Received: 22 March 2022 | Accepted: 26 April 2022

DOI: 10.1111/cts.13299

COMMENTARY

Feeling anxious yet? Interpreting findings on drug safety from large healthcare databases

Greta A. Bushnell | **Daniel B. Horton** | **Tobias Gerhard**

Center for Pharmacoepidemiology and Treatment Science, Institute for Health, Health Care Policy, and Aging Research, Rutgers University, New Brunswick, New Jersey, USA

Correspondence
Greta A. Bushnell, Center for Pharmacoepidemiology and Treatment Science, Institute for Health, Health Care Policy, and Aging Research, Rutgers University, New Brunswick, NJ 08901, USA.
Email: greta.bushnell@rutgers.edu

Funding information
National Institute on Drug Abuse: Greta A. Bushnell 5K01DA050769. This content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

COMMENTARY

In this issue of *Clinical and Translational Science*, Yun-Han Wang and colleagues report on a potential link between treatment with proton pump inhibitors (PPIs) and incident depression and anxiety disorders in children.[1] The proposed mechanism between PPI use and anxiety and depression is related to the "microbiota-gut-brain axis." The microbiota-gut-brain axis and associated health problems have been attributed to multiple inter-related pathways, including microbiota-derived production of neurotransmitters (e.g., serotonin); communications among afferent, efferent, autonomic, and limbic nerves; immune-mediated pathways (e.g., cytokines); and endocrine pathways (e.g., hypothalamic–pituitary–adrenal axis).[2] PPIs are known to dysregulate the microbiome, and microbiome dysregulation has been linked to mental health disorders in animal models and in adult populations.[2,3]

To investigate this pediatric drug safety question, Wang and colleagues utilized nationwide databases that cover medication dispensing and hospital and emergency department encounters in Sweden. The authors observed an increase in depression and anxiety in children who initiated PPI treatment, a finding consistent across multiple secondary and sensitivity analyses. Pharmacoepidemiological studies, such as the one by Wang et al.[4] can provide valuable insights on pediatric drug safety, particularly when randomized controlled trials (RCTs) are unavailable, underpowered, or infeasible to address outcomes of interest. Understanding the strengths and weaknesses of real-world data from large healthcare databases for pharmacoepidemiological research and associated threats to validity are key to interpreting findings.

Utility of large healthcare databases for pharmacoepidemiological research

Pharmacoepidemiological studies investigating drug safety often utilize large healthcare databases, similar to the data utilized by Wang et al. These types of data include national administrative data (sometimes termed "registries," including the Swedish registry data used by Wang et al.), electronic health record (EHR) data, and insurance claims. Although aspects of each data source vary, many share critical features that permit pharmacoepidemiologic research.

A key component to these data sources is the availability of detailed information on prescribed or pharmacy-dispensed medications and dates of prescribing or dispensing, with additional details such as days' supply and dosage. Details on medications in the context of

longitudinal data allow assessment of treatment patterns and identification of persons initiating treatment. The implementation of the "new user design" is a key feature of pharmacoepidemiologic studies that minimizes bias by restricting the study population to individuals who, analogous to RCTs, newly initiate the treatments of interest.[5]

Further, these data often capture diagnoses and procedures from healthcare encounters in inpatient, emergency, or outpatient settings with associated dates. This patient-level clinical information allows implementation of inclusion and exclusion criteria, characterization of the study population, adjustment for confounders, and identification of outcomes. Additionally, the large size of these datasets allows for assessment of rare outcomes, rare therapeutics, and restricted or vulnerable study populations lacking robust data from RCTs (e.g., children and pregnant women).

Variations and nuances across these databases are important to consider when selecting a data source for a particular research question and when interpretating results.[6] For example, certain databases may be restricted to primary care encounters, may capture only written prescriptions or only dispensed medications, and may lack data on medications received in inpatient settings.

POTENTIAL THREATS TO VALIDITY

Although large healthcare databases have key strengths for pediatric pharmacoepidemiological safety studies, threats to validity remain. Pharmacoepidemiologists attempt to address these concerns through study design and analysis. Three key topics in interpreting drug safety investigations which are relevant to the investigation by Wang et al. are confounding, misclassification, and protopathic bias.

