OXFORD UNIVERSITY PRESS | American Journal of Epidemiology

A simulation-based bias analysis to assess the impact of unmeasured confounding when designing non-randomized database studies

Journal:	*American Journal of Epidemiology*
Manuscript ID	AJE-00578-2023.R1
Manuscript Type:	Practice of Epidemiology
Key Words:	

A simulation-based bias analysis to assess the impact of unmeasured confounding when designing non-randomized database studies

Authors: Rishi J Desai,[1] Marie Bradley,[2] Hana Lee,[2] Efe Eworuke,[2*] Janick Weberpals,[1] Richard Wyss,[1] Sebastian Schneeweiss,[1] Robert Ball[2]

[1] Division of Pharmacoepidemiology and Pharmacoeconomics, Brigham and Women's Hospital, Harvard Medical School

[2] Office of Surveillance and Epidemiology, Center for Drug Evaluation and Research, US Food and Drug Administration

[3]Office of Biostatistics, Center for Drug Evaluation and Research, US Food and Drug Administration

* Currently with IQVIA

Word count: ~2,700

Figures: 5

Correspondence:

Rishi J Desai

Assistant Professor of Medicine,

Division of Pharmacoepidemiology and Pharmacoeconomics,

Department of Medicine, Brigham and Women's Hospital, Harvard Medical School,

1620 Tremont Street, Suite 3030-R, Boston, MA 02120, USA

Phone: 617-278-0932 | Fax: 617-232-8602 | Email: rdesai@bwh.harvard.edu

Disclaimer: The views are those of the author(s) and do not necessarily represent the official views of, nor an endorsement, by FDA/HHS, or the U.S. Government.

Abstract (240/250)

Background: Unmeasured confounding is often raised as a source of potential bias during the design of non-randomized studies but quantifying such concerns is challenging.

Methods: We developed a simulation-based approach to assess the potential impact of unmeasured confounding during the study design stage. The approach involved generation of hypothetical individual-level cohorts using realistic parameters including a binary treatment (prevalence 25%), a time-to-event outcome (incidence 5%), 13 measured covariates, a binary unmeasured confounder (u_1, 10%), and a binary measured 'proxy' variable (p_1) correlated with u_1. Strength of unmeasured confounding and correlations between u_1 and p_1 were varied in simulation scenarios. Treatment effects were estimated with, a) no adjustment, b) adjustment for measured confounders (Level 1), c) adjustment for measured confounders and their proxy (Level 2). We computed absolute standardized mean differences in u_1 and p_1 and relative bias with each level of adjustment.

Results: Across all scenarios, Level 2 adjustment led to improvement in balance of u_1, but this improvement was highly dependent on the correlation between u_1 and p_1. Level 2 adjustments also had lower relative bias than Level 1 adjustments (in strong u_1 scenarios: relative bias of 9.2%, 12.2%, 13.5% at correlations 0.7, 0.5, and 0.3, respectively versus 16.4%, 15.8%, 15.0% for Level 1, respectively).

Conclusion: An approach using simulated individual-level data was useful to explicitly convey the potential for bias due to unmeasured confounding while designing non-randomized studies and can be helpful in informing design choices.

Introduction

When relevant confounders of an exposure–outcome relationship are inadequately measured, missing or unknown, the resulting residual confounding leads to biased results in non-randomized etiologic studies.[1] Understanding the impact of unmeasured confounding is critical for generating reliable causal inference from non-randomized studies.[2] Unmeasured confounding is often raised as a source of potential bias when evaluating non- randomized study protocols but evaluating such concerns during their design remains challenging. This is pressingly important for the US Food and Drug Administration (FDA) given that findings from studies using real-world data from clinical practice are considered in regulatory decision making.[3]

Several sensitivity-analysis and bias-modeling approaches are available to evaluate the impact of unmeasured confounding in non-randomized studies.[4-11] Most commonly, such sensitivity analyses for unmeasured confounding are implemented *post hoc* once study results are known. Approaches that assess the potential impact of unmeasured confounding in the design stage remain underutilized.[12] A principled quantitative bias analysis (QBA) approach at the design stage could allow for a better understanding of the uncertainties associated with potential unmeasured confounding upfront and allow informed decisions on the feasibility and acceptability of a proposed non-randomized study protocol in an efficient manner. It may also reduce the tendency to select parameters for *post hoc* sensitivity analyses that result in findings aligned with a favorable outcome.

To that end, we developed and assessed a QBA approach for unmeasured confounding that can be implemented during the study design of non-randomized studies. This QBA approach leverages Monte Carlo simulations based on individual-level data .[13] It is specifically designed to consider the role of 'proxy' adjustment using correlated variables for unmeasured constructs, which is highly relevant when working with secondary real-world data to lessen the

threat of bias[14,15] but has received little attention when conducting QBA for unmeasured confounding.

The central contribution of the current paper is development of a flexible methodology that can allow researchers to characterize the bias arising from unmeasured confounding with a specified but modifiable structure during the study design. This can help identify whether real unmeasured confounders with those characteristics are likely to exist based on knowledge of the disease being studied and can lead to informed study design choices. The proposed methodology is intended to help decision-makers, for instance medical product regulators, in efficient review of protocols using real-world clinical practice data.

Methods

Simulation Design

We designed *de novo* individual-level simulations to mimic a safety study using an active comparator cohort study design, which is reflective of typical questions that may be addressed by non-randomized studies of healthcare data. **Figure 1** summarizes the simulation design using a directed acyclic graph. We simulated cohorts of patients initiating treatment with celecoxib or non-selective non-steroidal anti-inflammatory drugs (NSAIDs) to compare the risk of time to severe gastrointestinal bleeding. We generated 6 variables as measured confounders representing: sex (c1), age in years (c2), history of gastrointestinal bleeding (c3), concurrent use of gastrointestinal protective agents (c4), history of anti-coagulant use (c5), and number of generic medications used (c6). We generated 5 treatment-only predictors representing: peptic ulcer history (c7), osteoarthritis diagnosis (c8), rheumatoid arthritis diagnosis (c9), history of gastrointestinal protective agent use (c10), and any hospital admission at baseline (c11). We also generated two outcome risk factors representing: diagnosis of chronic obstructive pulmonary disease (c12) and corticosteroid use (c13). **Appendix** summarizes the prevalence

estimates and coefficients values we used for c1-c13 in the treatment and outcome models, based on earlier work.[16]

Next, we generated a binary unmeasured confounder (u_1) associated independently with both treatment and outcome. We further generated a binary measured 'proxy' variable (p_1) associated with treatment and outcome only through correlation with the generated unmeasured confounder. For example, if current smoking status is the suspected unmeasured confounder, indicators for smoking counseling visits recorded as procedure codes in healthcare data could be considered a proxy variable only associated with the treatment and outcome through correlation with smoking.

Simulation procedures

A binary treatment with prevalence of 25% was generated using a logistic regression model with 6 confounders (c1–c6), 5 treatment-predictors (c7–c11), and the unmeasured confounder (u_1). Time-to-event outcomes were generated from a Cox model assuming exponential distribution as well as a constant baseline hazard using the equation[17]

$$T = \frac{-logR}{\lambda * e^{\beta x}}$$

where R is uniformly distributed random variable, λ is the scale parameter (0.001), x is a vector of covariates, and β is a vector of coefficients relating the covariates to the hazard of outcome assuming proportional hazards. x included confounders (c1-c6), outcome risk factors (c12-c13), unmeasured confounder (u_1), and treatment (with coefficient of -0.25 considering protective effect of celecoxib on gastrointestinal bleeding outcome). Based on generated survival times for every patient, we simulated the outcome variable with incidence of 5% by considering patients with $T \leq 5^{th}$ percentile of distribution of T as those with observed outcomes and censored otherwise.

Binary covariates were generated using binomial distributions and continuous covariates (c2 and c6) were generated using truncated normal distributions. The unmeasured confounder and correlated proxy variable were generated using binomial distributions.[18]

Simulation scenarios

We considered a total of 10 simulation scenarios involving combinations of strength of unmeasured confounding and the correlation between the unmeasured confounder (u_1) and proxy variable (p1). We varied strength of the unmeasured confounding to represent "strong" (coefficients of 1.0 for u_1 in the treatment and outcome models) or "moderate" (coefficients of 0.5) scenarios. For the correlations between u_1 and p_1, we designed scenarios representing high (correlation coefficient of 0.7), medium (0.5), or weak (0.3) correlation.

Further, we generated 2 additional scenarios (2 with "strong" and 2 with "moderate" unmeasured confounding) where we simulated 3 proxies (p_1-p_3). The simulated proxies p_1-p_3 were generated to have weak independent correlations of 0.3 or 0.2 with u_1. These four scenarios were intended to mimic situations where investigators may be able to adjust for the unmeasured constructs with multiple "weak" proxies. In all scenarios, we generated 500 cohorts of 5,000 observations each.

Statistical analysis

Following data generation, three levels of adjustments were used to estimate treatment effects, a) no adjustment, b) Level 1: adjustment for measured confounders (c1-c6) through propensity score matching (PSM), c) Level 2: adjustment for measured confounders (c1-c6) and proxy variables (p_1 only in scenarios involving single proxy, p_1-p_3 in scenarios involving multiple proxies) through PSM. In all scenarios, we used logistic regression to model the PS and 1:1 nearest neighbor matching with a caliper of 0.2 of the standard deviation of the PS.[19]

Performance evaluation

To assess the degree of balance achieved in the unmeasured confounder with each level of adjustment, we plotted the distribution of average absolute standardized mean differences[20] (ASMD) for u_1 and proxy variables (p_1 only in scenarios involving single proxy, p_1-p_3 in scenarios involving multiple proxies) across 500 simulated cohorts in each scenario. Next, we calculated relative bias in treatment effect estimates with each level of adjustment on the hazard ratio scale as average of the difference between exponentiated estimated and exponentiated true coefficient (0.78) divided by exponentiated true coefficient (0.78) across 500 simulated cohorts in each scenario and presented as % bias.

Code availability

R codes to execute study-specific simulations are available on the following page: https://gitlab.partners.org/rjd48/unmeasured-confounding-simulations/. In the Appendix, we provide step by step instructions on how to use our R codes to conduct tailored simulation based bias analyses for unmeasured confounding.

Results

Figure 2 summarizes the ASMD for u_1 and proxy variables between treatment groups. Before any adjustment, u_1 was highly imbalanced between treatment groups in the "strong" unmeasured confounding scenarios (average ASMD 0.31), but less so in the "moderate" scenarios (ASMD 0.15). Owing to a high correlation with u_1 (0.7), p_1 was also imbalanced (ASMD 0.22 in "strong" and 0.11 in "moderate" scenarios). As correlations between u_1 and p_1 decreased, the imbalance in p_1 also decreased, as expected. Adjusting for measured confounding alone in Level 1 adjustment did not have any impact on imbalances in either u_1 or p_1. Across all scenarios, including proxies in PSM (Level 2) led to near perfect balance in p_1 and consequently improvement in balance of u_1 across the two treatment groups. However, the

improvement in balance of u_1 was highly dependent on the correlation between u_1 and p_1 (mean ASMD 0.16, 0.23, and 0.29 with correlations of 0.7, 0.5, and 0.3 respectively in "strong" scenarios; mean ASMD 0.08, 0.12, and 0.14 with correlations of 0.7, 0.5, and 0.3 respectively in "moderate" scenarios). When considering scenarios with multiple weak proxies, improvement in balance of u_1 was observed with Level 2 adjustment in both "strong" and "moderate" scenarios (**Figure 3**).

