

Using causal frameworks to reduce bias in observational TB research: a comparison of model-building approaches

L. Barcellini,¹ S. Sauer,² M. Romo,² C.D. Mitnick,^{2,3,4} H. Huerga,⁵ U. Khan,⁶ C. Hewison,⁷ M.L. Rich,^{3,4} M.F. Franke,² P.Y. Khan^{6,8}

¹Department of Paediatrics, “V. Buzzi” Children’s Hospital, ASST FBF Sacco, Milan, Italy; ²Department of Global Health and Social Medicine, Harvard Medical School, Boston, MA, USA; ³Partners In Health, Boston, MA, USA; ⁴Division of Global Health Equity, Brigham and Women’s Hospital, Boston, MA, USA; ⁵Field Epidemiology Department, Epicentre, Paris, France; ⁶Interactive Research and Development (IRD) Global, Singapore, Singapore; ⁷Medical Department, Médecins Sans Frontières, Paris, France; ⁸TB Centre, London School of Hygiene & Tropical Medicine, London, UK.

SUMMARY

BACKGROUND: Observational studies investigating causal (aetiological) questions often address confounding bias using data-driven predictive models where variables are selected regardless of their causal role, challenging interpretation. We compared the estimated effect of HIV co-infection on end-of-treatment outcomes among people with multidrug/rifampicin-resistant TB using a causal framework model and a data-driven predictive model.

METHODS: The causal framework guided confounder adjustment. Results were compared to those from models that applied alternative variable selection strategies (based on *P* values and change-in-estimate), which do not draw on domain knowledge to identify relevant variables.

RESULTS: The model informed by a causal diagram indicated that people living with HIV had a 31% lower probability of achieving a successful outcome compared

to those without HIV (adjusted relative risk [aRR] 0.69, 95% confidence interval [CI]: 0.41–0.98). In contrast, data-driven models produced attenuated associations (aRR 0.78, 95% CI: 0.50–1.06 and aRR 0.80, 95% CI: 0.5–1.09 for the *P* value and change-in-estimate strategies, respectively), reflecting omission of key confounders, and inappropriate inclusion of mediators and colliders.

DISCUSSION: When the research question is aetiological, using a causal approach to guide variable selection ensures proper adjustment, improves interpretability, and establishes a stronger foundation for future observational research.

KEY WORDS: tuberculosis; MDR/RR-TB; HIV; observational studies; epidemiologic methods; causal inference; causal structure

Observational studies may aim to describe, identify predictors, or estimate causal effects.¹ This ambiguity can lead to confusion in the analytical approach, including variable selection and interpretation of results.^{2,3} More critically, researchers often infer causality from observational data without anchoring their analyses in a causal framework. When our aim is to determine the effect of an exposure on an outcome, we need to address confounding, a bias introduced by variables influencing both the exposure and the outcome and distorting their relationship. To reduce this bias, we select an *adjustment set* – a group of variables that we control for in the analysis – to isolate the effect of interest. Variables should be selected based on hypothesised exposure–outcome relationships, supported by existing evidence or expert knowledge. These relationships can be summarised in a causal diagram,⁴ making explicit the assumptions underpinning the analysis. This approach facilitates the identification of all confounders while avoiding adjustment for

variables that could introduce bias, such as colliders (variables influenced by both the exposure and the outcome).⁵ This causal framework approach differs from prediction analyses, wherein variables are selected based on their ability to improve model performance.

We present a practical example, applying a causal framework to estimate the effect of HIV infection on treatment outcomes among individuals with multidrug-resistant/rifampicin-resistant TB (MDR/RR-TB). We compare results obtained from a model grounded in a causal framework with those from models selecting covariates without accounting for causal structure.

METHODS

We used data from the endTB (Expand New Drug Markets for TB) observational cohort study (NCT03259269), including patients who started treatment with a bedaquiline (BDQ) and/or delamanid (DLM)-containing regimen at an endTB site across 17 countries between 1 April 2015 and 30 September

MFF and PYK are joint last authors.

Correspondence to: Lucia Barcellini, Department of Paediatrics, Vittore Buzzi Children’s Hospital, via Castelvetro 32, 20154 Milan, Italy. E-mail: lucia.barcellini@gmail.com

Article submitted 5 February 2026. Final version accepted 24 April 2026.

2018.⁶ Details are provided elsewhere⁷ and in the Supplementary Data.

Outcome

MDR/RR-TB end-of-treatment outcomes were recorded using the WHO 2014 definitions.⁸ Successful treatment outcome was a composite, comprising cure and treatment completion.

Exposure of interest

HIV status of individuals starting MDR/RR-TB treatment, established by serology or self-report prior HIV diagnosis.

Covariates

Covariates included demographic, clinical, comorbidity, microbiological, and disease-related characteristics (detailed in the Supplementary Data).

