Quantitative bias analysis methods for epidemiologic research: a systematic review

Xiaoting Shi, BMedSci[1]; Zeyan Liew, PhD, MPH[2]; Kate Nyhan, MLS,[1,2] Mingfeng Zhang[3], MD, PhD; Wei Hua, PhD[3]; Jie Li, PhD[3]; Joo-Yeon Lee, PhD[4]; Sai Dharmarajan, PhD[4], Molly M. Jeffery, MD[5,6]; Joseph S. Ross, MD, MHS[7,8,9]; Joshua D. Wallach, PhD, MS[1]

[1] Department of Environmental Health Sciences, Yale School of Public Health, New Haven, Connecticut, USA

[2] Cushing/Whitney Medical Library, Yale University, New Haven, Connecticut, USA

[3] Office of Surveillance and Epidemiology, Center for Drug Evaluation and Research, U.S. Food and Drug Administration, Silver Spring, Maryland, USA

[4] Office of Biostatistics, Office of Translational Sciences, Center for Drug Evaluation and Research, USA Food and Drug Administration, Silver Spring, Maryland, USA

[5] Division of Health Care Delivery Research, Mayo Clinic, Rochester, Minnesota, USA

[6] Division of Emergency Medicine, Mayo Clinic, Rochester, Minnesota, USA

[7] Section of general Medicine and the National Clinician Scholars Program, Department of Internal Medicine, Yale University School of Medicine, New Haven, Connecticut, USA

[8] Department of Health Policy and Management, Yale School of Public Health, New Haven, Connecticut, USA

[9] Center for Outcomes Research and Evaluation, Yale-New Haven Health. New Haven, Connecticut, USA

Corresponding author:
Joshua D. Wallach, PhD, MS
Department of Environmental Health Sciences
Yale School of Public Health
New Haven, CT, USA
Joshua.wallach@yale.edu

Objectives:

1. To systematically identify all methodological papers describing, evaluating, and comparing quantitative bias analysis (QBA) methods for observational and experimental research.
2. To summarize key differences related to parameters, assumptions and requirements, and trade-offs between different QBA methods.

Introduction

When new therapeutics are approved by the U.S. Food and Drug Administration (FDA), the usual requirement is that drug sponsors conduct at least two well-controlled clinical trials (i.e., Phase III pivotal trials) that independently provide evidence of efficacy.[1] Randomized controlled trials (RCTs) are generally considered the gold standard for studying and estimating causal effect. However, RCTs often have strict inclusion and exclusion criteria, face recruitment and retention difficulties, and take a long time to complete. Moreover, RCTs may not be feasible for all clinical questions, including those related to rare exposures or underrepresented patient populations. The shortcomings and operational challenges of RCTs can lead to higher costs and lower generalizability to real-world clinical practice.

Over the past few years, FDA has developed a framework and guidance for utilizing real-world data and observational methods to inform regulatory decision-making, including the approval of new indications for approved drugs and monitoring postmarket safety and adverse events.[2] Although real-world data and observational research methods may be useful for describing how therapeutics are used in clinical practice,[3-5] observational studies are generally susceptible to various biases, including uncontrolled confounding, misclassification, and selection bias, that may undermine the validity of their findings. For observational studies to inform regulatory decision-making, analytical methods are needed to help assess the direction, magnitude, and uncertainty of various biases.

Quantitative bias analysis (QBA) methods have been developed to evaluate the potential impact of biases arising from systematic errors in observational studies.[6,7] Previous studies have categorized QBA methods into four overarching approaches - simple sensitivity analyses, multidimensional analyses, probabilistic analyses, and multiple bias modelling.[8] For example, the E-value was recently proposed as a simple sensitivity analysis method to quantify the minimum strength of association that an unmeasured confounder would need to have to fully explain away an exposure-outcome relationship.[9] More advanced probabilistic sensitivity analyses allow authors to estimate bias-adjusted exposure-outcome effect estimates by simulating distributions for the bias parameters.[10] Although numerous QBA methods have been proposed in the literature, recent empirical research suggests that there are relatively few applications of QBA methods in epidemiologic studies.[7] Moreover, studies that use QBA methods often include incomplete descriptions and interpretations of the methodology.[7] Indeed, there are a number of challenges that have limited the widespread application of QBA in observational studies. For instance, certain QBA methods require extensive statistical and modeling expertise, and investigators may struggle to establish reasonable bias parameters and priors across different fields. While relatively straightforward QBA methods have been proposed (e.g., the E-value),[9] these methods may be easier to misuse and can lead to widespread misinterpretation.[11] Lastly, certain QBA methods may require unrealistic assumptions. Given these challenges and concerns, it is necessary to have a robust understanding of the characteristics of different

QBA methods reported in the existing literature and to identify QBA methods that can be used for different observational and non-randomized experimental study design scenarios.

