The “MISSING CAUSE” approach to DETECT AND REDUCE BIAS due to UNMEASURED CONFOUNDING in PHARMACOEPIDEMIOLOGY

Michal Abrahamowicz¹*
James McGill Professor
(1) Department of Epidemiology & Biostatistics, McGill University, Montreal, CANADA

Lise Bjerre², Marie-Eve Beauchamp¹, Jacques LeLorier³, Rebecca Burne¹
(2) University of Ottawa, Canada
(3) Departments of Medicine & Pharmacology, University of Montreal, Canada

Support: Canadian Institutes for Health Research (CIHR) grant 81275
OBJECTIVES

- To Propose a New “Missing Cause” method to Detect & Reduce Bias due to Un-observed Confounding [1] in Non-experimental Studies of Drugs Effects
- To Explain the Missing Cause principle using a simple hypothetical example
- To Compare the New Method versus the Instrumental Variable approach in Simulations
- To Illustrate the method Application using real-life data

Confounding by Indication

• Confounding by Indication is a Major Source of Bias in Observational Studies of Drugs Effectiveness/Safety [Walker et al 1996]

• Such Studies usually rely on Large Administrative Databases (to ensure adequate Power & Precision, especially for Serious Adverse Events (e.g. Mortality or Hospitalization)) [Skegg et al 2001]

• In real-life Clinical Practice, (i) Choice of Drugs or Treatments depends on several Clinical and Socio-demographic patient’s characteristics, that (ii) may also Affect the Outcome, i.e. Act as Potential Confounders of the Treatment Effect

• Yet, Administrative Databases do NOT provide information on such Important Confounders as e.g. Disease Severity, Smoking, Obesity, Blood Pressure or Lab Tests/Biomarkers [e.g., Wolfe 2002]
Instrumental Variables (IV) approach based on (Subjective) Physician Preferences

- Brookhart et al adapted the ‘generic’ Instrumental Variables (IV) approach to control for Un-observed Confounding to Studies of Drugs [Brookhart et al Epidemiology 2006]
- **IV = Subjective Physician’s “Prescribing Preferences”** (well documented in research on drugs utilization) [Sondergaard et al 2006; Wazana et al 2000]
- **ADVANTAGE:** IV approach **Removes Bias due to Unobserved Confounding by Indication** [Brookhart et al 2006; Abrahamowicz et al 2011]
- **LIMITATIONS of IV approach:**
  - Serious **VARIANCE INFLATION** [Ionescu-Ittu et al 2009; 2012]
  - Relies on Linear Risk-Difference models to model Probability of Outcome
  - Hence, **Difficult to Adapt to Time-to-Event analyses of Cohort Studies**

Adapted to Pharmacoepidemiology the ‘generic’ Instrumental Variable (IV) method, popular to deal with Unmeasured Confounding in e.g. econometrics [Angrist, Imbens & Rubin, *J Am Stat Assoc (JASA)* 1996], and more recently in etiologic studies in epidemiology [e.g. Greenland, *Int J Epidemiology* 2000].

- The IV approach involves replacing, in the analyses, the indicator of the treatment actually received by individual patients by another variable; the ‘Instrument’ (IV), which affects the outcome exclusively through its association with the actual treatment and is Not associated with confounders [Angrist et al, *JASA* 1996].

- Brookhart et al define the IV as Prescribing Preferences
Rationale for using Physician Prescribing Preference as an Instrumental Variable

• There is ample evidence that Important differences between individual physicians’ prescribing patterns persist independently of patients’ characteristics

• Such subjective Physician-specific Prescribing Preferences may be due to different factors e.g.:
  ➢ Training and/or Past experience with alternative drugs
  ➢ access to recent Results of influential Clinical Trials
  ➢ (last not least) Visit from a Pharmaceutical company Representative
Rationale for using Physician Prescribing Preference as Instrumental Variable (cont-d)

Physician-specific Prescribing Preferences are likely to Meet Crucial IV Assumptions:

- As **Physician-specific Prescribing Preferences** affect Treatment of **All** Patients of a given physician, they should NOT be associated with individual patient’s characteristics (including measured & Unmeasured Confounders) [Brookhart et al. *Epidemiology* 2006].

- Subjective **Prescribing Preferences** affect the treatment choice for individual patients, over and above patient characteristics, **BUT – per se –** they cannot affect the clinical outcome independently of the actual treatment. Thus, **any association between Prescribing Preferences (IV) and the Outcome must be Mediated through the Treatment Actually Received** by the patient.

- **IF the Assumptions are met: IV estimates are Un-Biased, even in the presence of Strong Unmeasured Confounding** [Ionescu-Ittu, Delaney & Abrahamowicz. *Pharma & Drug Saf (PDS)* 2009]
Conceptual Framework:
Determinants of Treatment Choices

- **Recorded Patient Characteristics**
- **Un-Recorded Patient Characteristics**
- **Physician Prescribing Preferences**
- **Treatment Choice**
- **Outcome**
How to Define the IV in the Analyses?

