Pharmacoepidemiology and Drug Safety

WILEY

CORE CONCEPTS IN PHARMACOEPIDEMIOLOGY OPEN ACCESS

Core Concepts in Pharmacoepidemiology: Quantitative Bias Analysis

Jeremy P. Brown[1,2] | Jacob N. Hunnicutt[3] | M. Sanni Ali[2] | Krishnan Bhaskaran[2] | Ashley Cole[4] | Sinead M. Langan[2] | Dorothea Nitsch[2] | Christopher T. Rentsch[2,5] | Nicholas W. Galwey[6] | Kevin Wing[2] | Ian J. Douglas[2]

[1]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA | [2]Department of Non-Communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, UK | [3]Epidemiology, Value Evidence and Outcomes, R&D Global Medical, GSK plc, Collegeville, Pennsylvania, USA | [4]Real-World Analytics, Value Evidence and Outcomes, R&D Global Medical, GSK plc, Collegeville, Pennsylvania, USA | [5]Department of Internal Medicine, Yale School of Medicine, New Haven, Connecticut, USA | [6]R&D, GSK Medicines Research Centre, GSK plc, Stevenage, UK

Correspondence: Jeremy P. Brown (jeremybrown@hsph.harvard.edu)

Received: 10 May 2024 | **Revised:** 13 August 2024 | **Accepted:** 16 September 2024

Funding: J.P.B.'s PhD studentship was funded by GSK plc.

ABSTRACT

Pharmacoepidemiological studies provide important information on the safety and effectiveness of medications, but the validity of study findings can be threatened by residual bias. Ideally, biases would be minimized through appropriate study design and statistical analysis methods. However, residual biases can remain, for example, due to unmeasured confounders, measurement error, or selection into the study. A group of sensitivity analysis methods, termed quantitative bias analyses, are available to assess, quantitatively and transparently, the robustness of study results to these residual biases. These approaches include methods to quantify how the estimated effect would be altered under specified assumptions about the potential bias, and methods to calculate bounds on effect estimates. This article introduces quantitative bias analyses for unmeasured confounding, misclassification, and selection bias, with a focus on their relevance and application to pharmacoepidemiological studies.

1 | Introduction

Pharmacoepidemiological studies are an important source of evidence on the safety and effectiveness of medications. However, the validity of findings from pharmacoepidemiological studies can be threatened by potential sources of residual bias such as unmeasured confounders, measurement error, and the selection of study subjects [1–4].

The potential for residual bias is commonly identified qualitatively, as a study limitation, in the discussion section of a manuscript. However, the robustness of results to residual biases is typically not immediately clear, and a subjective assessment is prone to error [5]. A group of sensitivity analysis methods, termed quantitative bias analysis (QBA), is available for assessing the sensitivity of results to residual bias in a quantitative, transparent, and objective manner [6, 7]. Study investigators typically quantify uncertainty in results due to random error, but quantification of uncertainty due to potential systematic error (bias) has been less common. There is growing recognition of the utility of QBA methods, and they form part of several recent guidelines on the design of pharmacoepidemiological studies [8, 9].

In this article, we introduce commonly applied QBA methods for three of the most common residual biases present in pharmacoepidemiological studies, namely residual confounding due to unmeasured confounders, bias due to misclassification, and selection bias [10].

Summary

- The validity of findings from pharmacoepidemiological studies can be undermined by residual biases.
- Common sources of residual bias in pharmacoepidemiological studies include unmeasured confounders, measurement error, and selection into the study.
- Qualitative assessment of the potential effects of residual bias can be subjective and is prone to error.
- Quantitative bias analysis methods are a group of sensitivity analysis methods which allow study investigators to quantify the potential effect of residual biases under specified assumptions.
- Quantitative bias analysis enables study investigators to assess the robustness of study results to residual bias in a quantitative, transparent, and rigorous manner.

2 | Identifying Potential Residual Bias

An initial step in planning QBA is to identify potential residual biases in a planned study (Figure 1). Bias should be minimized through appropriate study design and statistical analyses. However, often the potential for residual bias remains, due to unmeasured confounders, measurement error, and/or selection of subjects into the study [11].

2.1 | Residual Confounding Due to Unmeasured Confounders

Confounding occurs in studies which aim to estimate causal effects when there are common causes of exposure and outcome which bias the estimated effect of exposure away from the causal effect [10]. Measured confounders can be accounted for by using study design (e.g., restriction, matching) and/or statistical analysis methods (e.g., regression adjustment, inverse probability of treatment weighting) [10, 12]. However, even after taking account of measured confounders, there may be unmeasured or mismeasured confounders, leading to residual confounding. As an example, in a US insurance claims database study of the association between use of selective serotonin reuptake inhibitors and hip fracture, the estimated harmful association was exaggerated without supplemental data collection on variables, such as smoking status, that are typically unmeasured or poorly measured in claims data [13].

2.2 | Measurement Bias

Measurement bias (also known as information bias) can occur when variables included in the study (i.e., exposure, outcome, or covariates) are measured with error (i.e., measurement error) [3]. For categorical variables, such measurement error is referred as misclassification. Measurement error, and similarly misclassification, can be categorized as either differential or non-differential [14]. Error is said to be differential when it depends on other analytical variables. For example, recollection of exposure may differ by outcome status in a case–control study of the association between medication use in pregnancy and congenital malformations with participants whose newborn child either did (case) or did not (control) have a congenital malformation surveyed on medication use in pregnancy [15]. Measurement error of multiple variables can be categorized as having dependent or independent errors. Errors are said to be dependent when measurement error for one variable is associated with error in another. For example, individuals from a particular electronic health record (EHR) system who receive care outside that system are more likely to be misclassified on multiple variables [16]. Non-differential misclassification of a binary exposure or outcome typically leads to bias toward the null in expectation, though with several exceptions [5]. One useful exception is that risk ratio estimates are unbiased in the presence of non-differential outcome misclassification when specificity is 100%. Differential misclassification can lead to bias toward or away from the null dependent on the misclassification mechanism. Misclassification in confounders typically leads to residual confounding [17].

2.3 | Selection Bias

Selection bias occurs when there is a systematic difference between estimates observed in the analytic sample and those that would have been obtained in the population of interest (i.e., the target population) [2, 18, 19]. When selection is associated both with exposure and outcome, estimates may differ systematically [18, 20] even if there is no causal association between treatment and outcome [18, 20]. For example, in a test-negative study of the effectiveness of influenza vaccination, the study population is by design restricted to the tested, which can introduce selection bias when health-seeking behavior is associated with both testing and vaccination [21]. Selection bias commonly arises in case–control studies, due to inappropriate control selection, and in cohort studies, due to differential loss to follow-up [18].

3 | Applying QBA Methods

Once we have identified potential residual biases, we can choose a QBA method to apply (see Table 1 for a summary of described methods). Residual bias due to confounding, measurement error, and selection can be seen as arising due to a lack of information [22]. With complete information on confounders, the true values for variables, and access to data for the entire target population, these biases could be eliminated. In practice, there is often incomplete information, for example, due to an unmeasured confounder, and to assess the validity of results using incomplete information, we can make quantitative assumptions about the potential sources of bias (e.g., the prevalence of the unmeasured confounder) and use these quantitative assumptions, termed bias parameters, to quantify how the effect estimate would be expected to differ in the absence of this bias.

