Causality, confounding, and counterfactuals: An overview of the conceptual framework

Erica E. M. Moodie

Department of Epidemiology, Biostatistics, & Occupational Health
McGill University
Montreal, QC, Canada

erica.moodie@mcgill.ca

Session goals

- What is a cause, and what is causal inference?
- How can causation be established?
- How can we make the statistical study of causation formal?

Road map

1. Causality and etiology
2. The counterfactual framework
   - A causal model
   - Causal estimands
3. Confounding and exchangeability
   - Randomized trials
   - Confounding
   - Exchangeability

Causality

- The goal of most, if not all, statistical inference is to uncover causal relationships.
- In health/medical research, e.g., Miettinen & Karp (2012, p. 35) write "Epidemiological research is, almost exclusively, concerned with *etiology* of illness", where etiology concerns the "causal origin (in the case of a disease or defect, specifically, the causation of the inception and/or progression of its pathogenesis)" of illness.

Causality

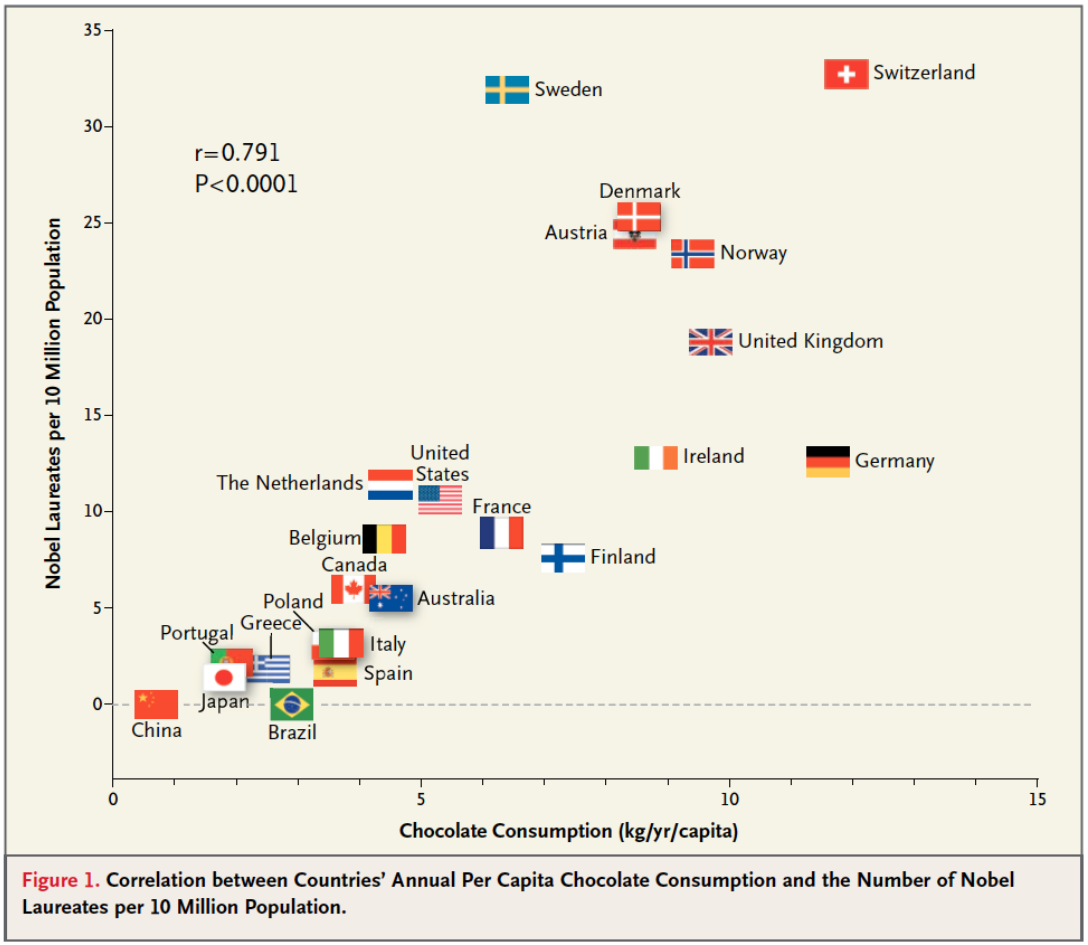
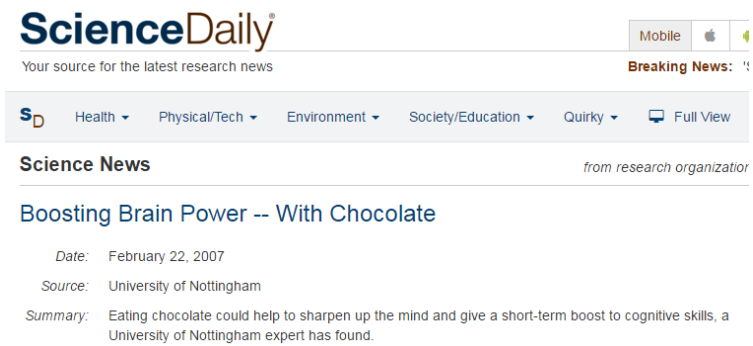


Figure 1. Correlation between Countries' Annual Per Capita Chocolate Consumption and the Number of Nobel Laureates per 10 Million Population.