- because *Prescribing Preferences are NOT Observable*:
  - Brookhart et al [*Epidemiology* 2006] define the IV for patient “j” of physician “i” as the Binary Indicator of the Treatment received by the Previous Patient (j-1) of the Same Physician (“BINARY IV”)
  - Alternative: define IV as the % of All Previous Patients (of the Same Physician) who received e.g. Drug “B” (vs Drug “A”) (“CONTINOUS IV”)

Limitation of the IV methods:
Variance Inflation ⇒ Unstable estimates

- If Prescribing Preferences are Weak:
  the IV estimates show high **Variance Inflation** implying:
  (a) very **wide confidence intervals** &
  (b) **Mean Squared Errors (MSE)** which may be **Higher**
  than MSE of biased but much more stable
  ‘conventional’ estimates that adjust only for the
  measured confounders

[Ionescu-Ittu, Delaney & Abrahamowicz, *Pharma & Drug Saf (PDS)* 2009]

(a) may make IV estimates practically **Un-interpretable**
(b) implies that, **in a given real-life study**, the IV estimate
     may be more ‘‘Biased’’ than the conventional one
**COX-2 vs NSAIDs Example:**

Inflated (very Wide) CI’s for the IV estimates

**Differences in risk (RD) of gastrointestinal events between COX-2 inhibitor vs. NSAID users**

<table>
<thead>
<tr>
<th>Model</th>
<th>Treatment RD (%) for COX-2 vs. NSAID users (95% CI) a</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unadjusted</td>
<td>6.4 (4.5, 8.3)</td>
</tr>
<tr>
<td>Conventional Multivariable (Fully Adjusted) Model</td>
<td>-1.9 (-3.5, -0.3)</td>
</tr>
<tr>
<td>Binary IV model (IV = Previous Pt RX)</td>
<td>3.3 (-12.7, 19.3)</td>
</tr>
<tr>
<td>Continuous IV model 3 (IV = % of Treated among All previous Pts.)</td>
<td>2.5 (-6.2, 11.3)</td>
</tr>
</tbody>
</table>
Objectives of our new “Missing Cause” approach

We propose a New Method that uses Prescribing Preferences in a Different way from the IV methods in order to, simultaneously:

(1) detect Bias due to Unmeasured Confounding;
(2) largely Reduce this Bias;
(3) ensure More Stable i.e. More Precise estimates of Treatment effect than the IV’s

Conceptual Framework:
the “MISSING CAUSE” Principle

- For individual patients, we cannot determine if unobserved confounders did or did not affect the choice of their treatment.

- However, the **“Missing Cause” principle** postulates that:
  If individual patient’s Treatment Choice cannot be explained by Observed Variables (including Physician Preferences) then it was likely influenced by some UN-Observed Characteristics.

- Conversely: the probability that a patient’s treatment was (partly) based on un-observed characteristics increases if this treatment appears inconsistent with the observed patient’s characteristics and his/her physician’s prescribing preferences.

- Thus, our new method attempts to use the information on the Discrepancies between (i) observed Versus (ii) expected treatment.
Conceptual Framework: Determinants of Treatment Choices

- Recorded Patient Characteristics
- Un-Recorded Patient Characteristics
- Physician Prescribing Preferences
- Treatment Choice
- Outcome
2 Implications of the “MISSING CAUSE” Principle

**Assumption:** Un-measured confounder U increases the probability of the patient being prescribed drug B. Then:

(I1) Among the **users of drug B**, the prevalence of **Un-observed Confounder U** will be higher for **“Discrepant patients”**, who appear less likely to receive this drug (i.e. their Actual Treatment) given their observed characteristics and their physicians preferences (*).

- According to the above Assumption: for these patients, U provides the “Missing Cause” for prescribing drug B.

(I2) In contrast, among the **users of drug A**, the prevalence of **U** will be **lower** for **“Discrepant patients”**, i.e. those who would be expected to get drug B but instead received drug A. (**)

** According to the above Assumption: **Presence of U should further increase the probability of getting drug B.**
Conceptual Framework: Treatment Discrepancy as a Source of Variation in the Estimated Treatment Effects

• Taken together, statements (I1) and (I2) imply that:
The difference in the prevalence of U between groups of patients prescribed (i) drug A versus (ii) drug B will increase for more discrepant patients (***)
[(***) Prevalence of U will (i) decrease among discrepant users of drug A, and (ii) increase among discrepant users of drug B.]