Alternatively, we can leave one or more bias parameters unspecified and quantify bounds on the estimates that would occur if the unspecified bias parameters were at their most extreme [23–26]. If study conclusions are robust in relation to this worst-case bias, this provides some reassurance with regard to the validity of study findings.

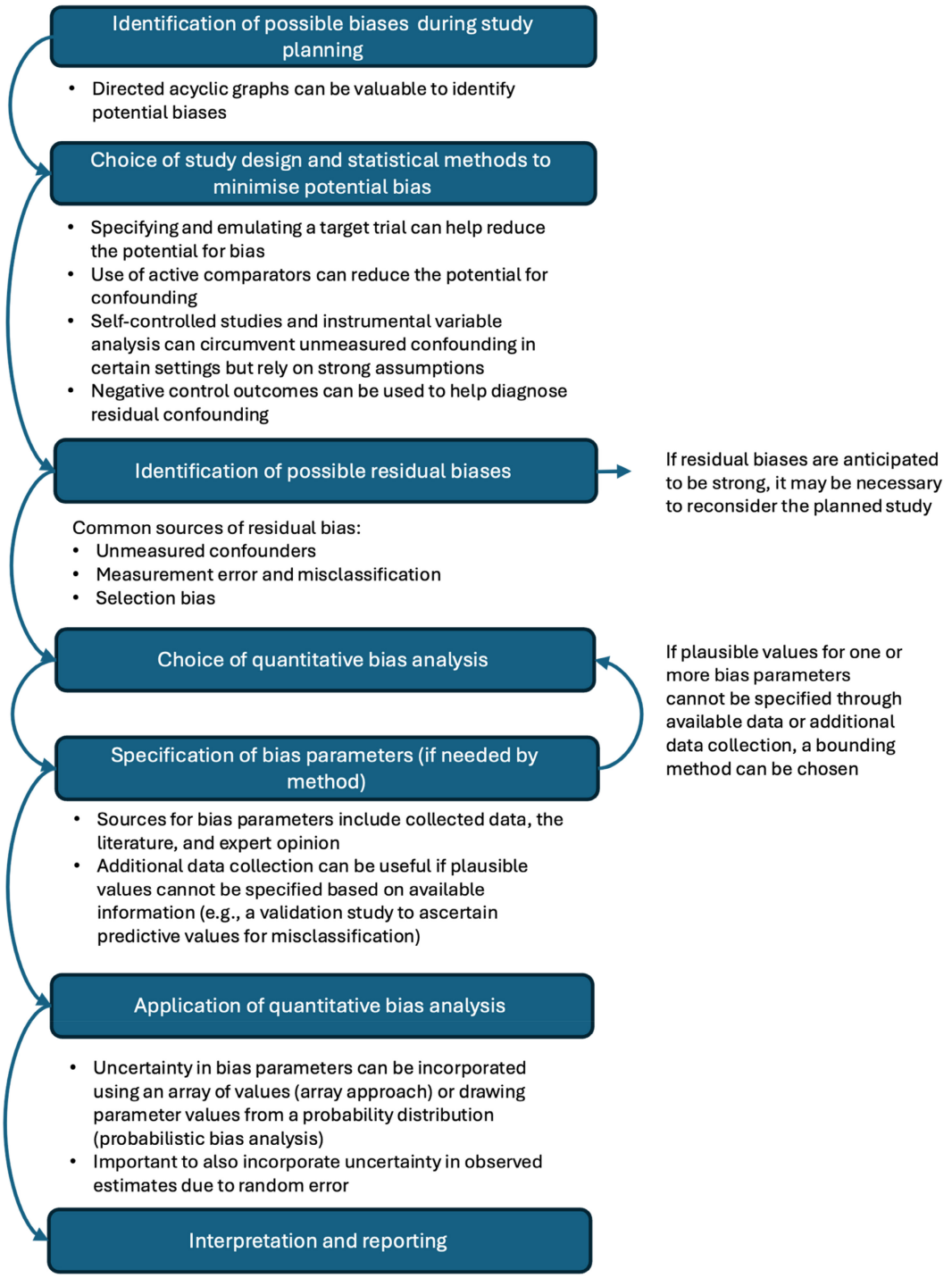

FIGURE 1 | Steps involved in conducting QBA.

3.1 | Specification of Bias Parameters

A key component to applying many QBA methods is the specification of bias parameters. The appropriate specification differs depending on the source of bias, namely unmeasured confounding, misclassification, or selection bias.

For unmeasured confounding, bias parameters include the strength of the association between unmeasured confounder and outcome, and the prevalence of the unmeasured confounder in the different treatment groups. Ideally, these bias parameters would be estimated in a subset of the study population, but when this is not possible, alternative sources of information include the literature and external data sources.

For misclassification, bias parameters include the sensitivity and specificity of classification, or positive and negative predictive values (defined in Table 2). Validation studies can be conducted to estimate these parameters [27]. Typically, in such studies a sample of individuals is identified, either from the study population (i.e., internal validation study) or from a similar population (i.e., external validation study), who are categorized as having

TABLE 1 | Summary of described methods.

QBA method	Biases addressed by method	Advantages	Disadvantages
Formulas	Uncontrolled confounding, misclassification, selection bias	Simplicity of application	Challenging to incorporate measured covariates
Probabilistic bias analysis	Uncontrolled confounding, misclassification, selection bias	Incorporation of uncertainty in bias parameters	Can be challenging to specify bias parameter distributions
Regression likelihood-based	Uncontrolled confounding, misclassification	Ease of incorporation of measured covariates	Statistical and computational complexity; limited to regression analyses; requirement of record-level data
Bounding methods	Uncontrolled confounding, misclassification, selection bias	Requires specification of fewer bias parameters; ease of application to covariate-adjusted associations	Less informative than methods that estimate a bias-corrected effect; can be difficult to interpret
Matrix methods	Misclassification	Ability to include misclassified variables with >2 categories and with dependent errors	Complexity of specification
Multiple imputation for measurement error	Measurement bias	Ease of incorporation of measured covariates and multiple mismeasured variables; applicability with continuous covariates	Statistical and computational complexity; requirement of record-level data
Inverse probability of selection weights	Selection bias	Ease of incorporation into standard analyses	Requirement of record-level data

TABLE 2 | Definition of misclassification terms.

Term	Description
Sensitivity	Probability categorized as having the event (e.g., outcome, exposure) given truly have the event
Specificity	Probability categorized as not having the event given truly do not have the event
Positive predictive value	Probability truly have the event given categorized as having the event
Negative predictive value	Probability don't have the event given categorized as not having the event

(and/or not having) the event of interest (e.g., type 2 diabetes recorded in insurance claims data). Additional data collection is conducted (e.g., by surveying the individual's clinician or through chart review), to ascertain whether the original categorization is correct and to estimate predictive values. When sampling in the validation study is conducted within levels of the misclassified variable, sensitivity and specificity cannot be directly estimated, but predictive values can. However, in this situation, sensitivity and specificity can be indirectly estimated from predictive values using Bayes' theorem and the prevalence of the misclassified variable in the study population (see Appendix).