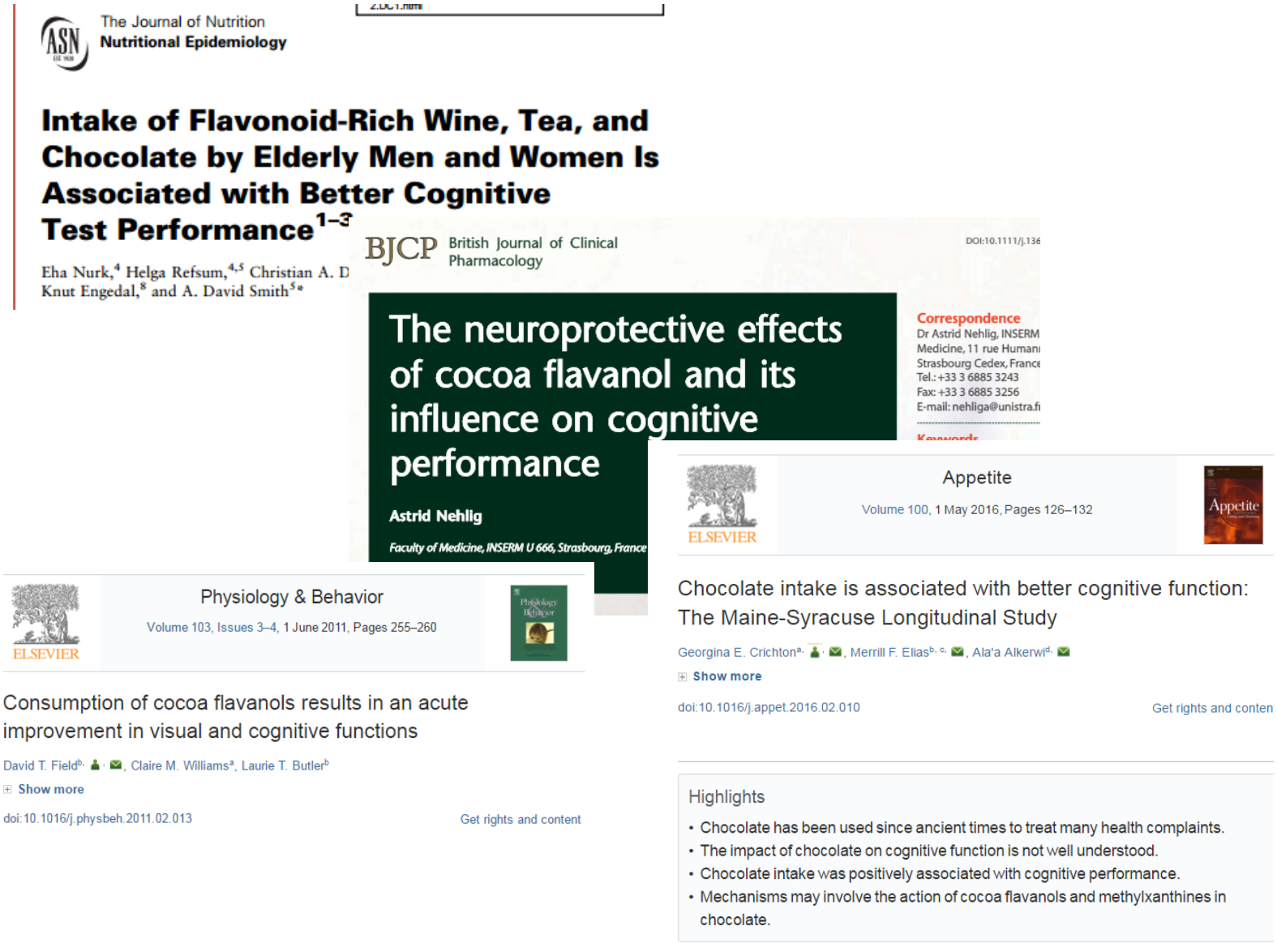
Messerli (2012), New England Journal of Medicine

Causality

HUFFPOST SCIENCE



Causality



The Journal of Nutrition
Nutritional Epidemiology

Intake of Flavonoid-Rich Wine, Tea, and Chocolate by Elderly Men and Women Is Associated with Better Cognitive Test Performance[1–3]

Eha Nurk,[4] Helga Refsum,[4,5] Christian A. D
Knut Engedal,[8] and A. David Smith[5]*

BJCP British Journal of Clinical Pharmacology

DOI:10.1111/j.136

The neuroprotective effects of cocoa flavanol and its influence on cognitive performance

Astrid Nehlig

Faculty of Medicine, INSERM U 666, Strasbourg, France

Correspondence
Dr Astrid Nehlig, INSERM
Medicine, 11 rue Humani
Strasbourg Cedex, France
Tel.: +33 3 6885 3243
Fax: +33 3 6885 3256
E-mail: nehliga@unistra.fr

ELSEVIER

Physiology & Behavior

Volume 103, Issues 3–4, 1 June 2011, Pages 255–260

Consumption of cocoa flavanols results in an acute improvement in visual and cognitive functions

