

UC Berkeley

UC Berkeley Previously Published Works

Title

E-values for effect heterogeneity and approximations for causal interaction

Permalink

<https://escholarship.org/uc/item/0jq176v4>

Journal

International Journal of Epidemiology, 51(4)

ISSN

0300-5771

Authors

Mathur, Maya B

Smith, Louisa H

Yoshida, Kazuki

et al.

Publication Date

2022-08-10

DOI

10.1093/ije/dyac073

Peer reviewed

Methods

E-values for effect heterogeneity and approximations for causal interaction

Maya B Mathur ^{1*}, Louisa H Smith ², Kazuki Yoshida,³ Peng Ding⁴ and Tyler J VanderWeele ⁵

¹Quantitative Sciences Unit and Department of Pediatrics, Stanford University, Palo Alto, CA, USA, ²Roux Institute, Northeastern University, Portland, ME, USA, ³Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA, ⁴Department of Statistics, University of California at Berkeley, Berkeley, CA, USA and ⁵Department of Epidemiology, Harvard University, Boston, MA, USA

*Corresponding author. Quantitative Sciences Unit and Department of Pediatrics, Stanford University, 1701 Page Mill Road, Palo Alto, CA 94304, USA. E-mail: mmathur@stanford.edu

Received 5 August 2021; Editorial decision 13 March 2022; Accepted 1 April 2022

Abstract

Background: Estimates of effect heterogeneity (i.e. the extent to which the causal effect of one exposure varies across strata of a second exposure) can be biased if the exposure–outcome relationship is subject to uncontrolled confounding whose severity differs across strata of the second exposure.

Methods: We propose methods, analogous to the E-value for total effects, that help to assess the sensitivity of effect heterogeneity estimates to possible uncontrolled confounding. These E-value analogues characterize the severity of uncontrolled confounding strengths that would be required, hypothetically, to ‘explain away’ an estimate of multiplicative or additive effect heterogeneity in the sense that appropriately controlling for those confounder(s) would have shifted the effect heterogeneity estimate to the null, or alternatively would have shifted its confidence interval to include the null. One can also consider shifting the estimate or confidence interval to an arbitrary non-null value. All of these E-values can be obtained using the R package EValue.

Results: We illustrate applying the proposed E-value analogues to studies on: (i) effect heterogeneity by sex of the effect of educational attainment on dementia incidence and (ii) effect heterogeneity by age on the effect of obesity on all-cause mortality.

Conclusion: Reporting these proposed E-values could help characterize the robustness of effect heterogeneity estimates to potential uncontrolled confounding.

Key words: Sensitivity analysis, bias analysis, confounding, interaction, effect modification, effect heterogeneity

Key Messages

- Effect heterogeneity estimates can be biased if the exposure–outcome relationship is subject to uncontrolled confounding whose severity differs across strata of the second exposure.
- We propose sensitivity analyses, analogous to the E-value, that characterize the severity of uncontrolled confounding strengths that would be required to ‘explain away’ an estimate of effect heterogeneity.
- We provide an R package, EValue, to conduct all proposed sensitivity analyses.

Background

Estimates of effect heterogeneity (i.e. the extent to which the causal effect of an exposure, X , varies across strata of another variable, Z) can be biased if the exposure–outcome relationship is subject to uncontrolled confounding whose severity differs across strata of Z .^{1,2} For example, suppose that in the $Z = 1$ stratum, the exposure–outcome relationship is biased upward due to uncontrolled confounding, but that in the $Z = 0$ stratum, the relationship is not biased. The uncontrolled confounding in the $Z = 1$ stratum could, for example, spuriously increase the magnitude of the observed ratio or difference between the two strata’s estimated exposure–outcome effects, thus increasing the observed estimate of effect heterogeneity. Such bias in effect heterogeneity estimates can be consequential in practice. For example, effect heterogeneity estimates on the additive scale can inform decisions of which individuals (i.e. which stratum of Z) to treat in order to most reduce the disease caseload if treatment resources are limited.¹ In this context, a biased estimate of effect heterogeneity could potentially suggest a treatment allocation scheme that fails to minimize, or even increases, the overall caseload.