Figure 4 provides relative bias across all scenarios. In the "strong" scenarios, relative bias of approximately 27-29% was observed before any adjustment, indicating bias towards the null. Level 2 adjustments including proxy p_1 lowered relative bias consistently more (9.2%, 12.2%, 13.5% at correlations 0.7, 0.5, and 0.3, respectively) than Level 1 adjustments only including measured confounders (16.4%, 15.8%, 15.0%). When including multiple weak proxies, relative bias was lowered with Level 2 adjustment to 12.9% when correlation was 0.3 and 14.4% when correlation was 0.2. In "moderate" scenarios, relative bias was approximately 15% prior to any adjustment. Adjusting for measured confounding in Level 1 adjustments substantially reduced relative bias (3.6-4.6%) and adjusting for u_1 using single or multiple proxies in Level 2 adjustments had little impact on relative bias.

Discussion

In this study, we developed a flexible simulation approach for use during study planning to quantify the expected bias due unmeasured confounding based on a user-specified confounding structure. We exemplified this approach to assess the impact of an unmeasured confounding variable when comparing celecoxib vs NSAID on GI bleeding while planning a cohort study. We demonstrated how to quantify residual pre-exposure covariate imbalances and remaining bias when adjusting for the unmeasured confounder through a correlated proxy variable, something that is highly relevant when working with secondary data from clinical practice where many proxies can be generated.[21]

Literature on QBA for unmeasured confounding is rich.[4-11] Numerous methods are available that allow for analytic 'correction' of observed treatment effect estimates based on assumptions related to prevalence of the unmeasured confounder, unmeasured confounder-exposure relationship, and unmeasured confounder-outcome relationship. Some examples of these methods include the Bross correction formula,[6] Rosenbaum-Rubin method based on maximum-likelihood estimation,[8] a regression-based approach of Lin et al.,[9] and the bounding factor approach of Ding and VanderWeele.[7] All these approaches use aggregate level data to estimate corrected treatment effects and make certain assumptions. For instance, Bross and Rosenbaum-Rubin approaches require the unmeasured confounder to be binary; Bross and Lin further assume no interactions between unmeasured confounder and treatment. The bounding factor approach is free from these assumptions; however, it is not designed to explicitly provide quantification of bias but rather bounds for the range of possible bias.

The primary motivation of the current study was to provide decision makers like regulators of medical products with a practical and flexible methodologic tool to aid with review of proposed study protocols and with the design of studies using real-world data. Our proposed approach has three distinct features that improves upon the literature: First, it is developed for implementation at the design stage and aid in decision making around the feasibility and acceptability of a proposed analysis in the presence of suspected unmeasured confounding. Second, it is based on individual-level simulations and is considerably more flexible in that it requires minimal simplifying assumptions. For instance, by explicitly including interactions between treatment and unmeasured confounder or simulating the confounder as continuous or categorical variable in the data-generating process, some of the assumptions that are inherent to the existing approaches can be relaxed. Third, our approach can explicitly quantify remaining bias after proxy adjustment, which is highly relevant when working with secondary healthcare

data, by simulating individual-level data with specified correlations between unmeasured confounders and their proxy variables.

During the planning of a non-interventional study, investigators typically conduct some sample size calculations or expected precision evaluations based on certain simplifying assumptions regarding factors such as the outcome incidence, effect size, and treatment prevalence. Similar to these feasibility evaluations, we recommend using the proposed approach in the design stage of non-randomized studies. A stepwise framework of how to use this method in practice is shown in **Figure 5**. Briefly, as the first step, researchers require a realistic simulation set up with explicit specification of the suspected unmeasured confounder including prevalence, associations with treatment and outcome, and whether any proxy measurements for the unmeasured confounder are available along with anticipated correlation. Next, conducting simulations with sufficiently large iterations will give quantification of imbalance and bias. Finally, based on the quantification, threat of unmeasured confounding can be summarized by making statements of the form: 'An unmeasured confounder with the strength specified in the simulations could lead to x% relative bias if completely ignored. However, adjusting for a proxy variable correlated with the unmeasured confounder may be able to reduce the relative bias to y%'. Such explicit characterization of the bias can lead to informed study design choices. For instance, knowing that an unmeasured confounder is likely to inflict large bias in the proposed analysis, a concerted effort to identify correlated proxy variables can be made from existing data or alternative data sources with some information on the unmeasured confounder may be considered. Even when unmeasured confounding is deemed not addressable with design modifications, knowing the likely strength and direction of bias upfront can be helpful in nuanced interpretation of the study when it is completed. On the other end of the spectrum, quantification of bias can also help with rationalizing use of limited resources by allowing identification of confounders that are unlikely to have an important impact on study

results or confounders that have readily available proxies which may be able to address the bias to a large extent. In such circumstances, resources may be better spent on alternative activities such as outcome algorithm validation using chart reviews to address bias due to outcome misclassification.

While we have focused our discussion on accounting for unmeasured confounding through proxy variables, individual-level simulation-based approach may also be leveraged to conduct similar exercises for measurement error in key confounding variables. In the specific case of measurement error, the error prone measurements serve as 'proxy' of the true underlying construct. Through specification of correlation between error prone measurement and true measurements, one can similarly derive estimates of remaining bias resulting from such measurement error by simulating true measurements in the data generation steps and using error-prone measurements in the analytic steps.

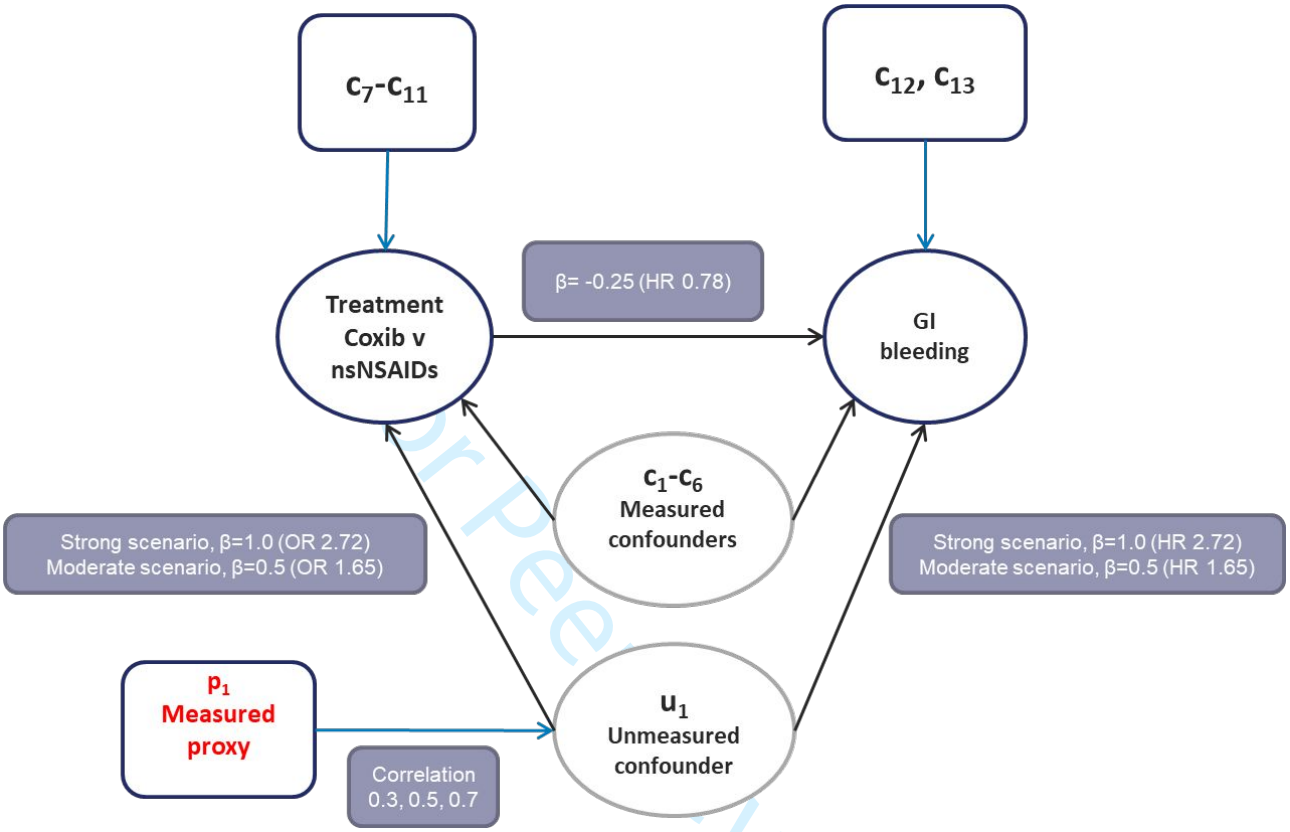
There are some important caveats to the proposed approach that deserve discussion. While simulation based QBA using individual-level data has more flexibility and can accommodate greater complexity than QBA based on aggregate-level data; it may still not fully capture the complexities observed in real-world data. Further, setting up realistic simulation scenarios at the design stage may be challenging as investigators are unlikely to be fully familiar with relationships among variables for their specific question of interest. For instance, in our simulations, we only generated 6 measured confounders and didn't simulate all correlations between measured confounders and unmeasured confounder. In real world data, many more confounders may exist with correlations among them and with unmeasured confounders. Therefore, results from the current simulation based QBA should be interpreted considering the simplifying assumptions made in the data generation steps. This issue can be addressed through more complex simulations that represent a more realistic description of the relationships among multiple confounders and outcomes. Investigators can use the 'plasmode' framework[22]

that builds on actual patient-level data as the basis of simulations to preserve the naturally occurring correlation patterns among study variables. However, investigators may not have access to empirical data during the planning of a study; therefore, our approach based on *de novo* simulations maybe more applicable at the design stage. We also recognize that reliable measurements for the correlation between measured proxies and unmeasured confounder may not be available. In such circumstances, investigators may need to run QBA under numerous scenarios to explicitly acknowledge this uncertainty.

In conclusion, a simulation based QBA approach based on covariate structures generated with individual-level data was useful to explicitly assess the potential for bias due to unmeasured confounding in a proposed study in the design stage. The flexibility, clarity, and relative ease of implementation of this approach makes it an important tool when planning studies.