• On the other hand, the strength of Un-measured Confounding Bias, due to a failure to adjust the Treatment effect for U, will also increase with increasing difference in the prevalence of U

• Thus, if U acts as an unobserved confounder, then the estimated treatment effect will vary systematically across different levels of discrepancy (between observed versus expected treatments), implying a Treatment*Discrepancy Interaction
Hypothetical Example to Illustrate the ‘Missing Cause’ concepts

• STUDY DESIGN:
  A study estimates the effect of a Binary Treatment: T=0 (“Old drug”) vs T=1 (“New drug”)
  Study patients are divided across 4 Physicians (MD’s): A-D; each with a Different `Preference` for old vs new drug, and each with 1,000 patients

COVARIATES:
• 1 Observed Binary Treatment Determinant (X), that is NOT a risk factor (thus, is Not Adjusted for in the analyses)
• 1 Un-observed Binary Confounder (U)

OUTCOME (Adverse Event):
• Binary Outcome (Y=1 vs Y=0), which depends on T and U only (but NOT on X or MD).
Hypothetical Example: Assumptions used to construct the data

- Prevalence of the measured covariate \( X \): \( P(X=1) = 0.4 \); 
- Prevalence of the Un-measured Confounder \( U \): \( P(U=1) = 0.3 \); 
- \( U \) and \( X \) are Independent of each other (OR = 1), and have the Same prevalence across patients of each of the 4 MD’s; 
- Probability of Treatment (\( T=1 \), i.e. of receiving a “new drug”) is a Linear function (consistent with the Risk Difference model [Brookhart et al, 2006]) & depends on \( X \), \( U \) and MD’s preferences:

\[
P(T=1) = 0.2 + 0.3 \, X + 0.3 \, U + 0\, A + 0.2\, B + 0.1\, C - 0.2\, D
\]

[Equation 1]

where: \( A \) – \( D \) are 4 dummy indicators of individual physicians.
Hypothetical Example (continued):
Assumptions used to construct the data

- Probability of Outcome (Y=1) depends only on T & U (but Not on X or MD’s preferences), through a Risk Difference (RD) model:

\[ P(Y=1) = 0.2 + 0.3 \times U - 0.1 \times T \]  

[Equation 2]

- Notice:
  - “U” is a Strong Un-observed Confounder, as it increases the Probability of Both T=1 [Eq 1] and Y=1 [Eq 2] by 0.3!
  - Treatment (T=1) Decreases Prob(Y=1) by 0.1 (true RD = -10%)
Let’s consider e.g. the Stratum $\{MD=C; X=1; U=1\}$

- Because $P(X=1)= 0.4$ & $P(U=1)= 0.3$ & $U$ & $X$ are independent of each other: $P(X=1 \& U=1) = 0.3 \times 0.4 = 0.12$
- Thus, among a Total of 1,000 pts for each MD, the stratum $\{X=1; U=1\}$ has $N = 0.12 \times 1000 = 120$ pts.
- Because coefficient for MD = C is + 0.1 in [Eq. 1], $P(T=1| MD=C; X=1; U=1) = 0.9$ from [Eq. 1]
- Thus, among $N=120$ pts. in stratum $\{MD=C; X=1; U=1\}$: $N(T=1) = 0.9 \times 120 = 108$ pts. receive $T=1$ while $N(T=0) = 0.1 \times 120 = 12$ pts. receive $T=0$
Similar calculations for Stratum \( \{MD=C; X=1; U=0\} \):

- \( P(X=1 \& U=0) = 0.4 \times 0.7 = 0.28 \)
- Thus, the stratum \( \{X=1; U=0\} \) has \( N = 0.28 \times 1000 = 280 \) pts.
- \( P(T=1| MD=C; X=1; U=0) = 0.6 \) from [Eq. 1]
- Thus, among \( N=280 \) pts. in stratum \( \{MD=C; X=1; U=0\} \):
  \( N (T=1) = 0.6 \times 280 = 168 \) pts. receive \( T=1 \) while
  \( N (T=0) = 0.4 \times 280 = 112 \) pts. receive \( T=0 \)
Re-constructing “Observed” Treatment Frequencies (when U is NOT Measured)

• Because U is NOT Measured, in the “Observed Data” (Available for research), 2 Strata (U=0 & U=1) for given \{MD;X\} will be Combined:

  - e.g. in \{MD=C;X=1\}, the Observed Frequencies of will be:
    \[ N(T=0) = 112 \times [N(T=0|U=0)] + 12 \times [N(T=0|U=1) = 124] \]
    \[ N(T=1) = 168 \times [N(T=1|U=0)] + 108 \times [N(T=1|U=1) = 276] \]

  - with the resulting (Un-Observed) Prevalence of \( U=1 \):
    \[ P(U=1|MD=C;X=1 & T=0) = \frac{12}{124} = 9.7\% \]
    \[ P(U=1|MD=C;X=1 & T=1) = \frac{108}{276} = 39.1\% \]