Although sensitivity analyses exist for causal interactions (i.e. effects of jointly manipulating X and Z vs each independently),³ we are not aware of comparable methods for effect heterogeneity. We propose sensitivity analyses that characterize the severity of uncontrolled confounding strengths that would be required, hypothetically, to ‘explain away’ an estimate of multiplicative or additive effect heterogeneity in the sense that controlling for those confounder(s) would have shifted the effect heterogeneity estimate to the null, or alternatively would have shifted its confidence interval to include the null. We recommend reporting E-values for both the point estimate and the confidence interval; the latter is especially important for effect heterogeneity estimates, for which statistical precision is often considerably lower than for total effects. We also discuss ‘non-null’ E-values required to shift the estimate or its confidence interval to any arbitrary value. These metrics are straightforward extensions of the standard E-value for total effects, which represents the minimum strength of

association, on the risk ratio (RR) scale, that uncontrolled confounder(s) would need to have with the exposure, the outcome, or both, conditional on any measured and controlled confounders, to explain away the total effect.^{4,5} As we describe below, it is equivalent to interpret the E-value as the minimum strengths of association that uncontrolled confounder(s) would need to have with both the exposure and the outcome if these two strengths of association are taken to be of equal magnitude.

Like the standard E-value, our proposed E-values for effect heterogeneity do not require assumptions on the nature of the uncontrolled confounder(s).^{4–6} That is, these metrics represent bounds under hypothetical worst-case confounding: they consider the *maximum* bias that could be generated by a given strength of confounder strengths, but actual uncontrolled confounders might not generate that much bias.^{4,5} We also give alternative E-values that can be applied under the assumption that uncontrolled confounding operates in the same direction in each stratum of Z (‘unidirectional confounding’). We provide software to calculate all of these E-values, discussed below.

Details and reporting guidelines for the standard E-value are discussed elsewhere.^{4–7} The E-value has limitations, which have been discussed and debated elsewhere.^{8–11} For example, to avoid making assumptions about the prevalence or distribution of uncontrolled confounder(s), the E-value does not make use of potential known information about these and hence may understate the amount of confounding required to explain away an effect. Additionally, the E-value does not account for biases or threats to inferential validity other than uncontrolled confounding, such as measurement error, selective reporting or uncontrolled multiple testing; we provide methods and recommendations regarding these issues elsewhere.^{6,12–14} The same considerations and limitations apply for the analogues we propose.

Setting and notation

All proofs, with formalized assumptions and definitions, appear in the [Supplementary material](#) (available as [Supplementary data](#) at *IJE* online). Let $Z \in \{0, 1\}$ define the strata between which the causal effect of $X \in \{0, 1\}$ might

vary, and let $Y_x \in \{0, 1\}$ be a potential outcome when intervening to set $X = x$. (If Z is categorical rather than binary, the same results can be applied to contrasts between two specified levels of Z .) We assume that the confounded estimate of effect heterogeneity is greater than the null (e.g. >1 for multiplicative measures or >0 for additive measures); otherwise, one can simply reverse the coding of Z before applying the results below. We also assume that, other than the omission of uncontrolled confounder(s), the estimator used to obtain the confounded estimate of effect heterogeneity is correctly specified; i.e. the estimate would have been unbiased for the causal estimand if the uncontrolled confounder(s) had in fact been controlled. For example, if the statistical interaction coefficient for $X \times Z$ in a regression model is taken to be the effect heterogeneity estimate, this requires assuming that X does not affect Z , even if there were no uncontrolled confounding.

We define two sensitivity parameters for each stratum, called the ‘within-stratum confounding strengths’. These are the same parameters as used for the standard E-value, but applied separately to each stratum of Z .⁵ Within stratum $Z = z$, let U_z denote the uncontrolled confounder(s), defined as a set of one or more variables that would suffice to control for confounding of the exposure–outcome relationship in this stratum.