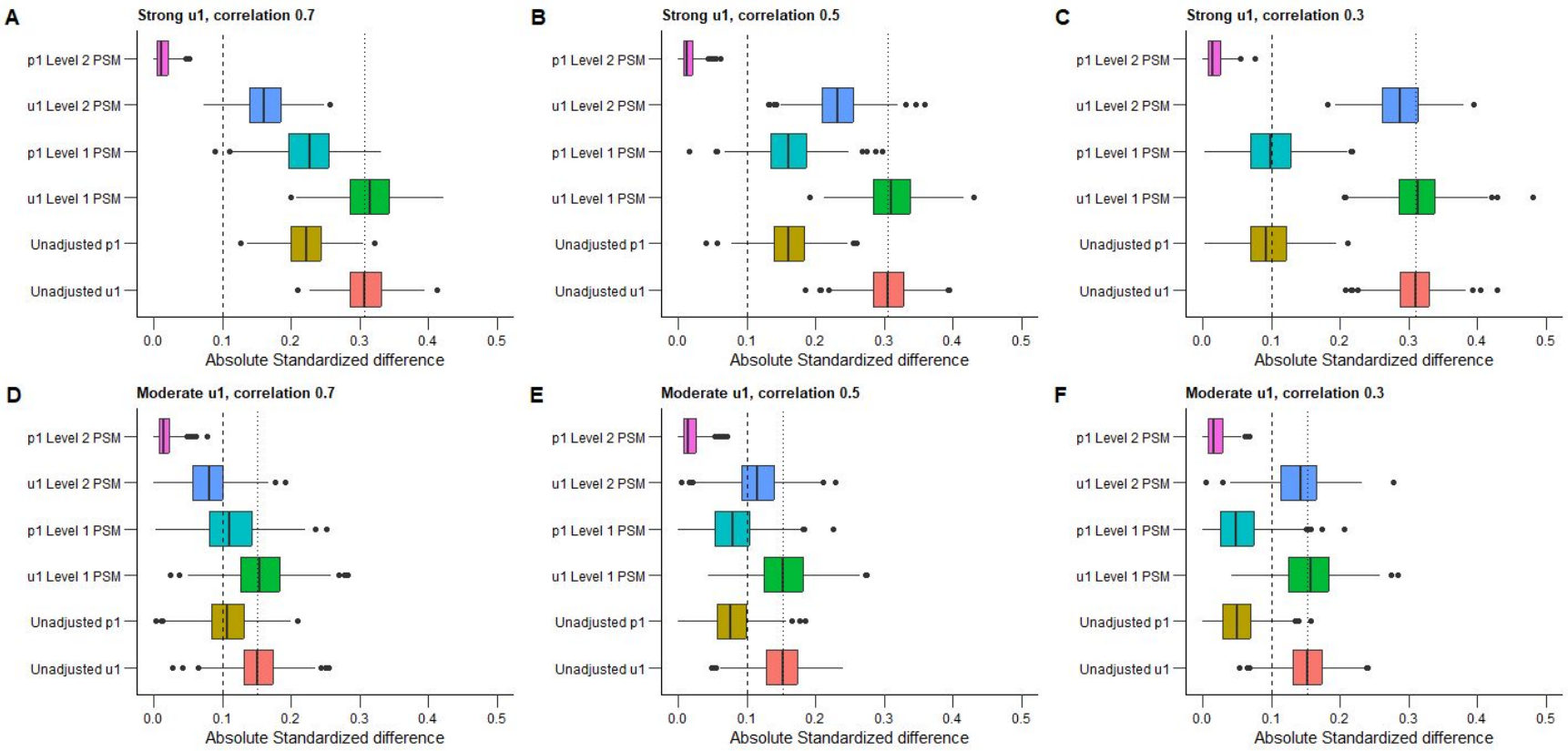
Figures

Figure 1: Data generating mechanism



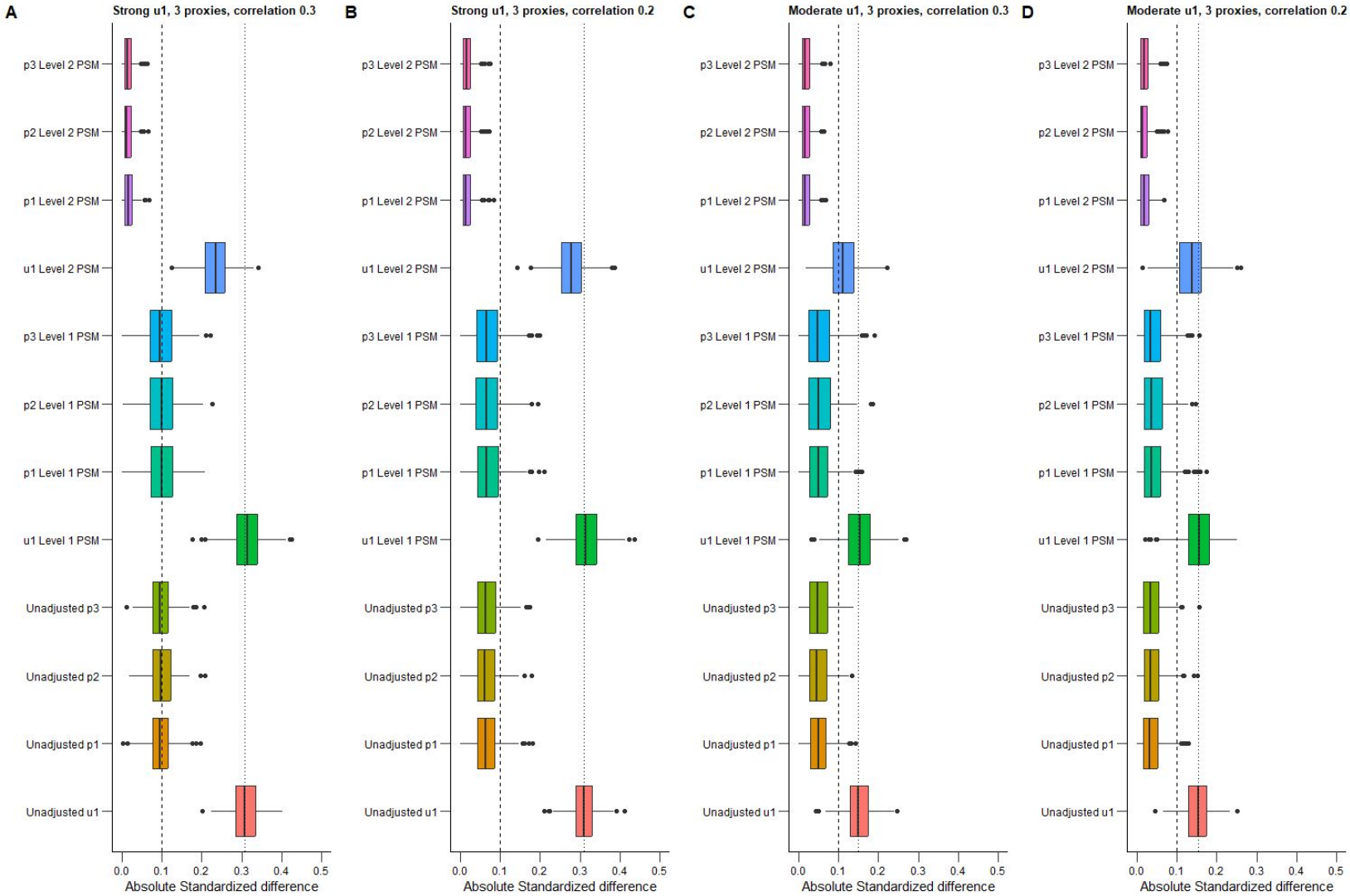
Abbreviations: GI- gastrointestinal; nsNSAID- nonselective nonsteroidal anti-inflammatory drugs

Figure 2: Balance in the unmeasured confounder and proxy variable with different levels of adjustments, one proxy variable scenarios



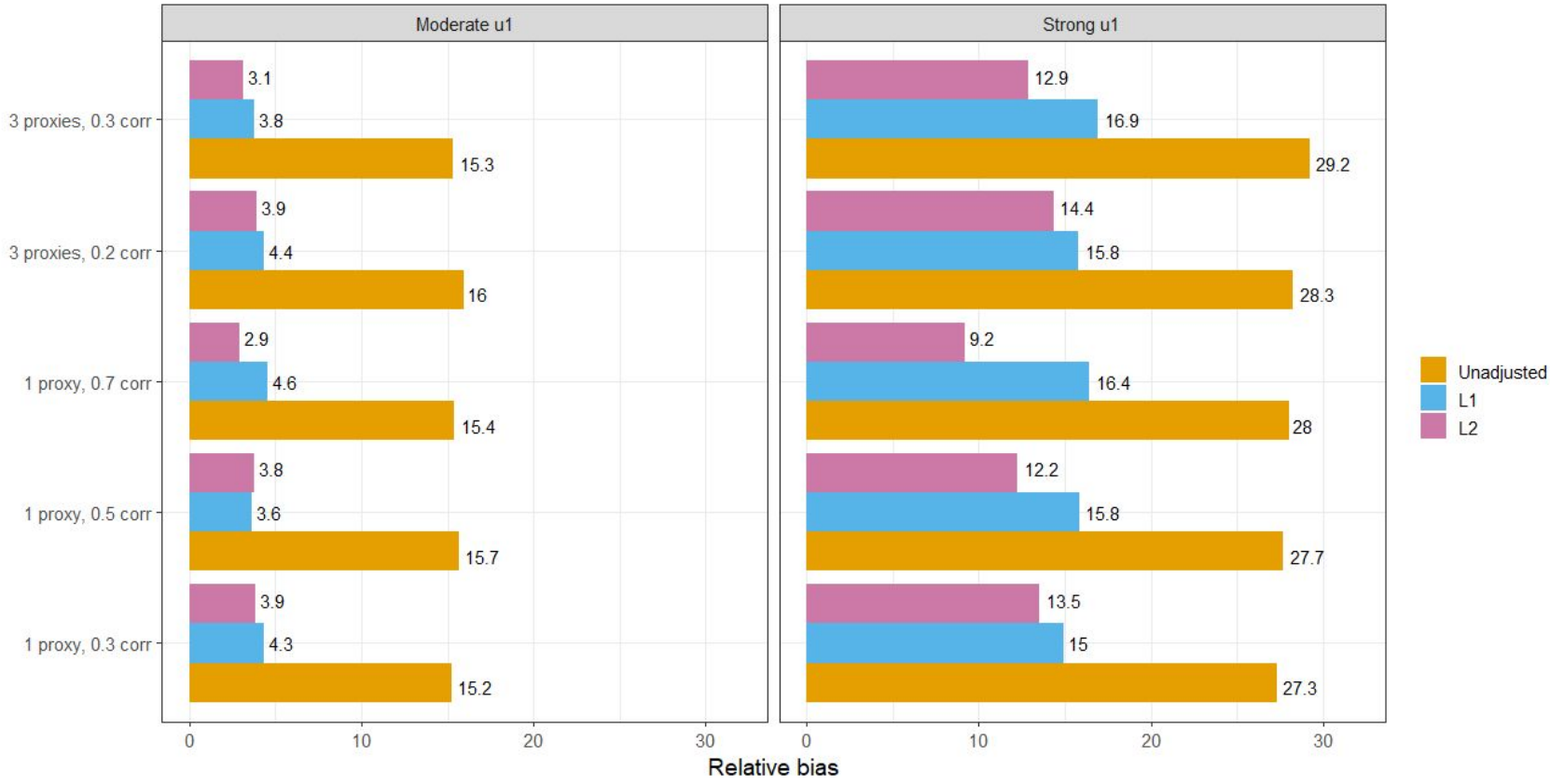
Footnote: Unadjusted analysis had no variables included for adjustment, Level 1 adjustment had only measured confounders included in the propensity score model (c1-c6), Level 2 adjustment had measured confounders + proxy for unmeasured confounder included in the propensity score model included (c1-c6+p1)

