Exercise at least 30 minutes a day
3. Keep diet score (described above) in a range corresponding to the top 2 quintiles of the observed data
4. Consume at least 5 grams of alcohol per day
5. Maintain body mass index (BMI) less than 25
6. Interventions 1 - 3 combined
7. Interventions 1 - 3 and 5 combined
8. Interventions 1 - 4 combined
9. Interventions 1 - 5 combined

Jun-09

Hypothetical interventions

49

For each of these interventions

- We estimated the 20-year CHD risk *were the entire population to follow the prescribed intervention beginning at start of follow-up in 1982*
- Applying the parametric g-formula
 - to the most updated NHS data
 - assuming **no** unmeasured confounding (sequential randomization), measurement error, and model misspecification
- Software available from
 - www.hsph.harvard.edu/causal

Jun-09

Hypothetical interventions

50

Estimates from the NHS Taubman et al. *Int J Epidemiol* 2009

Intervention	20-year Risk	Risk Ratio
(0) No intervention	3.68 (3.56, 4.09)	1
(1) Quit smoking	3.01 (2.86, 3.38)	0.82 (0.78, 0.85)
(2) Exercise at least 30 minutes per day	2.90 (2.47, 3.60)	0.79 (0.64, 0.92)
(3) Keep diet score in the top 2 quintiles	3.27 (3.08, 3.68)	0.89 (0.82, 0.95)
(4) Consume at least 5g alcohol per day	3.19 (2.84, 3.72)	0.87 (0.75, 0.98)
(5) Maintain BMI less than 25	3.62 (3.45, 4.11)	0.98 (0.93, 1.04)
(6) "Low-risk" lifestyle (1-3 combined)	2.22 (1.85, 2.74)	0.60 (0.48, 0.70)
(7) "Low-risk" lifestyle (1-3 and 5 combined)	2.17 (1.78, 2.69)	0.59 (0.47, 0.70)
(8) "Low-risk" lifestyle (1-3 and 4 combined)	1.88 (1.51, 2.38)	0.51 (0.40, 0.63)
(9) "Low-risk" lifestyle (1-5 combined)	1.89 (1.46, 2.41)	0.51 (0.39, 0.64)

Jun-09

Hypothetical interventions

51

Interpretation

- 49% of CHD cases attributable to these lifestyle interventions
- Compare with
 - 82% in Stampfer et al (2000)
 - 67% after applying Stampfer et al's analytic approach to updated NHS data
- Strong effect of lifestyle, though weaker than previously reported

Jun-09

Hypothetical interventions

52

How seriously should we take our estimates?

- Analytic approach is now consistent with the assumptions
- But assumptions are surely violated to some degree
 - Bias (of unknown direction and magnitude) because of unmeasured confounding, measurement error, and model misspecification

Jun-09

Hypothetical interventions

53

Unmeasured confounding

- The g-formula appropriately adjusts for *measured* time-varying confounding but...
- Surely there is residual confounding by *unmeasured* factors
 - e.g., access to preventive medicine, subclinical disease
- May result in upwards or downwards bias
 - e.g., unmeasured (subclinical) disease would make BMI reduction look worse, and physical activity increase look better

Jun-09

Hypothetical interventions

54

Measurement error

- Surely exposures measured with error
- On one hand:
 - possible attenuation of the effect
 - e.g., if measurement error for diet and average 2-year change in diet are of similar magnitude
- On the other hand:
 - Because past exposures are confounders, measurement error results in more residual confounding

Jun-09

Hypothetical interventions

55

Model misspecification

- Surely our models are misspecified
 - Alternative specifications result in 10% change in estimates
- Correct specification is almost an impossible task:
 - exposures and confounder measured simultaneously in the same questionnaire
 - Time sequence cannot be discerned
 - Common problem to all "interval" cohorts

Jun-09

Hypothetical interventions

56

And yet...

- the explicit specification of the hypothetical interventions is helpful
- ❖ To precisely state our assumptions
 - and anchor the discussion on confounding, measurement error, model misspecification
- ❖ To identify additional problems...

Jun-09

Hypothetical interventions

57

What we learn from making hypothetical interventions explicit

- Causal question is relatively straightforward for smoking
- Causal question too vague for all other risk factors
 - e.g., 5g of alcohol per day: you mean 5g every day, or 35g on Saturdays?
 - e.g., high dietary score: how is this translated into actual public health advice?
 - plus the intervention depends on the distribution of score (percentiles) in the study
 - e.g., maintain BMI < 25: how?

Jun-09

Hypothetical interventions

58

What we learn from making hypothetical interventions explicit

- Some interventions are extremely unrealistic
- For example, "Maintain BMI less than 25 starting in 1982"
 - Meaning that if your BMI was 30 in 1982, you instantaneously reduce it to 25? How?
 - Hernán and Taubman. *Int J Obesity* 2008

Jun-09

Hypothetical interventions

59

What we learn from making hypothetical interventions explicit

- All analytic methods somehow compare individuals
 - whose data are consistent with the interventions
 - with the others
- What if no study subjects actually followed the proposed interventions?
 - We can still get effect estimates
 - But they rely on model extrapolation

Jun-09

Hypothetical interventions

60

% of subjects whose data is not consistent with the intervention

Intervention	%
(0) No intervention	0
(1) Quit smoking	30
(2) Exercise at least 30 minutes per day	99
(3) Keep diet score in the top 2 quintiles	99
(4) Consume at least 5g alcohol per day	89
(5) Maintain BMI less than 25	73
(6) "Low-risk" lifestyle (1-3 combined)	100
(7) "Low-risk" lifestyle (1-3 and 5 combined)	100
(8) "Low-risk" lifestyle (1-3 and 4 combined)	100
(9) "Low-risk" lifestyle (1-5 combined)	100

Jun-09

Hypothetical interventions

61

Example #2 Conclusions

- We don't quite know what causal question we are asking
 - Our estimates are hard to interpret
 - Discussion of the merits of **any** analytic approach is premature
- Need to go back and try less vague/more realistic specification of hypothetical interventions that are sufficiently represented in the study

Jun-09

Hypothetical interventions

62

Overall conclusion

- The complexity of causal questions is often overlooked
 - especially when dealing with complex longitudinal data
- If confused when attending a lecture on fancy methods and sophisticated assumptions for causal inference, feel free to interrupt the speaker and ask:
What was your question again?

Jun-09

Hypothetical interventions

63