Within stratum $Z = z$ and for each stratum x of X , define an RR of U_z on Y , maximized across all possible contrasts of U_z :

$$RR_{UY|Z=z,X=x} = \frac{\max_u E[Y|Z=z, X=x, U_z=u]}{\min_u E[Y|Z=z, X=x, U_z=u]}, \quad x \in \{0, 1\}$$

The sensitivity parameters $RR_{UY|Z=1}$ and $RR_{UY|Z=0}$ are then:

$$RR_{UY|Z=z} = \max\{RR_{UY|Z=z,X=0}, RR_{UY|Z=z,X=1}\}, \quad z \in \{0, 1\}$$

Any confounders that are measured and controlled could also be conditioned throughout. These sensitivity parameters represent, within stratum $Z = z$ and considering both strata of X , the largest of the maximal RRs of that stratum’s uncontrolled confounders, U_z , on Y conditional on X . The sensitivity parameters $RR_{XU|Z=1}$ and $RR_{XU|Z=0}$ are defined as:

$$RR_{XU|Z=z} = \max_u \left\{ \frac{P(U_z = u | Z = z, X = 1)}{P(U_z = u | Z = z, X = 0)} \right\}, \quad z \in \{0, 1\}$$

Each parameter $RR_{XU|Z=z}$ represents, within stratum $Z = z$, the maximal RR of $U_z = u$ for $X = 1$ vs $X = 0$ across strata of U_z . As for the standard E-value, the same

sensitivity parameters and results accommodate the possibility that U_z includes one or more confounders of any type (e.g. binary, categorical or continuous) and distribution. For example, if U_z is binary, $RR_{UY|Z=z,X=x}$ is simply the RR relating U_z to Y within stratum $Z = z$. Precise interpretations of the sensitivity parameters when U_z contains multiple confounders or is continuous are given elsewhere.⁶

We define the E-value for an effect heterogeneity point estimate as the minimum magnitude that at least one of the four within-stratum confounding strengths must have, on the RR scale, such that fully controlling for confounding would have shifted the estimate to the null. Like the standard E-value, this E-value essentially sets the four confounding strengths equal to one another to obtain the required joint minimum for all of them. For the special case in which confounding is assumed to be unidirectional, we define E-values as the minimum value that at least one of the two confounding strengths in at least one stratum must have in order to explain away the effect heterogeneity.

To dispel a common misconception, we note that mathematically setting the confounding strengths equal to one another in this manner does not require *assuming* that, in reality, the confounding strengths actually are equal. The E-value is derived by considering *all* possible combinations of confounding strengths that could produce enough bias to explain away the effect heterogeneity, and then solving for the combination that minimizes the maximum of these confounding strengths (Supplementary material, Section 2.1, available as Supplementary data at IJE online). This unique combination, it turns out, is the one in which the confounding strengths are equal. If, in reality, there are uncontrolled confounders whose confounding strengths are not equal, the E-value still applies; it states that *at least one* of the confounding strengths must exceed the E-value in order to explain away the effect. Again, this interpretation is mathematically equivalent to considering confounding strengths of equal magnitude. Additionally, we note that setting the confounding strengths equal to one another does not require assuming that these associations arise from the same confounders in the two strata.

Multiplicative effect heterogeneity

We consider the causal estimand:

$$RR_{EH}^t = RR_{XY|Z=1}^t / RR_{XY|Z=0}^t$$

where $RR_{XY|Z=z}^t = E[Y_1 | Z = z] / E[Y_0 | Z = z]$. Its confounded counterpart is:

$$RR_{EH}^c = RR_{XY|Z=1}^c / RR_{XY|Z=0}^c$$

where $RR_{XY|Z=z}^c = E[Y|Z=z, X=1] / E[Y|Z=z, X=0]$. Again, any confounders that are measured and controlled can also be conditioned. (If multiplicative effect heterogeneity is measured on a different scale, such as the odds ratio or hazard ratio scale, then the estimate could be approximately converted to an RR before applying the results below, as for the standard E-value.⁴) Without making assumptions on the direction of confounding bias, the E-value required to shift RR_{EH}^c to the null is:

$$E\text{-value} = \sqrt{RR_{EH}^c} + \sqrt{RR_{EH}^c - \sqrt{RR_{EH}^c}} \quad (1)$$