Figure 3: Balance in the unmeasured confounder and proxy variables with different levels of adjustments, three proxy variable scenarios



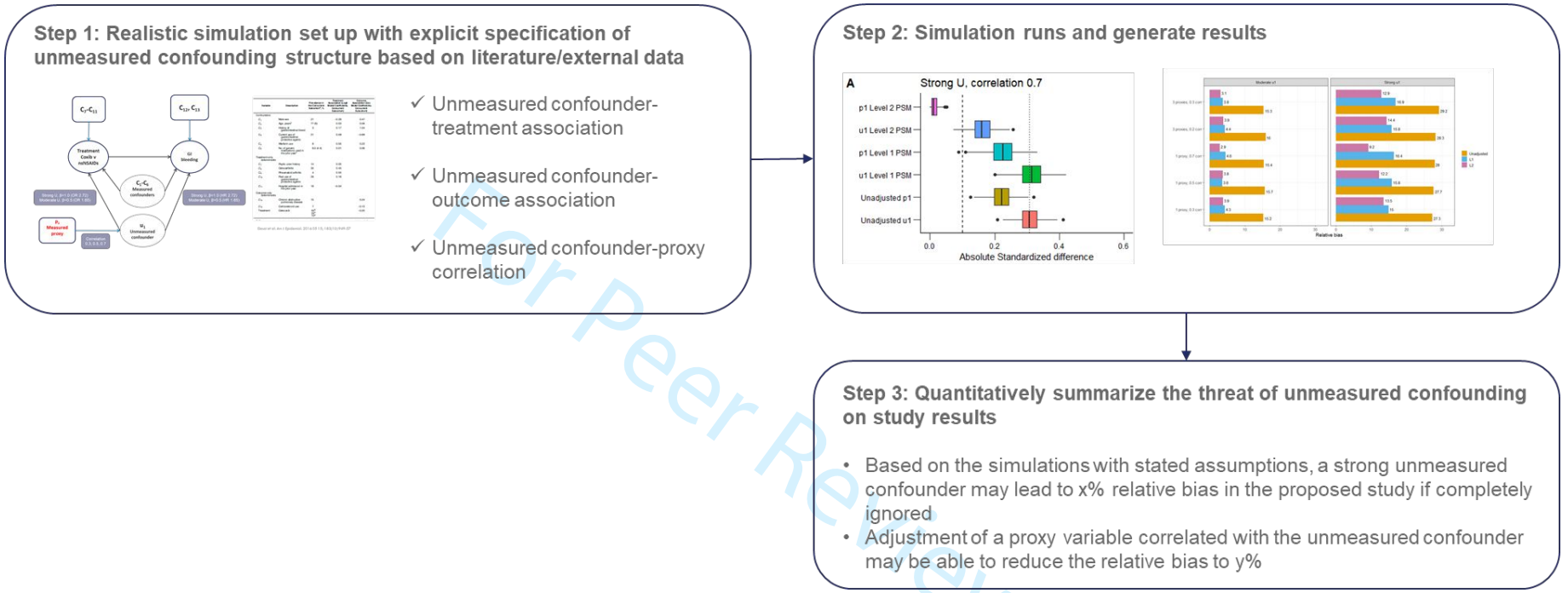
Footnote: Unadjusted analysis had no variables included for adjustment, Level 1 adjustment had only measured confounders included in the propensity score model (c1-c6), Level 2 adjustment had measured confounders + proxy for unmeasured confounder included in the propensity score model included (c1-c6+p1-p3)

Figure 4: Relative bias in treatment effect estimates



Footnote: Unadjusted analysis had no variables included for adjustment, L1 adjustment had only measured confounders included in the propensity score model (c1-c6), L2 adjustment had measured confounders + proxy for unmeasured confounder included in the propensity score model included (c1-c6+p1 in 1 proxy scenarios; c1-c6+p1-p3 in 3 proxy scenarios)

Figure 5: Application framework for the proposed simulation-based quantitative bias analysis



References

1. Vanderweele TJ, Arah OA. Bias formulas for sensitivity analysis of unmeasured confounding for general outcomes, treatments, and confounders. *Epidemiology*. Jan 2011;22(1):42-52. doi:10.1097/EDE.0b013e3181f74493
2. VanderWeele TJ. The sign of the bias of unmeasured confounding. *Biometrics*. Sep 2008;64(3):702-706. doi:10.1111/j.1541-0420.2007.00957.x
3. Considerations for the Use of Real-World Data and Real-World Evidence To Support Regulatory Decision-Making for Drug and Biological Products (2021).
4. Schneeweiss S. Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics. *Pharmacoepidemiology and drug safety*. May 2006;15(5):291-303. doi:10.1002/pds.1200
5. Lash TL, Fox MP, MacLehose RF, Maldonado G, McCandless LC, Greenland S. Good practices for quantitative bias analysis. *Int J Epidemiol*. Dec 2014;43(6):1969-85. doi:10.1093/ije/dyu149
6. Bross ID. Spurious effects from an extraneous variable. *Journal of chronic diseases*. 1966;19(6):637-647.
7. Ding P, VanderWeele TJ. Sensitivity analysis without assumptions. *Epidemiology (Cambridge, Mass)*. 2016;27(3):368.
8. Rosenbaum PR, Rubin DB. Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome. *Journal of the Royal Statistical Society: Series B (Methodological)*. 1983;45(2):212-218.
9. Lin DY, Psaty BM, Kronmal RA. Assessing the sensitivity of regression results to unmeasured confounders in observational studies. *Biometrics*. 1998:948-963.
10. Barberio J, Ahern TP, MacLehose RF, et al. Assessing techniques for quantifying the impact of bias due to an unmeasured confounder: an applied example. *Clinical Epidemiology*. 2021:627-635.
11. Groenwold RH, Nelson DB, Nichol KL, Hoes AW, Hak E. Sensitivity analyses to estimate the potential impact of unmeasured confounding in causal research. *International journal of epidemiology*. 2010;39(1):107-117.
12. Fox MP, Lash TL. Quantitative bias analysis for study and grant planning. *Ann Epidemiol*. Mar 2020;43:32-36. doi:10.1016/j.annepidem.2020.01.013
13. Banack HR, Hayes-Larson E, Mayeda ER. Monte Carlo simulation approaches for quantitative bias analysis: a tutorial. *Epidemiologic Reviews*. 2021;43(1):106-117.
14. Nørgaard M, Ehrenstein V, Vandenbroucke JP. Confounding in observational studies based on large health care databases: problems and potential solutions–a primer for the clinician. *Clinical epidemiology*. 2017:185-193.
15. Schneeweiss S, Rassen JA, Glynn RJ, Avorn J, Mogun H, Brookhart MA. High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. *Epidemiology (Cambridge, Mass)*. 2009;20(4):512.
16. Desai RJ, Glynn RJ, Wang S, Gagne JJ. Performance of Disease Risk Score Matching in Nested Case-Control Studies: A Simulation Study. *American journal of epidemiology*. May 15, 2016 2016;183(10):949-957. doi:10.1093/aje/kwv269
17. Bender R, Augustin T, Blettner M. Generating survival times to simulate Cox proportional hazards models. *Statistics in medicine*. Jun 15 2005;24(11):1713-23. doi:10.1002/sim.2059
18. Emrich LJ, Piedmonte MR. A method for generating high-dimensional multivariate binary variates. *The American Statistician*. 1991;45(4):302-304.
19. Austin PC. Some Methods of Propensity-Score Matching had Superior Performance to Others: Results of an Empirical Investigation and Monte Carlo simulations. *Biometrical journal*. 2009;51(1):171-184.

20. Austin PC. Using the standardized difference to compare the prevalence of a binary variable between two groups in observational research. *Communications in Statistics-Simulation and Computation*. 2009;38(6):1228-1234.
21. Wyss R, Plasek JM, Zhou L, et al. Scalable Feature Engineering from Electronic Free Text Notes to Supplement Confounding Adjustment of Claims-Based Pharmacoepidemiologic Studies. *Clin Pharmacol Ther*. Apr 2023;113(4):832-838. doi:10.1002/cpt.2826
22. Franklin JM, Schneeweiss S, Polinski JM, Rassen JA. Plasmode simulation for the evaluation of pharmacoepidemiologic methods in complex healthcare databases. *Computational statistics & data analysis*. Apr 2014;72:219-226. doi:10.1016/j.csda.2013.10.018

A simulation-based bias analysis to assess the impact of unmeasured confounding when designing non-randomized database studies

Authors: Rishi J Desai,[1] Marie Bradley,[2] Hana Lee,[2] Efe Eworuke,[2*] Janick Weberpals,[1] Richard Wyss,[1] Sebastian Schneeweiss,[1] Robert Ball[2]

[1] Division of Pharmacoepidemiology and Pharmacoeconomics, Brigham and Women's Hospital, Harvard Medical School

[2] Office of Surveillance and Epidemiology, Center for Drug Evaluation and Research, US Food and Drug Administration

[3]Office of Biostatistics, Center for Drug Evaluation and Research, US Food and Drug Administration

* Currently with IQVIA

Word count: ~2,700

Figures: 5

Correspondence:

Rishi J Desai

Assistant Professor of Medicine,

Division of Pharmacoepidemiology and Pharmacoeconomics,

Department of Medicine, Brigham and Women's Hospital, Harvard Medical School,

1620 Tremont Street, Suite 3030-R, Boston, MA 02120, USA

Phone: 617-278-0932 | Fax: 617-232-8602 | Email: rdesai@bwh.harvard.edu

Disclaimer: The views are those of the author(s) and do not necessarily represent the official views of, nor an endorsement, by FDA/HHS, or the U.S. Government.

Abstract (240/250)

Background: Unmeasured confounding is often raised as a source of potential bias during the design of non-randomized studies but quantifying such concerns is challenging.

Methods: We developed a simulation-based approach to assess the potential impact of unmeasured confounding during the study design stage. The approach involved generation of hypothetical individual-level cohorts using realistic parameters including a binary treatment (prevalence 25%), a time-to-event outcome (incidence 5%), 13 measured covariates, a binary unmeasured confounder (u_1, 10%), and a binary measured 'proxy' variable (p_1) correlated with u_1. Strength of unmeasured confounding and correlations between u_1 and p_1 were varied in simulation scenarios. Treatment effects were estimated with, a) no adjustment, b) adjustment for measured confounders (Level 1), c) adjustment for measured confounders and their proxy (Level 2). We computed absolute standardized mean differences in u_1 and p_1 and relative bias with each level of adjustment.

