A structural approach to bias:
Causal diagrams provide an internally coherent and transparent approach for observational studies

OR

What you never wanted but needed to know about confounding and didn't even know to ask

Ian Shrier MD, PhD, Dip Sport Med (FACSM)
Centre for Clinical Epidemiology and Community Studies, Jewish General Hospital, McGill University, Montreal Quebec, Canada.
Background

- Why Observational? Limited RCTs with respect to PICO
- Results/Interpretation = Data + Assumptions

“It’s a rather interesting phenomenon. Every time I press this lever, the graduate student breathes a sigh of relief”
Background

- Why Observational? Limited RCTs with respect to PICO
- Results/Interpretation = Data + Assumptions
- Randomized Trial: Does treatment Z reduce mortality?

Causes of Adherence

Randomization → Assigned Rx → Rx Received → Outcome

- Some participants do not adhere to their Rx assignment

“The perfect study exists only in the minds of those who do no research.” (Tim Noakes)
Background

- Why Observational? Limited RCTs with respect to PICO
- Results/Interpretation = Data + Assumptions
- Randomized Trial: Does treatment Z reduce mortality?

Intention to Treat (ITT): treatment assignment
  ⇒ Regulatory Agency: avoids overestimation of effect (vs. placebo…)

Causes of Adherence

Randomization → Assigned Rx → Rx Received → Outcome
IT'T Biased Towards No Effect?

Bias towards the null

Bias away from no effect
Background

• Why Observational? Limited RCTs with respect to PICO
• Results/Interpretation = Data + Assumptions
• Randomized Trial: Does treatment Z reduce mortality?

Randomization → Assigned Rx → Rx Received → Outcome

Causes of Adherence

• Intention to Treat (ITT): treatment assignment
  ⇒ Regulatory Agency: avoids overestimation of effect (vs. placebo…)
  ⇒ Health Policy: requires % adherence (& reasons) = target population

• Patient wants measure of treatment effectiveness
Background

- **Results/Interpretation = Data + Assumptions**
- **Results/Interpretation = Data + Assumptions**
- **Randomized Trial: Does treatment Z reduce mortality?**

Randomization $\rightarrow$ Assigned Rx $\rightarrow$ Rx Received $\rightarrow$ Outcome

- **ITT measures effect of treatment assignment**
  - Regulatory Agency: avoids overestimation of effect (vs. placebo…)
  - Health Policy: requires % adherence (& reasons) = target population

- **Patient wants measure of treatment effectiveness**
  - Analyses based on adherence-data have important assumptions
  - Analyses based on observational data have important assumptions
Causal diagrams and Individual Studies

- Confounding has *always* been focused on causes, not associations
- Similar to logic models, with more explicit assumptions

Cochrane Risk of Bias Tool (observational studies)

- Combining studies that use different regression models
- Bias-amplifying covariates
- Possible modifications
"STANDARD" CONFOUNDER

- Covariate

⇒ Must cause the exposure, or be a marker for a cause of the exposure

- Ex

⇒ Must cause the outcome (independently), or be a marker for an independent cause of the outcome

⇒ Must cause the exposure, or be a marker for a cause of the exposure
POTENTIAL CONFOUNDER?

\[ \begin{align*}
\text{U} & \rightarrow \text{Ex} \rightarrow \text{Outcome} \\
\downarrow & \hspace{2cm} \downarrow \\
\text{C} & \rightarrow \\
\end{align*} \]

\[ \begin{align*}
\text{U} & \rightarrow \text{Ex} \rightarrow \text{Outcome} \\
\downarrow & \hspace{2cm} \downarrow \\
\text{C} & \rightarrow \\
\end{align*} \]

(Hernán Am J Epid 2002)
Must cause the outcome, or be a marker for a cause of the outcome
Must cause the exposure, or be a marker for a cause of the exposure

(Hernán Am J Epid 2002)
Pearl’s Rules - Explanation

If one knows the value of the “collider”, the parents are associated.

If wet: the sprinkler is more likely to be on if there was no rain.

(Shrier & Platt, 2008)
Potential Confounder vs. Collider?

Must cause the outcome, or be a marker for a cause of the outcome
Must cause the exposure, or be a marker for a cause of the exposure

(Cole & Hernán Int J Epid 2002)
COMMON COLLIDER BIASES

Case-control selection bias
- Estrogen
  - Fracture
  - Myo. Infar.

Condition on common effect
- Emerg
  - Hospitalization

Complex Attrition bias
- Treatment
  - Disease
  - Side effects

Condition on common effect
- Drop Out
  - Death
Causal diagrams and Individual Studies

⇒ Epidemiology has *always* focused on causes, not associations

⇒ Similar to logic models, with more explicit assumptions

Cochrane Risk of Bias Tool (observational studies)

⇒ Combining studies that use different regression models

⇒ Allocation Concealment, Placebo Effect
Which measurements should be included in the model if we are interested in the relation between X and Outcome?
Complex Causal DAGs

Which measurements should be included in the model if we are interested in the relation between X and Outcome? Do $Z_1$ and $Z_2$ remove confounding?

Pearl’s Rules: 6-Step Simple Algorithm

(Pearl. Causality Book)
Which measurements should be included in the model if we are interested in the relation between X and Outcome? Do $Z_1$ and $Z_2$ remove confounding?