This E-value represents the minimum magnitude that at least one of the four within-stratum confounding strengths must have in order to explain away the effect heterogeneity (i.e. to have $RR_{EH}^t = 1$). Equivalently, this E-value represents the minimum magnitude of all four within-stratum confounding strengths that would be required to explain away the effect heterogeneity if all four confounding strengths are taken to be of equal magnitude. We use ‘magnitude’ to indicate that the confounding strengths on the RR scale are taken to be ≥ 1 regardless of the direction of association, as for the standard E-value.⁴ In fact, this bound is attained when the effects within each stratum have bias of the same magnitude, but in opposite directions. It might be quite unlikely in practice that any given uncontrolled confounder has all four confounding strengths of equal magnitude, or that confounding bias operates in different directions in each stratum. We return to these points in the section “Interpreting E-values in light of their mathematical conservatism”. In the [Supplementary material](#) Section 2.3 (available as [Supplementary data](#) at *IJE* online), we establish connections between this bound and classical bounds on bias in total effects due to uncontrolled confounding.^{15,16}

The E-value required to shift RR_{EH}^c to a non-null value, RR_{EH}^t , rather than to the null can be obtained by replacing RR_{EH}^c in [Equation \(1\)](#) with RR_{EH}^c / RR_{EH}^t . The E-value required to shift the confidence interval to include the null or another specified value can be obtained by replacing RR_{EH}^c above with the lower confidence interval limit. All of these E-values can be obtained using existing software that calculates E-values for total effects^{17,18} by simply performing the calculation using $\sqrt{RR_{EH}^c}$ rather than RR_{EH}^c itself, and likewise for the confidence interval limit or RR_{EH}^c / RR_{EH}^t , because doing so is equivalent to applying [Equation \(1\)](#).

If we instead assume unidirectional confounding, the E-value required to shift RR_{EH}^c to the null is:

$$E\text{-value} = RR_{EH}^c + \sqrt{RR_{EH}^c (RR_{EH}^c - 1)} \quad (2)$$

This represents the minimum magnitude that at least of one of the within-stratum confounding strengths must have in at least one stratum of Z in order to explain away the effect heterogeneity. Equivalently, this E-value represents the minimum magnitude that both within-stratum confounding strengths must have in at least one stratum of Z if both confounding strengths in that stratum are taken to be equal. This bound is attained when the effect within the other stratum is unbiased. The expression is in fact equivalent to the standard E-value for RR_{EH}^c .

Additive effect heterogeneity

Define the causal risk difference in stratum $Z = z$ and its confounded counterpart as:

$$RD_{XY|Z=z}^t = E[Y_1 | Z = z] - E[Y_0 | Z = z]$$

$$RD_{XY|Z=z}^c = E[Y | Z = z, X = 1] - E[Y | Z = z, X = 0]$$

We consider the additive interaction contrast $IC^t = RD_{XY|Z=1}^t - RD_{XY|Z=0}^t$ and its confounded counterpart $IC^c = RD_{XY|Z=1}^c - RD_{XY|Z=0}^c$.¹⁹ Let $f_z = E[X | Z = z]$ be the prevalence of X in stratum z , and let $p_{zx} = E[Y | Z = z, X = x]$ be the outcome probability within the joint stratum ($Z = z, X = x$). Then, without assumptions on the direction of confounding bias, the E-value required to shift IC^c to the null is:

$$E\text{-value} = g\left(\frac{1}{2(f_1 p_{10} + f_0 p_{01})}\left\{\sqrt{\gamma^2 + 4(p_{10} f_1 + p_{01} f_0)(p_{11}[1 - f_1] + p_{00}[1 - f_0])} - \gamma\right\}\right) \quad (3)$$

where

$$g(w) = w + \sqrt{w(w - 1)}$$

and

$$\gamma = p_{10}(1 - f_1) - p_{11}f_1 + p_{01}(1 - f_0) - p_{00}f_0$$

This bound is attained when the within-stratum effects are biased in opposite directions, but with potentially different amounts of absolute bias ($|RD_{XY|Z=z}^c - RD_{XY|Z=z}^t|$) in each stratum. (This asymmetry arises because, in the additive case, the amount of absolute bias produced by a given fixed set of sensitivity parameters depends on nuisance parameters, such as exposure prevalences and outcome probabilities, that can differ between strata of Z . In contrast, for the multiplicative case, the amount of multiplicative bias is independent of any such nuisance

parameters.) The [Supplementary material](#) (available as [Supplementary data](#) at *IJE* online) provides an E-value for the confidence interval, a generalization for shifting IC^c to a non-null value, and E-values that apply if the direction of the confounding bias is assumed to be unidirectional and positive, unidirectional and negative, or unidirectional with the direction unknown. All of these E-values for interaction contrasts can be obtained in R via `EValue :: evalues.IC`.¹⁷