Results: Across all scenarios, Level 2 adjustment led to improvement in balance of u_1, but this improvement was highly dependent on the correlation between u_1 and p_1. Level 2 adjustments also had lower relative bias than Level 1 adjustments (in strong u_1 scenarios: relative bias of 9.2%, 12.2%, 13.5% at correlations 0.7, 0.5, and 0.3, respectively versus 16.4%, 15.8%, 15.0% for Level 1, respectively).

Conclusion: An approach using simulated individual-level data was useful to explicitly convey the potential for bias due to unmeasured confounding while designing non-randomized studies and can be helpful in informing design choices.

Introduction

When relevant confounders of an exposure–outcome relationship are inadequately measured, missing or unknown, the resulting residual confounding leads to biased results in non-randomized etiologic studies.[1] Understanding the impact of unmeasured confounding is critical for generating reliable causal inference from non-randomized studies.[2] Unmeasured confounding is often raised as a source of potential bias when evaluating non- randomized study protocols but evaluating such concerns during their design remains challenging. This is pressingly important for the US Food and Drug Administration (FDA) given that findings from studies using real-world data from clinical practice are considered in regulatory decision making.[3]

Several sensitivity-analysis and bias-modeling approaches are available to evaluate the impact of unmeasured confounding in non-randomized studies.[4-11] Most commonly, such sensitivity analyses for unmeasured confounding are implemented *post hoc* once study results are known. Approaches that assess the potential impact of unmeasured confounding in the design stage remain underutilized.[12] A principled quantitative bias analysis (QBA) approach at the design stage could allow for a better understanding of the uncertainties associated with potential unmeasured confounding upfront and allow informed decisions on the feasibility and acceptability of a proposed non-randomized study protocol in an efficient manner. It may also reduce the tendency to select parameters for *post hoc* sensitivity analyses that result in findings aligned with a favorable outcome.

To that end, we developed and assessed a QBA approach for unmeasured confounding that can be implemented during the study design of non-randomized studies. This QBA approach leverages Monte Carlo simulations based on individual-level data .[13] It is specifically designed to consider the role of 'proxy' adjustment using correlated variables for unmeasured constructs, which is highly relevant when working with secondary real-world data to lessen the

threat of bias[14,15] but has received little attention when conducting QBA for unmeasured confounding.

The central contribution of the current paper is development of a flexible methodology that can allow researchers to characterize the bias arising from unmeasured confounding with a specified but modifiable structure during the study design. This can help identify whether real unmeasured confounders with those characteristics are likely to exist based on knowledge of the disease being studied and can lead to informed study design choices. The proposed methodology is intended to help decision-makers, for instance medical product regulators, in efficient review of protocols using real-world clinical practice data.

Methods

Simulation Design

We designed *de novo* individual-level simulations to mimic a safety study using an active comparator cohort study design, which is reflective of typical questions that may be addressed by non-randomized studies of healthcare data. **Figure 1** summarizes the simulation design using a directed acyclic graph. We simulated cohorts of patients initiating treatment with celecoxib or non-selective non-steroidal anti-inflammatory drugs (NSAIDs) to compare the risk of time to severe gastrointestinal bleeding. We generated 6 variables as measured confounders representing: sex (c1), age in years (c2), history of gastrointestinal bleeding (c3), concurrent use of gastrointestinal protective agents (c4), history of anti-coagulant use (c5), and number of generic medications used (c6). We generated 5 treatment-only predictors representing: peptic ulcer history (c7), osteoarthritis diagnosis (c8), rheumatoid arthritis diagnosis (c9), history of gastrointestinal protective agent use (c10), and any hospital admission at baseline (c11). We also generated two outcome risk factors representing: diagnosis of chronic obstructive pulmonary disease (c12) and corticosteroid use (c13). **Appendix** summarizes the prevalence

estimates and coefficients values we used for c1-c13 in the treatment and outcome models, based on earlier work.[16]

Next, we generated a binary unmeasured confounder (u_1) associated independently with both treatment and outcome. We further generated a binary measured 'proxy' variable (p_1) associated with treatment and outcome only through correlation with the generated unmeasured confounder. For example, if current smoking status is the suspected unmeasured confounder, indicators for smoking counseling visits recorded as procedure codes in healthcare data could be considered a proxy variable only associated with the treatment and outcome through correlation with smoking.

Simulation procedures

A binary treatment with prevalence of 25% was generated using a logistic regression model with 6 confounders (c1–c6), 5 treatment-predictors (c7–c11), and the unmeasured confounder (u_1). Time-to-event outcomes were generated from a Cox model assuming exponential distribution as well as a constant baseline hazard using the equation[17]

$$T = \frac{-logR}{\lambda * e^{\beta x}}$$

where R is uniformly distributed random variable, λ is the scale parameter (0.001), x is a vector of covariates, and β is a vector of coefficients relating the covariates to the hazard of outcome assuming proportional hazards. x included confounders (c1-c6), outcome risk factors (c12-c13), unmeasured confounder (u_1), and treatment (with coefficient of -0.25 considering protective effect of celecoxib on gastrointestinal bleeding outcome). Based on generated survival times for every patient, we simulated the outcome variable with incidence of 5% by considering patients with $T \leq 5^{th}$ percentile of distribution of T as those with observed outcomes and censored otherwise.

Binary covariates were generated using binomial distributions and continuous covariates (c2 and c6) were generated using truncated normal distributions. The unmeasured confounder and correlated proxy variable were generated using binomial distributions.[18]

Simulation scenarios

We considered a total of 10 simulation scenarios involving combinations of strength of unmeasured confounding and the correlation between the unmeasured confounder (u_1) and proxy variable (p1). We varied strength of the unmeasured confounding to represent "strong" (coefficients of 1.0 for u_1 in the treatment and outcome models) or "moderate" (coefficients of 0.5) scenarios. For the correlations between u_1 and p_1, we designed scenarios representing high (correlation coefficient of 0.7), medium (0.5), or weak (0.3) correlation.

Further, we generated 2 additional scenarios (2 with "strong" and 2 with "moderate" unmeasured confounding) where we simulated 3 proxies (p_1-p_3). The simulated proxies p_1-p_3 were generated to have weak independent correlations of 0.3 or 0.2 with u_1. These four scenarios were intended to mimic situations where investigators may be able to adjust for the unmeasured constructs with multiple "weak" proxies. In all scenarios, we generated 500 cohorts of 5,000 observations each.

Statistical analysis

Following data generation, three levels of adjustments were used to estimate treatment effects, a) no adjustment, b) Level 1: adjustment for measured confounders (c1-c6) through propensity score matching (PSM), c) Level 2: adjustment for measured confounders (c1-c6) and proxy variables (p_1 only in scenarios involving single proxy, p_1-p_3 in scenarios involving multiple proxies) through PSM. In all scenarios, we used logistic regression to model the PS and 1:1 nearest neighbor matching with a caliper of 0.2 of the standard deviation of the PS.[19]

Performance evaluation

To assess the degree of balance achieved in the unmeasured confounder with each level of adjustment, we plotted the distribution of average absolute standardized mean differences[20] (ASMD) for u_1 and proxy variables (p_1 only in scenarios involving single proxy, p_1-p_3 in scenarios involving multiple proxies) across 500 simulated cohorts in each scenario. Next, we calculated relative bias in treatment effect estimates with each level of adjustment on the hazard ratio scale as average of the difference between exponentiated estimated and exponentiated true coefficient (0.78) divided by exponentiated true coefficient (0.78) across 500 simulated cohorts in each scenario and presented as % bias.

Code availability

R codes to execute study-specific simulations are available on the following page: https://gitlab.partners.org/rjd48/unmeasured-confounding-simulations/. In the Appendix, we provide step by step instructions on how to use our R codes to conduct tailored simulation based bias analyses for unmeasured confounding.

Results

Figure 2 summarizes the ASMD for u_1 and proxy variables between treatment groups. Before any adjustment, u_1 was highly imbalanced between treatment groups in the "strong" unmeasured confounding scenarios (average ASMD 0.31), but less so in the "moderate" scenarios (ASMD 0.15). Owing to a high correlation with u_1 (0.7), p_1 was also imbalanced (ASMD 0.22 in "strong" and 0.11 in "moderate" scenarios). As correlations between u_1 and p_1 decreased, the imbalance in p_1 also decreased, as expected. Adjusting for measured confounding alone in Level 1 adjustment did not have any impact on imbalances in either u_1 or p_1. Across all scenarios, including proxies in PSM (Level 2) led to near perfect balance in p_1 and

consequently improvement in balance of u_1 across the two treatment groups. However, the improvement in balance of u_1 was highly dependent on the correlation between u_1 and p_1 (mean ASMD 0.16, 0.23, and 0.29 with correlations of 0.7, 0.5, and 0.3 respectively in "strong" scenarios; mean ASMD 0.08, 0.12, and 0.14 with correlations of 0.7, 0.5, and 0.3 respectively in "moderate" scenarios). When considering scenarios with multiple weak proxies, improvement in balance of u_1 was observed with Level 2 adjustment in both "strong" and "moderate" scenarios (**Figure 3**).

Figure 4 provides relative bias across all scenarios. In the "strong" scenarios, relative bias of approximately 27-29% was observed before any adjustment, indicating bias towards the null. Level 2 adjustments including proxy p_1 lowered relative bias consistently more (9.2%, 12.2%, 13.5% at correlations 0.7, 0.5, and 0.3, respectively) than Level 1 adjustments only including measured confounders (16.4%, 15.8%, 15.0%). When including multiple weak proxies, relative bias was lowered with Level 2 adjustment to 12.9% when correlation was 0.3 and 14.4% when correlation was 0.2. In "moderate" scenarios, relative bias was approximately 15% prior to any adjustment. Adjusting for measured confounding in Level 1 adjustments substantially reduced relative bias (3.6-4.6%) and adjusting for u_1 using single or multiple proxies in Level 2 adjustments had little impact on relative bias.

Discussion

In this study, we developed a flexible simulation approach for use during study planning to quantify the expected bias due unmeasured confounding based on a user-specified confounding structure. We exemplified this approach to assess the impact of an unmeasured confounding variable when comparing celecoxib vs NSAID on GI bleeding while planning a cohort study. We demonstrated how to quantify residual pre-exposure covariate imbalances and remaining bias when adjusting for the unmeasured confounder through a correlated proxy

variable, something that is highly relevant when working with secondary data from clinical practice where many proxies can be generated.[21]

Literature on QBA for unmeasured confounding is rich.[4-11] Numerous methods are available that allow for analytic 'correction' of observed treatment effect estimates based on assumptions related to prevalence of the unmeasured confounder, unmeasured confounder-exposure relationship, and unmeasured confounder-outcome relationship. Some examples of these methods include the Bross correction formula,[6] Rosenbaum-Rubin method based on maximum-likelihood estimation,[8] a regression-based approach of Lin et al.,[9] and the bounding factor approach of Ding and VanderWeele.[7] All these approaches use aggregate level data to estimate corrected treatment effects and make certain assumptions. For instance, Bross and Rosenbaum-Rubin approaches require the unmeasured confounder to be binary; Bross and Lin further assume no interactions between unmeasured confounder and treatment. The bounding factor approach is free from these assumptions; however, it is not designed to explicitly provide quantification of bias but rather bounds for the range of possible bias.