In this manuscript we focus on using sensitivity and specificity, because positive and negative predictive values can vary strongly depending on the underlying prevalence of the true variable, and therefore may vary more between populations [27]. Furthermore, even if misclassification were non-differential, if exposure and outcome are associated, we would need to stratify the predictive values for outcome misclassification by exposure, and similarly the values for exposure misclassification by outcome [6].

For selection bias, bias parameters include the probabilities of selection into the study sample, stratified by exposure group and outcome status. Choice of values for these bias parameters may be informed by data from the target population, if available, or from literature estimates and external data sources [28].

3.2 | QBA Methods for Unmeasured Confounding

3.2.1 | Bias Formulas

There are several approaches that can be taken for QBA for an unmeasured confounder. One simple approach to calculate bias-adjusted estimates is to apply algebraic formulas, termed bias formulas [7, 29]. As an example, consider a bias formula for the risk ratio between a binary treatment X (e.g., prescription of a retinoid during the first trimester of pregnancy) and a binary outcome Y (e.g., diagnosis of a congenital malformation in the infant) [30]. Using the bias parameters of the prevalence of the unmeasured confounder in the treated and comparator groups, and the risk ratio between unmeasured confounder U and

TABLE 3 | Notation used in formulas.

Term	Description
$RR_{XY}^{BiasAdj}$	Risk ratio between treatment and outcome adjusted for unmeasured confounder
RR_{XY}^{Obs}	Observed risk ratio between treatment and outcome
$RR_{UY\|X}$	Risk ratio between unmeasured confounder and outcome adjusted for treatment
$\Pr(U=1\|X=1)$	Prevalence of the unmeasured confounder among the treatment group
$\Pr(U=1\|X=0)$	Prevalence of the unmeasured confounder among the comparator group
$OR_{XY}^{BiasAdj}$	Odds ratio between treatment and outcome corrected for selection bias
OR_{XY}^{Obs}	Observed odds ratio between treatment and outcome
S_{xy}	Probability of selection for an individual with exposure x and outcome y
Se_{X1}	Sensitivity of outcome classification in treatment group
Se_{X0}	Sensitivity of outcome classification in comparator group
Sp_{X1}	Specificity of outcome classification in treatment group
Sp_{X0}	Specificity of outcome classification in comparator group

outcome Y (adjusted for treatment X), we can calculate a bias-adjusted effect estimate (see Table 3 for definitions of formula terms and Figure 2 for an example application) [7].

$$RR_{XY}^{BiasAdj} = RR_{XY}^{Obs} \Big/ \frac{\Pr(U=1|X=1)\left(RR_{UY|X}-1\right)+1}{\Pr(U=1|X=0)\left(RR_{UY|X}-1\right)+1}$$

We can apply the same formula to an odds ratio when the outcome is rare, and hence the odds ratio approximates the risk ratio, and to the hazard ratio when hazards are proportional and the outcome is rare, and hence the hazard ratio approximates the risk ratio [12, 31]. While this formula assumes no interaction between treatment and unmeasured confounder, alternative formulas are available that do not make this assumption [29]. This formula can be applied to the point estimate and to the limits of the confidence interval.

This formula applies to the crude risk ratio. Bias formulas that incorporate measured covariates are available, though specification becomes more challenging, and generally requires the specification of bias parameters within strata of measured covariates [29]. Similarly, bias formulas are available for categorical unmeasured confounders with >2 categories, propensity-score

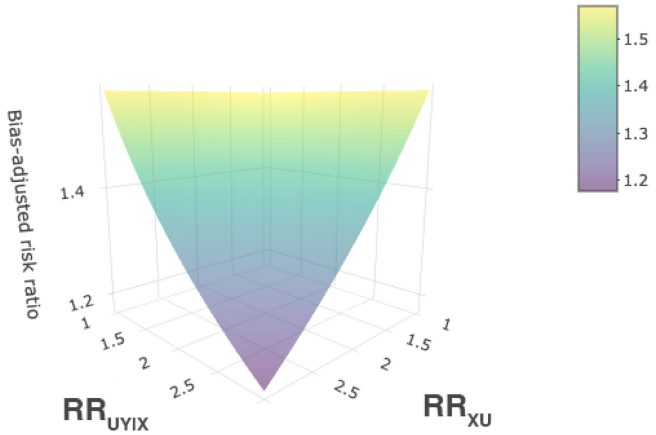

FIGURE 2 | Bias-adjusted risk ratio for an observed risk ratio of 1.57 when varying the risk ratio between unmeasured confounder and outcome conditional on exposure ($RR_{UY|X}$) and the risk ratio between exposure and unmeasured confounder (RR_{XU}). Prevalence of the binary unmeasured confounder among the comparator group is assumed to be 10% in this figure.

weighted estimates, and for the risk difference [32]. When there are data on the unmeasured confounder for a subset of participants, an alternative approach to correct propensity score adjusted estimates is to use propensity score calibration [32–35].

3.2.2 | Likelihood-Based Approaches

Regression analyses, such as logistic regression, are typically conducted using maximum likelihood estimation. In this approach, the estimated values of parameters (e.g., log odds ratios between treatment and outcomes) are the values that maximize the likelihood of the observed data for a given statistical model (e.g., a logistic model) [36]. An alternative to bias formulas for estimating a bias-adjusted association, applicable when patient record-level data are available and a regression model for the outcome is fitted (e.g., logistic regression adjusting for measured covariates), is to modify this likelihood in order to obtain a bias-adjusted effect estimate (e.g., a bias-adjusted odds ratio) [37, 38]. An advantage of this approach is ease of incorporation of measured covariates. However, this approach is more complicated computationally and statistically because it involves model specification and numerical likelihood maximization.

When data relating to the confounder are available for a subset of study participants, then, under certain assumptions, multiple imputation can be used [39]. Though the focus of this article is on binary and time-to-event outcomes, which are common in pharmacoepidemiological studies, QBA methods are also available for linear regression with a continuous outcome [40].

3.3 | QBA Methods for Misclassification

3.3.1 | Algebraic Formulas

A simple approach to QBA for exposure or outcome misclassification when both these variables are binary is to apply algebraic formulas to calculate a bias-adjusted 2×2 table [6]. These

TABLE 4 | Observed 2×2 table of the treatment-outcome association.

	$Y=1$	$Y=0$
$X=1$	a	b
$X=0$	c	D

Note: Treatment is denoted X, outcome as Y.

TABLE 5 | Corrected 2×2 table of treatment-outcome association.

	$Y=1$	$Y=0$
$X=1$	$A=\frac{a-(a+b)(1-Sp_{X1})}{Se_{X1}+Sp_{X1}-1}$	$B=a+b-A$
$X=0$	$C=\frac{c-(c+d)(1-Sp_{X0})}{Se_{X0}+Sp_{X0}-1}$	$D=c+d-C$

formulas require specification of the bias parameters of sensitivity and specificity or positive and negative predictive values [6]. For example, consider a scenario in which there is outcome misclassification, and we specify values for the bias parameters of sensitivity and specificity. We have the observed 2×2 table (Table 4), and from this we generate a corrected 2×2 table (Table 5) under assumed values for bias parameters. If we assume misclassification is non-differential, then sensitivity and specificity will not vary depending on exposure status, and we can use a single value for each of these bias parameters.