  ** this Big Difference (39.1% vs 9.7%) in the Prevalence of important Un-measured Risk Factor (U) produces CONFOUNDING BIAS **
Hypothetical Example: Calculating Treatment Discrepancies (D)

- We calculate **Treatment Discrepancy (D)**, for a given patient, as the **Absolute Difference** between 
  [Treatment Actually Received (T=0 or T=1)] and 
  [Probability of receiving T=1]:

\[ D = |T - \text{Prob}(T=1)| \]

- e.g. for 420 (124+276) pts. in the **stratum** \{MD=C;X=1\}: previous slide shows that \( \text{Prob}(T=1) = \frac{276}{420} = 0.69 \)
  - for each of 124 pts with \( T=0 \): \( D = |0 - 0.69| = 0.69 \)
  - for each of 276 pts with \( T=1 \): \( D = |1 - 0.69| = 0.31 \)
- **NOTE: Next Slide** is based on calculations similar to the previous & this slide, which yielded: (a) Observed Prevalence of U (y-axis) & (b) Treatment Discrepancy (“D” on x-axis) for all strata of \{MD,X,T\}. Pts. were then grouped by 7 intervals of D (0.05-0.25;0.25-0.35...etc.)
Hypothetical Example: **Difference in Prevalence of U between Treatment groups** (T=0: Green vs T=1: Red) Increases with Increasing Discrepancy D!
Hypothetical Example:
Calculating strata-specific Risk of Outcome: P(Y=1)

We can use [Eq. 2] \( P(Y=1|U;T) = 0.2 + 0.3 \times U - 0.1 \times T \)
to calculate “Risk” \( P(Y=1) \) and, then **Number of pts. with Y=1**
\( (P(Y=1)*N) \) for each subgroup \((T=0 \text{ or } T=1)\) in different strata.

- e.g. in subgroup \{MD=C;X=1; T=0\}:
  - for 112 pts. with \( T=0, U=0 \): \( P(Y=1)=0.2 \rightarrow N(Y=1)=0.2 \times 112=22 \)
  - for 12 pts. with \( T=0, U=1 \): \( P(Y=1)=0.5 \rightarrow N(Y=1)=0.5 \times 12=6 \)
  - \( T=0: 124 \text{ pts.}, 28(22+6) \text{ with } Y=1 \Rightarrow P(Y=1|T=0)=28/124=22.6\% \)

- similar, in subgroup \{MD=C;X=1; T=1\}:
  - for 168 pts. with \( T=1, U=0 \): \( P(Y=1)=0.1 \rightarrow N(Y=1)=0.1 \times 168=17 \)
  - for 108 pts. with \( T=1, U=1 \): \( P(Y=1)=0.4 \rightarrow N(Y=1)=0.4 \times 108=43 \)
  - \( T=1: 276 \text{ pts.}, 60(17+43) \text{ with } Y=1 \Rightarrow P(Y=1|T=1)=60/276=21.7\% \)

- We then assigned the observed Numbers of Pts. with Y=1, separately,
  for \( T=0 \) and \( T=1 \), to Discrepancy \(|D|\) calculated as on a previous slide,
  and calculated the resulting RISK DIFFERENCE (\( T=1 \) vs \( T=0 \)) for
different categories of \(|D|\)
Hypothetical Example: Observed (BIASED) Risk Difference (T=1 vs T=0) Increases with Increasing D => Treatment*Discrepancy (T*D) Interaction
using previous calculations of the Number of pts. with $Y=1$ ($P(Y=1)\times N$) for each subgroup in stratum \{MD=C; X=1\}, we can now show how the Confounding BIAS is a direct function of difference in Prevalence of U between subgroups with $T=1$ vs $T=0$

- we calculated before that:

  $T=0$: 124 pts., 28(22+6) with $Y=1$ => $P(Y=1|T=0)=28/124 = 22.6\%$
  $T=1$: 276 pts., 60(17+43) with $Y=1$ => $P(Y=1|T=1)=60/276= 21.7\%$

- thus:

  **RISK DIFFERENCE ($T=1$ vs $T=0$): 21.7% - 22.6% = - 0.9 %**

  **Notice that -0.9% is close to Risk Difference (RD) expected given:

  (i) True RD = -10%,
  (ii) 29.4% diff. In Prev. of U (39.1% in T=1 -9.7% in T=0),
  (iii) U increases $P(Y=1)$ by 30% in [Eq. 2]:

  $-10\% + 0.294\times 30\% = - 1.2\%$ ** (very close to -0.9%)**
Under the ‘Missing Cause’ assumption: Un-observed Confounder(s) should induce an artificial *Interaction* between (a) the Treatment actually received and (b) a measure of Treatment Discrepancy (observed *versus* ‘expected’ treatment) [Abrahamowicz et al, *Statistics-in-Medicine* 2016]

