Observational Studies: Can we drawn causal inference?

Erin Moshier, MS

Managing Director, *TCI Biostatistics and Clinical Informatics Shared Resource Facility*

Marcio A. Diniz, PhD

Co-Director, *TCI Biostatistics and Clinical Informatics Shared Resource Facility*

November, 30th

Icahn School of Medicine at Mount Sinai

Biostatistics and Clinical Informatics Core at Tisch Cancer Institute

Biostatistics and Clinical Informatics Core

Aims of the Biostatistics team in Tisch Cancer Institute are to:

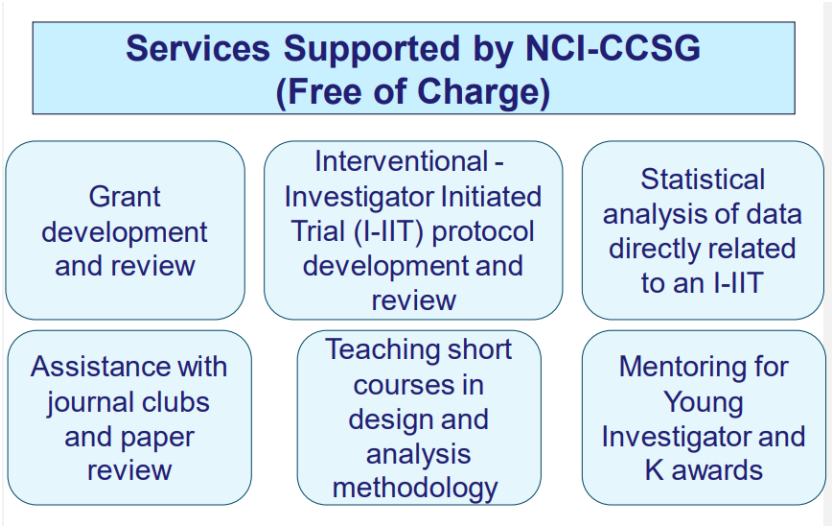
- Establish a scientific and administrative structure that supports investigators from a broad background and creates a collegial environment;

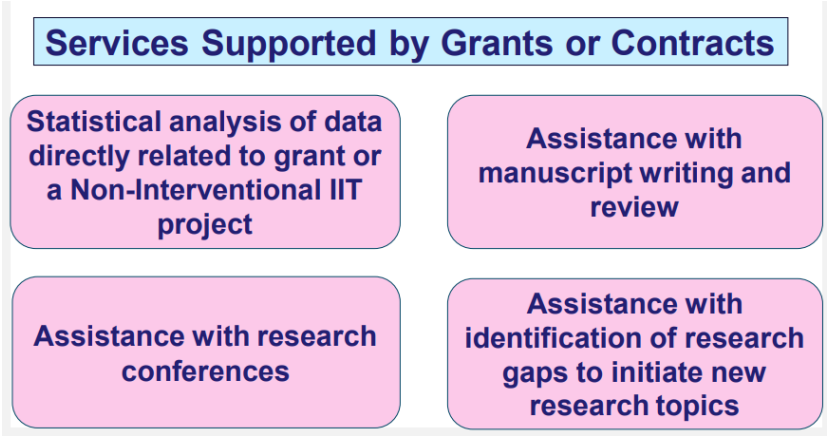
- Provide high-quality consultation in research design and biostatistical analysis;

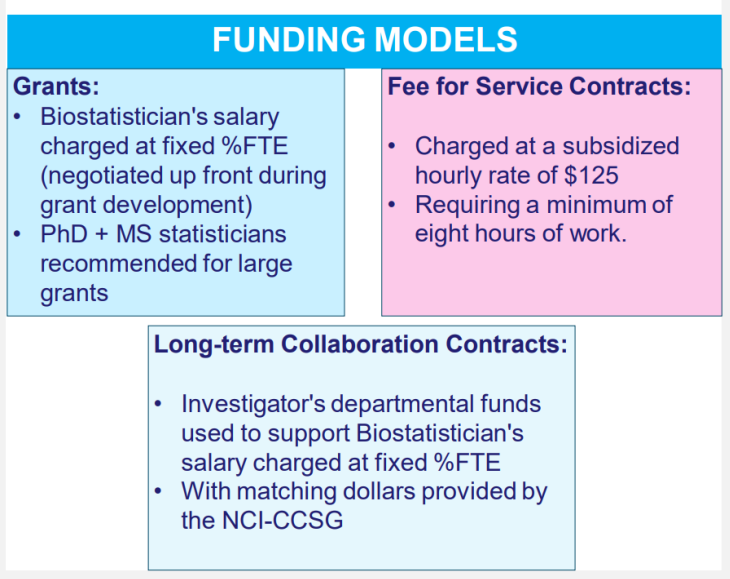
- Train laboratory and clinical investigators in the quantitative aspects of research;

