Letter to the Editor

**Correction of ‘Bias factor, maximum bias and the E-value’**

Arvid Sjölander 1* and Tyler J VanderWeele 2

1 Department of Medical Epidemiology and Biostatistics, Karolinska Institute, Stockholm, Sweden and 2 Epidemiology and Biostatistics, Harvard School of Public Health, Boston, MA, USA

*Corresponding author. Karolinska Institute, Stockholm, Sweden. E-mail: arvid.sjolander@ki.se

We would like to correct an incorrect statement in the paper ‘Bias factor, maximum bias and the E-value: insight and extended applications’, recently published in the *International Journal of Epidemiology* by Cusson and Infante-Rivard. 1 In the box of Key Messages, the authors state:

‘As currently available, the E-value cannot be used with all possible direction combinations between the unmeasured confounder and both the exposure and the outcome; new formulae we propose make this possible.’

This is incorrect. The E-value, as it was originally proposed (i.e. ‘as currently available’), can in fact be used regardless of the direction of association between the unmeasured confounders and the exposure and the outcome.

Before tracing the source of this confusion, we recap the main results of Ding and VanderWeele, 2,3 who originally introduced the E-value. In Cusson and Infante-Rivard’s notation, Ding and VanderWeele defined the observed risk ratio as

\[
RR_{obs} = \frac{p(Y = 1|X = 1)}{p(Y = 1|X = 0)} ,
\]

and the ‘true’ (i.e. causal) risk ratio as

\[
RR_{true} = \frac{\sum_k p(Y = 1|X = 1, U = k)p(U = k)}{\sum_k p(Y = 1|X = 0, U = k)p(U = k)} ,
\]

where \( U \) is assumed to be sufficient for confounding control. Ding and VanderWeele further defined the sensitivity parameters as

\[
RR_{UX} = \max_k \left\{ \frac{p(U = k|X = 1)}{p(U = k|X = 0)} \right\} , \quad (1)
\]

and

\[
RR_{UY} = \max_x \left\{ \max_k p(Y = 1|X = x, U = k) \right\} \min_k p(Y = 1|X = x, U = k) , \quad (2)
\]

Finally, they defined the bias factor as

\[
B = \frac{RR_{UX} RR_{UY}}{RR_{UX} + RR_{UY} - 1} , \quad (3)
\]

which is identical to Cusson and Infante-Rivard’s Equation (5). Ding and VanderWeele proved that, if \( RR_{obs} \geq 1 \), then

\[
RR_{true} \geq \frac{RR_{obs}}{B} , \quad (4)
\]

and they defined the E-value as the common value \( RR_{UX} = RR_{UY} \) such that the lower bound in Equation (1) is equal to 1 or, alternatively and more formally, as:

\[
\min_{RR_{UX}, RR_{UY}; B \geq RR_{obs}} \{ \max(RR_{UX}, RR_{UY}) \}
\]

The E-value can then be shown to be equal to:\n
\[
E - value = RR_{obs} + \sqrt{RR_{obs}(RR_{obs} - 1)} , \quad (5)
\]

which is identical to Cusson and Infante-Rivard’s Equation (7’). Equations (4) and (5) do not make any assumptions
about, and are thus valid regardless of, the direction of association between \( U \) and \((X, Y)\).

The source of confusion appears to be that Cusson and Infante-Rivard have, contrary to Ding and VanderWeele, defined the sensitivity parameters \( RR_{XU} \) and \( RR_{UY} \) as

\[
RR_{XU} = \frac{p(U = 1|X = 1)}{p(U = 1|X = 0)}, \quad (6)
\]

and

\[
RR_{UY} = \frac{p(Y = 1|X = U = 1)}{p(Y = 1|X, U = 0)}, \quad (7)
\]

where \( U \) is assumed binary and \( RR_{UY} \) is assumed constant across levels of \( X \). Under these definitions, \( RR_{XU} \) and \( RR_{UY} \) are \( < 1 \) if the direction of association is negative between \( U \) and \( X \), and between \( U \) and \( Y \), respectively. In contrast, Equations (4) and (5) require \( RR_{XU} \) and \( RR_{UY} \) to be \( \geq 1 \) by definition. Under Cusson and Infante-Rivard’s redefinitions in Equations (6) and (7), negative associations between \( U \) and \((X, Y)\) do indeed invalidate Equations (4) and (5), which may explain Cusson and Infante-Rivard’s statement about the E-value in their Key Messages.

However, under the original definitions in Equations (1) and (2), \( RR_{XU} \) and \( RR_{UY} \) are always \( \geq 1 \), regardless of the direction of association between \( U \) and \((X, Y)\); see Sjölander\(^5\) for further discussion and proof. Thus, again, under the original definitions, Equations (4) and (5) are always valid, regardless of the direction of association between \( U \) and \((X, Y)\).

Cusson and Infante-Rivard are, of course, free to use any definitions they like. But it is incorrect then to assume that their conclusions hold with a different set of definitions as employed by the E-value. By failing to mention that they have redefined important quantities, they incorrectly make the E-value framework look more restrictive than it actually is. The problem is compounded by the fact that it is hard to tell from their text what definitions they have actually used. When introducing the parameters \( RR_{XU} \) and \( RR_{UY} \) in the ‘Introduction’ section, Cusson and Infante-Rivard only define them verbally, as ‘the risk ratio between exposure \( X \) and \( U \)’ and ‘the risk ratio between \( U \) and the outcome \( Y \)’. Much later [below their Equation (3)], they briefly mention that \( RR_{XU} \) is ‘defined as \( p(U|X = 1)/p(U|X = 0) \)’, but without specifying the value of \( U \) in this expression. We could not find a formal definition of \( RR_{UY} \) anywhere in the paper. This omission of important details makes the paper hard to properly interpret, and invites misunderstandings and, in the end, faulty conclusions concerning the E-value.

Finally, we wish to note that Cusson and Infante-Rivard’s alternative formulae for negative associations [their Equations (12), (13) and (14)] rely on the alternative definitions of \( RR_{XU} \) and \( RR_{UY} \) in Equations (6) and (7), and hence also on the assumption that \( U \) is binary and that \( RR_{XY} \) in Equation (7) does not depend on \( X \). These assumptions are fairly strong, and would often be violated in real scenarios, in which case Cusson and Infante-Rivard’s alternative formulae would not apply. The E-value approach, in contrast, does not require either of these assumptions.

**Conflict of interest**

The authors have no conflict of interest.

**References**

1. Cusson A, Infante-Rivard C. Bias factor, maximum bias and the E-value: insight and extended applications. *Int J Epidemiol* 2020; 49:1509–16.
2. Ding P, VanderWeele TJ. Sensitivity analysis without assumptions. *Epidemiology* 2016;27:368–77.
3. VanderWeele TJ, Ding P. Sensitivity analysis in observational research: introducing the E-value. *Ann Intern Med* 2017;167:268–74.
4. VanderWeele TJ, Ding P, Mathur M. Technical considerations in the use of the E-Value. *Journal of Causal Inference* 2019;7.
5. Sjölander A. A note on a recently proposed sensitivity analysis, and the related E-value. *J Causal Inference* 2020;8:229–48.