The primary motivation of the current study was to provide decision makers like regulators of medical products with a practical and flexible methodologic tool to aid with review of proposed study protocols and with the design of studies using real-world data. Our proposed approach has three distinct features that improves upon the literature: First, it is developed for implementation at the design stage and aid in decision making around the feasibility and acceptability of a proposed analysis in the presence of suspected unmeasured confounding. Second, it is based on individual-level simulations and is considerably more flexible in that it requires minimal simplifying assumptions. For instance, by explicitly including interactions between treatment and unmeasured confounder or simulating the confounder as continuous or categorical variable in the data-generating process, some of the assumptions that are inherent to the existing approaches can be relaxed. Third, our approach can explicitly quantify remaining

bias after proxy adjustment, which is highly relevant when working with secondary healthcare data, by simulating individual-level data with specified correlations between unmeasured confounders and their proxy variables.

During the planning of a non-interventional study, investigators typically conduct some sample size calculations or expected precision evaluations based on certain simplifying assumptions regarding factors such as the outcome incidence, effect size, and treatment prevalence. Similar to these feasibility evaluations, we recommend using the proposed approach in the design stage of non-randomized studies. A stepwise framework of how to use this method in practice is shown in **Figure 5**. Briefly, as the first step, researchers require a realistic simulation set up with explicit specification of the suspected unmeasured confounder including prevalence, associations with treatment and outcome, and whether any proxy measurements for the unmeasured confounder are available along with anticipated correlation. Next, conducting simulations with sufficiently large iterations will give quantification of imbalance and bias. Finally, based on the quantification, threat of unmeasured confounding can be summarized by making statements of the form: 'An unmeasured confounder with the strength specified in the simulations could lead to x% relative bias if completely ignored. However, adjusting for a proxy variable correlated with the unmeasured confounder may be able to reduce the relative bias to y%'. Such explicit characterization of the bias can lead to informed study design choices. For instance, knowing that an unmeasured confounder is likely to inflict large bias in the proposed analysis, a concerted effort to identify correlated proxy variables can be made from existing data or alternative data sources with some information on the unmeasured confounder may be considered. Even when unmeasured confounding is deemed not addressable with design modifications, knowing the likely strength and direction of bias upfront can be helpful in nuanced interpretation of the study when it is completed. On the other end of the spectrum, quantification of bias can also help with rationalizing use of limited resources by

allowing identification of confounders that are unlikely to have an important impact on study results or confounders that have readily available proxies which may be able to address the bias to a large extent. In such circumstances, resources may be better spent on alternative activities such as outcome algorithm validation using chart reviews to address bias due to outcome misclassification.

While we have focused our discussion on accounting for unmeasured confounding through proxy variables, individual-level simulation-based approach may also be leveraged to conduct similar exercises for measurement error in key confounding variables. In the specific case of measurement error, the error prone measurements serve as 'proxy' of the true underlying construct. Through specification of correlation between error prone measurement and true measurements, one can similarly derive estimates of remaining bias resulting from such measurement error by simulating true measurements in the data generation steps and using error-prone measurements in the analytic steps.

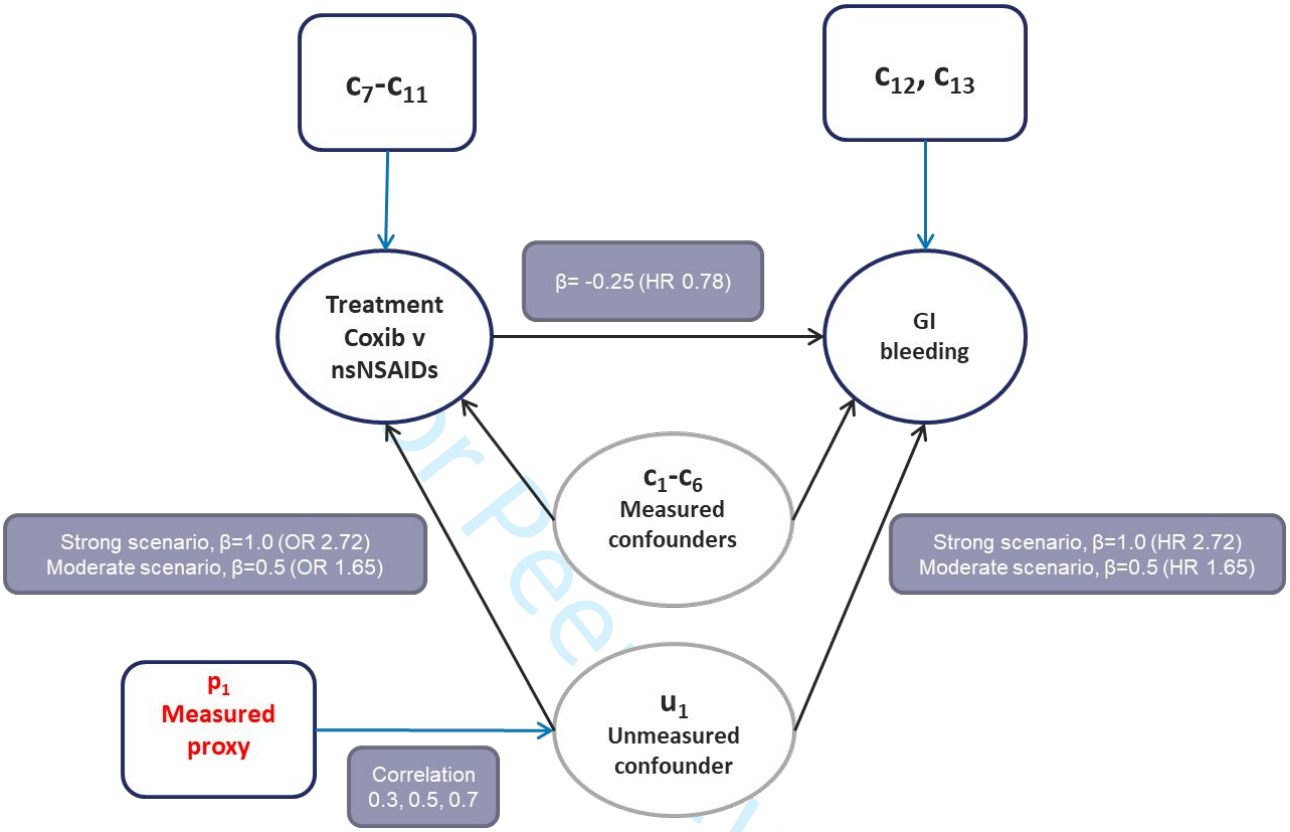
There are some important caveats to the proposed approach that deserve discussion. While simulation based QBA using individual-level data has more flexibility and can accommodate greater complexity than QBA based on aggregate-level data; it may still not fully capture the complexities observed in real-world data. Further, setting up realistic simulation scenarios at the design stage may be challenging as investigators are unlikely to be fully familiar with relationships among variables for their specific question of interest. For instance, in our simulations, we only generated 6 measured confounders and didn't simulate all correlations between measured confounders and unmeasured confounder. In real world data, many more confounders may exist with correlations among them and with unmeasured confounders. Therefore, results from the current simulation based QBA should be interpreted considering the simplifying assumptions made in the data generation steps. This issue can be addressed through more complex simulations that represent a more realistic description of the relationships

among multiple confounders and outcomes. Investigators can use the 'plasmode' framework[22] that builds on actual patient-level data as the basis of simulations to preserve the naturally occurring correlation patterns among study variables. However, investigators may not have access to empirical data during the planning of a study; therefore, our approach based on *de novo* simulations maybe more applicable at the design stage. We also recognize that reliable measurements for the correlation between measured proxies and unmeasured confounder may not be available. In such circumstances, investigators may need to run QBA under numerous scenarios to explicitly acknowledge this uncertainty.

In conclusion, a simulation based QBA approach based on covariate structures generated with individual-level data was useful to explicitly assess the potential for bias due to unmeasured confounding in a proposed study in the design stage. The flexibility, clarity, and relative ease of implementation of this approach makes it an important tool when planning studies.