To incorporate measured covariates, for example as part of adjustment for measured confounders, we can stratify the 2×2 tables by these covariates and perform a stratified bias analysis (e.g., estimating a Mantel–Haenszel odds ratio). However, limiting analyses to 2×2 tables is restrictive, and a more flexible approach, termed record-level correction, is to use these corrected 2×2 tables to impute a corrected value of the relevant variable in each data record, and analyze the data as planned with this imputed variable [41, 42].

3.3.2 | Likelihood-Based Approaches

Similar to QBA for unmeasured confounding, another approach to QBA for misclassification is to modify the regression likelihood [6, 43, 44]. This approach has the advantage that it can incorporate measured covariates, but it is more statistically and computationally involved.

3.3.3 | Other Methods

Matrix methods are available when there is simultaneous misclassification of exposure and outcome, including situations with dependent errors (i.e., when the misclassification in exposure and outcome are not independent), or when categorical variables (with >2 categories) are misclassified [6, 45]. In pharmacoepidemiological studies with active comparators, the occurrence of non-adherence can make exposure classification categorical (e.g., exposed to treatment of interest, exposed to comparator, not exposed to either treatment) rather than binary. Ross et al. outline bias formulas in an active comparator study for non-differential misclassification, while matrix methods may be used if misclassification is differential [46].

Further alternative methods are available, including predictive value weighting and reparametrized imputation for measurement error (RIME) [6, 43, 45–47]. While we focus in this article on misclassification, methods for measurement error correction for continuous variables are also available, including regression calibration and simulation-extrapolation (SIMEX) [3, 48, 49].

3.3.4 | Multiple Imputation for Measurement Error

When internal validation data are available (i.e., data on a random or stratified random subset of participants) then misclassification can be treated as a missing data problem, with the true variable considered as a variable with missing values, and multiple imputation for measurement error applied [3, 39, 50, 51].

4 | QBA Methods for Selection Bias

4.1 | Bias Formulas

A simple formula to take account of selection bias applies to the odds ratio [52]. To apply this formula, we must specify the bias parameters of probabilities of selection into the study for different categories of participant, namely participants in treatment group with outcome (S_{11}), treatment group without outcome (S_{10}), comparator group with outcome (S_{01}), and comparator group without outcome (S_{00}). The bias-adjustment formula is then:

$$\mathrm{OR}_{XY}^{BiasAdj}=\mathrm{OR}_{XY}^{Obs}/\frac{S_{11}S_{00}}{S_{10}S_{01}}$$

If selection is independent of covariates given exposure and outcome, the formula can be applied to the covariate-adjusted odds ratio. Assuming the outcome is rare, it can also be applied as an approximation to a risk ratio or a hazard ratio [31]. To incorporate random error, this formula can be applied to both the point estimate and the limits of the confidence interval.

4.2 | Inverse Probability of Selection Weights

With record-level data, an alternative and more flexible approach is to use inverse probability of selection weights. Each record can be weighted by the inverse probability of selection, and weighted analyses conducted. If selection depends on covariates, these covariates can be incorporated by specifying different selection probabilities for each combination of covariates, exposure, and outcome, though accurate estimates of these stratified probabilities are often less readily available.

5 | Incorporating Uncertainty in Bias Parameters

There will typically be uncertainty in bias parameter values, and several approaches are available to account for that uncertainty. Rather than specify a single value for each bias parameter, a

range of plausible values can be specified. Alternatively, a probability distribution can be specified for each bias parameter, bias parameters can be drawn repeatedly from these distributions (e.g., 10000 times), and the previously described bias analyses conducted repeatedly with these different values of the bias parameters. This approach is termed probabilistic bias analysis or Monte Carlo sensitivity analysis. Common choices for distributions, chosen to reflect investigator uncertainty in bias parameters, include trapezoidal, log-normal, and beta distributions (see Figure 3 for an illustrative distribution). Correlation between bias parameters can be specified as part of the distribution specification [6]. Probabilistic bias analysis generates a distribution of bias-adjusted estimates, which can be summarized by reporting the median and 95% simulation interval (see Boxes 1 and 2). If a fully Bayesian approach is desired, methods are available that can incorporate uncertainty in bias parameters through probability distributions as part of a Bayesian analysis [6, 53–56].

Uncertainty in the bias-adjusted estimates arises not only due to the uncertainty in bias parameter values, but also due to random error affecting the observed data. It is important to incorporate this uncertainty, for example, by bootstrapping the bias analysis [57].

6 | Bounding Methods

One approach to QBA that can circumvent the need to specify one or more bias parameter values is to use bounding methods [26, 58]. These methods can be used to quantify bounds on estimates and to ascertain values of one or more bias parameters that would be necessary to reduce observed estimates to the null (or another specified value) if unspecified bias parameters were at their most extreme.

For unmeasured confounding, two popular bounds are *E*-values and Cornfield conditions [26]. These methods specify the minimum strength of association with an unmeasured confounder that is necessary, if other bias parameters are at their worst possible values, to potentially reduce an observed association to the null or to some other specified value. These formulas can be applied to the point estimate for the risk ratio, and to its confidence interval limits, to account for random error (see Figure 4 for an example application). As with the bias formulas for unmeasured confounding, these formulas can also be applied to the odds ratio and hazard ratio when the outcome is rare. The formulas for applying these boundary methods are as follows:

Cornfield conditions:

$$RR_{UY|X,C} \geq RR^{Obs}_{XY|C}$$

$$RR_{XU|C} \geq RR^{Obs}_{XY|C}$$

E-value formula:

$$\max\left(RR_{UY|X,C}, RR_{XU|C}\right) \geq RR^{Obs}_{XY|C} + \sqrt{RR^{Obs}_{XY|C}\left(RR^{Obs}_{XY|C} - 1\right)}$$

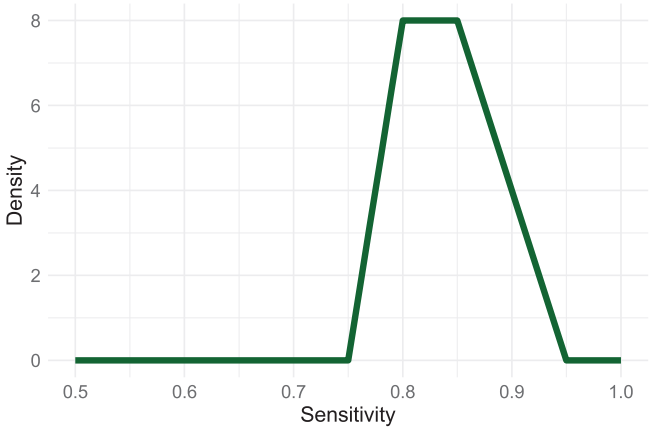

FIGURE 3 | Trapezoidal distribution for a bias parameter of sensitivity of classification with an upper bound of 95%, a lower bound of 75%, and most likely values between 80% and 85%.