If $X$ is disconnected from Outcome (d-separation), there is no confounding.
Which measurements should be included in the model if we are interested in the relation between X and Outcome? Do Z₁, Z₂ and Z₃ remove confounding?

(Shrier & Platt, 2008)
Which measurements should be included in the model if we are interested in the relation between X and Outcome? Do Z₁, Z₂ and Z₃ remove confounding?

X is NOT disconnected from Outcome

Including Previous Injury” Introduces Bias!

(Shrier & Platt, 2008)
Which measurements should be included in the model if we are interested in the relation between X and Outcome?

- Coach
- Team motivation, aggression
- Fitness Level
- Genetics
- Connective Tissue Disorder
- Neuromuscular fatigue
- Tissue Weakness
- Pre-game Proprioception
- Contact Sport
- Intra-game proprioception
- Warm-up Exercises (X)
- Injury (Outcome)

Basis for combining covariate sets?

(With different covariate sets!)

Unbiased Covariate Sets?
BIAS-AMPLIFYING COVARIATES

To decrease bias:

\[ \frac{c_4}{c_3} \geq \frac{c_2 c_1}{1 - c_3^2} \]

i.e. best predictor of Exp is most likely to increase bias

<table>
<thead>
<tr>
<th>Allocation generation</th>
<th>Describe the method used to generate the allocation sequence in sufficient detail to allow an assessment of whether it should produce comparable groups.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Was the allocation sequence adequately generated? <strong>For obs. studies, was the allocation based on the indications for treatment, or presence of outcome (introduces bias)?</strong></td>
</tr>
</tbody>
</table>

No changes for:
1. **Allocation Concealment**
2. **Blinding (investigator, participant, assessor)**
3. **Incomplete Outcome Data**
4. **Selective Outcome Reporting**

(Shrier. Res Synth Methods 2012)
<table>
<thead>
<tr>
<th>Other sources of bias.</th>
<th>State any important concerns about bias not addressed in the other domains in the tool. If particular questions/entries were pre-specified in the review’s protocol, responses should be provided for each question/entry.</th>
<th>Was the study apparently free of other problems that could put it at a high risk of bias? In particular, were there any other “co-interventions” by design or association through clustering that could explain the results?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Analytical Procedures</strong></td>
<td><strong>Describe the statistical methods used to minimize bias.</strong></td>
<td><strong>Were appropriate statistical analyses used to minimize bias? A causal diagram outlining the theoretical causal relationships between variables of interest would be beneficial</strong></td>
</tr>
</tbody>
</table>

(Shrier. Res Synth Methods 2012)
Observational studies address treatment effectiveness: patient-oriented analysis

Epidemiology has always focused on causes

Causal diagrams greatly enhance transparency when combining studies that use different adjustment sets

Risk of Bias tool may lead to double-counting of bias, and inappropriate inferences

“Placebo effect” assumes treatment allocation does not affect outcome

Current Risk of Bias tool appropriate for observational studies with slight modifications But still not as good as 2014 version!
REFERENCES

• **Causal Diagrams**

• **Bias Modelling**
Blinding: Placebo Effect

RCT: Uncertainty?

Allocation Process → Exposure Group → True Exposure → Outcome

(Placebo effect)

Participant Beliefs of Exposure Effects

Blinding: Placebo Effect (Shrier, Epidemiology 2013)
Sequence Generation

Unmeasured Factor
(e.g. month of birth)

Sequence Generation

Randomized → Exposure Group → True Exposure → Outcome
Allocation Concealment

Randomized → Exposure Group → True Exposure → Outcome

Poor Research Training → Allocation Not Concealed

Allocator Knowledge of Participant Prognosis → Allocator Cheats

Causal Outcome Factors

(Chaimani et al Effects of study precision and risk of bias in networks of interventions: a network meta-epidemiological study 2013)
Blinding: Assessor

- **E***
- **O***

**Study Methods**
- Allocation Process
- Exposure Group
- True Exposure

**Participant Actions**
- Un-Blinded Assessor
- Outcome

**Bias**
- (Detection bias, Misclassification Bias)
- (Misclassification/recall bias in Case-Control)
Incomplete Outcome Data

Allocation Process → Exposure Group → True Exposure → Outcome

Side Effects Causing Loss to Follow-Up

(Hernán Am J Epid 2002)
Selective Outcome Reporting

- Study Results
- Choice of Study Outcome
- Publication
- Study Biases
- Study Quality
- Outcome (Meta-Analysis Results)
Other Biases: Cluster Effects

Cluster by Time (pre-post), Location

- Clustering Factor
- Co-intervention
- Exposure Group
- True Exposure
- Outcome

Regression to the Mean

- Clustered Subjects
- Decision to do Study
- Exposure Group
- True Exposure
- Outcome

(Probability function)
(chance association)
“STANDARD” CONFOUNDER?

• A variable may (i.e. potential confounder) affect the magnitude or direction of the estimated effect if it is associated with exposure and outcome:
  - Associated with Exposure:
    - is not caused by exposure (e.g. lie along the causal path)
    - is not a marker for a variable caused by exposure
  - Associated with Outcome:
    - is not caused by the outcome
    - is not a marker for a variable caused by the outcome