- Support development of innovative statistical methods and promote application of novel analytical methods to collaborative projects.

Biostatistics and Clinical Informatics Core







Statisticians Supporting Contracts

PI	Group	Biostatistician
Cardinale Smith	Hematology & Medical Oncology	Marcio Diniz/Simon Sheng
Joseph Sparano	Hematology Oncology Fellowship	Marcio Diniz/Simon Sheng
Che-Kai Tsao	GU	Himanshu Joshi
Sundar Jagannath	Multiple Myeloma	Erin Moshier/Mayuri Jain
Amy Tiersten	Breast Oncology	Erin Moshier/Asem Berkalieva
Elisa Port	Breast Surgery	Erin Moshier
Kenneth Rosenzweig	Radiation Oncology	Erin Moshier/Mayuri Jain
John Mascarenhas	Myeloproliferative Disorders	Erin Moshier/Grace Van Hyfte
Fred Hirsch	Lung	Marcio Diniz

Access to Biostatistics Core – Cancer Institute Request

https://labs.icahn.mssm.edu/tcibci/

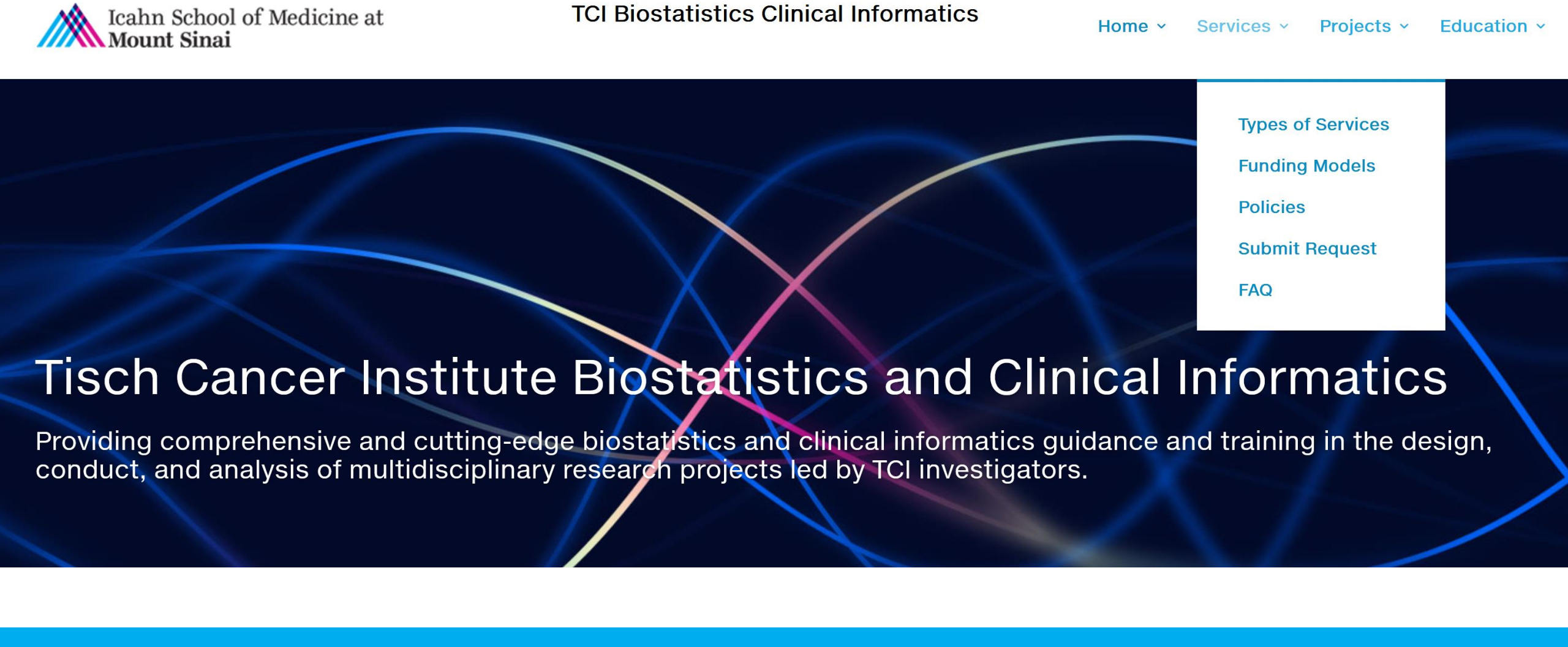
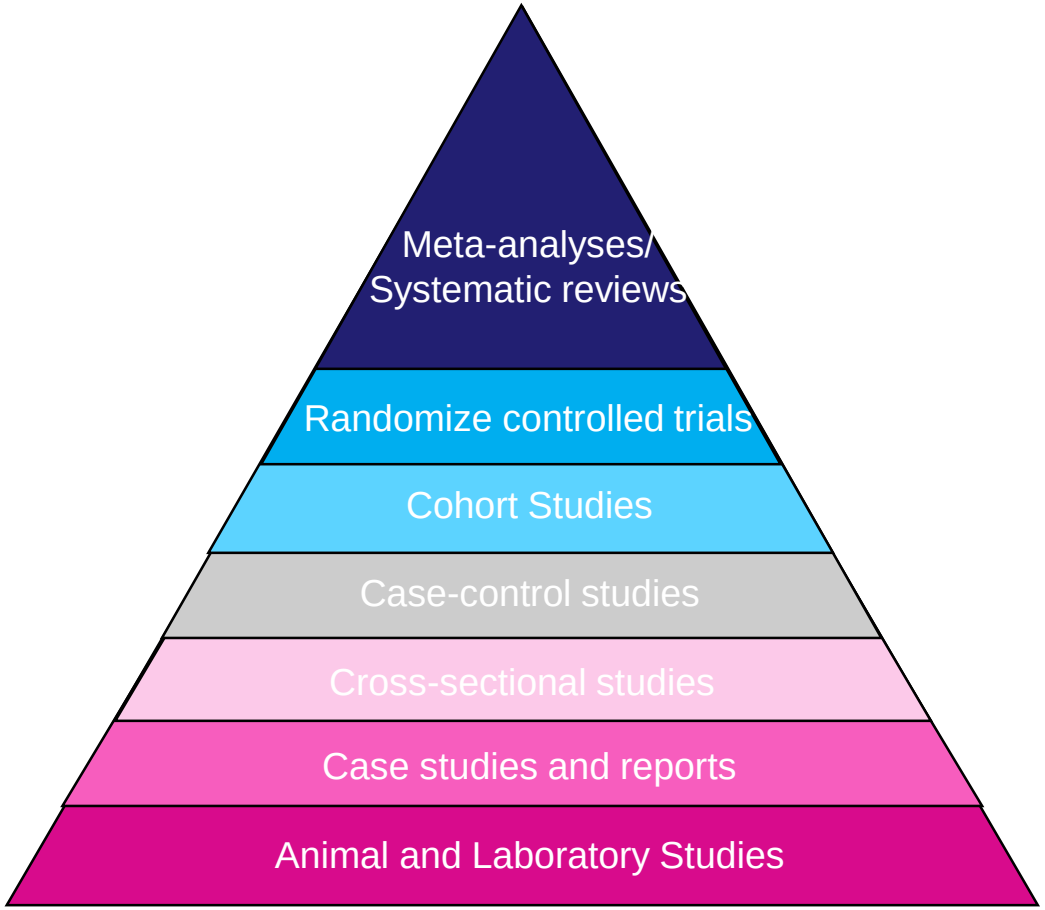


Table of Contents

The challenge of causal inference

Introduction



- In evidence-based medicine, not all evidence is the same;
- Randomized controlled clinical trials and their summary through meta-analyses/systematic reviews are the highest level of evidence to guide treatments to patients;
- Cohorts studies, case-control studies and cross-sectional studies are observational studies with conclusions that should be carefully interpreted.