BOX 1 | Example application of probabilistic bias analysis.

Suarez et al. applied probabilistic bias analysis to examine the robustness of the association between buprenorphine vs. methadone use in pregnancy and pregnancy outcomes [72]. Non-differential outcome misclassification was assumed, and sensitivity and specificity values were drawn repeatedly from triangular probability distributions. The crude risk ratio for preterm birth with buprenorphine was 0.58 (95% CI 0.54–0.62), which in this study was close to the adjusted risk ratio (0.58, 95% CI 0.53–0.62). The median estimate from probabilistic bias analysis applied to the crude risk ratio was 0.47 (95% simulation interval 0.37–0.55), providing support for the robustness to outcome misclassification of the finding of reduced risk with buprenorphine.

BOX 2 | Example application of bounding methods.

Zhang et al. examined the potential impact of unmeasured confounding by smoking on the association between fluoroquinolone use and risk of aortic aneurysm, using *E*-values [73]. For a covariate-adjusted risk ratio of two between fluoroquinolone use and aortic aneurysm, the estimated *E*-value was 3.41, indicating that in order to potentially explain the observed association, either smoking must be associated with aortic aneurysm, or fluoroquinolone use must be associated with smoking, with a risk ratio greater than 3.41. While risk ratios of this size were considered unlikely by the study investigators based on the literature, additional unmeasured confounders, such as severity of infection, could contribute to the observed association [74].

Note: These associations are adjusted for one or more measured covariates *C*.

The Cornfield condition is a lower threshold, which both the treatment-unmeasured confounder risk ratio, and the unmeasured confounder-outcome risk ratio, must exceed, in order to potentially explain the observed result; whereas the *E*-value is

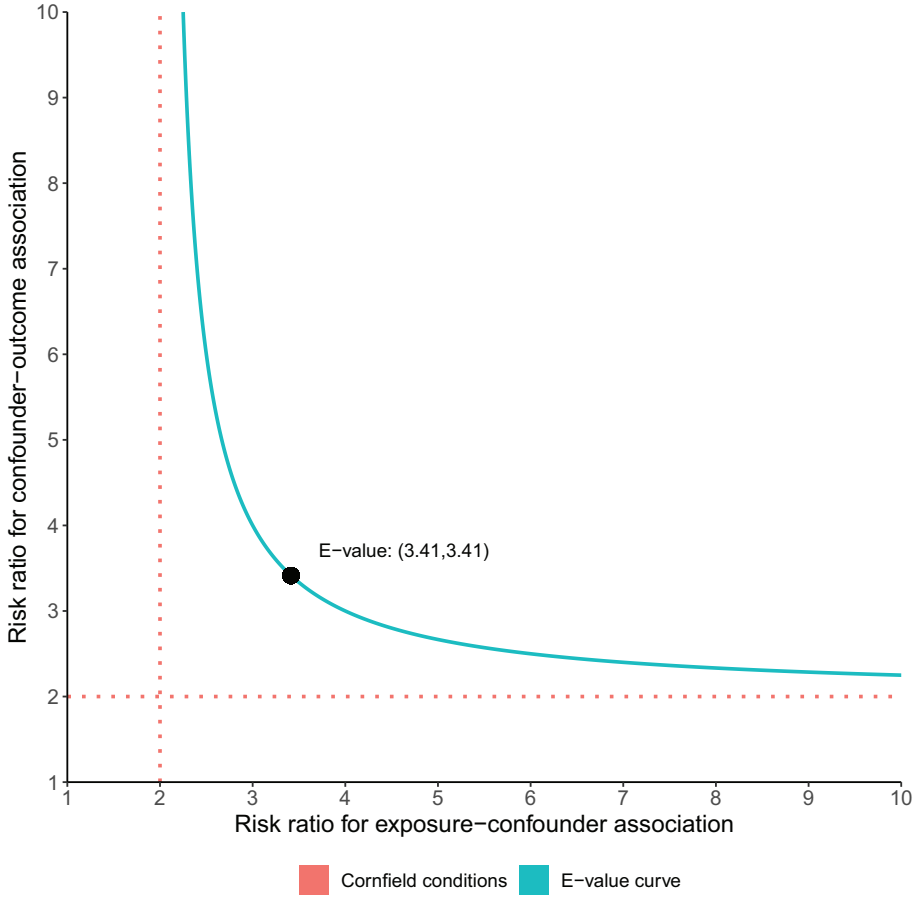

FIGURE 4 | *E*-value plot for an observed risk ratio of 2. An unmeasured confounder with strength of association with both exposure and outcome below the *E*-value curve could not fully explain the observed association (i.e., reduce it to a null risk ratio of 1).

a higher threshold which only one of these two risk ratios must exceed. While *E*-values are a popular method, there has been criticism of their application in the literature, given common misinterpretation and use without careful consideration of specific unmeasured confounders [59–61].

Bounding methods for unmeasured confounding are available for the risk difference, and also methods that bound the *p*-value [58]. Bounding methods are also available for measurement error and selection bias [24, 62]. While bounding methods can be useful and are simple to implement, they do not provide an estimate of a bias-corrected association. Instead, a bound on the maximum bias possible under given assumptions can rule out a particular bias as an explanation of findings, but does not provide evidence, if bias cannot be ruled out, that the bias would necessarily explain the observed findings. Because of these limitations, bounding methods provide less information than other QBA methods.

6.1 | Multiple Bias Analysis

To conduct QBA for multiple biases we can apply sequentially the bias correction methods introduced earlier. Typically, we should apply these methods in the reverse order to that in which the biases arose [6]. For example, if we have a study in which we selected a sample, then interviewed participants, we can consider biases to have arisen in the order: confounding (before the study commenced), selection (during the specification of the sample) then misclassification (when the interview results were recorded). Bounding methods are also available for multiple biases [63].

6.2 | Application of QBA Using Statistical Software

QBA can be implemented manually or through user-written code, but there are also a number of software packages available that can make implementation of QBA easier [64–68]. Statistical software for QBA includes R, Stata, and SAS code for algebraic and probabilistic bias analysis [65, 66, 69]. An R package and SAS macro are available to implement *E*-values, with the R "EValue" package also implementing bounds for selection bias, misclassification, and multiple biases [67, 68]. Spreadsheets to implement QBA are available that accompany the book "*Applying Quantitative Bias Analysis to Epidemiologic Data*" [6].

6.3 | Interpretation and Reporting of QBA

The validity of QBA depends inherently on accurate specification of bias parameters and on the assumptions made in the conduct of the bias analysis (e.g., that misclassification is non-differential) [70]. It is important that these assumptions are made explicit when reporting, and that clear justification is given for

choices of bias parameters. Furthermore, caution is warranted that QBA results are not overinterpreted, given the possibility of additional residual biases that have not been accounted for, and given the reliance of the methods on assumptions made, such as assumed bias parameter values [71].

