

## PEER REVIEW HISTORY

BMJ Medicine publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

### ARTICLE DETAILS

<b>TITLE (PROVISIONAL)</b>	Quantifying the impact of unmeasured confounding in observational studies using the E-value
<b>AUTHORS</b>	Ehrenstein, Vera; Gaster, Tobias; Eggertsen, Christine Marie; Støvring, Henrik; Petersen, Irene

### VERSION 1 - REVIEW

<b>REVIEWER 1</b>	Riley, Richard; University of Birmingham, Institute of Applied Health Research. Competing Interest: None
<b>REVIEW RETURNED</b>	27-Sep-2022

<b>GENERAL COMMENTS</b>	<p>This is a very well-written piece explaining the use of E-value method for examining the potential impact of unmeasured confounding in observational research. The paper would be best as a methods primer, about the e-value method, rather than focusing as much on the case study. In terms of the methods, I have a number of suggestions for improvement and clarification:</p> <ol style="list-style-type: none"><li>1) Given the focus on the e-value method, I think this should be mentioned in the title, and the paper considered as a methods primer for the journal.</li><li>2) My strongest comment is to explain how this method should relate to causal inference methods that base inferences on (assumed) causal pathways and DAGs. At the moment, this paper does not consider the causal pathway premise, but rather focuses on regression adjustment. How do all these approaches tally together, as a whole?</li><li>3) When introducing the e-value, there is no quantification of the potential range of values. What is a small e-value, what is large, etc? An immediate example would be welcome at this early stage of introducing the e-value.</li><li>4) The same applies to when B is introduced – what values could this take?</li><li>5) “This gives the investigators an easy tool to assess the strength of total unmeasured confounding needed to explain away the observed association” – going back to my earlier point about causal pathway and DAGs, surely there is more to our appreciation of the (causal) situation than just calculating the e-value to reveal what would be needed to explain away the association?</li><li>6) The use of the meta-analysis example is helpful, but the meta-analysis angle makes this more complicated than considered previously in the article, as we now have the issue of potential heterogeneity across studies. Looking at the plot, the estimates seem reasonable consistent with each other – but can this be quantified by an estimate of the between-study variance (<math>\tau^2</math>)?</li><li>7) Another complication of meta-analysis is regards to what adjustment factors were used in each study, and whether they were</li></ol>
-------------------------	--

	<p>a consistent set, or if they differed across studies. If the latter, then how can the e-value be interpreted?</p> <p>8) How can the e-value method be applied when using odds ratios or hazard ratios in observational research? Do these need to be transformed to risk ratios first? Some explanation would be welcome.</p> <p>In summary, as a methods primer article this will be a useful addition to those new to the e-value method, but there is a need for clarification in a number of places to help improve the article for the BMJ Medicine reader, especially in regards to the causal pathway context. I hope these comments are helpful to the authors moving forwards.</p>
--	--

<b>REVIEWER 2</b>	Davies, Neil; University of Bristol. Competing Interest: None
<b>REVIEW RETURNED</b>	University of Bristol

<b>GENERAL COMMENTS</b>	<p>I enjoyed reading this paper explaining the use of the E-value. It explains the concepts clearly, and I think it will be useful for applied researchers. The authors use a nice applied example, anti-depressants in pregnancy to illustrate the use of the method and the paper is pitched at the right level.</p> <p>It might be helpful to explain why the E-value makes no assumption about the number of unmeasured confounders, and also to provide some explanation of how the E-value can be interpreted as either requiring strong or weak confounding and hence the plausibility of confounding explaining the observed associations.</p> <p>Specific comments</p> <p>Page 6 - "In fact, the E-value estimates the overall strength of potential unmeasured confounding rather than the impact of individual confounding factors." This isn't obvious to me why this should be the case. How can the E-value indicate the total strength of confounding needed from all confounders?</p> <p>Page 7 - the bounding factor - this is all discussed in the context of a singular confounder - does the same apply as above? Is the bounding factor indicating the required confoundign for all confounders?</p> <p>Page 7 - what data types can the bounding factor be used for? Is it only valid for binary variables as per your examples, or can it also be used for continuous variables?</p> <p>Page 8 - "Since the E-value is relatively small, it is likely that unmeasured confounding could be of a sufficient magnitude to explain the apparent effect." Can you provide any guidance or indication about what should be considered a small or big E-value?</p> <p>Page 9 line 16 - would it be worth calculating/reporting the bounding factor here? It might also be worth clarifying the relationship between the bounding factor and the E-value.</p> <p>Page 10 line 35 - something strange has happened to the decimal place formatting here.</p> <p>Page 11 - "However, information about the distribution of the</p>
-------------------------	--