Analysis

We conducted multivariable logistic regression to assess the association between HIV status and successful MDR/RR-TB treatment outcome. We used marginal standardisation^{9,10} to calculate predicted probabilities and risk ratios (RRs) with 95% confidence intervals (CIs). We constructed three multivariable models to illustrate the influence of different variable selection methods on adjustment sets and effect estimates.

Building the causal framework

To determine the appropriate adjustment set, we constructed a causal diagram (Figure 1) that synthesised existing evidence and expert reasoning to map the assumed relationships among HIV, MDR/RR-TB end-of-treatment outcome, and other variables. We identified age, sex, country, and year of enrolment as potential confounders. Age and sex influence care-seeking and treatment adherence¹¹ while country and year reflect variation in care standards and delivery that may affect both diagnosis and treatment. To account for socio-economic vulnerability, we included marital status (proxy for social support), homelessness or refugee status, incarceration history, and intravenous drug use (IVDU) (proxy for social exclusion).^{12,13} We classified as confounders comorbidities that adversely influence treatment outcomes but are not downstream consequences of HIV, including diabetes, cirrhosis, chronic obstructive pulmonary disease (COPD), and chronic hepatitis B or C. Diabetes is associated with delayed sputum conversion and a higher risk of treatment failure or relapse.¹⁴ Cirrhosis and COPD may heighten the risk of treatment-related adverse events and restrict the use of certain anti-TB drugs.¹⁵ Hepatitis B and C have been linked to an elevated risk of drug-induced liver injury during TB treatment.¹⁶ These comorbidities frequently co-occur with HIV infection,

either through biological mechanisms or shared behavioural and social risk factors. We considered drug resistance patterns to lie on the causal pathway, as more resistance can lead to poorer treatment outcomes¹⁷ by limiting effective therapeutic options. People living with HIV (PLHIV) may have higher levels of drug resistance due to increased TB exposure within health care settings and vulnerability to active disease (primary resistance) and to treatment-related factors such as impaired drug absorption and complex polypharmacy (secondary resistance).¹⁸ We considered comorbidities affected by HIV to be on the causal pathway. PLHIV are at increased risk of mental health disorders,¹⁹ cardiovascular and renal comorbidities due to HIV-related immune activation, and systemic inflammation, and prone to antiretroviral therapy (ART) toxicity.^{20,21} We proposed that HIV infection worsens MDR/RR-TB treatment outcomes both directly by compromising the immune response²² and indirectly, by delaying diagnosis and leading to greater disease severity and poorer general condition.²³ PLHIV often present with extra-pulmonary disease, atypical pulmonary disease, or paucibacillary forms,²⁴ which can hinder timely diagnosis.²⁵ We considered baseline anaemia and low body mass index (BMI) as proxies for poor general condition, possibly attributable to HIV,^{26,27} placing them on the causal pathway. This assumption is debated as anaemia and low BMI can also arise from poor diet and/or food insecurity, conditions which may co-occur with but not be caused by HIV, and which lead to unfavourable outcomes via impaired cell-mediated immunity response to TB.²⁸ In the latter scenario, these conditions would be confounders requiring adjustment. In the former, they are a consequence of HIV disease and may induce bias if adjusted.

Causal model

Model 1 used a structural approach to identify confounders, selecting adjustment variables based on the causal framework. The final adjustment set included socio-demographic factors (age, sex, country, year, marital status, homelessness, refugee status, history of incarceration, IVDU, and alcohol use) and comorbid conditions that were not downstream of HIV (i.e., COPD, diabetes, hepatitis B virus [HBV], and hepatitis C virus [HCV]). We conducted sensitivity analyses considering low BMI and anaemia to occur with, rather than result from, HIV, and therefore were confounders requiring adjustment. To assess the robustness of the causal estimate to potential unmeasured confounding, we computed the E-value, which represents the minimum strength of association (on the RR scale) that an unmeasured confounder would need to have with both the exposure and the outcome, conditional on measured covariates, to fully explain away the observed effect of the primary exposure (HIV infection).²⁹