The challenge of causal inference

- Zeus is a patient with prostate cancer. After six months receiving a new drug, a biopsy shows a complete response;
- Let's imagine that we could somehow – by divine revelation – know that had Zeus received the standard of care, a 6-month biopsy would have shown progressive disease.

- Another patient, Hera also received the new drug with a 6-month biopsy showing complete response.
- Divine revelation also tells us that had Hera not received the new drug, she would still have a 6-month biopsy showing complete response.

- What is the human reasoning about causal effect?
- We compare (usually only mentally) the outcome when an action is taken versus the outcome when the action is withheld. If the two outcomes differ, we say that action has a causal effect, causative or preventive, on the outcome. Otherwise, we say that the action has no causal effect on the outcome.

The challenge of causal inference

Patient	Y(0)	Y(1)	L	A	Y
Rheia	0	1	0	0	0
Kronos	1	0	0	0	1
Demeter	0	0	0	0	0
Hades	0	0	0	0	0
Hestia	0	0	0	1	0
Poseidon	1	0	0	1	0
Hera	0	0	0	1	0
Zeus	0	1	0	1	1
Artemis	1	1	1	0	1
Apollo	1	0	1	0	1
Leto	0	1	1	0	0
Ares	1	1	1	1	1
Athena	1	1	1	1	1
Hephaestus	0	1	1	1	1
Aphrodite	0	1	1	1	1
Polyphemus	0	1	1	1	1
Persephone	1	1	1	1	1
Hermes	1	0	1	1	0
Hebe	1	0	1	1	0
Dionysus	1	0	1	1	0

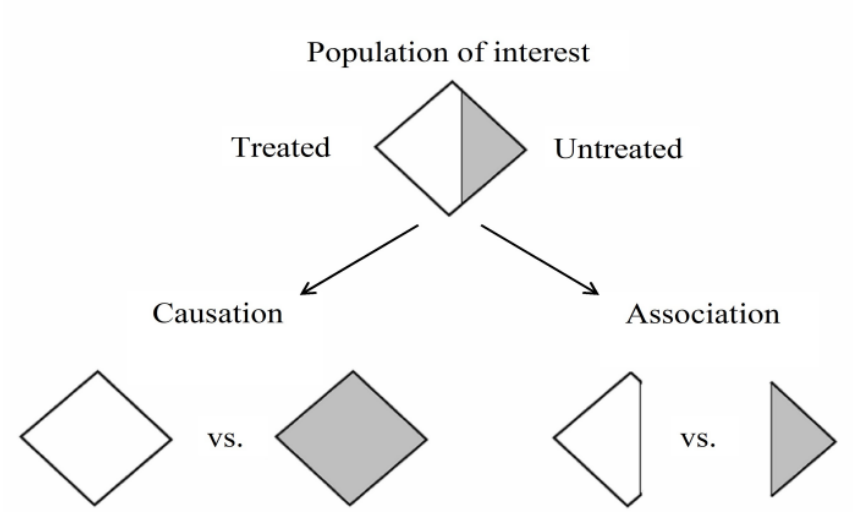
- Researchers are interested in average causal effects instead of individual causal effects.
- If we were able to know the potential outcomes when patients either receive the new drug or standard of treatment, we could state:
  - If all patients had received the new drug, Y(1), the proportion of complete response (Y = 1) is 0.5.
  - If all patients had received the standard of care, Y(0), the proportion of complete response (Y = 1) is 0.5.
- Therefore, the new drug does NOT have a treatment effect on RECIST response criteria.

The challenge of causal inference

Patient	Y(0)	Y(1)	L	A	Y
Rheia	0	1	0	0	0
Kronos	1	0	0	0	1
Demeter	0	0	0	0	0
Hades	0	0	0	0	0
Hestia	0	0	0	1	0
Poseidon	1	0	0	1	0
Hera	0	0	0	1	0
Zeus	0	1	0	1	1
Artemis	1	1	1	0	1
Apollo	1	0	1	0	1
Leto	0	1	1	0	0
Ares	1	1	1	1	1
Athena	1	1	1	1	1
Hephaestus	0	1	1	1	1
Aphrodite	0	1	1	1	1
Polyphemus	0	1	1	1	1
Persephone	1	1	1	1	1
Hermes	1	0	1	1	0
Hebe	1	0	1	1	0
Dionysus	1	0	1	1	0

- In the real world, we are not able to know the potential outcome of patients when they either receive the new drug of standard of treatment;
- Among the patients that received the new treatment (A = 1), the proportion of complete response (Y = 1) is 7/13 = 0.538;
- Among the patients that received the standard of care (A = 0), the proportion of complete response (Y = 1) is 3/7= 0.429;
- Apparently, there is a treatment effect.