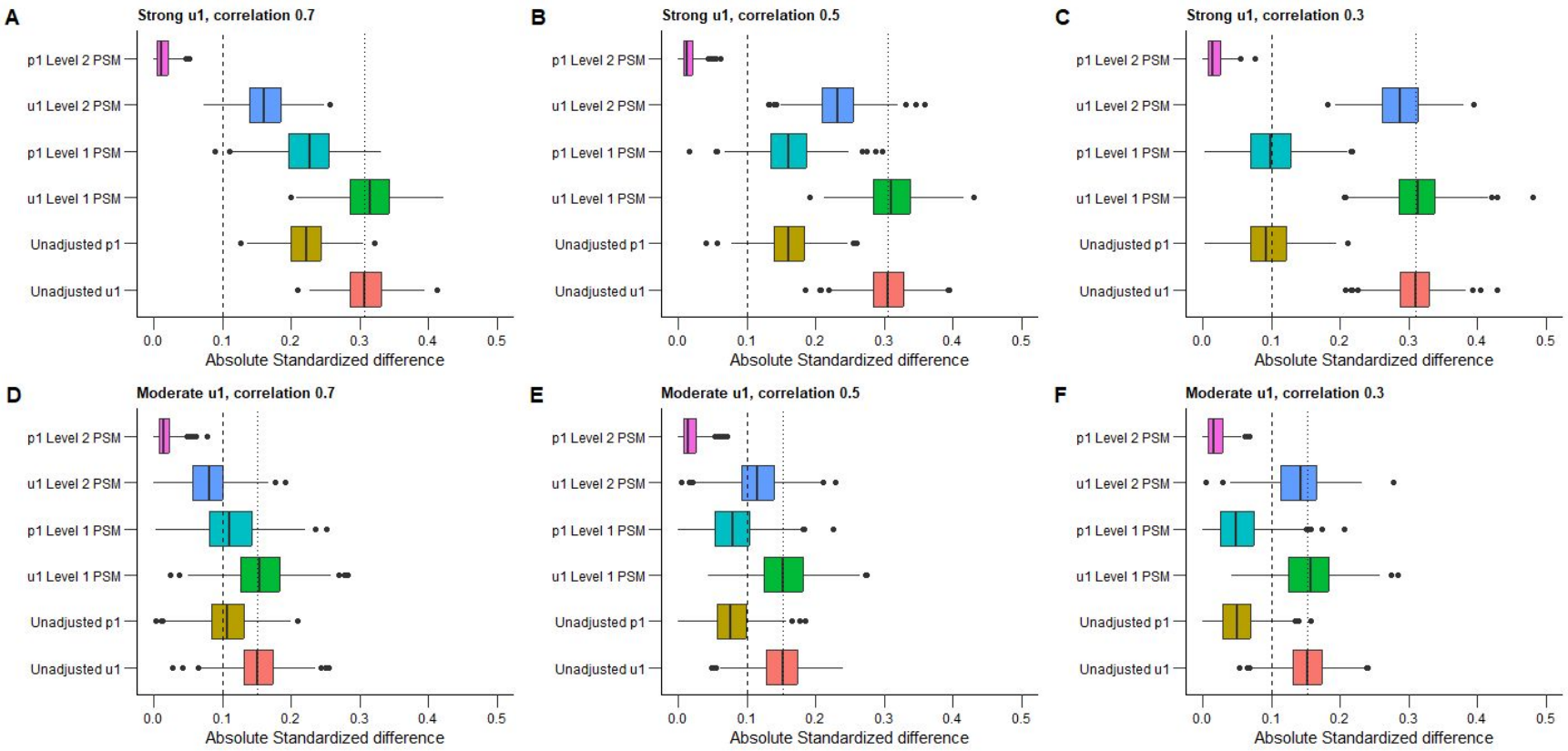
Figures

Figure 1: Data generating mechanism



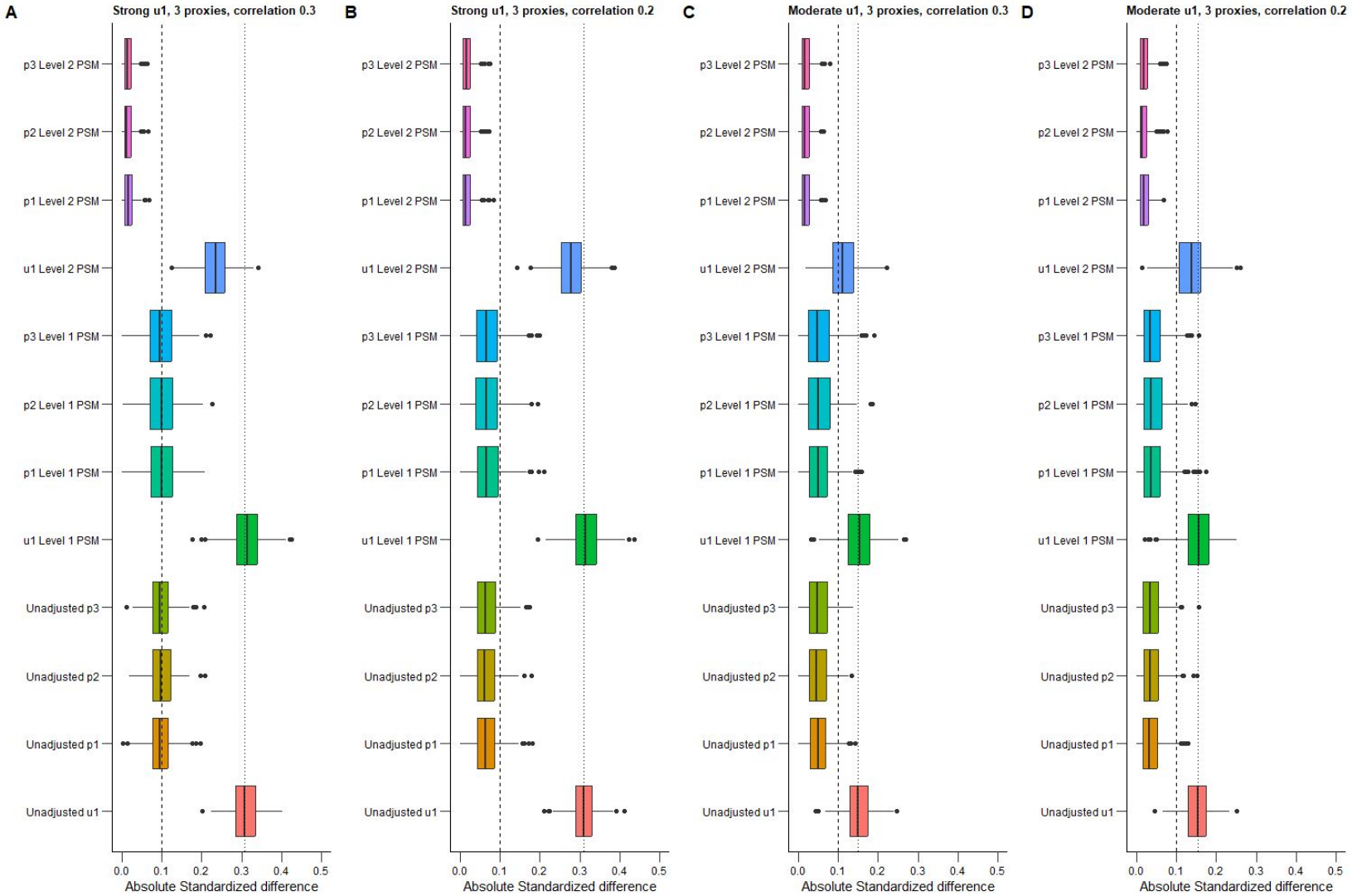
Abbreviations: GI- gastrointestinal; nsNSAID- nonselective nonsteroidal anti-inflammatory drugs

Figure 2: Balance in the unmeasured confounder and proxy variable with different levels of adjustments, one proxy variable scenarios



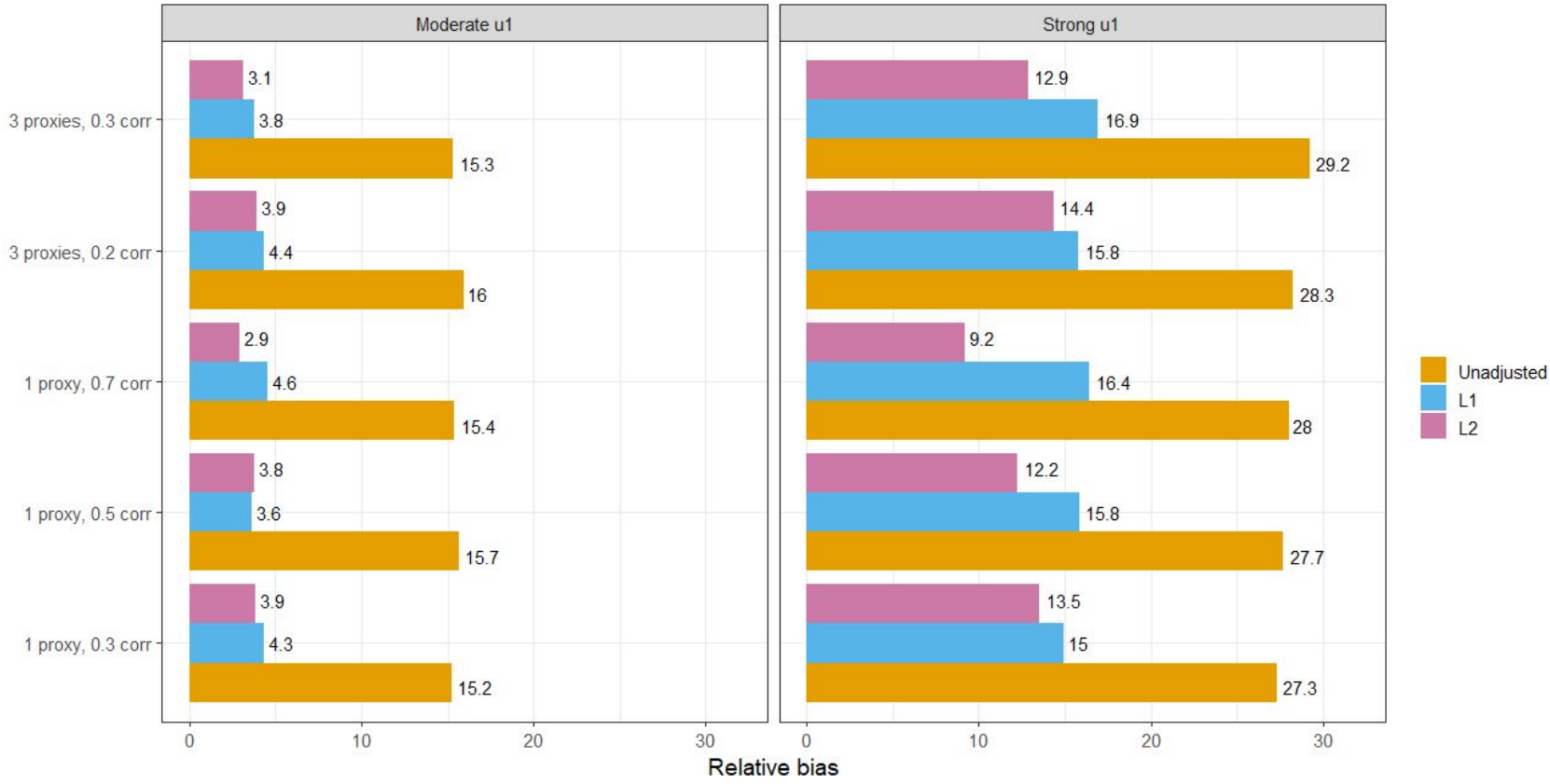
Footnote: Unadjusted analysis had no variables included for adjustment, Level 1 adjustment had only measured confounders included in the propensity score model (c1-c6), Level 2 adjustment had measured confounders + proxy for unmeasured confounder included in the propensity score model included (c1-c6+p1)

Figure 3: Balance in the unmeasured confounder and proxy variables with different levels of adjustments, three proxy variable scenarios