7 | Conclusions

Residual bias is a common concern in pharmacoepidemiological studies and can threaten the validity of study findings. Using QBA, the sensitivity of study results to potential residual biases can be quantified, enabling a transparent assessment of the robustness of study findings. When QBA is not applied, interpretation of estimates as causal makes the assumption of absence of unmeasured confounding, measurement bias, or selection bias, which is difficult to justify in practice.

Acknowledgments

J.P.B.'s PhD studentship was funded by GSK plc. The opinions expressed by GSK employees in this article are the author's own and do not reflect the view of GSK plc.

Ethics Statement

No written consent has been obtained from the patients as there is no patient identifiable data included.

Conflicts of Interest

The authors declare no conflicts of interest.

References

1. R. Sendor and T. Stürmer, "Core Concepts in Pharmacoepidemiology: Confounding by Indication and the Role of Active Comparators," *Pharmacoepidemiology and Drug Safety* 31, no. 3 (2022): 261–269, https://doi.org/10.1002/pds.5407.

2. H. Lu, S. R. Cole, C. J. Howe, and D. Westreich, "Toward a Clearer Definition of Selection Bias When Estimating Causal Effects," *Epidemiology* 33, no. 5 (2022): 699–706, https://doi.org/10.1097/ede.0000000000001516.

3. R. J. Carroll, D. Ruppert, L. A. Stefanski, and C. M. Crainiceanu, *Measurement Error in Nonlinear Models: A Modern Perspective* (New York: Chapman and Hall/CRC, 2006).

4. E. K. Acton, A. W. Willis, and S. Hennessy, "Core Concepts in Pharmacoepidemiology: Key Biases Arising in Pharmacoepidemiologic Studies," *Pharmacoepidemiology and Drug Safety* 32, no. 1 (2022): 9–18, https://doi.org/10.1002/pds.5547.

5. J. J. Yland, A. K. Wesselink, T. L. Lash, and M. P. Fox, "Misconceptions About the Direction of Bias From Nondifferential Misclassification," *American Journal of Epidemiology* 191, no. 8 (2022): 1485–1495, https://doi.org/10.1093/aje/kwac035.

6. M. P. Fox, R. F. MacLehose, and T. L. Lash, *Applying Quantitative Bias Analysis to Epidemiologic Data* (New York: Springer, 2022).

7. S. Schneeweiss, "Sensitivity Analysis and External Adjustment for Unmeasured Confounders in Epidemiologic Database Studies of Therapeutics," *Pharmacoepidemiology and Drug Safety* 15, no. 5 (2006): 291–303, https://doi.org/10.1002/pds.1200.

8. R. J. Desai, S. V. Wang, S. K. Sreedhara, et al., "Process Guide for Inferential Studies Using Healthcare Data From Routine Clinical Practice to Evaluate Causal Effects of Drugs (PRINCIPLED): Considerations From the FDA Sentinel Innovation Center," *BMJ* 384 (2024): e076460, https://doi.org/10.1136/bmj-2023-076460.

9. National Institute for Health and Care Excellence, NICE Real-World Evidence Framework (2022).

10. M. A. Hernán and J. M. Robins, *What If* (Boca Raton: Chapman & Hall/CRC, 2020).

11. S. Suissa, "Immortal Time Bias in Pharmacoepidemiology," *American Journal of Epidemiology* 167, no. 4 (2008): 492–499, https://doi.org/10.1093/aje/kwm324.

12. K. J. Rothman, T. L. Lash, T. J. VanderWeele, and S. Haneuse, *Modern Epidemiology*, vol. 3 (Philadelphia, USA: Wolters Kluwer, 2020).

13. S. Schneeweiss and P. S. Wang, "Association Between SSRI Use and Hip Fractures and the Effect of Residual Confounding Bias in Claims Database Studies," *Journal of Clinical Psychopharmacology* 24, no. 6 (2004): 632–638, https://doi.org/10.1097/01.jcp.0000145344.76288.39.

14. M. A. Hernán and S. R. Cole, "Invited Commentary: Causal Diagrams and Measurement Bias," *American Journal of Epidemiology* 170, no. 8 (2009): 959–962, https://doi.org/10.1093/aje/kwp293.

15. M. Rockenbauer, J. Olsen, A. E. Czeizel, L. Pedersen, H. T. Sørensen, and Group E, "Recall Bias in a Case-Control Surveillance System on the Use of Medicine During Pregnancy," *Epidemiology* 12, no. 4 (2001): 461–466, https://doi.org/10.1097/00001648-200107000-00017.

16. K. J. Lin, R. J. Glynn, D. E. Singer, S. N. Murphy, J. Lii, and S. Schneeweiss, "Out-of-System Care and Recording of Patient Characteristics Critical for Comparative Effectiveness Research," *Epidemiology* 29, no. 3 (2018): 356–363, https://doi.org/10.1097/ede.0000000000000794.

17. E. L. Ogburn and T. J. VanderWeele, "On the Nondifferential Misclassification of a Binary Confounder," *Epidemiology* 23, no. 3 (2012): 433–439, https://doi.org/10.1097/ede.0b013e31824d1f63.

18. M. A. Hernán, S. Hernández-Díaz, and J. M. Robins, "A Structural Approach to Selection Bias," *Epidemiology* 15, no. 5 (2004): 615–625, https://doi.org/10.1097/01.ede.0000135174.63482.43.

19. M. A. Hernán, "Invited Commentary: Selection Bias Without Colliders," *American Journal of Epidemiology* 185, no. 11 (2017): 1048–1050, https://doi.org/10.1093/aje/kwx077.

20. W. Liu, M. A. Brookhart, S. Schneeweiss, X. Mi, and S. Setoguchi, "Implications of M Bias in Epidemiologic Studies: A Simulation Study," *American Journal of Epidemiology* 176, no. 10 (2012): 938–948, https://doi.org/10.1093/aje/kws165.

21. S. G. Sullivan, E. J. T. Tchetgen, and B. J. Cowling, "Theoretical Basis of the Test-Negative Study Design for Assessment of Influenza Vaccine Effectiveness," *American Journal of Epidemiology* 184, no. 5 (2016): 345–353, https://doi.org/10.1093/aje/kww064.

22. C. J. Howe, L. E. Cain, and J. W. Hogan, "Are All Biases Missing Data Problems?," *Current Epidemiology Reports* 2, no. 3 (2015): 162–171, https://doi.org/10.1007/s40471-015-0050-8.

23. S. Greenland, "Bounding Analysis as an Inadequately Specified Methodology," *Risk Analysis* 24, no. 5 (2004): 1085–1092, https://doi.org/10.1111/j.0272-4332.2004.00509.x.

24. L. H. Smith and T. J. VanderWeele, "Bounding Bias Due to Selection," *Epidemiology (Cambridge, Mass.)* 30, no. 4 (2019): 509–516, https://doi.org/10.1097/ede.0000000000001032.

25. P. Ding and T. J. Vanderweele, "Generalized Cornfield Conditions for the Risk Difference," *Biometrika* 101, no. 4 (2014): 971–977, https://doi.org/10.1093/biomet/asu030.