The challenge of causal inference



- Causation is a contrast between the average outcome between **all patients had them received** the new drug and **all patients had them received** the standard of care.
- Association is a contrast between the average outcome between **patients that actually received** the new drug and **patients that actually received** the standard of care.
- Randomization makes treatment groups **exchangeable**, except for the treatment they have actually received.
- Consequently, the average outcome observed in the group receiving the new drug would be the same in the entire population;
- Why is there a difference between the conclusions?

The challenge of causal inference

Patient	Y(0)	Y(1)	L	A	Y
Rheia	0	1	0	0	0
Kronos	1	0	0	0	1
Demeter	0	0	0	0	0
Hades	0	0	0	0	0
Hestia	0	0	0	1	0
Poseidon	1	0	0	1	0
Hera	0	0	0	1	0
Zeus	0	1	0	1	1
Artemis	1	1	1	0	1
Apollo	1	0	1	0	1
Leto	0	1	1	0	0
Ares	1	1	1	1	1
Athena	1	1	1	1	1
Hephaestus	0	1	1	1	1
Aphrodite	0	1	1	1	1
Polyphemus	0	1	1	1	1
Persephone	1	1	1	1	1
Hermes	1	0	1	1	0
Hebe	1	0	1	1	0
Dionysus	1	0	1	1	0

- Cancer staging can be considered a prognostic factor (L = 1 high grade, 0 low grade) measured before treatment was assigned;
- Patients could not have been marginally randomized because 69% (9/13) of patients receiving the new drug had high grade cancer stage while only 43% (3/7) of patients that received the standard of care had high grade cancer stage, in other words, the groups (new drug, standard of care) are **not exchangeable**.
- If they had been marginally randomized, we would expect similar proportion of high grade cancer stage in both groups.

The challenge of causal inference

Patient	Y(0)	Y(1)	L	A	Y
Rheia	0	1	0	0	0
Kronos	1	0	0	0	1
Demeter	0	0	0	0	0
Hades	0	0	0	0	0
Hestia	0	0	0	1	0
Poseidon	1	0	0	1	0
Hera	0	0	0	1	0
Zeus	0	1	0	1	1
Artemis	1	1	1	0	1
Apollo	1	0	1	0	1
Leto	0	1	1	0	0
Ares	1	1	1	1	1
Athena	1	1	1	1	1
Hephaestus	0	1	1	1	1
Aphrodite	0	1	1	1	1
Polyphemus	0	1	1	1	1
Persephone	1	1	1	1	1
Hermes	1	0	1	1	0
Hebe	1	0	1	1	0
Dionysus	1	0	1	1	0

- We could consider that patients had been conditionally randomized;
- Among high grade cancer patients, 75% (9/12) of them were randomized to the new drug, while 50% (4/8) of low grade cancer patients were randomized to the new drug;

Hernán, M.A. and Robins, J.M., 2010. Causal inference. Available on https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/

The challenge of causal inference

Patient	Y(0)	Y(1)	L	A	Y
Rheia	0	1	0	0	0
Kronos	1	0	0	0	1
Demeter	0	0	0	0	0
Hades	0	0	0	0	0
Hestia	0	0	0	1	0
Poseidon	1	0	0	1	0
Hera	0	0	0	1	0
Zeus	0	1	0	1	1
Artemis	1	1	1	0	1
Apollo	1	0	1	0	1
Leto	0	1	1	0	0
Ares	1	1	1	1	1
Athena	1	1	1	1	1
Hephaestus	0	1	1	1	1
Aphrodite	0	1	1	1	1
Polyphemus	0	1	1	1	1
Persephone	1	1	1	1	1
Hermes	1	0	1	1	0
Hebe	1	0	1	1	0
Dionysus	1	0	1	1	0

- The groups of new drug and standard of care could be considered exchangeable within each strata (high, low grade cancer stage) assuming that cancer stage was the only unbalanced characteristic that could affect the outcome;
- We say that treatment groups are conditionally exchangeable.
- Association:
  - $P(Y = 1|L = 1, A= 1) = 6/9 = 2/3$
  - $P(Y = 1|L = 1, A = 0) = 2/3$
- Entire Population:
  - $P(Y(1) = 1|L = 1) = 8/12 = 2/3$
  - $P(Y(0) = 1|L = 1) = 8/12 = 2/3$
- Now, the conclusions match among patients with high grade cancer stage.