David T. Field[b], Claire M. Williams[a], Laurie T. Butler[b]

Show more

doi:10.1016/j.physbeh.2011.02.013

Get rights and content

ELSEVIER

Appetite

Volume 100, 1 May 2016, Pages 126–132

Chocolate intake is associated with better cognitive function: The Maine-Syracuse Longitudinal Study

Georgina E. Crichton[a], Merrill F. Elias[b, c], Ala'a Alkerwi[d]

Show more

doi:10.1016/j.appet.2016.02.010

Get rights and conten

Highlights

- Chocolate has been used since ancient times to treat many health complaints.
- The impact of chocolate on cognitive function is not well understood.
- Chocolate intake was positively associated with cognitive performance.
- Mechanisms may involve the action of cocoa flavanols and methylxanthines in chocolate.

Causality

- Some of the earliest, and best known, ideas in epidemiology on causality are the criteria for causality given by Sir Austin Bradford-Hill in 1965:
  1. Strength
  2. Consistency (reproducibility)
  3. Specificity
  4. Temporality
  5. Biological gradient
  6. Plausibility
  7. Coherence
  8. Experiment
  9. Analogy
- A group of minimal conditions needed to establish causality.
- But are they all equally convincing? Sufficiently formal?

Causality

- There is no agreement on the definition of causality, or even whether it exists in the objective physical reality.
- Some statisticians have adopted a determinist position originating from Laplace, where the present state of the universe is *determined* by its past states, the present state then being an effect of the past states and a cause of the future states.
- Determinism can be linked to the information-based interpretation of probability, whereby randomness exists due to incomplete information.
- Determinism is not compatible with other interpretations of causality based on manipulable interventions or potential outcomes [to be defined], since determinism rules out free will and hence interventions.

Causality

- Pearl (2009, p. 25-26), a computer scientist and a leader in the field of modern causality, does not explicitly define causality at all, thought refers to causal relationships as “stable” and “ontological”.
- Nicolai Meinshausen (along with J. Peters and P. Buehlmann, 2016) takes a similar view, viewing causal relationships to be present when multiple data sources produce invariant prediction.

Causality

- Miettinen & Karp (2012, p. 43) note that causality is inherently unobservable, making causal inferences particularly challenging:

> *Research on etiogenesis of morbidity – or of illness per se – is, by the very nature of this genre of causation in medicine, generally bound to be non-experimental; but a much greater added challenge is that causation is not a phenomenon, subject to observation; it is a 'conception a priori,' a noumenon (Kant), needing to be inferred from phenomenal patterns.*

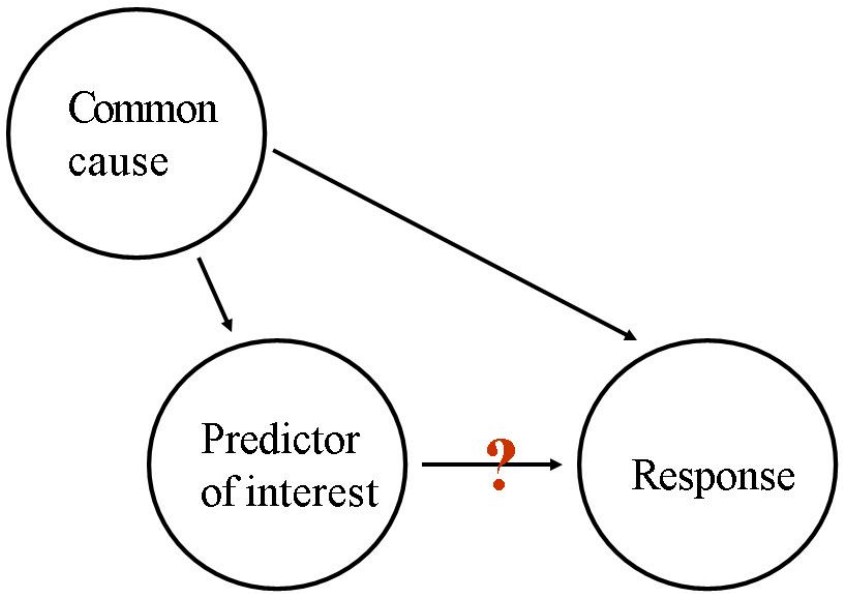
- In practice, of course, the major challenge in any observational study of etiology is confounding.

Causality

- The rest of this course deals with methods for controlling of confounding, either through study design or through modelling.
- Before this, however, we need to understand expressly what confounding is, and to do this we need to introduce a causal model.
- Through this causal model, we can define an explicit causal parameter and causal contrast, which is necessary to interpret confounding.