Therefore, we propose to:

use the

Treatment-by-Discrepancy Interaction

as an analytical tool to

Detect and Control for Unobserved Confounding
Implementation of the “Missing Cause” method: Data Pre-Processing

• Similar to Prescribing Preference-based IV approach of Brookhart et al [Epidemiology 2006],

to Implement our method, data analyst must first:

  ➢ Assign each Patient to a Single Prescribing Physician (who wrote the “Index Rx”)
  ➢ ensure the dataset includes several physicians who prescribed (one of or both) drugs being compared to n>2 patients (preferably n>10 pts/physician)
1. Estimating expected Probability of Treatment:

\[ P\left(T_{ij} = 1 \mid X_{ij,1}, \ldots, X_{ij,p}, M_{ij}\right) = \sum_{k=1}^{p} \beta_k X_{ij,k} + \sum_{i=1}^{m} \gamma_i M_{ij} \]  

where:

- \( T_{ij} \) is the binary indicator of Treatment actually assigned to the \( j^{th} \) patient \((j=1, \ldots, n_j)\) of the \( i^{th} \) physician \((i=1, \ldots, m)\)
- \( X_{ij,k} \) is the value of the observed covariate \( X_k \) \((k=1, \ldots, p)\) for this patient
- \( M_{ij} \) are dummy indicators of individual Prescribing Physicians
- \( \gamma_i \) estimates the Preference of \( i^{th} \) physician for \( T_{ij} = 1 \), independent of his/her patients’ characteristics.
2. Estimating Treatment Discrepancy (between Treatment Actually Prescribed VS ‘Expected Treatment’) :

\[ D_{ij} = \left| T_{ij} - P(T_{ij} = 1) \right| \]  
(2)

where:

- \( T_{ij} \) indicates the treatment actually received (0 or 1),
- \( P(T_{ij} = 1) \) is the Expected Prob. of \( T_{ij} = 1 \) estimated in equation (1)
- \( D_{ij} \) can range from 0 to 1
  - for patients with \( T_{ij} = 1 \), \( D_{ij} \) increases as \( P(T_{ij} = 1) \) decreases
  - for patients with \( T_{ij} = 0 \), \( D_{ij} \) increases with increasing \( P(T_{ij} = 1) \), i.e. decreasing probability of receiving their actual treatment \([P(T_{ij} = 0) = 1 - P(T_{ij} = 1)]\)
  - thus, \( D_{ij} \) quantifies the Treatment Discrepancy for patients prescribed Either Treatment.
3. Estimating Treatment-by-Discrepancy Interaction:

- We expand the multivariable model for the Outcome to include the INTERACTION between (i) Treatment actually prescribed ($T_{ij}$) and (ii) treatment Discrepancy $D_{ij}$, estimated in (2).

- Thus, we fit the following RD regression model:

$$P(Y_{ij} = 1 | X_{ij,1}, ..., X_{ij,p}, T_{ij}, D_{ij}) = \alpha_0 + \sum_{k=1}^{p} \alpha_k X_{ij,k} + \omega T_{ij} + \eta D_{ij} + \theta\left(T_{ij} D_{ij}\right)$$ (3)

Where $P(Y_{ij} = 1 | X_{ij,1}, ..., X_{ij,p}, T_{ij})$ is the individual patient’s probability of $Y_{ij} = 1$, conditional on $p$ covariates ($X_{ij,1}, ..., X_{ij,p}$) (But NOT on Physician’s Preference) and on the received treatment $T_{ij}$.

- Our model in eq. (3) differs from the conventional RD model only in that it includes the treatment*discrepancy interaction $T_{ij} D_{ij}$ and the main effect of $D_{ij}$.
Statements (A), (B) & (C) imply that:
variation of estimated treatment effect across different levels of
treatment discrepancy (Treatment*Discrepancy Interaction) will
occur if and only if this effect is truly affected by un-observed
confounder(s).

Thus: **we use the Test of the Significance of the**
**Treatment*Discrepancy Interaction as an Empirical Criterion to**
**Verify if the Treatment Effect estimate in model in (3) is affected**
**by Un-observed Confounding**

If we reject the H0: $\theta = 0$ in (3) at $\alpha = 0.05$ then we conclude that
Un-observed Confounding is likely (***)

(*** the underlying assumption is that Treatment Effect does Not depend on
important Determinants of Treatment Choice (i.e. No Un-accounted for
Interactions with determinants)
Notice that in our Interaction-based model in (3): the Treatment Effect $\omega$ represents the hypothetical effect (adjusted for all the observed covariates) for the subjects with No Treatment Discrepancy ($D_{ij} = 0$) between their Observed Treatment ($T_{ij}$) and ‘Expected Treatment’ ($P(T_{ij}) = 1$).