The challenge of causal inference

Patient	Y(0)	Y(1)	L	A	Y
Rheia	0	1	0	0	0
Kronos	1	0	0	0	1
Demeter	0	0	0	0	0
Hades	0	0	0	0	0
Hestia	0	0	0	1	0
Poseidon	1	0	0	1	0
Hera	0	0	0	1	0
Zeus	0	1	0	1	1
Artemis	1	1	1	0	1
Apollo	1	0	1	0	1
Leto	0	1	1	0	0
Ares	1	1	1	1	1
Athena	1	1	1	1	1
Hephaestus	0	1	1	1	1
Aphrodite	0	1	1	1	1
Polyphemus	0	1	1	1	1
Persephone	1	1	1	1	1
Hermes	1	0	1	1	0
Hebe	1	0	1	1	0
Dionysus	1	0	1	1	0

- The groups of new drug and standard of care could be considered exchangeable within each strata (high, low grade cancer stage) assuming that cancer stage was the only unbalanced characteristic that could affect the outcome;
- We say that treatment groups are conditionally exchangeable.
- Association:
  - $P(Y = 1|L = 0, A = 1) = 1/4$
  - $P(Y = 1|L = 0, A = 0) = ¼$
- Entire population:
  - $P(Y(1) = 1|L = 0) = 2/8 = 1/4$
  - $P(Y(0) = 1|L = 0) = 2/8 = 1/4$
- Now, the conclusions match among patients with low grade cancer stage.

The challenge of causal inference

- In practice, logistic, ethical, or financial constraints can make it difficult or impossible to externally assign treatments, and simple estimates of the treatment effect based on differences can be biased when selection into treatment and control group is not random.
- The treatment effect comparing the new drug with the standard of care is entangled with unbalanced cancer stage. Cancer stage is a confounding factor.
- How can still obtain valid causal inferences?
- We analyze observational studies based on the hope that they can seen as be conditionally randomized experiments.

Statistical Tools

- In observational studies, we often have several confounding factors making stratification by each factor separately not feasible.

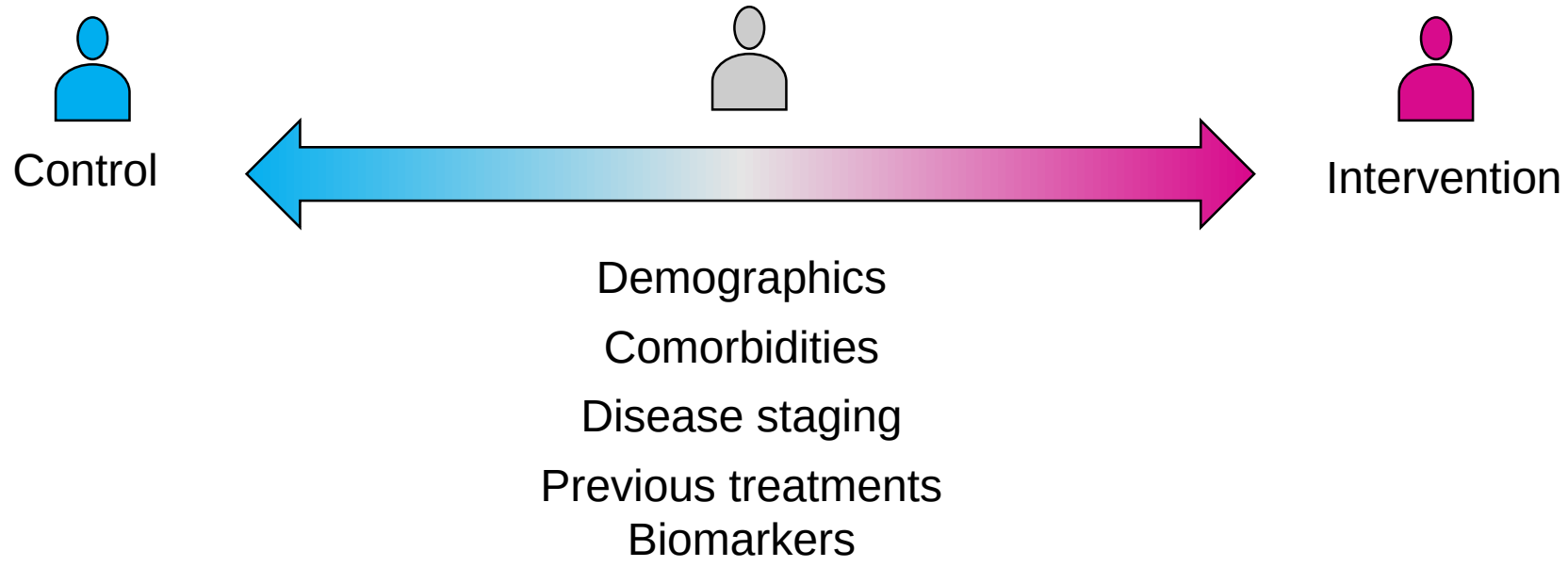
- Regression models can accommodate several confounding factors calculating a conditional treatment effect mimicking conditionally randomized experiment without requiring stratification.