	unmeasured confounder can be useful when assessing how likely an unmeasured confounder is.” But isn’t the problem here that we don’t know the distribution or relationships with the unmeasured confounders? Playing devil’s advocate - the E-value just tells us that if an association is strong it requires a strong unmeasured confounder, if an association is weak, it requires a weak unmeasured confounder. It doesn’t tell us anything about the likelihood or plausibility of this form of confounding.
--	---

<b>REVIEWER 3</b>	Ho, Frederick. Competing Interest: None
<b>REVIEW RETURNED</b>	25-Oct-2022

<b>GENERAL COMMENTS</b>	<p>This is a Methods Primer paper focus around the use of E-values in observational studies. I agree this is a method that is underutilised and read the paper with interest. I have a few comments:</p> <ol style="list-style-type: none"> <li>1. I wonder if E-value is more around an analysis around *residual* confounding rather than *unmeasured* confounding. The former could be stemmed from improper handling of confounders in analysis (e.g. by ignoring the true functional relationship) as well as unmeasured confounding.</li> <li>2. Some parts of the paper feels a bit too technical to general readers and some restructuring might help. E.g. when introducing the joint bounding factor B, it might be easier to follow if the rationale, meaning, and use of B at the beginning. Without that the readers might be confused why an additional quantity was introduced.</li> <li>3. The use of example potential confounding variables (tobacco and alcohol) in contextualising the E-values are excellent. I actually think this is an important (yet underused) step in interpreting the E-values. However, on the previous page (p7/11), the last line already stated 'Since the E-value is relatively small, it is likely that unmeasured confounding could be of a sufficient magnitude to explain the apparent effect.' seemed to have jumped to the conclusion without explaining. I think it might be more justified to say we'd need to examine the strengths of potential confounders before actually assessing whether that number is small.</li> <li>4. Sometimes ',' was used as decimal point i think those should be consistent (e.g. on p9/11)</li> <li>5. I wonder if the summary points and the conclusion could be phrased differently. I'm glad that the nuances in E-values and its utilisation were pointed out. But at those areas it might be more useful for the readers to get a 'take-home message' in how to use the E-values? e.g. in one summary point you could outline how E-values could be used and interpreted?</li> </ol>
-------------------------	--

### VERSION 1 – AUTHOR RESPONSE

Response to reviewer 1:

Comments to the Author: This is a very well-written piece explaining the use of E-value method for examining the potential impact of unmeasured confounding in observational research. The paper

would be best as a methods primer, about the e-value method, rather than focusing as much on the case study. In terms of the methods, I have a number of suggestions for improvement and clarification  
Response: Thank you very much for your time and your helpful comments on this manuscript. We read them with interest and addressed them one by one.

1) Given the focus on the e-value method, I think this should be mentioned in the title, and the paper considered as a methods primer for the journal.

Response: We revised the title as follows: Quantifying the impact of unmeasured confounding in observational studies using the E-value

2) My strongest comment is to explain how this method should relate to causal inference methods that base inferences on (assumed) causal pathways and DAGs. At the moment, this paper does not consider the causal pathway premise, but rather focuses on regression adjustment. How do all these approaches tally together, as a whole?

Response: To address this comment, we explicitly connected absence of baseline confounding with a causal interpretation of an association (page 2, Introduction). In addition we explicitly referred to Figure 1 as representing DAGs, with appropriate citations.

3) When introducing the e-value, there is no quantification of the potential range of values. What is a small e-value, what is large, etc? An immediate example would be welcome at this early stage of introducing the e-value.

Response: Thank you for pointing this out. We have added a sentence to discuss this matter in the section of the E-value method: "The E-value does not have a specific range, and whether it is considered large or small depends on the given exposure and outcome and the amount of controlled confounding" (page 4).

4) The same applies to when B is introduced – what values could this take?