Applied examples

Education and dementia

Letenneur *et al.*²⁰ investigated the effect of low vs high education (≤ 7 years of schooling vs ≥ 12 years) on dementia incidence, additionally estimating effect heterogeneity by sex. They pooled data from population-based longitudinal studies of aged women ($n = 3,352$) and men ($n = 2,395$). In analyses that adjusted for baseline confounders (age, smoking, myocardial infarction, stroke and study centre), the authors estimated a strong association between low education and dementia incidence in women [$RR = 3.78$ (95% CI: 1.64, 8.72)], but an apparently much weaker association in men [$RR = 1.09$ (0.61, 1.94)]. The authors suggested that uncontrolled confounding might have produced this apparent effect heterogeneity. For example, they speculated that socio-economic status, an uncontrolled confounder, might produce different biases for each sex.

On the multiplicative scale, we estimated $RR_{EH}^c = 3.47$ (1.26, 9.57); $P = 0.02$. Without making assumptions on the direction of uncontrolled confounding bias for each sex, the E-values for this estimate and its lower confidence interval limit were thus 3.13 and 1.49, respectively [from [Equation \(1\)](#)]. Thus, at least one of the four confounding strengths would need to be at least 3.13 on the RR scale in order to explain away the effect heterogeneity and would need to be at least 1.49 to shift the confidence interval to include the null. If we assume that uncontrolled confounding operated in the same direction for each sex, then the E-values for the point estimate and its lower confidence interval limit increase to 6.39 and 1.82, respectively [from [Equation \(2\)](#)]. Given these studies' control of several known confounders, it may be somewhat implausible that uncontrolled confounders, such as socio-economic status, were strong enough to attain these confounding strengths, although such judgments would need to be informed by domain expertise.

On the additive scale, we estimated risk differences of 0.04 (0.02, 0.05) for women and 0.01 (−0.01, 0.02) for men, such that $IC^c = 0.03$ (0.01, 0.05); $P = 0.01$. This analysis does not adjust for confounders because Letenneur *et al.*²⁰

reported only unadjusted prevalences. Without assumptions on the direction of confounding bias, the E-values for the interaction contrast and its lower confidence interval limit were thus 2.54 and 1.44, respectively [from [Equation \(3\)](#)]. If we assume that uncontrolled confounding operated in the same unspecified direction for both sexes, then these E-values respectively become 3.30 and 1.63 ([Supplementary material](#), Section 3.2, available as [Supplementary data](#) at *IJE* online).

Obesity and all-cause mortality

Winter *et al.*²¹ meta-analysed eight longitudinal studies to investigate the extent to which the effects of body mass index (BMI) on mortality differed by age (< 65 vs ≥ 65 years). The authors estimated that being obese ($BMI \geq 30$) vs normal weight ($20 \leq BMI < 25$) was associated with increased mortality among participants aged < 65 years [$HR = 1.42$ (95% CI: 1.22, 1.65)], but not among participants aged ≥ 65 years [$HR = 1.04$ (95% CI: 0.91, 1.19)]. Confounding control in the eight studies was quite limited: e.g. five studies did not control for comorbid health conditions, two studies did not control for smoking, five did not control for physical activity, and none controlled for diet or caloric intake.