Confounding

Confounding, a distortion that alters the observed effect of an exposure on an outcome due to a factor or group of factors associated with the exposure and a risk factor for the outcome, is a potential source of bias in nonexperimental studies. In pharmacoepidemiology, confounding by indication if often a chief concern. This is because, in clinical practice, prescribing of medications is done for a reason (the indication), with the aim of achieving or preventing an outcome or improving an ongoing condition. When the indication for treatment is independently related to the outcome, confounding by indication arises.[5] The implementation of the active-comparator design can help mitigate confounding by indication in the design stage of the study. In this approach, researchers compare patients initiating the treatment of interest with those initiating an alternative treatment for the same indication.[5] However, in certain situations, an active-comparator design may not be feasible or appropriate, including when alternative treatment is not widely used, when alternative treatment is not captured in the data, or when alternative treatment increases the risk for the outcome of interest. For studies of PPIs, H2-receptor antagonists could be considered as active comparators given their common indications and lesser impacts on microbiota, but studying these drugs may be problematic due to their over-the-counter availability.[7]

In addition, many potential confounders that are captured in large healthcare databases (e.g., age, sex, prescription medication use, procedures, and healthcare utilization) can be adjusted for in the data analysis phase of the study. However, depending on the specific database used, there commonly are many patient-level factors not routinely collected or poorly collected in these data, such as smoking, substance use, blood pressure, weight, race, ethnicity, and physical activity. In the study by Wang et al., psychosocial stressors, which may independently be associated with abdominal discomfort and incident anxiety, are such examples of potential unmeasured confounders. Depending on the research question, these unmeasured factors may introduce residual confounding and bias the findings.

Wang and colleagues elegantly utilized multiple techniques to reduce residual confounding, including high-dimensional propensity score methods, a negative control outcome, stratification, and cohort restriction.[6] Additional methods to control for confounding include external adjustment of confounders and bias analyses.

Misclassification

Key to pharmacoepidemiological studies is the classification of medication use. Exposure (drug) misclassification can occur in cases of medication nonadherence (subjects classified as exposed may actually be unexposed) and when medications are available over-the-counter or accessed through other means (subjects classified as unexposed may actually be exposed), potentially biasing results. Large healthcare databases are usually limited to prescription or dispensing data, with no certainty of if or when medication is consumed.

Similarly, drug safety outcomes are also subject to misclassification in large healthcare databases, which can bias effect estimates. Generally, outcomes that require immediate attention and necessitate intervention

(e.g., hip-fracture and non-fatal heart attack) can be captured accurately. In contrast, outcomes that do not routinely result in medical encounters or are associated with delays in medical care (e.g., a fall and low back pain) are more difficult to evaluate in these types of data. Depression and anxiety may present similar challenges for measurement. For chronic illnesses, the date of diagnosis rarely represents the date of onset. For certain outcomes, utilizing procedure codes or prescriptions instead of or combined with diagnostic codes may improve outcome validity.

Persons with a recorded diagnosis (diagnostic code) are assumed to have the condition, but a code could represent a "rule-out" diagnosis or error in coding or data entry. Persons without a recorded diagnosis—depression, for example—may not have depression, may meet diagnostic criteria for depression but have not been diagnosed or sought medical care, or may have symptoms at a subthreshold level. For mental illnesses, a minority of youths with severe mental health impairment are estimated to use outpatient mental health services annually,[8] demonstrating the potential extent of missing diagnoses.

Detection bias is a particular problematic form of outcome misclassification, which can result in substantial bias, including bias away from the null hypothesis. In the study by Yun-Han Wang and colleagues, detection bias, could, for example, occur if individuals who seek treatment for reflux symptoms are also more likely to seek treatment for—and subsequently be diagnosed with—anxiety or depression.

The potential impact of various types of exposure and outcome misclassification can be quantified through sensitivity analyses.[9]

Protopathic bias occurs when treatment is initiated due to symptoms or clinical manifestations of the (yet undiagnosed) outcome (reverse causality). For example, abdominal discomfort may be an early somatic symptom of childhood anxiety or depression that prompts treatment with PPIs. Typically, those with the outcome present at baseline would be excluded from cohort entry; however, in large healthcare databases, we can only exclude those with documented prior diagnoses or related prescriptions. Protopathic bias is important to consider when interpreting studies that evaluate outcomes associated with long delays between symptom onset and treatment-seeking; in the United States, only one-third of those with mood disorders and 11% of those with anxiety disorder present for treatment within 1 year of symptom onset.[10] Adding lag time between exposure and outcome assessment, as implemented in a sensitivity analysis in Wang et al., can help lessen the potential for protopathic bias.