It can be demonstrated that for those subjects, their Observed Treatment would be Fully Explained by their Physician Preferences and, thus, the estimated Treatment Effect would be NOT Affected by any potential Un-observed Confounders.

Thus: we use $\omega$ in (3) as the estimator of the ‘corrected’ treatment effect (RD) that would be expected in the Absence of Un-observed Confounding.

[Abrahamowicz et al. *Statistics in Medicine* 2016]
Objectives of Simulations:
To Compare our Missing Cause-based Interaction model with:
- Conventional model, which adjust only for observed covariates
- Instrumental Variables (IVs), with IV defined as either:
  a. Treatment received by the previous patient of the same physician (Binary IV model)
  b. Proportion of all previous patients of the same physician prescribed treatment $T=1$ (Continuous IV model)
Simulations Design

- Cohort study comparing Time to 1st Adverse Event between 2 Drugs A ($T_{ij} = 0$) vs B ($T_{ij} = 1$);
- $m = 500$ prescribing physicians, 10-30 patients/physician
- Total N = about 10,000 patients
- 14 Individual Patients’ Characteristics:
  > 11 “Observed” covariates (recorded in the Database)
  + 3 UN-Observed Confounders (associated with Both Treatment & Outcome)
- Binary Treatment indicators ($T_{ij}$) generated conditional on all relevant Patients’ Characteristics (observed & unobserved) and Physician’s Prescribing Preference
Table 1. Simulation results: Empirical power (%) for detecting unmeasured confounding

<table>
<thead>
<tr>
<th>Selected scenarios</th>
<th>N (# events)</th>
<th>Unmeasured confounding</th>
<th>Relative bias of conventional treatment estimate (Model 1)</th>
<th>95% CI of Binary IV Model 2&lt;sup&gt;a&lt;/sup&gt;</th>
<th>95% CI of Continuous IV Model 3&lt;sup&gt;a&lt;/sup&gt;</th>
<th>95% CI of Interaction Model 4&lt;sup&gt;a&lt;/sup&gt;</th>
<th>p-value &lt; 0.05 for Interaction test (Model 4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Column 1</td>
<td>Column 2</td>
<td>Column 3</td>
<td>Column 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>17,993</td>
<td>Moderate</td>
<td>0.589</td>
<td>10.9</td>
<td>30.7</td>
<td>51.4</td>
<td>55.1</td>
</tr>
<tr>
<td></td>
<td>(6,556)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>29,986</td>
<td>Moderate</td>
<td>0.589</td>
<td>11.0</td>
<td>48.2</td>
<td>72.6</td>
<td>75.8</td>
</tr>
<tr>
<td></td>
<td>(10,918)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Percentage of samples where the 95% confidence interval for treatment effect of the given model excluded the conventional Model 1 estimate.
### Empirical Power for detecting unmeasured confounding

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Outcome</th>
<th>N (# events)</th>
<th>Unmeasured confounding</th>
<th>Relative bias of conventional treatment estimate (Model 1)</th>
<th>Power (%) for detecting unmeasured confounding based on criteria: c</th>
<th>95% CI of binary IV Model 2 d</th>
<th>95% CI of continuous IV Model 3 d</th>
<th>95% CI of interaction Model 4 d</th>
<th>p-value &lt; 0.05 for two-tailed interaction test (Model 4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Column 1</td>
<td>Column 2</td>
<td>Column 3</td>
<td>Column 4</td>
<td>Column 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Binary</td>
<td>11,992 (1,370)</td>
<td>NO</td>
<td>0.001</td>
<td>4.8 c</td>
<td>4.7 c</td>
<td>3.9 c</td>
<td>4.9 c</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Binary</td>
<td>12,016 (2,964)</td>
<td>NO</td>
<td>0.002</td>
<td>4.4 c</td>
<td>3.4 c</td>
<td>5.7 c</td>
<td>5.7 c</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Binary</td>
<td>12,000 (3,170)</td>
<td>Weak</td>
<td>0.349</td>
<td>5.2</td>
<td>11.1</td>
<td>15.7</td>
<td>18.5</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Binary</td>
<td>29,975 (7,922)</td>
<td>Weak</td>
<td>0.353</td>
<td>9.9</td>
<td>24.9</td>
<td>39.2</td>
<td>42.6</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Binary</td>
<td>17,993 (6,556)</td>
<td>Moderate</td>
<td>0.589</td>
<td>10.9</td>
<td>30.7</td>
<td>51.4</td>
<td>55.1</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Binary</td>
<td>29,986 (10,918)</td>
<td>Moderate</td>
<td>0.589</td>
<td>11.0</td>
<td>48.2</td>
<td>72.6</td>
<td>75.8</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Binary</td>
<td>30,020 (9,459)</td>
<td>Moderate</td>
<td>NA c</td>
<td>15.5</td>
<td>51.2</td>
<td>74.8</td>
<td>77.5</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Binary</td>
<td>11,989 (3,156)</td>
<td>Weak</td>
<td>-0.335</td>
<td>6.4</td>
<td>11.4</td>
<td>16.4</td>
<td>18.7</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Continuous</td>
<td>11,994</td>
<td>Weak</td>
<td>-0.235</td>
<td>7.5</td>
<td>20.5</td>
<td>32.1</td>
<td>34.8</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Continuous</td>
<td>29,996</td>
<td>Weak</td>
<td>-0.234</td>
<td>11.2</td>
<td>42.8</td>
<td>64.0</td>
<td>67.2</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Continuous</td>
<td>11,993</td>
<td>Moderate</td>
<td>-0.468</td>
<td>13.7</td>
<td>51.6</td>
<td>73.3</td>
<td>77.0</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Continuous</td>
<td>30,000</td>
<td>Moderate</td>
<td>-0.470</td>
<td>30.4</td>
<td>89.7</td>
<td>99.0</td>
<td>99.1</td>
<td></td>
</tr>
</tbody>
</table>