On the multiplicative scale, we estimated $HR_{EH}^c = 1.37$ (1.12, 1.67); $P = 0.002$. E-values can be applied to meta-analysis point estimates, in which case they represent average confounding strengths across studies.^{13,22} Without making assumptions on the direction of uncontrolled confounding bias for each age group, the E-values for this estimate and its lower confidence interval limit were thus 1.61 and 1.30, respectively. If we assume that uncontrolled confounding operated in the same direction for each age group, then the E-values for the point estimate and its lower confidence interval limit increase to 2.07 and 1.48, respectively. Given these studies' limited control of confounders whose associations with both obesity and mortality may be quite strong, it may be plausible that uncontrolled confounders could have the strengths of association indicated by the E-values. Furthermore, it may be plausible that the strength of confounding bias could differ by age group. For example, older individuals might be less physically resilient to comorbid conditions than younger individuals, such that having comorbid conditions might be more strongly associated with BMI or with mortality among older individuals. We could not conduct analyses on the additive scale given the statistics reported in the meta-analysis.²¹

Practical interpretation and reporting of E-values

Elsewhere, we have provided recommendations on reporting E-values for total effects^{4,7} and provided caveats about

potential misinterpretations.^{11,23,24} These considerations apply to the present analogues as well; we comment here on only a subset of these considerations that are particularly pertinent to the setting of effect heterogeneity.

Assessing the plausibility that confounding strengths attain the E-value

As we have emphasized in the context of total effects, E-values must be interpreted in light of the quality of a study's existing control for confounding: it will be more plausible that there exists uncontrolled confounding as strong as is indicated by the E-value in a poorly controlled study than a well-controlled study.^{4,7} As a limitation of sensitivity analyses in general, it can be challenging to assess plausible strengths of association of uncontrolled confounder(s). This may be particularly so in the context of effect heterogeneity, for which the relevant confounding strengths are defined within the strata of Z rather than marginally.

In the context of total effects, we have suggested listing specific variables that are thought to be uncontrolled confounders and, to help benchmark intuitions, reporting measured confounders' strengths of association with X and with Y . Because E-values consider joint associations produced by potentially multiple uncontrolled confounder(s), above and beyond controlled confounder(s), it may be particularly informative to report confounding strengths for each measured confounder as well as for all measured confounders jointly.⁶ However, these empirical benchmarks must be interpreted carefully to avoid common misconceptions.^{7,24}

When applying E-values for effect heterogeneity, one could similarly report measured confounding strengths within strata of Z ; e.g. one could report the associations of each measured confounder, and of all measured confounders jointly, with X , and with Y for each stratum of Z . We acknowledge, though, that comparing the E-value to such benchmarks may be difficult because confounders, whether measured or not, can differ substantially in their associations with the exposure and outcome as well as in the strengths of these associations in each stratum. When additional information about uncontrolled confounders is available, the E-value could be supplemented or replaced by more precise sensitivity analyses; we return to this point in the next section.

Interpreting E-values in light of their mathematical conservatism

As noted above, like the standard E-value for total effects,⁴ our proposed analogues avoid making assumptions about the

prevalence or distribution of uncontrolled confounder(s) by not incorporating any sensitivity parameters regarding these properties. Therefore, E-values might understate the amount of confounding required to explain away the effect or effect heterogeneity. For example, if U_z is a binary variable, then given its confounding strengths, the bias it produces in stratum $Z = z$ would be maximized if $E[U_z | Z = z, X = 1] = 1$ and $E[U_z | Z = z, X = 0] = 0$ (or vice versa). The E-value is conservative in that it allows for this extreme, and sometimes implausible, possibility.

In some settings, one might wish to conduct a sensitivity analysis for specific uncontrolled confounder(s) whose prevalence conditional on X or conditional on Z is known. One could then obtain a more precise sensitivity analysis by applying methods that do incorporate information about prevalences. For effect heterogeneity, one could apply existing sensitivity analyses for uncontrolled confounding that incorporate prevalences or other external information (reviewed in²⁵) to bound the bias in each stratum of Z separately. Analogous bounds incorporating prevalences exist for causal interaction.³

Of course, the disadvantage of incorporating prevalences is that, if the specified prevalences are incorrect or if there exist other uncontrolled confounder(s) with prevalences other than those specified, bounds obtained by incorporating prevalences may give a false impression of robustness to uncontrolled confounding. Additionally, methods that involve specifying numerous sensitivity parameters could introduce additional 'researcher degrees of freedom', such that researchers could potentially search for combinations of parameters that produce attractive results of sensitivity analyses.⁴

For these reasons, in the context of total effects, we suggested that when external information is available regarding uncontrolled confounder(s), one might first report the E-value, because despite its conservatism, the E-value might nevertheless be large enough to suggest robustness to uncontrolled confounding, even without making use of external information.²⁵ One could then supplement this mathematically conservative analysis with additional sensitivity analyses that do incorporate external information.²⁵ The same considerations also apply when considering effect heterogeneity.