- Multivariable regression models are not the same as multivariate regression models: Hidalgo B, Goodman M. Multivariate or multivariable regression?. American journal of public health. 2013 Jan;103(1):39-40.

- Propensity score can also mimic conditionally randomized experiments. Patients are matched, weighted or stratified based on their propensity to be assigned to intervention/control. The propensity score is often estimated using multivariable regression models.

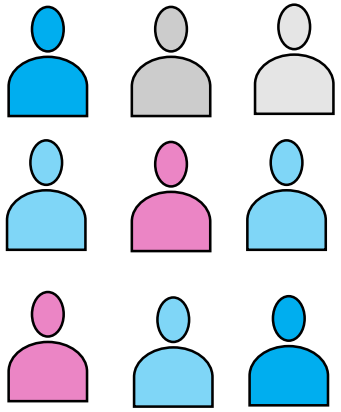
Propensity Score

- Based on data available, we can calculate the propensity a patient has to receive intervention or control in clinical practice (not randomized).



Propensity Score - Matching

Control

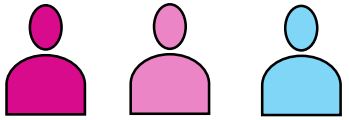


Matching

Control

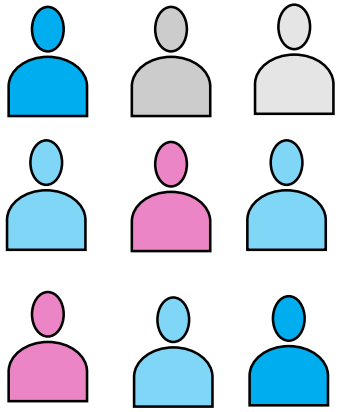
Intervention

Intervention



Propensity Score - Weighting

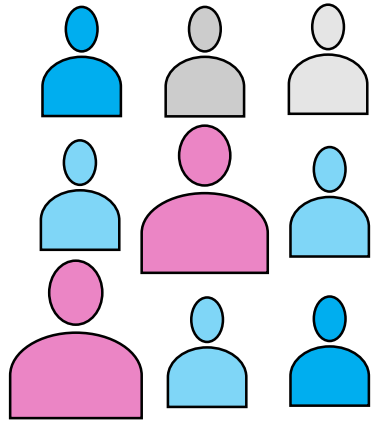
Control



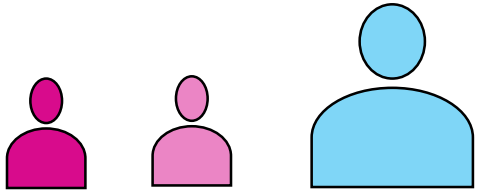
Intervention

Weighting

Control



Intervention



Thank You/Any Questions?

More information, visit:

https://labs.icahn.mssm.edu/tcibci/

Icahn School of Medicine at Mount Sinai