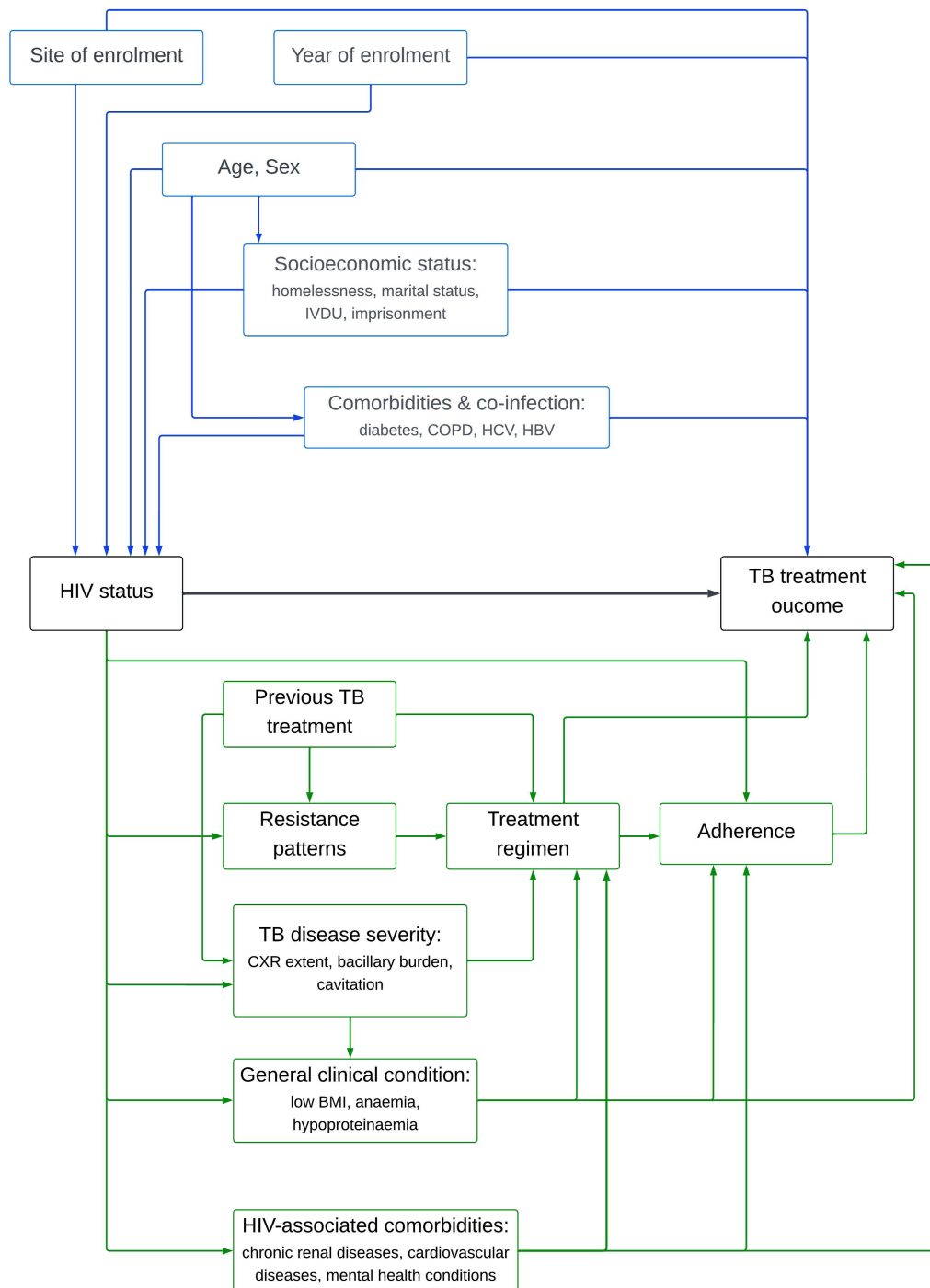


Figure 1. Graphical theoretical framework of the focal relationship of interest – HIV status (exposure of interest) and successful MDR/RR-TB treatment outcome (outcome of interest; composite: cured and completed). HIV status is defined without accounting for antiretroviral therapy. Variables and their measurements are represented as nodes connected by unidirectional arrows (i.e., directed), which depict hypothesised causal relationships based on prior knowledge and theory. An arrow between two nodes indicates the assumed existence and direction of a causal relationship. A path consists of one or more arrows connecting two nodes and may be open or closed; open paths transmit statistical associations, whereas closed paths do not. A causal path is defined as a sequence of arrows all pointing in the same direction from the exposure to the outcome. With respect to this focal relationship, a confounder is a common cause of both the exposure and the outcome; a mediator lies on the causal pathway between the exposure and outcome. A direct causal effect refers to the effect not operating through one or more specified mediators. The total causal effect of HIV on treatment outcome is the combined effect transmitted through all causal paths connecting the exposure and the outcome. MDR/RR-TB = multidrug-resistant/rifampicin-resistant TB; IVDU = intravenous drug use; COPD = chronic obstructive pulmonary disease; HCV = hepatitis C virus; HBV = hepatitis B virus; CXR = chest X-ray; BMI = body mass index.

Building data-driven models

Models 2 and 3 applied data-driven strategies to variable selection. Model 2 was developed using an automated backwards stepwise selection, with a *P*-value threshold of 0.1 for variable removal and 0.05 for variable addition. Model 3 started with a parsimonious model (HIV, age, sex) and added variables incrementally based on their influence on the HIV estimate. Variables were retained in the model until changes in the estimate for HIV were less than 0.02 or the model became saturated (change-in-estimate [CIE] approach).³⁰

Missingness

We used a two-step approach to handle missing covariate data. In the causal model, we defined the adjustment set a priori based on the theoretical framework. In Models 2 (backward selection) and 3 (CIE), the missing indicator method was used during variable selection. Multiple imputation was then applied to all final models to address missingness in the selected covariates.