A second form of conservatism arises specifically in the context of E-values for effect heterogeneity. Namely, as noted above, the general bound in Equation (1) is attained when the effects in each stratum of Z have bias of the same magnitude, but in opposite directions. In some scientific contexts, it might not be plausible that uncontrolled confounder(s) could in fact produce bias in opposite directions,

and in these settings, the general bound in Equation (1) might again understate the amount of confounding required to explain away the effect heterogeneity. However, in other contexts, it might be quite plausible that confounding bias could operate in different directions, because either the uncontrolled confounder(s) themselves, or alternatively the direction of their effects on X or on Y , could differ between strata. Consider a hypothetical study examining effect heterogeneity between men and women (Z) in the effects of smoking (X) on all-cause mortality (Y). Suppose the study excluded heavy drinkers but did not control for moderate vs low alcohol consumption (U_2)—a variable that is associated with increased smoking for both sexes. Moderate alcohol consumption is thought to reduce the risk of cardiovascular disease and diabetes, especially among individuals with predisposing risk factors, but is thought to increase the risk of cancer and other chronic diseases.²⁶ In some populations, then, moderate consumption could plausibly have a protective net effect on all-cause mortality for men (with their relatively higher burden of cardiovascular and metabolic diseases and related risk factors), yet could have a detrimental net effect for women.²⁷ If this is the case, alcohol consumption could produce confounding bias in different directions for men vs women.

Conclusion

Reporting these proposed E-values could help characterize the robustness of effect heterogeneity estimates to potential uncontrolled confounding. These results apply to effect heterogeneity rather than causal interaction; the [Supplementary material](#) Section 4 (available as [Supplementary data](#) at *IJE* online) provides E-values for causal interaction that are approximate, ‘weak’ bounds in a sense detailed there. The above E-values for effect heterogeneity could also be applied for causal interaction if one exposure is assumed to be unconfounded (e.g. because it was randomized).

Ethics approval

Not applicable.

Data availability

All code, materials and data required to reproduce the applied example are publicly available and documented (<https://osf.io/79scv/>). Code for the R package EValue and the online tool are open-source (https://github.com/mayamathur/evalue_package and https://github.com/mayamathur/evalue_website).

Supplementary data

[Supplementary data](#) are available at *IJE* online.

Author contributions

M.B.M. conceived of the research, led theoretical developments, led writing, analysed the applied example and wrote code for the R package. L.H.S. contributed to theoretical developments. L.H.S. and all other authors contributed critical intellectual content to writing.

Funding

M.B.M. and T.V.W. were supported by R01 CA222147 and R01 LM013866. M.B.M. was supported by: (i) the NIH-funded Biostatistics, Epidemiology and Research Design (BERD) Shared Resource of Stanford University’s Clinical and Translational Education and Research (UL1TR003142); (ii) the Biostatistics Shared Resource (BSR) of the NIH-funded Stanford Cancer Institute (P30CA124435); and (iii) the Quantitative Sciences Unit through the Stanford Diabetes Research Center (P30DK116074). K.Y. was supported by Brigham and Women’s Hospital Department of Medicine Fellowship Award and K23AR076453 (NIAMS). The funders had no role in the design, conduct or reporting of this research.

Acknowledgements

None.

Conflict of interest

None declared.