INTERPRETING EVIDENCE ABOUT DRUG SAFETY

Pharmacoepidemiological evaluations can improve outcomes by informing clinical practice and facilitating shared decision-making. Given that RCTs cannot address all drug safety questions in a timely or feasible manner, such evaluations play a vital role in creating real-world evidence about drug safety. The well-conducted study by Yun-Han Wang and colleagues provides intriguing evidence that PPIs are associated with newly diagnosed anxiety and depression in children. Whether these associations are causal, however, remains in question, given limitations such as protopathic bias.

Observational studies using large healthcare databases, such as the Swedish registries, come with multiple strengths as well as potential concerns over confounding, misclassification, or other threats to validity. The impacts of such threats depend on the study question, dataset specifications, and the design and analytic approaches taken by investigators. When interpreting results from studies in these real-world data, one must consider potential sources of bias and the extent to which study investigators mitigated these concerns. One should also consider the plausibility of the proposed mechanism(s) (e.g., disruption of the microbiota-gut-brain axis) and supportive evidence outside the study at hand. Whereas a single, high-quality pharmacoepidemiological study cannot confirm a proposed mechanism, it can serve as a vital starting point to guide future investigations.

CONFLICT OF INTEREST

The authors declared no competing interests for this work.

ORCID

Greta A. Bushnell https://orcid.org/0000-0002-4056-8405
Daniel B. Horton https://orcid.org/0000-0002-1831-1339
Tobias Gerhard https://orcid.org/0000-0002-8598-5771

REFERENCES

1. Wang YH, Wintzell V, Ludvigsson JF, Svanstrom H, Pasternak B. Proton pump inhibitor use and risk of depression and anxiety in children: nationwide cohort study. *Clin Transl Sci*. 2022;15(5):1112-1122.
2. Malan-Muller S, Valles-Colomer M, Raes J, Lowry CA, Seedat S, Hemmings SMJ. The gut microbiome and mental health: implications for anxiety- and trauma-related disorders. *OMICS*. 2018;22(2):90-107. doi:10.1089/omi.2017.0077
3. Huang WS, Bai YM, Hsu JW, et al. Use of proton pump inhibitors and risk of major depressive disorder: a nationwide population-based study. *Psychother Psychosom*. 2018;87(1):62-64. doi:10.1159/000485190

4. Horton DB, Blum MD, Burcu M. Real-world evidence for assessing treatment effectiveness and safety in pediatric populations. *J Pediatr*. 2021;238:312-316. doi:10.1016/j.jpeds.2021.06.062
5. Lund JL, Richardson DB, Sturmer T. The active comparator, new user study design in pharmacoepidemiology: historical foundations and contemporary application. *Curr Epidemiol Rep*. 2015;2(4):221-228. doi:10.1007/s40471-015-0053-5
6. Strom BL, Kimmel SE. In: Hennessy S, ed. *Pharmacoepidemiology*. Sixth ed. Wiley-Blackwell; 2019.
7. Maier L, Pruteanu M, Kuhn M, et al. Extensive impact of non-antibiotic drugs on human gut bacteria. *Nature*. 2018;555(7698):623-628. doi:10.1038/nature25979
8. Olfson M, Druss BG, Marcus SC. Trends in mental health care among children and adolescents. *N Engl J Med*. 2015;373(11):1079. doi:10.1056/NEJMc1507642
9. Funk MJ, Landi SN. Misclassification in administrative claims data: quantifying the impact on treatment effect estimates. *Curr Epidemiol Rep*. 2014;1(4):175-185. doi:10.1007/s40471-014-0027-z
10. Wang PS, Angermeyer M, Borges G, et al. Delay and failure in treatment seeking after first onset of mental disorders in the World Health Organization's World Mental Health Survey Initiative. *World Psychiatry*. 2007;6(3):177-185.