Ethical statement

The endTB observational study protocol was approved by central ethics review committees for each consortium partner, and local ethics approval was obtained in all endTB countries. Participants provided written informed consent for inclusion in the observational cohort.

RESULTS

The analysis sample included 2,220 individuals with MDR/RR-TB (Supplementary Data Figure S1), of whom 262 (11.8%) were PLHIV (Supplementary Data Table S1). Participants with missing HIV status

(*n* = 1) or missing treatment outcome (*n* = 92) were excluded from the analysis sample (Supplementary Data Figure S1). Among PLHIV, 198 were on ART (198/262, 75.6%). Overall, 64.1% (168/262) of PLHIV experienced a successful treatment outcome compared to 79.8% (1,555/1,958) of HIV-negative individuals (Supplementary Data Table S2). Missingness across key variables is reported in Supplementary Data Table S3.

Comparison of modelling approaches

Adjustment sets and effect estimates are shown in Table 1 and Figure 2.

Model 1: Causal model

Confounders were selected based on the theoretical causal framework described above. The adjusted relative risk (aRR) for the effect estimate of HIV on successful treatment outcome was 0.69 (95% CI: 0.41–0.98). In a sensitivity analysis additionally adjusting for BMI and anaemia, the aRR was 0.73 (95% CI: 0.47–1.13) (Supplementary Data Table S4). The calculated E-value for the point estimate was 2.26 (Figure 3).

Model 2: Backward selection

Model 2 included a larger number of covariates such as extensive disease (high smear grade with cavitary disease), anaemia, chronic renal disease, depression, psychiatric disorders, and low BMI. When confronted with the causal framework used to build model 1, extensive disease acts as a collider, while the other variables are mediators on the causal pathway between HIV and successful treatment outcomes (Table 2). The aRR for HIV from model 2 was 0.78 (95% CI: 0.50–1.06).

Table 1. Adjustment set, predicted probabilities, and relative risk (RR) of the effect of HIV status on multidrug-resistant/rifampicin-resistant TB treatment outcomes using different model-building strategies.

Model-building strategy	Adjustment set	Analysis sample	Predicted probability of successful outcome (95% CI)	Adjusted RR	95% CI
Unadjusted RR	–	2,220		0.78	0.76–0.79
Model 1: causal framework	Age, sex, site, year, history of incarceration, marital status, IVDU, alcohol use, homelessness, COPD, diabetes, HCV, HBV	2,217 ^A	HIV-negative: 0.79 (0.77–0.80) PLHIV: 0.71 (0.64–0.78)	0.69	0.41–0.98
Model 2: stepwise backwards selection	Age, sex, site, chronic renal disease, diabetes, bilateral lung involvement, depression, marital status, history of incarceration, homelessness, positive baseline culture, psychiatric disorders, HCV, HBV, low BMI	2,217 ^A	HIV-negative: 0.78 (0.77–0.82) PLHIV: 0.73 (0.67–0.80)	0.78	0.50–1.06
Model 3: forward selection using change-in-estimate	Age, sex, anaemia, extensive disease, HCV, site, HBV, IVDU, low BMI, chronic renal disease, year	2,217 ^A	HIV-negative: 0.78 (0.76–0.80) PLHIV: 0.74 (0.67–0.80)	0.80	0.51–1.09

CI = confidence interval; IVDU = intravenous drug use; COPD = chronic obstructive pulmonary disease; HCV = hepatitis C virus; HBV = hepatitis B virus; PLHIV = people living with HIV; BMI = body mass index.

^AAnalysis sample *n* = 2,220, dropped Kenya (3 observations).