Response: We have now added the following comment on page 4: "The joint bounding factor B could take an infinite number of different values depending on RRUD and RREU"

5) "This gives the investigators an easy tool to assess the strength of total unmeasured confounding needed to explain away the observed association" – going back to my earlier point about causal pathway and DAGs, surely there is more to our appreciation of the (causal) situation than just calculating the e-value to reveal what would be needed to explain away the association?

Response: Please see response to an earlier comment.

6) The use of the meta-analysis example is helpful, but the meta-analysis angle makes this more complicated than considered previously in the article, as we now have the issue of potential heterogeneity across studies. Looking at the plot, the estimates seem reasonable consistent with each other – but can this be quantified by an estimate of the between-study variance (tau squared)?

Response: Following the suggestion, we computed the tau, Q and I squared measures, which made clear that we have one "outlier" study, which is that of Johansen et al. Redoing all analyses after omission of this study do not seem to affect the calculated RR substantially. We added the following sentence to the figure 3 legend: "Computing measures of heterogeneity ( $Tau^2=0,039$  and  $Q=223,28$ ) showed low heterogeneity apart from the study by Johansen et. However, leaving out this study did not change the overall estimate of RR substantially."

7) Another complication of meta-analysis is regards to what adjustment factors were used in each study, and whether they were a consistent set, or if they differed across studies. If the latter, then how can the e-value be interpreted?

Response: We added the following clarification on page 6 "The E-value can similarly be applied to results of any individual study."

8) How can the e-value method be applied when using odds ratios or hazard ratios in observational research? Do these need to be transformed to risk ratios first? Some explanation would be welcome.

Response: The method works just the same for OR and HRs, and they should not be transformed to RRs. We added the following sentence: "A similar approach is possible for odds ratios and hazard ratios" on page 4

In summary, as a methods primer article this will be a useful addition to those new to the e-value method, but there is a need for clarification in a number of places to help improve the article for the

BMJ Medicine reader, especially in regards to the causal pathway context. I hope these comments are helpful to the authors moving forwards.

Best wishes, Prof Richard Riley  
Chief Statistics Editor for BMJ Medicine

Response to reviewer 2:

Comments to the Author: I enjoyed reading this paper explaining the use of the E-value. It explains the concepts clearly, and I think it will be useful for applied researchers. The authors use a nice applied example, anti-depressants in pregnancy to illustrate the use of the method and the paper is pitched at the right level.

It might be helpful to explain why the E-value makes no assumption about the number of unmeasured confounders, and also to provide some explanation of how the E-value can be interpreted as either requiring strong or weak confounding and hence the plausibility of confounding explaining the observed associations.

Response: Thank you very much for these comments. We edited the text of the Discussion, second paragraph, page 6, to explain the distinction between confounding and confounders.

Specific comments

Page 6 - "In fact, the E-value estimates the overall strength of potential unmeasured confounding rather than the impact of individual confounding factors." This isn't obvious to me why this should be the case. How can the E-value indicate the total strength of confounding needed from all confounders?

Response: Since E-value is calculated without any assumption about the nature or the number of unmeasured confounders, it estimates the amount of potential confounding needed to explain away an association. Whether or not such confounders exist is a matter of plausibility and subject matter knowledge. We added an explanation on page 4, following the quoted passage.

Page 7 - the bounding factor - this is all discussed in the context of a single confounder - does the same apply as above? Is the bounding factor indicating the required confounding for all confounders?

Response: In the same paragraph on page 4, we clarified: "B does not require assumptions about the structure of the unmeasured confounding."

Page 7 - what data types can the bounding factor be used for? Is it only valid for binary variables as per your examples, or can it also be used for continuous variables?

Response: Explained as above. Bounding factor requires no assumptions about confounding structure.

Page 8 - "Since the E-value is relatively small, it is likely that unmeasured confounding could be of a sufficient magnitude to explain the apparent effect." Can you provide any guidance or indication about what should be considered a small or big E-value?

Response: Thank you for pointing this out. We have added a sentence to discuss this matter in the section of the E-value method (page 4): "The E-value does not have a specific range, and whether it is considered large or small depends on the given exposure and outcome."

Page 9 line 16 - would it be worth calculating/reporting the bounding factor here? It might also be worth clarifying the relationship between the bounding factor and the E-value.