To address these knowledge gaps, we will conduct a systematic review focused on comprehensively identifying and summarizing QBA methods that have been proposed in the biomedical literature. Our goal is to provide a robust understanding of the types of QBA methods that can be used when conducting observational and experimental studies. This can help guide researchers and regulators on selecting appropriate QBA methods for different study designs when conducting these types of studies or making decisions with imperfect data.

Methods

We plan to conduct a systematic review of papers describing, evaluating, or comparing QBA methods for observational and experimental studies. This review will be conducted following the Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA) statement and Synthesis without meta-analysis (SWiM).[12] The study protocol will be registered on International prospective register of systematic reviews (PROSPERO) and posted on Open Science Framework (OSF) before our study is conducted.

Eligibility criteria

Records will be considered eligible for inclusion if they are English-language publications describing, evaluating, and/or comparing QBA methods for: observational study designs (cohort, case-control, self-control, and cross-sectional studies) and non-randomized experimental study designs (non-randomized interventional or non-randomized pragmatic studies, including single-arm studies with or without external controls, and quasi-experimental studies). No date limits will be applied. We will include methodological papers that describe significant or slight modifications of the same QBA methods. In particular, we will compare and record the differences across the articles that discuss the same QBA methods.

Articles only testing or comparing the performance of existing QBA methods using observational or experimental data as their primary objective will be used to identify potential trade-offs between QBA methods. We will screen article abstracts and exclude articles where QBA methods were only used in sensitivity analyses, as a previous systematic review already focused on the application of QBA in observational studies.[7] We will also exclude all conference abstracts, commentaries/viewpoints, and published studies with non-human subjects.

Literature search and study selection

Working with an experienced librarian (KN), we will systematically search MEDLINE, Embase, Scopus, and Web of Science Core Collection using the key words related to QBA. We will include the terms that capture the broad concepts of bias analysis, uncertainty analysis, sensitivity analysis, confounding factors and epidemiologic methods.

To retrieve additional potential articles on QBA methods that may not be identified through our literature search, we will: (1) review the references of the studies identified during the full text screening; (2) review prominent epidemiology and bias analysis text books (i.e., *Modern Epidemiology, Essentials of Epidemiology and Public Health*, *Applying Quantitative Bias Analysis to Epidemiologic Data,* and *Quantitative Data Analysis A Companion for Accounting and Information Systems Research*), and (3) manually review the references included in a previous article on the applications of QBA methods.[7]

All the articles retrieved by our comprehensive searches will be pooled and deduplicated in EndNote (https://endnote.com/). All remaining articles will be uploaded to Covidence (https://www.covidence.org/) for additional deduplication and screening. Two independent investigators will review the articles at the title-abstract and then full-text level. Discrepancies will be discussed and addressed with the arbitration by two additional reviewers (ZL and JDW).

Data collection

For all included articles, two investigators will abstract the following article characteristics: study title, first author last name, publication year, journal name, and digital object identifier (DOI). Next, two authors (XS and JDW) will classify the eligible articles into the following categories:

Methods articles:

1. Methodological articles that describe a **new**[a] QBA method (with or without testing the method using empirical or simulated data).
2. Methodological articles that describe a **new**[a] QBA method and compare the performance of the **new** method to at least one other QBA method using empirical or simulated data.
3. Methodological articles that describe **modifications** to previously developed QBA methods (with or without testing the methods using empirical or simulated data).

[a] To determine whether a methodological article describes a new approach, we will evaluate whether the article includes references to previous methods or language suggesting that the approach builds upon previously described approaches.

Application-based articles:

4. Articles where the primary objective is to test or compare **existing QBA methods** using empirical or simulated data

For all methods articles (categories 1-3), we will record:

- the name of the QBA method
- the QBA classification (select all that apply: simple sensitivity analysis; multidimensional analysis; probabilistic analysis; multiple bias modelling; other (please specify); unclear)
- the applicable study design scenarios (select all that apply: observational: prospective and/or retrospective cohort, case-control, self-controlled, and/or cross-sectional; experimental: single-arm trial or non-randomized trial with external controls; and/or meta-analysis)
- the result of interest (explain-away (yes/no), corrected estimate, multiple approaches)
- the data structure/format of the eligible exposure/intervention (select all that apply: continuous, categorical, ordinal, count, time-to-event, other (please specify))
- the data structure/format of eligible outcome (select all that apply: continuous, categorical, ordinal, count, time-to-event, other (please specify))
- the data structure/format of the eligible covariate/confounder (select all that apply: continuous, categorical, original, time-to-event, other (please specify))
- the measures of effect (select all that apply: risk ratio, rate ratio, odds ratio, hazard ratio, risk difference, rate difference, mean difference, other (please specify))
- the source(s) of bias addressed (select all that apply: selection bias; incompletely controlled confounding; uncontrolled/ residual confounding; misclassification (specify the variable of concern: e.g., exposure, outcome, event); missing data; other (please specify); unclear)
- the type of data required to conduct the QBA method (raw/individual patient-level data vs. summary statistics/published data)
- the information/parameters needed to conduct the QBA method (please record)
- the output of QBA method and its interpretation
- the additional required data features, assumptions, or restrictions (please record [e.g., no interactions between exposure and confounder])
- the available and/or recommended software developed for the QBA method (select all that apply: R; SAS; STAT; Excel; other(s); unclear)
- any considerations related to the interpretation of the results (please record)
- for category 3, we will also record whether the authors were the same as the original article (i.e., whether any of the authors overlap)