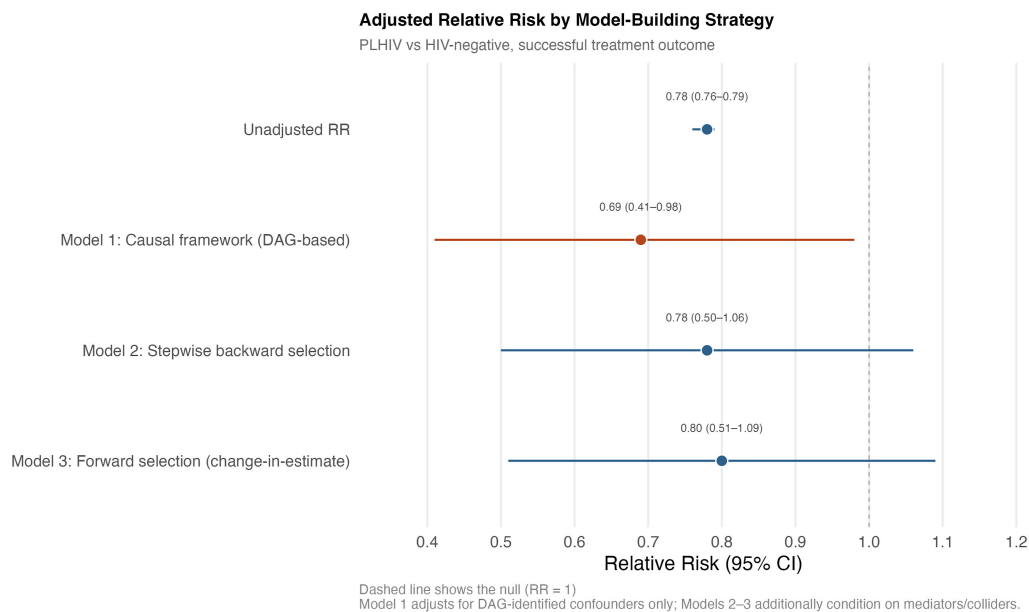


Figure 2. Forest plot of adjusted relative risk of successful MDR/RR-TB treatment outcome among people living with HIV compared with HIV-negative patients, by model-building strategy. MDR/RR-TB = multidrug-resistant/rifampicin-resistant TB; PLHIV = people living with HIV; CI = confidence interval.

Model 3: CIE approach

In Model 3, the adjustment set included age, sex, site, year, anaemia, chronic renal disease, IVDU, low BMI, HBV, HCV, and extensive disease. This resulted in the inclusion of colliders (extensive disease) and mediators (anaemia, low BMI, chronic renal disease), while omitting confounders (Table 2). This model produced an aRR of 0.80 (95% CI: 0.51–1.09) for HIV.

DISCUSSION

This analysis illustrates the application of a causal framework to address an aetiological question, namely estimating the total effect of HIV on MDR/RR-TB

treatment outcomes. Using causal theory to inform selection of adjustment variables, we found that PLHIV had a 31% lower probability of achieving a successful treatment outcome compared to those without HIV (aRR 0.69, 95% CI: 0.41–0.98). In contrast, the data-driven models estimated lower probabilities of 22% (aRR 0.78, 95% CI: 0.50–1.06) and 20% (aRR 0.80, 95% CI: 0.51–1.09). This attenuation is consistent with inappropriate adjustment for variables that lie on the causal pathway between HIV and treatment outcome. By conditioning on mediators, these models block part of the total causal effect of HIV on treatment outcomes, biasing the estimate toward the null.

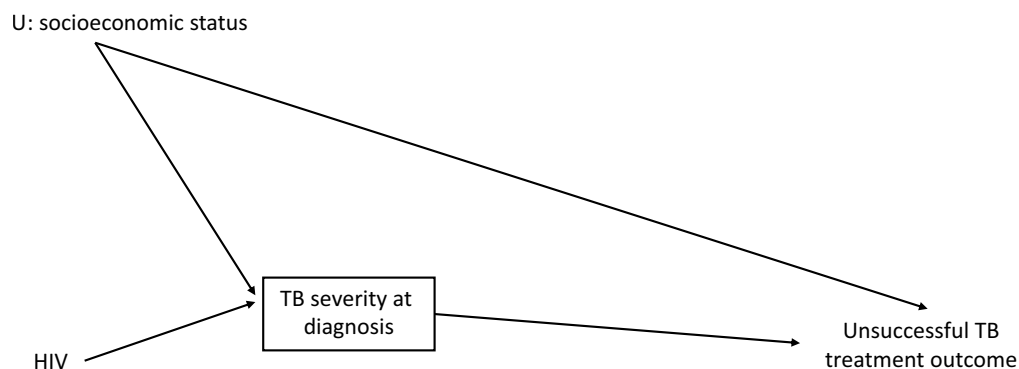


Figure 3. Causal diagram illustrating collider bias due to conditioning on TB severity. TB disease severity is influenced by both HIV status and other factors and, in turn, affects study inclusion and treatment outcomes. TB severity therefore lies on a collider path between HIV status and treatment outcome. A collider path is a closed path between an exposure and an outcome that passes through a variable receiving two or more arrows and, by default, does not transmit statistical associations. Conditioning on TB severity (indicated by the square around the node) opens this path, creating a non-causal backdoor path between HIV and treatment outcome through severity, and can induce a spurious association between exposure and outcome, even in the absence of a true causal effect.

Table 2. Comparison of variable selection strategies across models and implications for bias. Confounders, mediators, and colliders are based on the proposed causal framework.