Response: If the joint bounding factor should be able to explain away the association it should be at least equal to the RR. Thus, the bounding factor is already reported in the example (1.41). The relationship between the bounding factor and the E-value already clarified in section explaining the E-value method (page 3-4).

Page 10 line 35 - something strange has happened to the decimal place formatting here.

Response: Thank you, we corrected this mistake.

Page 11 - "However, information about the distribution of the unmeasured confounder can be useful when assessing how likely an unmeasured confounder is." But isn't the problem here that we don't know the distribution or relationships with the unmeasured confounders? Playing devil's advocate - the E-value just tells us that if an association is strong it requires a strong unmeasured confounder, if an association is weak, it requires a weak unmeasured confounder. It doesn't tell us anything about

the likelihood or plausibility of this form of confounding.

Response: The reference to the distribution of confounder refers to the prevalence of a confounder. A strong confounder that is rare is not likely to explain away an association, as explained in the Discussion section, page xx.

Response to reviewer 3:

Comments to the Author: This is a very helpful, clearly written article that explains how the E value can be used for sensitivity analyses in observational studies. I have a few minor comments:

Response: Thank you for the favourable assessment of our work.

Page 4: Could some brief examples or further explanations be given of external adjustment, instrumental variables, positive/negative controls?

Response: We added a brief description for each method in the second paragraph of the Introduction, page 3. We also cited self-controlled designs.

Where you describe the equation for the E value on page 4, perhaps it would be helpful to state that because of the way it is calculated, larger relative risks always have larger E values, and therefore that the strength of unmeasured confounding required to explain away an effect has to be greater for larger effect sizes (which is intuitive).

Response: We agree. We revised the sentence after the equation (page 4):

“Consequently, a strong association would have a large E-value, which suggests that the unmeasured confounding must be strongly associated with both the exposure and outcome to fully explain the association.”

Has the meta-analysis that is used as an example been published? It would be helpful to have some more details about the methods for this and for the search, e.g. a protocol or some other description of methods (I realise this is not central to this article)

Response: Thank you for your question. The meta-analysis has not been published since a recent similar meta-analysis regarding antidepressants during pregnancy and risks for adverse perinatal outcomes was published in 2020 by Xing et al. Due to word count restriction we will therefore not include methods for the meta-analysis, but we have added the methods in an appendix.

Page 7: You state that the E value is relatively small. How do you determine what small is in this context? A RR of 2.2 would not always be regarded as small!

Response: Thank you for your comment. We agree that a RR of 2.2 is not always small and have therefore removed the sentence.

Why did some of the studies dismiss that alcohol and tobacco are likely to bias the result

Response: We have removed the sentence to avoid confusion. The intention was to highlight that some of the studies do not have access to data on smoking and alcohol and thereby are likely to bias the result.

Did any of the studies include alcohol in their estimation of the RR or are these all unadjusted RRs?

Response: The majority adjusted for alcohol, but they did not consider the residual confounding.

Page 10: You state that a low E-value does not necessarily mean that an unmeasured confounder could fully explain an association. Should this be a "single" unmeasured confounder? Could there be some description of how the E value can be interpreted when there are multiple unmeasured confounders?

Response: We clarified the distinction between confounding and confounders in the second paragraph of the Discussion, page 9.

Could you expand more on why "even if an unmeasured confounder matches the requirements of the E-value it does not automatically mean that it can explain the effect"?

Response: Yes, e.g. if the unmeasured confounder is rare, it is unlikely that it can fully explain the effect even if it matches the requirements of the E-value. We revised the sentence: “Some unmeasured confounders (e.g., those with low prevalence) that fulfil the requirements of the E-value may not explain away the observed effect. Thus, the prevalence of a given confounder should be considered carefully.” on page 7.

Could you discuss whether the E value can be used in the context where there is a null or spurious

effect which you think is due to confounding by indication? i.e. there is a true effect but this can't be identified due to unmeasured confounders?

Response: We added the following caveat in the last paragraph of the Discussion, on page x. "Moreover, the E-value addresses confounding that overestimates an association, and cannot address confounding that masks a true one."