**a** See Table A.1 in Supplementary Materials for details about each scenario.

**b** Mean total sample size N, across the 1,000 simulated samples, and mean number of events for the binary outcome in scenarios 1-8.

**c** For scenarios 1 and 2, the proportion of samples where the null hypothesis was rejected represents in fact the type I error because there is no unmeasured confounding.

**d** Percentage of samples where the 95% confidence interval for treatment effect of the given model excluded the conventional Model 1 estimate.

**e** True effect = 0.
Table 2. Simulation results: Bias, standard deviation (SD), and root mean square error (RMSE) of treatment effect estimates

<table>
<thead>
<tr>
<th>Scenario a</th>
<th>Outcome</th>
<th>True effect of A b</th>
<th>Bias Conventional Model 1</th>
<th>Bias Interaction Model 4</th>
<th>SD ratio (relative to SD of interaction Model 4)</th>
<th>RMSE ratio (relative to RMSE of interaction Model 4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Binary</td>
<td>0.10</td>
<td>0.00</td>
<td>0.00</td>
<td>0.30</td>
<td>3.83</td>
</tr>
<tr>
<td>2</td>
<td>Binary</td>
<td>0.10</td>
<td>0.00</td>
<td>0.003</td>
<td>0.23</td>
<td>2.36</td>
</tr>
<tr>
<td>3</td>
<td>Binary</td>
<td>0.10</td>
<td>0.035</td>
<td>0.007</td>
<td>0.29</td>
<td>3.51</td>
</tr>
<tr>
<td>4</td>
<td>Binary</td>
<td>0.10</td>
<td>0.035</td>
<td>0.006</td>
<td>0.29</td>
<td>3.67</td>
</tr>
<tr>
<td>5</td>
<td>Binary</td>
<td>0.10</td>
<td>0.059</td>
<td>0.011</td>
<td>0.29</td>
<td>3.71</td>
</tr>
<tr>
<td>6</td>
<td>Binary</td>
<td>0.10</td>
<td>0.059</td>
<td>0.011</td>
<td>0.29</td>
<td>3.57</td>
</tr>
<tr>
<td>7</td>
<td>Binary</td>
<td>0.00</td>
<td>0.059</td>
<td>0.011</td>
<td>0.29</td>
<td>3.57</td>
</tr>
<tr>
<td>8</td>
<td>Binary</td>
<td>0.10</td>
<td>-0.036</td>
<td>-0.008</td>
<td>0.29</td>
<td>3.78</td>
</tr>
<tr>
<td>9</td>
<td>Continuous</td>
<td>-5.00</td>
<td>1.173</td>
<td>0.191</td>
<td>0.28</td>
<td>3.59</td>
</tr>
<tr>
<td>10</td>
<td>Continuous</td>
<td>-5.00</td>
<td>1.172</td>
<td>0.218</td>
<td>0.28</td>
<td>3.50</td>
</tr>
<tr>
<td>11</td>
<td>Continuous</td>
<td>-5.00</td>
<td>2.340</td>
<td>0.462</td>
<td>0.30</td>
<td>3.89</td>
</tr>
<tr>
<td>12</td>
<td>Continuous</td>
<td>-5.00</td>
<td>2.348</td>
<td>0.436</td>
<td>0.29</td>
<td>3.56</td>
</tr>
</tbody>
</table>

a Columns 2 to 4 in Table 1 describe the main features of each scenario. See Table A.1 in Supplementary Materials for details about each scenario.

b Risk difference for the binary outcome in scenarios 1 to 8, difference in the mean values of a continuous outcome for scenarios 9 to 12.
Results of 1,000 Simulations (N=18,000; Moderate Confounding): relative Bias (%)
Results of 1,000 Simulations (N=18,000; Moderate Confounding): **Standard Deviation**