Model	Variable selection approach	Types of variables included	Rationale/notes
Model 1	Theory-driven (causal framework)	Confounders only – all identified confounders included based on theoretical reasoning ^A	Based on theoretical framework; aimed at estimating the total causal effect of HIV on MDR/RR-TB treatment outcome
Model 2	Stepwise backwards selection (<i>P</i> value based)	Mix of confounders, mediators (low BMI), and colliders (e.g., ‘extensive disease’ – see TB severity causal model in Figure 2) Confounders omitted: IVDU, alcohol use, homelessness, COPD, diabetes	Variable selection driven by statistical significance; ignores causal structure and may introduce bias by including mediators or colliders and omitting key confounders
Model 3	Change-in-estimate (forward selection)	Mix of confounders, mediators (anaemia, low BMI), and colliders (e.g., ‘extensive disease’ – see TB severity causal model in Figure 2) Confounders left out of model (history of incarceration, marital status, homelessness, COPD, diabetes)	Variable selection driven by statistical significance; ignores causal structure and may introduce bias by including mediators or colliders and omitting key confounders

MDR/RR-TB = multidrug-resistant/rifampicin-resistant TB; BMI = body mass index; IVDU = intravenous drug use; COPD = chronic obstructive pulmonary disease.
^AUnder a structural definition of confounders, a sufficient adjustment set for a specific *estimand* (i.e., the causal effects being sought to determine) refers to a set of variables that, when fully controlled for, leads to an unbiased estimate by closing non-causal paths and keeping causal paths open (Hernán MA, et al. Causal Knowledge as a Prerequisite for Confounding Evaluation: An Application to Birth Defects Epidemiology. American Journal of Epidemiology. 2002; 155(2): 176-184).

This finding underscores the importance of selecting adjustment variables based on causal reasoning rather than statistical associations alone. Data-driven selection strategies may exclude important confounders or adjust for mediators or colliders, introducing bias and obscuring the true relationship between the exposure and outcome. Variables that distort causal inference may remain strong predictors of the outcome, which explains their persistence in data-driven models. This illustrates the ambiguity inherent in many observational studies, where variables informative for description or prediction may be inappropriate for estimating causal effects, underscoring the need to explicitly define the analytical objective. This is also true for non-data-driven approaches, such as simply including known ‘risk factors’ for the outcome in the adjustment set. This latter approach requires further vetting of how the risk factors relate to the exposure and outcome of interest to avoid the same pitfalls of data-driven selection strategies.

In practice, failure to encode causal structure can bias estimates through three main mechanisms: 1) Omission of confounders: Variable selection based solely on *P* values may overlook confounders that have weak statistical associations with the outcome in the sample but strong theoretical links to both exposure and outcome.³¹ In our analysis, IVDU was excluded from the data-driven model despite being a known confounder. 2) Inclusion of mediators: Failing to consider the causal structure may lead to adjustment for mediators. In model 3, the adjustment set includes low BMI and anaemia, both of which are likely mediators. Because mediators lie along the causal pathway through which the exposure exerts its effect, adjusting for them blocks part of that pathway and attenuates the estimated association, leading to an underestimation of the

total effect.³² 3) Inclusion of colliders: Collider stratification bias is a type of selection bias that occurs when we adjust for a variable influenced by both the exposure and outcome. This can occur at both design and analysis stages, through restricting the study population or statistical adjustment.³³ Both data-driven models adjusted for proxies of poor medical condition, a likely collider. HIV worsens medical conditions at diagnosis due to diagnostic delays and immune dysfunction, while socioeconomic factors may influence both condition and MDR/RR-TB treatment outcomes. Adjusting for poor condition can therefore induce spurious associations and bias the estimated effect (Figure 3).

This illustrative example shows how causal frameworks guide analysis and interpretation; further refinements could strengthen the estimation of the causal effect of HIV on MDR/RR-TB outcomes. Causal model assumptions: Placing certain variables within the causal framework is debatable and inherently complex – for instance, the role of BMI and anaemia. We conducted sensitivity analyses in which low BMI and anaemia were treated as upstream risk factors that co-occur with HIV rather than as downstream consequences, resulting similar estimates. Importantly, ART was not accounted for in this analysis. The positions of several variables within the causal framework would likely differ between people on ART with viral suppression and those not receiving ART. Addressing this would constitute a different causal question and would require explicitly accounting for these differences. Selection bias: TB disease severity influences both study inclusion and treatment outcomes (Figure 2). Because the underlying cohort study was limited to microbiologically confirmed MDR/RR-TB, PLHIV with pulmonary disease were more likely to be included, whilst those with