Some concepts, cross-sectionally

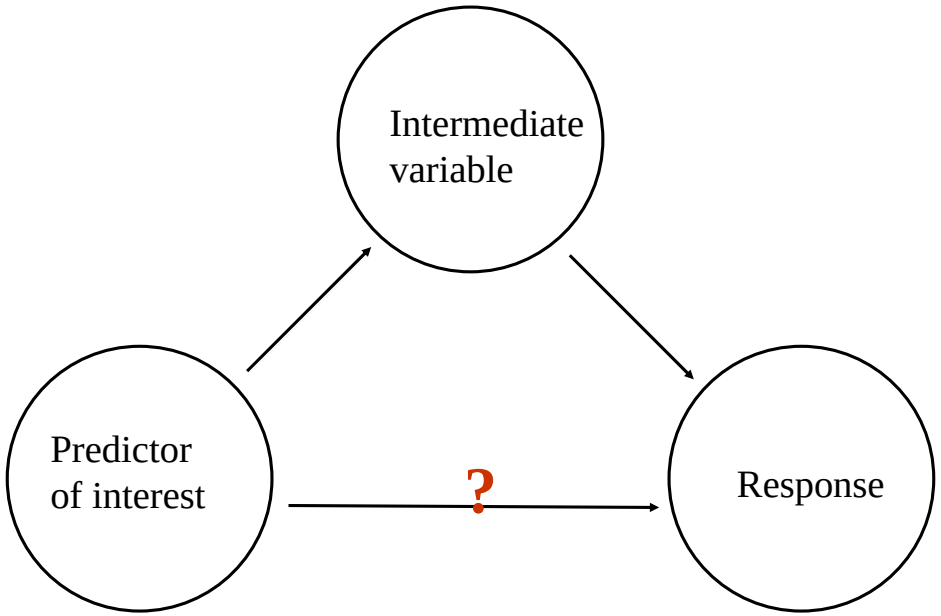
In a cross-sectional setting, confounding is loosely defined as the exposure and response having a *common cause*:



Including the common cause in a regression model "breaks" the confounding, allows estimation of the true exposure effect.

Some concepts, cross-sectionally

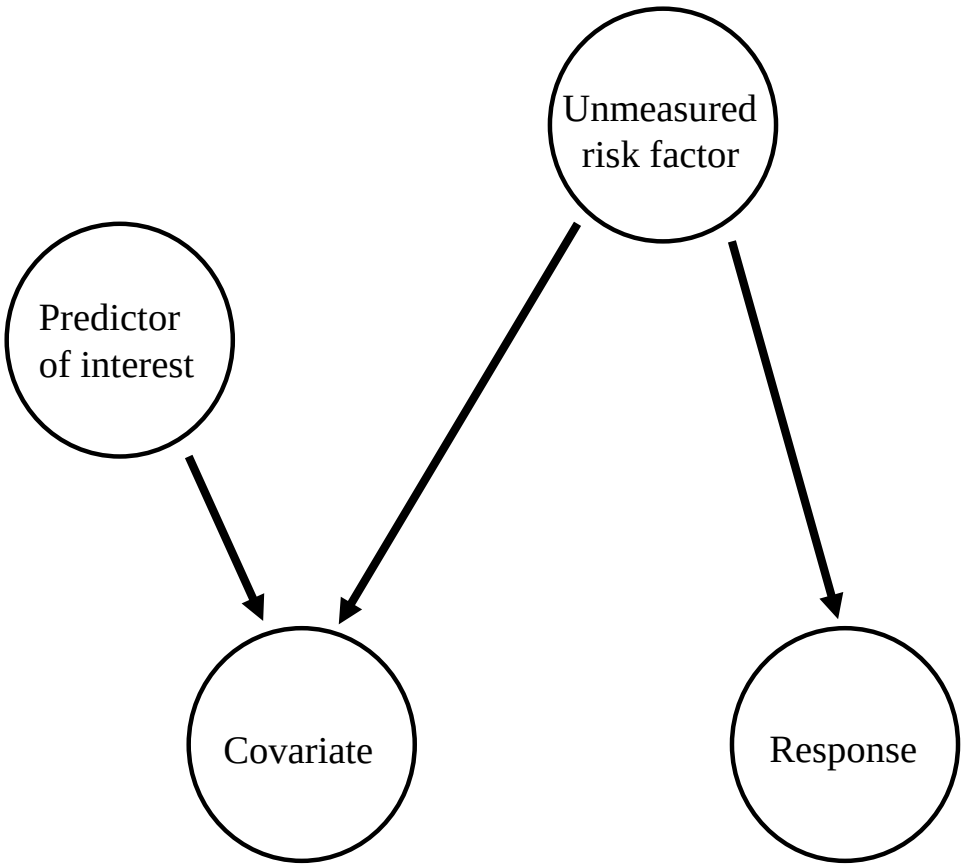
An intermediate (mediating, intervening) variable is caused by the exposure of interest and causes the response:



Including such a variable in a regression model affects the interpretation of the exposure parameter: *total* effect (direct and indirect) of exposure cannot be estimated.