References

1. VanderWeele TJ. *Explanation in Causal Inference: Methods for Mediation and Interaction*. Oxford: Oxford University Press, 2015.
2. VanderWeele TJ. On the distinction between interaction and effect modification. *Epidemiology* 2009;20:863–71.
3. VanderWeele TJ, Mukherjee B, Chen J. Sensitivity analysis for interactions under unmeasured confounding. *Stat Med* 2012;31:2552–64.
4. VanderWeele TJ, Ding P. Sensitivity analysis in observational research: introducing the E-value. *Ann Intern Med* 2017;167:268–74.
5. Ding P, VanderWeele TJ. Sensitivity analysis without assumptions. *Epidemiology* 2016;27:368–77.
6. VanderWeele TJ, Ding P, Mathur M. Technical considerations in the use of the E-value. *J Causal Inference* 2019;7.
7. VanderWeele TJ, Mathur MB. Commentary: developing best-practice guidelines for the reporting of E-values. *Int J Epidemiol* 2020;49:1495–97.
8. Poole C. Commentary: Continuing the E-value’s post-publication peer review. *Int J Epidemiol* 2020;49:1497–500.
9. Greenland S. Commentary: An argument against E-values for assessing the plausibility that an association could be explained away by residual confounding. *Int J Epidemiol* 2020;49:1501–03.
10. Ioannidis JP, Tan YJ, Blum MR. Limitations and misinterpretations of E-values for sensitivity analyses of observational studies. *Ann Intern Med* 2019;170:108–11.

11. VanderWeele TJ. Are Greenland, Ioannidis, and Poole opposed to the Cornfield conditions? A defense of the E-value. *Int J Epidemiol* 2022;51:364–71.
12. VanderWeele TJ, Mathur MB, Chen Y. Outcome-wide longitudinal designs for causal inference: a new template for empirical studies. *Stat Sci* 2020;35:437–66.
13. Mathur MB, VanderWeele TJ. Sensitivity analysis for unmeasured confounding in meta-analyses. *J Am Stat Assoc* 2020;115:163–72.
14. Smith LH, Mathur MB, VanderWeele TJ. Multiple-bias sensitivity analysis using bounds. *Epidemiology* 2021;32:625–34.
15. Cornfield J, Haenszel W, Hammond EC, Lilienfeld AM, Shimkin MB, Wynder EL. Smoking and lung cancer: recent evidence and a discussion of some questions. *J Natl Cancer Inst* 1959;22:173–203.
16. Schlesselman JJ. Assessing effects of confounding variables. *Am J Epidemiol* 1978;108:3–8.
17. Mathur MB, Ding P, Riddell CA, VanderWeele TJ. Website and R package for computing E-values. *Epidemiology* 2018;29:e45–47.
18. Linden A, Mathur MB, VanderWeele TJ. Conducting sensitivity analysis for unmeasured confounding in observational studies using E-values: the EValue package. *Stata J* 2020;20:162–75.
19. VanderWeele TJ, Lash TL, Rothman KJ, Chapter 26, Analysis of interaction. In: *Modern Epidemiology*, 4th edn. Philadelphia, PA: Wolters Kluwer, 2021, pp. 619–54.
20. Letenneur L, Launer J, Andersen K *et al.*; for the EURODEM Incidence Research Group. Education and risk for Alzheimer's disease: sex makes a difference. EURODEM pooled analyses. *Am J Epidemiol* 2000;151:1064–71.
21. Winter J, MacInnis R, Nowson C. The influence of age on the BMI and all-cause mortality association: a meta-analysis. *J Nutr Health Aging* 2017;21:1254–58.
22. Mathur MB, VanderWeele TJ. Methods to address confounding and other biases in meta-analyses: review and recommendations. *Annu Rev Public Health* 2021;43:19–35.
23. VanderWeele TJ, Mathur MB, Ding P. Correcting Misinterpretations of the E-Value. *Ann Intern Med* 2019;170:131–32.
24. Mathur MB, VanderWeele TJ. How to report E-values for meta-analyses: recommended improvements and additions to the new GRADE approach. *Environ Int* 2022;160:107032.
25. Zhang X, Stamey JD, Mathur MB. Assessing the impact of unmeasured confounders for credible and reliable real-world evidence. *Pharmacoepidemiol Drug Saf* 2020;29:1219–27.
26. Griswold MG, Fullman N, Hawley C *et al.* Alcohol use and burden for 195 countries and territories, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet* 2018;392:1015–35.
27. White IR, Altmann DR, Nanchahal K. Alcohol consumption and mortality: modelling risks for men and women at different ages. *BMJ* 2002;325:191.