paucibacillary or atypical presentations, which is nearly half of HIV-associated TB,³⁴ may have been missed. If present, this selection bias could lead to an underrepresentation of advanced HIV-associated MDR/RR-TB and an over-estimation of treatment success among PLHIV, resulting in upward bias in the HIV effect estimate. In our cohort, this limitation common to most MDR/RR-TB programmatic cohorts remains; the use of both Xpert and culture-based inclusion criteria was intended to mitigate this source of selection bias. Residual confounding: This was a secondary analysis of data originally collected to evaluate the safety and effectiveness of BDQ and DLM-containing regimens. Not all potential confounders of the HIV and MDR/RR-TB treatment outcome relationship were recorded. Unmeasured factors that influence both HIV and treatment success could bias the estimates. We computed the E-value for the causal model estimate. The E-value of 2.26 for the point estimate suggests that a moderately strong unmeasured confounder would be required to fully explain away the observed effect. Finally, some argue that causal questions should focus on modifiable exposures.³⁵ In this analysis, HIV status is defined at the point of MDR/RR-TB treatment initiation, and the causal effect we are trying to estimate can be framed as the effect of initiating MDR/RR-TB treatment whilst living with HIV versus initiating treatment without HIV. While HIV status itself is not directly modifiable in the absence of cure, key determinants along the HIV care pathway are, including prevention of HIV acquisition and provision of effective ART to ameliorate immune dysfunction. For these reasons, causal questions about HIV therefore remain valuable: by quantifying its effects on specific health outcomes, it can inform and strengthen the case for investment in preventive and therapeutic interventions. Future studies with more granular data, such as CD4 cell counts, HIV viral load, and ART history, would enable a more nuanced evaluation of how variations in immune status, stage of HIV disease, and viral control modify MDR/RR-TB treatment outcomes.

CONCLUSION

This study demonstrates the practical application of a causal framework to estimate the effect of HIV status on MDR/RR-TB treatment outcome. It shows how variable selection strategies grounded in causal reasoning can alter adjustment sets and effect estimates, highlighting the value of explicit causal thinking in strengthening the credibility and interpretability of epidemiological analyses.

Acknowledgements

The authors thank the patients who participated in the endTB observational study and the clinicians and programme staff of participating national TB programmes. The authors also thank the

endTB staff at Partners in Health, Doctors Without Borders, Epicentre, and Interactive Research and Development. Funding: the endTB project, including the observational study, is funded by Unitaids. Bedaquiline was donated by Janssen to the Global Drug Facility and delamanid was donated by Otsuka for initial participants enrolled in the endTB Observational Study. These companies did not have any role in study design, data analyses, data interpretation, or manuscript writing. Some of the data included in this analysis are managed in countries governed by the European Union General Data Protection Regulation (GDPR). The data contain sensitive and potentially identifying information and cannot be sufficiently anonymised to meet GDPR standards and retain their utility. Pseudo-anonymised data will be available upon request through the endTB Data Sharing Initiative (eDSI), hosted by Médecins Sans Frontières (MSF) and coordinated by the endTB consortium. For more details, please visit www.endTB.org/data-sharing-initiative.

Conflicts of interest: MRO reports research funding from the National Institutes of Health with payments made to their institution. CDM, MRi, and MFF report research funding from Unitaids, the National Institutes of Health. CDM also reports participation in scientific advisory boards for Akagera (one payment made to Partners In Health as an honorarium for service on the advisory board) and Otsuka (no payment). The remaining authors declare no potential conflicts of interest.

References

1. Ramspek CL, et al. Prediction or causality? A scoping review of their conflation within current observational research. *Eur J Epidemiol.* 2021;36(9):889-898.
2. Carlin JB, Moreno-Betancur M. On the uses and abuses of regression models: a call for reform of statistical practice and teaching. *Stat Med.* 2025;44(13-14):e10244.
3. Romo ML, et al. Conflation of prediction and causality in the TB literature. *IJTLD Open.* 2025;2(7):388-396.
4. Digitale JC, Martin JN, Glymour MM. Tutorial on directed acyclic graphs. *J Clin Epidemiol.* 2022;142:264-267.
5. Griffith GJ, et al. Collider bias undermines our understanding of COVID-19 disease risk and severity. *Nat Commun.* 2020;11(1):5749.
6. Khan U, et al. The endTB observational study protocol: treatment of MDR-TB with bedaquiline or delamanid containing regimens. *BMC Infect Dis.* 2019;19(1):733.
7. Huerga H, et al. Safety and effectiveness outcomes from a 14-country cohort of patients with multi-drug resistant tuberculosis treated concomitantly with bedaquiline, delamanid, and other second-line drugs. *Clin Infect Dis.* 2022;75(8):1307-1314.
8. World Health Organization. Definitions and reporting framework for tuberculosis - 2013 revision. Geneva: WHO, 2013.
9. Muller CJ, MacLehose RF. Estimating predicted probabilities from logistic regression: different methods correspond to different target populations. *Int J Epidemiol.* 2014;43(3):962-970.
10. Klein D. MIMRGNS: Stata module to run margins after mi estimate. Statistical Software Components S457795. Boston, MA: Boston College Department of Economics, 2022.
11. Galdas PM, Cheater F, Marshall P. Men and health help-seeking behaviour: literature review. *J Adv Nurs.* 2005;49(6):616-623.
12. Leung Soo C, et al. Socioeconomic factors impact the risk of HIV acquisition in the township population of South Africa: a Bayesian analysis. *PLOS Glob Public Health.* 2023;3(1):e0001502.
13. Kamara RF, et al. Social and health factors associated with adverse treatment outcomes among people with multidrug-resistant tuberculosis in Sierra Leone: a national, retrospective cohort study. *Lancet Glob Health.* 2022;10(4):e543-e554.
14. Noubiap JJ, et al. Global prevalence of diabetes in active tuberculosis: a systematic review and meta-analysis of data from