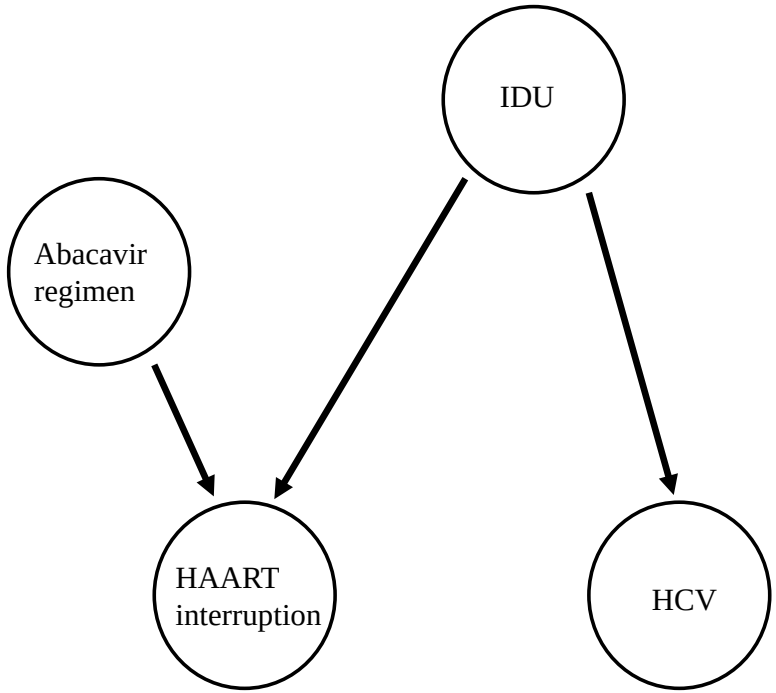
Some concepts, cross-sectionally

What if a variable is caused by the exposure, but does not affect the response?



Conditioning on the "descendant" of exposure leads to collider-stratification bias.

Some concepts, cross-sectionally



Selecting or conditioning on HAART interruption asks: *"Among people who are (or are not) on a HAART interruption*, is there an association between being prescribed in ABC-containing regimen and acquiring HCV?"

The central causal question

- As we have discussed, in most research domains, the objective of an investigation is to quantify the effect on a measurable outcome of changing one of the conditions under which the outcome is measured.
- The central statistical challenge is that, unless the condition of interest is changed independently, the inferred effect may be subject to the influence of other variables.

- Key point: It is essential to think about causation, or at least the underlying data-generating mechanism, in order to decide what variables to include in a regression model and how to analyze the data.

  Knowing that a variable is associated with the exposure and with the outcome is not enough to determine whether it is a confounding variable or an intermediate variable.

The language of causal inference

We seek to quantify the effect on an outcome of changes in the value of an exposure or treatment.

- Outcome: could be
  - binary;
  - integer-valued;
  - continuous-valued.
- Exposure: could be
  - binary;
  - integer-valued;
  - continuous-valued.
- Study: could be
  - cross-sectional (single time point);
  - longitudinal (multiple time points), with single or multiple exposures.

We consider the impact of an *intervention* to change exposure status.

Notation

We adopt the following notation: let

- i index individuals included in the study;
- Y_i denote the outcome for individual i;
- Z_i denote the exposure for individual i;
- X_i denote the values of other predictors or covariates.

For a cross-sectional study, Y_i and Z_i will be scalar-valued; for the longitudinal case, Y_i and Z_i may be vector valued. X_i is typically vector-valued at each measurement time point.

We will treat these variables as *random* quantities, and regard them as samples from an infinite population, rather than a finite population.

The counterfactual framework

- Although other causal models exist, here we will concentrate on the well known *Rubin's causal model* (Rubin 1978; Rosenbaum & Rubin 1983; Holland 1986).
- This is equivalent to the use of what is known as potential outcomes or counterfactual notation.
- Let's focus on an intervention Z_i that is a binary indicator of whether treatment (1) or placebo (0) is administered to individual i:
  - $z_i = 1$, representing the treatment or exposure, and $z_i = 0$, representing the *reference* level (e.g. no treatment, placebo).
- Now suppose that each individual has two potential outcomes $(Y_i(0), Y_i(1))$, corresponding to each possible treatment assignment (if there were more than two treatment alternatives, more potential outcomes would be defined accordingly).

The counterfactual framework

- Some take a philosophical view that potential or counterfactual outcomes may be thought to "exist" ("many-worlds" interpretation of quantum mechanics in physics).
- ...but it could be argued that they are best interpreted as useful tools (mental constructs) corresponding to what-if types of structures in language.
- Only one of the outcomes is realized and observed (observable), given by

$$Y_i = Y_i(0)(1 - Z_i) + Y_i(1)Z_i.$$

The counterfactual framework

If exposure is multi-valued, the potential outcomes

$$\{Y_i(z_1), Y_i(z_2), \ldots, Y_i(z_d)\}$$

represent the outcomes that would result for individual i if that subject exposed to exposure level $z_1, z_2, \ldots, z_d$ respectively.

The observed outcome, Y_i, may then be written in terms of the potential outcomes and the observed exposure, Z_i, as

$$Y_i = \sum_{j=1}^{d} \mathbb{1}_{z_j}(Z_i) Y_i(z_j).$$

where $\mathbb{1}_{\mathcal{A}}(Z)$ is the indicator random variable for the set $\mathcal{A}$, with $\mathbb{1}_{\mathcal{A}}(Z) = 1$ if $Z \in \mathcal{A}$, and zero otherwise.

The counterfactual framework

If exposure is continuous-valued, the potential outcomes

$$\{Y_i(z), z \in \mathcal{Z}\}$$

represent the outcomes that would result for individual i if that subject exposed to exposure level z which varies in the set $\mathcal{Z}$.