Response to reviewer 4, dr. Frederick Ho

Comments to the Author: This is a Methods Primer paper focus around the use of E-values in observational studies. I agree this is a method that is underutilised and read the paper with interest. I have a few comments:

Response: Thank you for your time and your comments.

1. I wonder if E-value is more around an analysis around \*residual\* confounding rather than \*unmeasured\* confounding. The former could be stemmed from improper handling of confounders in analysis (e.g. by ignoring the true functional relationship) as well as unmeasured confounding.

Response: We clarified the terminology between residual and unmeasured confounding in the second paragraph of the Discussion, p. 9. Essentially, residual confounding is a form of unmeasured confounding.

2. Some parts of the paper feels a bit too technical to general readers and some restructuring might help. E.g. when introducing the joint bounding factor B, it might be easier to follow if the rationale, meaning, and use of B at the beginning. Without that the readers might be confused why an additional quantity was introduced.

Response: Thank you for your comment. We agree with you that it might be easier to follow if the joint bounding factor is introduced first. However, we believe the current layout focuses more on the E-value which is the main theme of the manuscript.

3. The use of example potential confounding variables (tobacco and alcohol) in contextualising the E-values are excellent. I actually think this is an important (yet underused) step in interpreting the E-values. However, on the previous page (p7/11), the last line already stated 'Since the E-value is relatively small, it is likely that unmeasured confounding could be of a sufficient magnitude to explain the apparent effect.' seemed to have jumped to the conclusion without explaining. I think it might be more justified to say we'd need to examine the strengths of potential confounders before actually assessing whether that number is small.

Response: We understand your concern and we have removed the sentence to avoid confusion.

4. Sometimes ',' was used as decimal point i think those should be consistent (e.g. on p9/11)

Response: Revised accordingly.

5. I wonder if the summary points and the conclusion could be phrased differently. I'm glad that the nuances in E-values and its utilisation were pointed out. But at those areas it might be more useful for the readers to get a 'take-home message' in how to use the E-values? e.g. in one summary point you could outline how E-values could be used and interpreted?

Response: We added the following sentence to the key messages:

"We seek to raise awareness about the E-value and encourage researchers to adopt the E-value method in their standard toolbox as a consistent way of doing sensitivity analyses"

### VERSION 2 – REVIEW

<b>REVIEWER 1</b>	Riley, Richard; University of Birmingham, Institute of Applied Health Research. Competing Interest: None
<b>REVIEW RETURNED</b>	19-Dec-2022

<b>GENERAL COMMENTS</b>	<p>This is an excellent revision and response, thank you. I have some further questions that have arisen due to the new version and clarity it brought.</p> <p>1) When introducing the e-value, the authors refer to the 'unmeasured confounder' – would this not be better to say 'unmeasured confounder(s)'?</p>
-------------------------	--

	<p>2) “A similar approach applies to odds ratios and hazard ratios” – as the authors know, an effect could in principle be measured using any of RR, OR or HR. So, if the formula applies equally to any of these effect measures, then surely the e-value will change depending on what effect measure is chosen? This, to me, sounds alarming – can the authors address this issue please?</p> <p>3) “The E-value does not have a specific range, and whether it is considered large or small depends on the given exposure and outcome and the amount of controlled confounding” – some examples would be welcome here, as this is too vague at the moment and leaves the reader needing / wanting more.</p> <p>4) “RR_EU denotes the strength of association between the unmeasured confounder and the exposure” – in a model, an association effect size depends on the correlation with other factors included (adjusted for) in the model. So why can RR_EU be fixed, or even ‘known’ as mentioned. Is this the true causal effect of the unmeasured confounder? And would this be the same regardless of the main factor under study or other causal factors? Please clarify thank you.</p> <p>5) A related point is that the RR for other unmeasured confounders is (in the example and I assume typically) obtained from previous studies – but again these would adjust for particular variables themselves, and so what RR is needed for these confounders? An unadjusted RR? Or an adjusted version and if so, adjusted for what?</p> <p>6) Similarly, if there are multiple confounders, then is the RR-EU obtained by adding up or multiplying (perhaps on the log scale) the RR for each of the confounders, to get the RR_EU? Do they need to each be adjusted for the same other variables?</p> <p>I’m sure the authors can address these final queries, and I look forward to seeing their revision.</p>
--	---