![Bar chart showing simulation results: Standard Deviations](chart_image)

- **Models**:
  - Conventional
  - IV Binary
  - IV Continuous
  - Interaction

- **Standard Deviation**
  - IV Binary has the highest standard deviation
  - Standard deviation values for other models are lower.
Results of 1,000 Simulations (N=18,000; Moderate Confounding): \textbf{root Mean Squared Error (RMSE)}

![Bar Chart: Simulation results: Root Mean Squared Error](image-url)
Example of Application: Risk of Gastrointestinal (GI) events in COX-2 vs. NSAID users

- When introduced onto the market, COX-2 inhibitors were marketed as having the same anti-inflammatory and analgesic effects as traditional NSAIDs, but with lower risks of GI events.

- **Cohort of elderly from Québec (RAMQ data) followed from Oct. 1999 for 6 months**
  - 33.8% (1,513/4,475) of COX-2 vs. 27.4% (1,184/4,318) of NSAID users had a GI event.
  - Long list of known confounders adjusted for, including several risk factors for GI events (prior GI events, GI prophylaxis etc.)

- **However, confounding by additional unmeasured characteristics possible, as** patients switched from NSAIDs to COX-2 might have been Sicker [Wolfe et al, *J Rheum* 2002]

- All models adjusted for the same a priori selected observed potential confounders
### Applied Example: Results

**Differences in risk of gastrointestinal events between COX-2 inhibitor vs. NSAID users**

<table>
<thead>
<tr>
<th>Model</th>
<th>Treatment RD (%) for COX-2 vs. NSAID users (95% CI)(^a)</th>
<th>Interaction (95% CI)(^c)</th>
<th>Interaction p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Column 1</td>
<td>Column 2</td>
<td>Column 4</td>
<td>Column 5</td>
</tr>
<tr>
<td>Unadjusted</td>
<td>6.4 (4.5, 8.3)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Conventional Model 1</strong></td>
<td><strong>-1.9 (-3.5, -0.3)</strong></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Binary IV Model 2</strong></td>
<td><strong>3.3 (-12.7, 19.3)</strong></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Continuous IV Model 3</strong></td>
<td><strong>2.5 (-6.2, 11.3)</strong></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Interaction-based Model 4</strong></td>
<td><strong>-2.6 (-8.4, 3.2)</strong></td>
<td><strong>2.5 (-12.9, 17.8)</strong></td>
<td><strong>0.751</strong></td>
</tr>
</tbody>
</table>
Conclusion

- As in simulations, our interaction-based Model 4 had much more precise CI’s than IV Models 2 and 3.
- For the Interaction-based Model 4, RD = -2.6% is very similar to Conventional model (RD= -1.9%), and p-value = 0.75 for the Interaction test indicate the Absence of Unmeasured Confounding, suggesting that the extensive list of measured confounders might have accounted for possible confounding by indication.
New (yet, Un-published) Application:

• QUESTION:
  Does Influenza Vaccination prevent Influenza and Flu Symptoms in Rheum. Arth. (RA)?

DATABASE Study:
• 15,724 incident RA patients, identified after 1\textsuperscript{st} January 2000.
• 12,433 patients received flu vaccination during follow-up.
• During follow-up: 572 patients had influenza and 6726 patients had flu symptoms.
Flu Vaccination: Prelim. Results

- for Influenza:
  - Conventional model: \( RD = -0.02 \) \((-0.05, 0.02)\), \( p>0.20 \) (NO effect)
  - Missing Cause model: \( RD = -2.38 \) \((-3.66, -1.10)\), \( p<0.001 \) (Significant Risk Decrease)
  - Interaction test of Un-measured Confounding: \( p <0.001 \)

- for Flu Symptoms:
  - Conventional model: \( RD = +0.46 \) \((0.34, 0.57)\), \( p<0.001 \) (Significant Risk Increase)
  - Missing Cause model: \( RD = -3.73 \) \((-7.99, 0.54)\), \( p<0.001 \) (Significant Risk Decrease)
  - Interaction test of Un-measured Confounding: \( p <0.01 \)
(Selected…) Future Challenges

- Assess the advantages of our method in Real-life analyses
- Develop & Validate Empirical Criteria to Assess the Strength of Physician Preferences in real-life applications
- Systematically Compare the performance of our method with IV-based methods under a variety of assumptions about the true data structure
- Attempt to develop more ‘sensitive’ criteria and/or tests to Detect Unobserved Confounding based on different analyses of real-life data
References

- Kollhorst B, Abrahamowicz M, Pigeot I. The proportion of all previous patients was a potential instrument for patients' actual prescriptions of NSAIDs in Germany. *J Clinical Epidemiology*; doi: 10.1016/j.jclinepi.2015.08.008.
THANK YOU - MERCI

Michal.Abrahamowicz@McGill.ca