Binary Exposures

For a binary exposure, we define the causal effect of exposure by considering contrasts between $Y_i(0)$ and $Y_i(1)$; for example, we might consider

- Additive contrasts

$$Y_i(1) - Y_i(0)$$

- Multiplicative contrasts

$$Y_i(1)/Y_i(0)$$

- However only one of the potential outcomes may be observed, and so this individual level effect is not identifiable.
- Holland (1986, p. 947) calls this the fundamental problem of causal inference.
- In practice, we consider the causal effect of exposure to be defined by *contrasts* in (expected) potential outcomes corresponding to *different* exposure levels.

Continuous Exposures

For a continuous exposure, we might consider the path tracing how $Y_i(z)$ changes as z changes across some relevant set of values.

This leads to a *causal dose-response* function.

Example: Occlusion Therapy for Amblyopia

We might seek to study the effect of occlusion therapy (patching) on vision improvement of amblyopic children. Patching 'doses' are measured in terms of time for which the normal functioning ("fellow") eye is patched.

As time is measured continuously, we may consider how vision improvement changes for any relevant dose of occlusion.

Expected counterfactuals

In general, we are interested in population or subgroup, rather than individual level causal effects. The potential outcomes are random quantities. Therefore, we more typically consider expected potential outcomes

$$\mathbb{E}[Y_i(z)]$$

or contrasts of these quantities.

We might also consider subgroup conditional expected quantities

$$\mathbb{E}[Y_i(z)|i \in \mathcal{S}]$$

where $\mathcal{S}$ is some stratum of interest in the general population.

We typically assume that subject i is randomly sampled from the population or stratum, so that these individual-level expectations are representative of the population.

Expected counterfactuals: population-level contrasts

- Miettinen (2011, p. 84) notes that

> *The causal contrast - cause versus its alternative - does not have as its referent instances that differ in this respect (cause present in some, the alternative in others). Instead, the contrast has to do with all of the instances of the study domain (and study base) in the same way: the contrast is between all instances of the domain with the cause present (hypothetically) versus all of them with the alternative present (hypothetically). The contrast has to do with two mutually exclusive possibilities in each instance of the study domain, at least one of them a hypothetical (counterfactual).*

- Thus Miettinen's definitions for causal concepts should be understood in terms of Rubin's causal model.

Expected counterfactuals: binary exposure

For a binary exposure, we might consider the average effect of exposure (or average treatment effect, ATE) defined as

$$\mathbb{E}[Y_i(1) - Y_i(0)] = \mathbb{E}[Y_i(1)] - \mathbb{E}[Y_i(0)]$$

If the outcome is also binary, we note that

$$\mathbb{E}[Y_i(z)] \equiv \Pr[Y_i(z) = 1]$$

so may also consider odds or odds ratios quantities

$$\frac{\Pr[Y_i(z) = 1]}{\Pr[Y_i(z) = 0]} \qquad \frac{\Pr[Y_i(1) = 1] / \Pr[Y_i(1) = 0]}{\Pr[Y_i(0) = 1] / \Pr[Y_i(0) = 0]}.$$

Expected counterfactuals: binary exposure (cont.)

We may also consider quantities such as the

average treatment effect on the treated, ATT

defined as

$$\mathbb{E}[Y_i(1) - Y_i(0)|Z_i = 1]$$

although such quantities can be harder to interpret.

The utility of the potential outcomes formulation is evident in this definition.

Expected counterfactuals: binary exposure (cont.)

Example: Antidepressants and autism

Antidepressants are quite widely prescribed and for a variety of mental health concerns. However, patients can be reluctant to embark on a course of antidepressants during pregnancy. We might wish to investigate, in a population of users (and potentials users) of antidepressants, the incidence of autism-spectrum disorder in early childhood and to assess the causal influence of antidepressant use on this incidence.

- Outcome: binary, recording the a diagnosis of autism-spectrum disorder in the child by age 5;
- Exposure: antidepressant use during 2nd or 3rd trimester of pregnancy.

Then we may wish to quantity

$$\mathbb{E}[Y_i(\text{antidepressant}) - Y_i(\text{no antidepressant}) | \text{Antidep. actually used}].$$

The randomized study

The approach that intervenes to set exposure equal to z for all subjects, however, does not facilitate comparison of APOs for different values of z.

Therefore consider a study design based on *randomization*; consider from simplicity the binary exposure case. Suppose that a random sample of size $2n$ is obtained, and split into two equal parts.

- the first group of n are assigned the exposure and form the 'treated' sample,
- the second half are left 'untreated'.

The randomized study (cont.)

For both the treated and untreated groups we may use the previous logic, and estimate the ATE

$$\mathbb{E}[Y_i(1) - Y_i(0)] = \mathbb{E}[Y_i(1)] - \mathbb{E}[Y_i(0)]$$

by the difference in means in the two groups, that is

$$\frac{1}{n}\sum_{i=1}^{n} Y_i - \frac{1}{n}\sum_{i=n+1}^{2n} Y_i.$$

The key idea here is that the two halves of the original sample are exchangeable with respect to their properties; the only systematic difference between them is due to exposure assignment.

The randomized study (cont.)