For application-based articles (category 4), which are not the focus of the systematic review, we will collect the following:

- the QBA methods compared
- the study design
- the exposure(s), outcome(s) and covariate(s) and/or confounders and their data structure/format (select all that apply: continuous, categorical, ordinal, count, time-to-event, other (please specify))
- the measures of effect (e.g., risk ratio, rate ratio, odds ratio, hazard ratio, risk difference, rate difference, mean difference, other (please specify))
- the source(s) of bias addressed (select all that apply: selection bias; uncontrolled/ residual confounding; misclassification (specify the variable of concern: e.g., exposure, outcome, event); missing data; other (please specify))
- the changes to the original effect estimates and study conclusions after applying QBA method(s)
- the major strengths and/or limitations related to the QBA methods and specific application described by the authors
- the methods used to control for observed confounders (please record [e.g., propensity score, inverse probability treatment weighting, etc.])

Statistical analysis

After we summarize the characteristics across all the identified QBA methods in a tabular form, we aim to develop a decision tree that can be utilized to help guide researchers and regulators when it comes to the selection of QBA methods available for different study characteristics. Key components of the decision tree will be established once the data are collected. For example, the root node will be study design and the subsequent nodes will address the exposure and outcome characteristics, source(s) of bias, data structure, and other required parameters. The final node of the tree will include the recommended QBA methods, available software, and any known strengths and weaknesses.

References

1. US Food and Drug Administration. Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products. | FDA. 2021; https://www.fda.gov/regulatory-information/search-fda-guidance-documents/providing-clinical-evidence-effectiveness-human-drug-and-biological-products. Accessed October 15th, 2021.
2. US Food and Drug Administration. Use of Real-World Evidence to Support Regulatory Decision-Making for Medical Devices | FDA. 2021; https://www.fda.gov/regulatory-information/search-fda-guidance-documents/use-real-world-evidence-support-regulatory-decision-making-medical-devices. Accessed October 14th, 2021.
3. Dhruva SS, Ross JS, Desai NR. Real-world evidence: promise and peril for medical product evaluation. *Pharmacy and Therapeutics.* 2018;43(8):464.
4. Bartlett VL, Dhruva SS, Shah ND, Ryan P, Ross JS. Feasibility of Using Real-World Data to Replicate Clinical Trial Evidence. *JAMA Netw Open.* 2019;2(10):e1912869.
5. Cowie MR, Blomster JI, Curtis LH, et al. Electronic health records to facilitate clinical research. *Clin. Res. Cardiol.* 2017;106(1):1-9.
6. Lash TL, Fox MP, Fink AK. *Applying quantitative bias analysis to epidemiologic data.* Springer Science & Business Media; 2011.
7. Petersen JM, Ranker LR, Barnard-Mayers R, MacLehose RF, Fox MP. A systematic review of quantitative bias analysis applied to epidemiologic research. *Int. J. Epidemiol.* 2021.
8. Lash TL, Fox MP, MacLehose RF, Maldonado G, McCandless LC, Greenland S. Good practices for quantitative bias analysis. *Int. J. Epidemiol.* 2014;43(6):1969-1985.
9. VanderWeele TJ, Ding P. Sensitivity Analysis in Observational Research: Introducing the E-Value. *Ann. Intern. Med.* 2017;167(4):268-274.
10. Orsini N, Bellocco R, Bottai M, Wolk A, Greenland S. A Tool for Deterministic and Probabilistic Sensitivity Analysis of Epidemiologic Studies. *The Stata Journal.* 2008;8(1):29-48.
11. Ioannidis JPA, Tan YJ, Blum MR. Limitations and Misinterpretations of E-Values for Sensitivity Analyses of Observational Studies. *Ann. Intern. Med.* 2019;170(2):108-111.
12. Campbell M, McKenzie JE, Sowden A, et al. Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. *BMJ (Clinical research ed.).* 2020;368:l6890-l6890.