26. P. Ding and T. J. VanderWeele, "Sensitivity Analysis Without Assumptions," *Epidemiology* 27, no. 3 (2016): 368–377, https://doi.org/10.1097/ede.0000000000000457.

27. M. P. Fox, T. L. Lash, and L. M. Bodnar, "Common Misconceptions About Validation Studies," *International Journal of Epidemiology* 49, no. 4 (2020): 1392–1396, https://doi.org/10.1093/ije/dyaa090.

28. E. Garne, H. Dolk, M. Loane, P. A. Boyd, and EUROCAT, "Eurocat Website Data on Prenatal Detection Rates of Congenital Anomalies," *Journal of Medical Screening* 17, no. 2 (2010): 97–98, https://doi.org/10.1258/jms.2010.010050.

29. T. J. VanderWeele and O. A. Arah, "Bias Formulas for Sensitivity Analysis of Unmeasured Confounding for General Outcomes, Treatments, and Confounders," *Epidemiology* 22, no. 1 (2011): 42–52, https://doi.org/10.1097/ede.0b013e3181f74493.

30. I. D. J. Bross, "Spurious Effects From an Extraneous Variable," *Journal of Chronic Diseases* 19, no. 6 (1966): 637–647, https://doi.org/10.1016/0021-9681(66)90062-2.

31. T. J. VanderWeele, "Optimal Approximate Conversions of Odds Ratios and Hazard Ratios to Risk Ratios," *Biometrics* 76, no. 3 (2020): 746–752, https://doi.org/10.1111/biom.13197.

32. K. E. Rudolph and E. A. Stuart, "Using Sensitivity Analyses for Unobserved Confounding to Address Covariate Measurement Error in Propensity Score Methods," *American Journal of Epidemiology* 187, no. 3 (2017): 604–613, https://doi.org/10.1093/aje/kwx248.

33. T. Stürmer, S. Schneeweiss, J. Avorn, and R. J. Glynn, "Adjusting Effect Estimates for Unmeasured Confounding With Validation Data Using Propensity Score Calibration," *American Journal of Epidemiology* 162, no. 3 (2005): 279–289, https://doi.org/10.1093/aje/kwi192.

34. F. Wan, "A Cautionary Note on Using Propensity Score Calibration to Control for Unmeasured Confounding Bias When the Surrogacy Assumption Is Absent," *American Journal of Epidemiology* 193, no. 2 (2023): 360–369, https://doi.org/10.1093/aje/kwad189.

35. D. R. V. Domelen and R. H. Lyles, "A Look at the Unique Identifiability of Propensity Score Calibration," *American Journal of Epidemiology* 188, no. 7 (2019): 1397–1399, https://doi.org/10.1093/aje/kwz072.

36. S. R. Cole, H. Chu, and S. Greenland, "Maximum Likelihood, Profile Likelihood, and Penalized Likelihood: A Primer," *American Journal of Epidemiology* 179, no. 2 (2014): 252–260, https://doi.org/10.1093/aje/kwt245.

37. G. W. Imbens, "Sensitivity to Exogeneity Assumptions in Program Evaluation," *American Economic Review* 93, no. 2 (2003): 126–132, https://doi.org/10.1257/000282803321946921.

38. G. W. Imbens and D. B. Rubin, *Causal Inference in Statistics, Social, and Biomedical Sciences* (Cambridge, UK: Cambridge University Press, 2015).

39. J. R. Carpenter, J. W. Bartlett, T. P. Morris, A. M. Wood, M. Quartagno, and M. G. Kenward, *Multiple Imputation and Its Application* (Chichester, UK: John Wiley & Sons, 2023).

40. C. Cinelli and C. Hazlett, "Making Sense of Sensitivity: Extending Omitted Variable Bias," *Journal of the Royal Statistical Society, Series B (Statistical Methodology)* 82, no. 1 (2020): 39–67, https://doi.org/10.1111/rssb.12348.

41. M. Corbin, S. Haslett, N. Pearce, M. Maule, and S. Greenland, "A Comparison of Sensitivity-Specificity Imputation, Direct Imputation and Fully Bayesian Analysis to Adjust for Exposure Misclassification When Validation Data Are Unavailable," *International Journal of Epidemiology* 46, no. 3 (2017): 1063–1072, https://doi.org/10.1093/ije/dyx027.

42. M. P. Fox, T. L. Lash, and S. Greenland, "A Method to Automate Probabilistic Sensitivity Analyses of Misclassified Binary Variables," *International Journal of Epidemiology* 34, no. 6 (2005): 1370–1376, https://doi.org/10.1093/ije/dyi184.

43. R. H. Lyles and J. Lin, "Sensitivity Analysis for Misclassification in Logistic Regression via Likelihood Methods and Predictive Value Weighting," *Statistics in Medicine* 29, no. 22 (2010): 2297–2309, https://doi.org/10.1002/sim.3971.

44. R. H. Lyles, L. Tang, H. M. Superak, et al., "Validation Data-Based Adjustments for Outcome Misclassification in Logistic Regression," *Epidemiology* 22, no. 4 (2011): 589–597, https://doi.org/10.1097/ede.0b013e3182117c85.

45. M. J. Morrissey and D. Spiegelman, "Matrix Methods for Estimating Odds Ratios With Misclassified Exposure Data: Extensions and Comparisons," *Biometrics* 55, no. 2 (1999): 338–344, https://doi.org/10.1111/j.0006-341x.1999.00338.x.

46. R. K. Ross, I. H. Su, M. Webster-Clark, and M. J. Funk, "Nondifferential Treatment Misclassification Biases Toward the Null? Not a Safe Bet for Active Comparator Studies," *American Journal of Epidemiology* 191, no. 11 (2022): 1917–1925, https://doi.org/10.1093/aje/kwac131.

47. J. K. Edwards, S. R. Cole, and M. P. Fox, "Flexibly Accounting for Exposure Misclassification With External Validation Data," *American Journal of Epidemiology* 189, no. 8 (2020): 850–860, https://doi.org/10.1093/aje/kwaa011.

48. R. H. Keogh, P. A. Shaw, P. Gustafson, et al., "STRATOS Guidance Document on Measurement Error and Misclassification of Variables in Observational Epidemiology: Part 1—Basic Theory and Simple Methods of Adjustment," *Statistics in Medicine* 39, no. 16 (2020): 2197–2231, https://doi.org/10.1002/sim.8532.

49. R. H. Keogh, P. A. Shaw, P. Gustafson, et al., "STRATOS Guidance Document on Measurement Error and Misclassification of Variables in Observational Epidemiology: Part 1—Basic Theory and Simple Methods of Adjustment," *Statistics in Medicine* 39, no. 16 (2020): 2197–2231, https://doi.org/10.1002/sim.8532.

50. S. R. Cole, H. Chu, and S. Greenland, "Multiple-Imputation for Measurement-Error Correction," *International Journal of Epidemiology* 35, no. 4 (2006): 1074–1081, https://doi.org/10.1093/ije/dyl097.

51. J. K. Edwards, S. R. Cole, M. A. Troester, and D. B. Richardson, "Accounting for Misclassified Outcomes in Binary Regression Models Using Multiple Imputation With Internal Validation Data," *American Journal of Epidemiology* 177, no. 9 (2013): 904–912, https://doi.org/10.1093/aje/kws340.