In a slightly modified design, suppose that we obtain a random sample of size n from the study population, but then assign exposure *randomly* to subjects in the sample: subject i receives treatment with probability p.

- if $p = 1/2$, then there is an equal chance of receiving treatment or not;
- we may choose any value of $0 < p < 1$.

In the final sample, the number treated, n_1, is a realization of a random variable N_1 where

$$N_1 \sim \text{Binomial}(n, p).$$

The randomized study (cont.)

This suggests the estimators

$$\widehat{\mathbb{E}}[Y(z)] = \frac{\sum_{i=1}^{n} \mathbb{1}_z(Z_i) Y_i}{\sum_{i=1}^{n} \mathbb{1}_z(Z_i)} \qquad z = 0, 1 \tag{1}$$

where the indicators $\mathbb{1}_z(Z_i)$ identify those individuals that received treatment z.

- If the treatment assignment is completely randomized, it may be assumed that

$$(Y_i(0), Y_i(1)) \perp\!\!\!\perp Z_i.$$

The randomized study (cont.)

- With this assumption, the difference of the sample means in the treated and untreated groups:

$$\frac{\sum_{i=1}^{n} Z_i Y_i}{\sum_{i=1}^{n} Z_i} - \frac{\sum_{i=1}^{n} (1 - Z_i) Y_i}{\sum_{i=1}^{n} (1 - Z_i)}$$

$$= \frac{\sum_{i=1}^{n} Z_i [Y_i(0)(1 - Z_i) + Y_i(1) Z_i]}{\sum_{i=1}^{n} Z_i} - \frac{\sum_{i=1}^{n} (1 - Z_i) [Y_i(0)(1 - Z_i) + Y_i(1) Z_i]}{\sum_{i=1}^{n} (1 - Z_i)}$$

$$\approx \frac{1}{n} \sum_{i=1}^{n} \frac{Z_i Y_i(1)}{\Pr(Z_i = 1)} - \frac{1}{n} \sum_{i=1}^{n} \frac{(1 - Z_i) Y_i(0)}{\Pr(Z_i = 0)},$$

consistently estimates the average causal effect.

The randomized study (cont.)

- This is because

$$\mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n}\frac{Z_i Y_i(1)}{\Pr(Z_i=1)}\right]$$

$$= \frac{1}{n}\sum_{i=1}^{n}\int_{y_0,y_1}\sum_{z_i=0}^{1}\frac{z_i y_1}{\Pr(Z_i=1)} f_{Y(0),Y(1)|Z}(y_0,y_1 \mid z_i) f_Z(z_i)\,\mathrm{d}y_0\,\mathrm{d}y_1$$

$$= \frac{1}{n}\sum_{i=1}^{n}\int_{y_0,y_1}\frac{y_1}{\Pr(Z_i=1)} f_{Y(0),Y(1)|Z}(y_0,y_1 \mid Z_i=1)\Pr(Z_i=1)\,\mathrm{d}y_0\,\mathrm{d}y_1$$

$$\equiv \frac{1}{n}\sum_{i=1}^{n}\int_{y_0,y_1} y_1 f_{Y(0),Y(1)}(y_0,y_1)\,\mathrm{d}y_0\,\mathrm{d}y_1$$

$$= \frac{1}{n}\sum_{i=1}^{n}\int_{y_1} y_1 f_{Y(1)}(y_1)\,\mathrm{d}y_1$$

$$= \mathbb{E}[Y_i(1)].$$

The randomized study (cont.)

- Similarly,

$$\mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n}\frac{(1-Z_i)Y_i(0)}{\Pr(Z_i=0)}\right] = \mathbb{E}[Y_i(0)].$$

- Note that the outcome here could just as easily have been binary, the causal contrast of interest then being for a difference in probabilities or a risk difference.

The randomized study (cont.)

Note that for the denominator,

$$\sum_{i=1}^{n} \mathbb{1}_1(Z_i) \sim \text{Binomial}(n, p)$$

so we may consider replacing the denominators by their expected values

$$np \qquad \text{and} \qquad n(1-p)$$

respectively for $z = 0, 1$. This yields the estimators

$$\widehat{\mathbb{E}}[Y(1)] = \frac{1}{np} \sum_{i=1}^{n} \mathbb{1}_1(Z_i) Y_i \qquad \widehat{\mathbb{E}}[Y(0)] = \frac{1}{n(1-p)} \sum_{i=1}^{n} \mathbb{1}_0(Z_i) Y_i. \tag{2}$$

The randomized study (cont.)

Note

The estimators in (1) are *more efficient* than the estimators in (2), that is, they have *lower variances*.

It is more efficient to use an estimated value of p

$$\widehat{p} = \frac{N_1}{n}$$

than p itself.

Confounding & exchangeability

- Many writers reasoning in terms of Rubin's causal model (e.g. Holland 1986, p. 946 and Miettinen 2011, p. 110) make the further restriction that a cause must be something that is (at least in principle) manipulable, that is, both exposure to the cause and its alternative must be possible in each instance.
- "No causation without manipulation" !
- The idea of manipulation or intervention also raises the idea of whether there is but one version of the intervention, or several.