<b>REVIEWER 2</b>	Davies, Neil; University of Bristol. Competing Interest: None
<b>REVIEW RETURNED</b>	04-Jan-2023

<b>GENERAL COMMENTS</b>	Thank you for addressing my comments, this paper will make a nice contribution to the literature.
-------------------------	---

<b>REVIEWER 3</b>	Harron, Katie. Computing Interest: None
<b>REVIEW RETURNED</b>	05-Jan-2023

<b>GENERAL COMMENTS</b>	I am satisfied with the authors' responses to my previous comments.
-------------------------	---

### VERSION 2 – AUTHOR RESPONSE

Response to reviewer 1:

Comments to the Author: This is an excellent revision and response, thank you. I have some further questions that have arisen due to the new version and clarity it brought.

1) When introducing the e-value, the authors refer to the ‘unmeasured confounder’ – would this not be better to say ‘unmeasured confounder(s)’?



Response: We agree and revised the sentence in the section of the E-value method as follows: “In a study with a given exposure and outcome, the E-value presents the minimum strength of association the unmeasured confounders should have with both the exposure and the outcome to explain away the observed association.”

2) “A similar approach applies to odds ratios and hazard ratios” – as the authors know, an effect could in principle be measured using any of RR, OR or HR. So, if the formula applies equally to any of these effect measures, then surely the e-value will change depending on what effect measure is chosen? This, to me, sounds alarming – can the authors address this issue please?

Response: We understand your concern. The calculation of the E-value using OR and HR depends on the outcome. For rare outcomes the formula (1) provided in the article can be used, yet for common outcomes the OR and HR should be modified e.g. by replacing RR in formula (1) with the square root of the OR. VanderWeele et al. provided guidelines for calculating the E-value using different effect measures.

We have revised the sentence in the section of the E-value method as follows: “The same approach applies to odds ratios (OR) and hazard ratios (HR) when the outcome is rare. For common outcomes, OR and HR must be transformed, e.g. by replacing RR in formula (1) with the square root of the OR.”

3) “The E-value does not have a specific range, and whether it is considered large or small depends on the given exposure and outcome and the amount of controlled confounding” – some examples would be welcome here, as this is too vague at the moment and leaves the reader needing / wanting more.

Response: We added an example of a large e-value in the section of the E-value method: “For example, when examining glucocorticoid use and risk of suicide among patients with cancer, the incidence rate ratio (IRR) was high (IRR=7.2) and there was a dose-response pattern with the highest cumulative dose associated with a 20-fold risk increase compared with non-use. In this example, the calculated E-value indicated that to explain away the association, a hypothetical confounder would need be associated with a 14-fold higher use of glucocorticoids and a 3-fold greater risk of suicide. The authors judged that such a confounder is not likely to exist given the confounding already adjusted in the analysis. Thus, labeling the E-value large or small depends on the subject matter knowledge, the strength of the observed association, and the amount of removed confounding.

4) “RR\_EU denotes the strength of association between the unmeasured confounder and the exposure” – in a model, an association effect size depends on the correlation with other factors included (adjusted for) in the model. So why can RR\_EU be fixed, or even ‘known’ as mentioned. Is this the true causal effect of the unmeasured confounder? And would this be the same regardless of the main factor under study or other causal factors? Please clarify thank you.

Response: Thank you for this important comment. We added the following to the discussion of E-value limitations, in the last paragraph of the paper: “Another important caveat, applicable to any setting, is that any estimate of RREU and RRUD used in computing the E-value is itself subject to a no-bias assumption.”

5) A related point is that the RR for other unmeasured confounders is (in the example and I assume typically) obtained from previous studies – but again these would adjust for particular variables themselves, and so what RR is needed for these confounders? An unadjusted RR? Or an adjusted version and if so, adjusted for what?

Response: Please see our response to the previous comment.

6) Similarly, if there are multiple confounders, then is the RR-EU obtained by adding up or multiplying (perhaps on the log scale) the RR for each of the confounders, to get the RR\_EU? Do they need to each be adjusted for the same other variables?

Response: We addressed this by adding the following sentence to the penultimate paragraph of the revised manuscript: "Although the E-value does not allow partitioning of confounding sources, it provides an estimate of joint unmeasured confounding, needed for causal inference."