- 2·3 million patients with tuberculosis. *Lancet Glob Health*. 2019; 7(4):e448-e460.
15. Samuels JP, et al. Comorbidities and treatment outcomes in multidrug resistant tuberculosis: a systematic review and meta-analysis. *Sci Rep*. 2018;8(1):4980.
 16. Kim WS, et al. Hepatitis C and not hepatitis B virus is a risk factor for anti-tuberculosis drug induced liver injury. *BMC Infect Dis*. 2016;16(1):50.
 17. Sharling L, et al. Rifampin-resistant tuberculosis in the United States, 1998–2014. *Clin Infect Dis*. 2020;70(8):1596–1605.
 18. Khan PY, et al. Transmission of drug-resistant tuberculosis in HIV-endemic settings. *Lancet Infect Dis*. 2019;19(3):e77–e88.
 19. Gooden TE, et al. The risk of mental illness in people living with HIV in the UK: a propensity score-matched cohort study. *Lancet HIV*. 2022;9(3):e172–e181.
 20. Islam FM, et al. Relative risk of cardiovascular disease among people living with HIV: a systematic review and meta-analysis. *HIV Med*. 2012;13(8):453–468.
 21. Ekrikpo UE, et al. Chronic kidney disease in the global adult HIV-infected population: a systematic review and meta-analysis. *PLOS ONE*. 2018;13(4):e0195443.
 22. Lawn SD, Butera ST, Shinnick TM. Tuberculosis unleashed: the impact of human immunodeficiency virus infection on the host granulomatous response to *Mycobacterium tuberculosis*. *Microbes Infect*. 2002;4(6):635–646.
 23. Kwan Candice K, Ernst Joel D. HIV and tuberculosis: a deadly human syndrome. *Clin Microbiol Rev*. 2011;24(2):351–376.
 24. Aaron L, et al. Tuberculosis in HIV-infected patients: a comprehensive review. *Clin Microbiol Infect*. 2004;10(5):388–398.
 25. Getahun H, et al. Diagnosis of smear-negative pulmonary tuberculosis in people with HIV infection or AIDS in resource-constrained settings: informing urgent policy changes. *Lancet*. 2007;369(9578):2042–2049.
 26. Paton NI, Ng Y-M. Body composition studies in patients with wasting associated with tuberculosis. *Nutrition*. 2006;22(3):245–251.
 27. van Lettow M, et al. Micronutrient malnutrition and wasting in adults with pulmonary tuberculosis with and without HIV co-infection in Malawi. *BMC Infect Dis*. 2004;4(1):61.
 28. Cegielski J, McMurray D. The relationship between malnutrition and tuberculosis: evidence from studies in humans and experimental animals. *Int J Tuberc lung Dis*. 2004;8(3):286–298.
 29. VanderWeele T, Ding P. Sensitivity analysis in observational research: introducing the E-value. *Ann Intern Med*. 2017;167(4):268–274.
 30. Greenland S, Daniel R, Pearce N. Outcome modelling strategies in epidemiology: traditional methods and basic alternatives. *Int J Epidemiol*. 2016;45(2):565–575.
 31. Hernán MA, et al. Causal knowledge as a prerequisite for confounding evaluation: an application to birth defects epidemiology. *Am J Epidemiol*. 2002;155(2):176–184.
 32. Schisterman EF, Cole SR, Platt RW. Overadjustment bias and unnecessary adjustment in epidemiologic studies. *Epidemiology*. 2009;20(4):488–495.
 33. Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. *Epidemiology*. 2004;15(5):615–625.
 34. Gupta RK, et al. Prevalence of tuberculosis in post-mortem studies of HIV-infected adults and children in resource-limited settings: a systematic review and meta-analysis. *AIDS*. 2015; 29(15):1987–2002.
 35. Holland PW. Statistics and causal inference. *J Am Stat Assoc*. 1986;81(396):945–960.

IJTLD OPEN welcomes the submission of research articles on all aspects of TB and respiratory diseases such as asthma, bronchiectasis, COVID-19, COPD, child lung health and the hazards of tobacco and air pollution.

This is an open access article published by The Union under the terms of the Creative Commons Attribution License CC-BY.

For information on IJTLD OPEN see: <https://theunion.org/our-work/journals/ijtld-open> or contact: journal@theunion.org



If you found this article interesting, you can explore related courses offered by The Union.

Learn more here: <https://theunion.org/our-work/union-courses>