Confounding

- However, this restriction is particular to Rubin's causal model, which originates from randomized controlled trial settings, and does not necessarily correspond to a common (non-statistical) meaning of cause.
- As noted by Pearl (2009, p. 361)

> *Surely we have causation without manipulation. The moon causes tides, race causes discrimination, and sex causes the secretion of certain hormones and not others. Nature is a society of mechanisms that relentlessly sense the values of some variables and determine the values of others; it does not wait for a human manipulator before activating those mechanisms.*

Confounding & exchangeability

- A further assumption, called positivity, is related to the idea of manipulation.
- It says that both

$$\Pr(Z_i = 1) > 0 \text{ and } \Pr(Z_i = 0) > 0$$

  for all $i = 1, \ldots, n$.
- This prevents division by zero in the above equations; more importantly, it prevents extrapolation.
- Also called "experimental treatment assignment".

Confounding

- Confounding is now easy to understand in terms of Rubin's causal model; it is likely appear when the assumption $(Y_i(0), Y_i(1)) \perp\!\!\!\perp Z_i$ does not hold.
- This is the case for instance if the value of Z_i has not been determined through complete randomization, but has been recorded in a non-experimental ("observational") study.
- It may then be that $(Y_i(0), Y_i(1))$ and Z_i have common determinants, which in turn means that Z_i is informative of the pair of potential outcomes.
- Suppose that the values of all of such common determinants, denoted as X_i, have also been recorded in the study.

Confounding

- The previous independence assumption may then be restated as the conditional independence

$$(Y_i(0), Y_i(1)) \perp\!\!\!\perp Z_i \mid X_i,$$

  conditioning on X_i then being interpreted as controlling for confounding.
- Many different names have been given to this assumption in the literature; to avoid confusion, let us call it no unmeasured confounding (NUC).
- NUC of course implies that X_i are measured confounders.
- It should be noted right away that this assumption is empirically untestable based on observed data (X_i, Z_i, Y_i), $i = 1, \ldots, n$, alone, and must therefore be based on prior information on the causal mechanisms involved.

Confounding

- If adjustment is not made, the estimator based on simple averages in the exposed and unexposed is biased in estimation of the average causal effect, and the resulting estimate is (likely to be) confounded, since the estimated effect has some other than causal (w.r.t. Z_i) explanation.
- It should be noted that even randomized controlled studies have *imbalance* w.r.t. the confounder distributions between the treated and untreated groups due to the random assignment, the more the smaller is n.
- It depends on definition if the effects of this imbalance are also called confounding.
  - Miettinen (2011, p. 110) would view this as confounding, whereas statisticians working from the basis of a causal diagram would not.

Exchangeability

Note

No unmeasured confounding is closely related to the concept of exchangeability, due to de Finetti (1974).

- Exchangeability requires that the joint distribution $p(Y_1, \ldots, Y_n)$ is constant for all permutations ρ of the indices $\{1, \ldots, n\}$.
- I.e., $p(Y_1, \ldots, Y_n) = p(Y_{\rho(1)}, \ldots, Y_{\rho(n)})$.
- Loosely, this means that individual units are in some sense 'similar', so that we can learn something on further similar instances without having to carry out further measurements, and the individual indices are non-informative, that is, one does not have any other relevant information to tell apart the different individuals.

Exchangeability

- Suppose, however, that varying dose of the drug, represented by z_i, has been administered to different individuals.
- Then the marginal exchangeability of the responses would no longer hold, but exchangeability of the units would be restored by including the dose (covariate) information in the exchangeability statement as
  $p(Y_1, Z_1, \ldots, Y_n, Z_n) = p(Y_{\rho(1)}, Z_{\rho(1)}, \ldots, Y_{\rho(n)}, Z_{\rho(n)})$.
- This observation is the basis for Bayesian causal reasoning and causality without counterfactuals, employed primarily by Phil Dawid and Vanessa Didelez.

Foundations for causal inference: Summary

- There are many philosophical views on what defines causality, and what can be viewed as a 'cause'.
- Statisticians tend to formalize causality as the estimation of parameters corresponding to population-level contrasts that are not subject to confounding.
- Rubin's counterfactual framework provides a causal model that helps to formalize the definition of confounding, and clarify the estimands of interest.
- Causal estimation can be accomplished by design in a randomized trial setting.
- Causal estimation in the presence of confounding (typically arising from the absence of randomization) requires modelling [next section].

Key points: Summary

- Counterfactuals are "what if" quantities that can help to formalize the definition of a causal estimand.
- Only one counterfactual (of two in the case of a binary exposure, or perhaps infinitely many for continuous exposures) for each individual is ever observable.
- Estimation of causal quantities does not require observation of all counterfactuals, and is based only on observed data.

Selected references

- Holland (1986) Statistics and causal inference. *Journal of the American Statistical Association* **81**:945–960.
- Jaynes (2003) Probability theory: the logic of science. Cambridge University Press, Cambridge.
- Miettinen (2011) Epidemiological research: Terms and concepts. Springer, Dordrecht.
- Miettinen and Karp (2012) Epidemiological research: An introduction. Springer, Dordrecht.
- Pearl (2nd ed., 2009) Causality: models, reasoning, and inference. Cambridge University Press, Cambridge.