52. S. Greenland, "Basic Methods for Sensitivity Analysis of Biases," *International Journal of Epidemiology* 25, no. 6 (1996): 1107–1116, https://doi.org/10.1093/ije/25.6.1107-a.

53. K. Steenland and S. Greenland, "Monte Carlo Sensitivity Analysis and Bayesian Analysis of Smoking as an Unmeasured Confounder in a Study of Silica and Lung Cancer," *American Journal of Epidemiology* 160, no. 4 (2004): 384–392, https://doi.org/10.1093/aje/kwh211.

54. S. Greenland, "A Commentary on 'A Comparison of Bayesian and Monte Carlo Sensitivity Analysis for Unmeasured Confounding'," *Statistics in Medicine* 36, no. 20 (2017): 3278–3280, https://doi.org/10.1002/sim.7370.

55. L. C. McCandless and P. Gustafson, "A Comparison of Bayesian and Monte Carlo Sensitivity Analysis for Unmeasured Confounding," *Statistics in Medicine* 36, no. 18 (2017): 2887–2901, https://doi.org/10.1002/sim.7298.

56. R. F. MacLehose and P. Gustafson, "Is Probabilistic Bias Analysis Approximately Bayesian?," *Epidemiology* 23, no. 1 (2012): 151–158, https://doi.org/10.1097/ede.0b013e31823b539c.

57. G. James, D. Witten, T. Hastie, and R. Tibshirani, "An Introduction to Statistical Learning, With Applications in R," *Springer Texts in Statistics* (2021): 59–128, https://doi.org/10.1007/978-1-0716-1418-1_3.

58. G. W. Imbens and D. B. Rubin, Causal Inference for Statistics, Social, and Biomedical Sciences (2015), https://doi.org/10.1017/cbo9781139025751.

59. C. Poole, "Commentary: Continuing the E-Value's Post-Publication Peer Review," *International Journal of Epidemiology* 49, no. 5 (2020): 1497–1500, https://doi.org/10.1093/ije/dyaa097.

60. J. P. A. Ioannidis, Y. J. Tan, and M. R. Blum, "Limitations and Misinterpretations of E-Values for Sensitivity Analyses of Observational Studies," *Annals of Internal Medicine* 170, no. 2 (2019): 108–111, https://doi.org/10.7326/m18-2159.

61. M. P. Fox, O. A. Arah, and E. A. Stuart, "Commentary: The Value of E-Values and Why They Are Not Enough," *International Journal of Epidemiology* 49, no. 5 (2020): 1505–1506, https://doi.org/10.1093/ije/dyaa093.

62. T. J. VanderWeele and Y. Li, "Simple Sensitivity Analysis for Differential Measurement Error," *American Journal of Epidemiology* 188, no. 10 (2019): 1823–1829, https://doi.org/10.1093/aje/kwz133.

63. L. H. Smith, M. B. Mathur, and T. J. VanderWeele, "Multiple-Bias Sensitivity Analysis Using Bounds," *Epidemiology* 32, no. 5 (2021): 625–634, https://doi.org/10.1097/ede.0000000000001380.

64. E. Kawabata, K. Tilling, R. H. H. Groenwold, and R. A. Hughes, "Quantitative Bias Analysis in Practice: Review of Software for Regression With Unmeasured Confounding," *BMC Medical Research Methodology* 23, no. 1 (2023): 111, https://doi.org/10.1186/s12874-023-01906-8.

65. M. P. Fox, R. F. MacLehose, and T. L. Lash, "SAS and R Code for Probabilistic Quantitative Bias Analysis for Misclassified Binary Variables and Binary Unmeasured Confounders," *International Journal of Epidemiology* 52 (2023): 1624–1633, https://doi.org/10.1093/ije/dyad053.

66. D. Haine and M. D. Haine, Package 'Episensr' (2023).

67. M. B. Mathur, L. H. Smith, P. Ding, T. J. VanderWeele, and M. M. B. Mathur, Package 'EValue' (2021).

68. A. Linden, M. B. Mathur, and T. J. VanderWeele, "Conducting Sensitivity Analysis for Unmeasured Confounding in Observational Studies Using E-Values: The Evalue Package," *Stata Journal* 20, no. 1 (2020): 162–175, https://doi.org/10.1177/1536867x20909696.

69. N. Orsini, R. Bellocco, M. Bottai, A. Wolk, and S. Greenland, "A Tool for Deterministic and Probabilistic Sensitivity Analysis of Epidemiologic Studies," *Stata Journal* 8, no. 1 (2008): 29–48, https://doi.org/10.1177/1536867x0800800103.

70. T. L. Lash, T. P. Ahern, L. J. Collin, M. P. Fox, and R. F. MacLehose, "Bias Analysis Gone Bad," *American Journal of Epidemiology* 190, no. 8 (2021): 1604–1612, https://doi.org/10.1093/aje/kwab072.

71. R. H. H. Groenwold, J. A. C. Sterne, D. A. Lawlor, K. G. M. Moons, A. W. Hoes, and K. Tilling, "Sensitivity Analysis for the Effects of Multiple Unmeasured Confounders," *Annals of Epidemiology* 26, no. 9 (2016): 605–611, https://doi.org/10.1016/j.annepidem.2016.07.009.

72. E. A. Suarez, K. F. Huybrechts, L. Straub, et al., "Buprenorphine Versus Methadone for Opioid Use Disorder in Pregnancy," *New England Journal of Medicine* 387, no. 22 (2022): 2033–2044, https://doi.org/10.1056/nejmoa2203318.

73. M. Zhang, M. Falconer, and L. Taylor, "A Quantitative Bias Analysis of the Confounding Effects due to Smoking on the Association Between Fluoroquinolones and Risk of Aortic Aneurysm," *Pharmacoepidemiology and Drug Safety* 29, no. 8 (2020): 958–961, https://doi.org/10.1002/pds.5019.

74. J. P. Brown, K. Wing, C. Leyrat, et al., "Association Between Fluoroquinolone Use and Hospitalization With Aortic Aneurysm or Aortic Dissection," *JAMA Cardiology* 8 (2023): 865–870, https://doi.org/10.1001/jamacardio.2023.2418.

Appendix A

A.1 | Formulas

Formulas to calculate sensitivity and specificity from predictive values and misclassified (i.e., observed) prevalence of variable:

$$\text{Sensitivity} = \frac{\text{PPV} \times \text{misclassified prevalence}}{(\text{PPV} \times \text{misclassified prevalence}) + ((1 - \text{NPV}) \times (1 - \text{misclassified prevalence}))}$$

$$\text{Specificity} = \frac{\text{NPV} \times (1 - \text{misclassified prevalence})}{\big(\text{NPV} \times (1 - \text{misclassified prevalence})\big) + ((1 - \text{PPV}) \times \text{misclassified prevalence})}$$

Note these formulas assume sensitivity, specificity, predictive values, and prevalence are specified as proportions between zero and one (e.g., prevalence of 0.70).