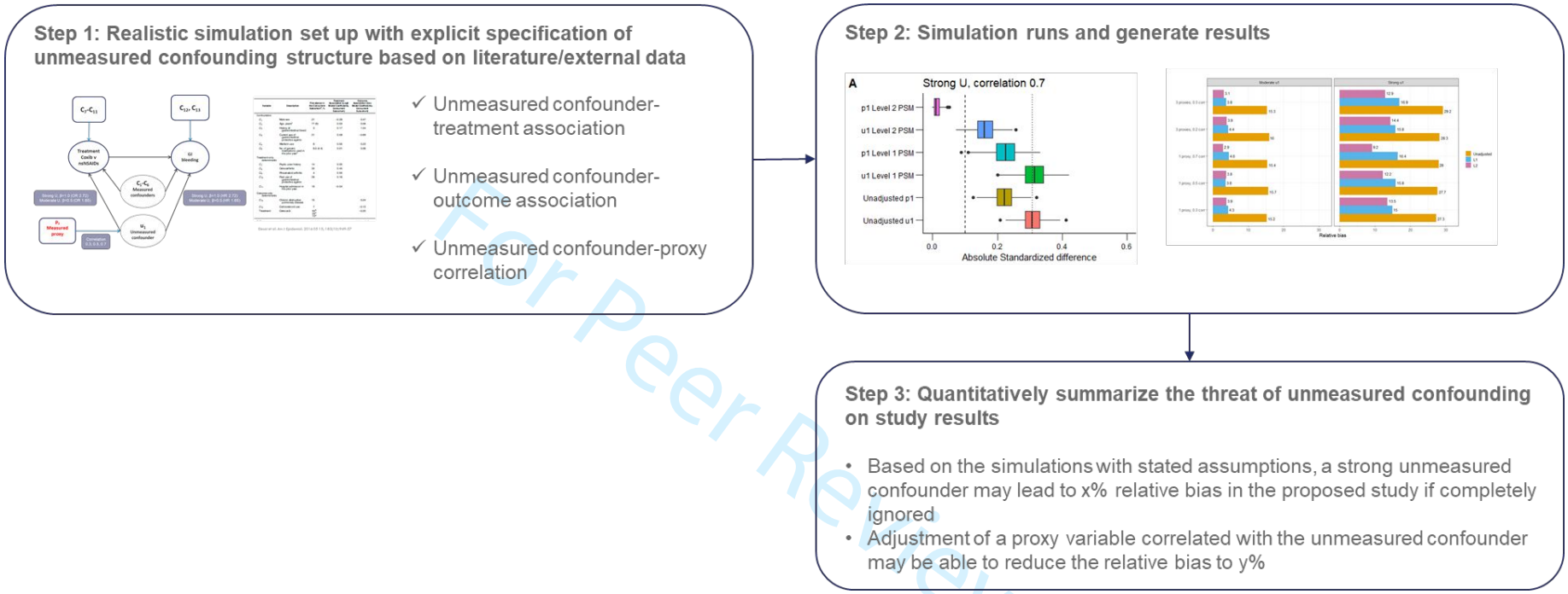
Footnote: Unadjusted analysis had no variables included for adjustment, Level 1 adjustment had only measured confounders included in the propensity score model (c1-c6), Level 2 adjustment had measured confounders + proxy for unmeasured confounder included in the propensity score model included (c1-c6+p1-p3)

Figure 4: Relative bias in treatment effect estimates



Footnote: Unadjusted analysis had no variables included for adjustment, L1 adjustment had only measured confounders included in the propensity score model (c1-c6), L2 adjustment had measured confounders + proxy for unmeasured confounder included in the propensity score model included (c1-c6+p1 in 1 proxy scenarios; c1-c6+p1-p3 in 3 proxy scenarios)

Figure 5: Application framework for the proposed simulation-based quantitative bias analysis



References

1. Vanderweele TJ, Arah OA. Bias formulas for sensitivity analysis of unmeasured confounding for general outcomes, treatments, and confounders. *Epidemiology*. Jan 2011;22(1):42-52. doi:10.1097/EDE.0b013e3181f74493
2. VanderWeele TJ. The sign of the bias of unmeasured confounding. *Biometrics*. Sep 2008;64(3):702-706. doi:10.1111/j.1541-0420.2007.00957.x
3. Considerations for the Use of Real-World Data and Real-World Evidence To Support Regulatory Decision-Making for Drug and Biological Products (2021).
4. Schneeweiss S. Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics. *Pharmacoepidemiology and drug safety*. May 2006;15(5):291-303. doi:10.1002/pds.1200
5. Lash TL, Fox MP, MacLehose RF, Maldonado G, McCandless LC, Greenland S. Good practices for quantitative bias analysis. *Int J Epidemiol*. Dec 2014;43(6):1969-85. doi:10.1093/ije/dyu149
6. Bross ID. Spurious effects from an extraneous variable. *Journal of chronic diseases*. 1966;19(6):637-647.
7. Ding P, VanderWeele TJ. Sensitivity analysis without assumptions. *Epidemiology (Cambridge, Mass)*. 2016;27(3):368.
8. Rosenbaum PR, Rubin DB. Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome. *Journal of the Royal Statistical Society: Series B (Methodological)*. 1983;45(2):212-218.
9. Lin DY, Psaty BM, Kronmal RA. Assessing the sensitivity of regression results to unmeasured confounders in observational studies. *Biometrics*. 1998:948-963.
10. Barberio J, Ahern TP, MacLehose RF, et al. Assessing techniques for quantifying the impact of bias due to an unmeasured confounder: an applied example. *Clinical Epidemiology*. 2021:627-635.
11. Groenwold RH, Nelson DB, Nichol KL, Hoes AW, Hak E. Sensitivity analyses to estimate the potential impact of unmeasured confounding in causal research. *International journal of epidemiology*. 2010;39(1):107-117.
12. Fox MP, Lash TL. Quantitative bias analysis for study and grant planning. *Ann Epidemiol*. Mar 2020;43:32-36. doi:10.1016/j.annepidem.2020.01.013
13. Banack HR, Hayes-Larson E, Mayeda ER. Monte Carlo simulation approaches for quantitative bias analysis: a tutorial. *Epidemiologic Reviews*. 2021;43(1):106-117.
14. Nørgaard M, Ehrenstein V, Vandenbroucke JP. Confounding in observational studies based on large health care databases: problems and potential solutions–a primer for the clinician. *Clinical epidemiology*. 2017:185-193.
15. Schneeweiss S, Rassen JA, Glynn RJ, Avorn J, Mogun H, Brookhart MA. High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. *Epidemiology (Cambridge, Mass)*. 2009;20(4):512.
16. Desai RJ, Glynn RJ, Wang S, Gagne JJ. Performance of Disease Risk Score Matching in Nested Case-Control Studies: A Simulation Study. *American journal of epidemiology*. May 15, 2016 2016;183(10):949-957. doi:10.1093/aje/kwv269
17. Bender R, Augustin T, Blettner M. Generating survival times to simulate Cox proportional hazards models. *Statistics in medicine*. Jun 15 2005;24(11):1713-23. doi:10.1002/sim.2059
18. Emrich LJ, Piedmonte MR. A method for generating high-dimensional multivariate binary variates. *The American Statistician*. 1991;45(4):302-304.
19. Austin PC. Some Methods of Propensity-Score Matching had Superior Performance to Others: Results of an Empirical Investigation and Monte Carlo simulations. *Biometrical journal*. 2009;51(1):171-184.

20. Austin PC. Using the standardized difference to compare the prevalence of a binary variable between two groups in observational research. *Communications in Statistics-Simulation and Computation*. 2009;38(6):1228-1234.
21. Wyss R, Plasek JM, Zhou L, et al. Scalable Feature Engineering from Electronic Free Text Notes to Supplement Confounding Adjustment of Claims-Based Pharmacoepidemiologic Studies. *Clin Pharmacol Ther*. Apr 2023;113(4):832-838. doi:10.1002/cpt.2826
22. Franklin JM, Schneeweiss S, Polinski JM, Rassen JA. Plasmode simulation for the evaluation of pharmacoepidemiologic methods in complex healthcare databases. *Computational statistics & data analysis*. Apr 2014;72:219-226. doi:10.1016/j.csda.2013.10.018

Appendix: Software instructions to execute customized, study specific simulations

We provide an R package, simBA, to accompany this manuscript to allow users to execute simulations tailored to their study questions and needs.

Step 1: Install the R functions on the following link:

https://gitlab.partners.org/rjd48/unmeasured-confounding-simulations.git

Two R scripts posted on the page are both needed and should be loaded in the environment.

Step 2: Prepare an input table to base the data generation in the following format

Keep column headings unchanged. The table can be populated with user desired values for all columns. For binary variables being simulated, provide desired prevalence and for continuous variables, provide desired mean and standard deviation. Variables with empty cells next to them in the coeff_treatment_model and coeff_outcome_model columns will not be included in those models while generating the data. True confounders should have coefficients supplied for both these models in last two columns. Under 'type', only permissible values are 'binary' or 'continuous'.

Variable	Description	Type	prevalence	mean	sd	coeff_treatment_model	coeff_outcome_model
c1	Gender	binary	0.21			-0.2783	0.46667
c2	Age	continuous		77	7.6	0.0321	0.07803
c3	GI complication history	binary	0.03			0.166	1.03895
c4	Concurrent GI protective agent use	binary	0.21			0.091	-0.68778
c5	warfarin use history	binary	0.08			0.5499	0.23121
c6	Number of hospitalizations	continuous		8	4.4	0.011	0.05539
c7	gastric ulcer history	binary	0.14			0.0488	
c8	OA diagnosis	binary	0.35			0.4507	
c9	RA diagnosis	binary	0.04			0.5488	
c10	GI protective agent use history	binary	0.28			0.1569	
c11	Any hospital admission in the CAP	binary	0.18			-0.0435	

c12	COPD	binary	0.15				0.24222
c13	Corticosteroid use history	binary	0.07				-0.13412
u1	unmeasured	binary	0.1			1	1

Step 3: Supply arguments to the simBA_de_novo function to run simulation based bias analysis for unmeasured confounding

Following is the function call, all the arguments are required and described in detail below

```
Scenario1 <- simBA_de_novo(iterations=500,
parameter_file_path = "C:\\Users\\rjd48\\SimBA input table.xlsx",
size=5000, treatment_prevalence=0.25, treatment_coeff= -0.25,
outcome_prevalence=0.05, dist= 'E',
unmeasured_conf="u1", n_proxies=2, proxy_type='binary', corr=0.5)
```

Argument	**Details and possible values**
iterations	Number of simulation runs
parameter_file_path	Path for the Excel file input table containing desired simulation modeling choices (as described in step 2)
size	Number of observations per simulation run
treatment_prevalence	Desired treatment prevalence (between 0 & 1)
treatment coefficient	Desired treatment effect
outcome_prevalence	Desired outcome prevalence (between 0 & 1)
dist	Distribution of the survival time (default= 'E' for exponential; other option 'W' for Weibull)
unmeasured_conf	Identify the unmeasured confounder in the simulations (variable name for unmeasured confounder provided in the parameter_file by the user in quotes e.g "u1")
n_proxies	Number of proxy variables
proxy_type	'binary' or 'continuous'
corr	Desired correlation between unmeasured confounder and proxy variable (between 0 and 1- if more than 1 proxies than

	independent correlation equal to the value of corr will be generated for all proxy variables)

As described in the manuscript, these simulations generate time-to-event outcomes from a Cox model assuming exponential distribution with a constant baseline hazard by default and has an option to relax this assumption with Weibull distribution. For analysis, 1:1 propensity score matching is used for adjustment and Cox proportional hazard models give hazard ratios.

Step 4: Evaluate results

The function in step 3 will output two tables in a list.

The first table will have average standardized mean differences in the unmeasured confounder and proxies (when n_proxies>0) across simulation iterations at up to 3 levels of adjustment: 1) crude (no adjustment), 2) L1 (adjustment for all the measured confounders), 3) L2 (when n_proxies>0, adjustment for proxies in addition to all measured confounders).

The second table will have average hazard ratios across simulation iterations at up to 3 levels of adjustment: 1) crude (no adjustment), 2) L1 (adjustment for all the measured confounders), 3) L2 (when n_proxies>0, adjustment for proxies in addition to all measured confounders)