bmjmedicine

Quantifying the impact of unmeasured confounding in observational studies with the E value

Tobias Gaster,[1] Christine Marie Eggertsen,[1] Henrik Støvring ,[2,3] Vera Ehrenstein ,[4] Irene Petersen [4,5]

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/bmjmed-2022-000366).

[1]Aarhus University, Aarhus, Denmark
[2]Steno Diabetes Centre Aarhus, Aarhus, Denmark
[3]Clinical Pharmacology, Pharmacy and Environmental Medicine, University of Southern Denmark, Odense, Denmark
[4]Department of Clinical Epidemiology, Aarhus University and Aarhus University Hospital, Aarhus, Denmark
[5]Department of Primary Care and Population Health, University College London, London, UK
Correspondence to: Dr Irene Petersen, Department of Clinical Epidemiology, Aarhus University, 8000 Aarhus, Denmark; i.petersen@ucl.ac.uk

Cite this as: *BMJMED* 2023;2:e000366. doi:10.1136/bmjmed-2022-000366

TG and CME contributed equally.

Received: 12 September 2022
Accepted: 21 March 2023

ABSTRACT

The E value method deals with unmeasured confounding, a key source of bias in observational studies. The E value method is described and its use is shown in a worked example of a meta-analysis examining the association between the use of antidepressants in pregnancy and the risk of miscarriage.

KEY MESSAGES

- ⇒ Unmeasured or poorly measured confounding is a major challenge in observational studies, especially those based on routinely collected data
- ⇒ Although the E value method is easy to implement and does not require assumptions to assess the minimum strength of unmeasured confounding needed to explain an association, its use in interpreting the results of observational studiesis is not widespread
- ⇒ The use of E value for interpreting the results of studies examining the association between the use of antidepressants in pregnancy and risk of miscarriage is illustrated by examining the distribution of an unmeasured confounder that might fully and plausibly explain an observed association
- ⇒ This article aims to raise awareness of the E value and encourage researchers to adopt the E value method in their standard toolbox as a consistent way of performing sensitivity analyses

Introduction

Randomised controlled trials are the gold standard of causal inference of the benefits and risks of drug treatments. Successful randomisation removes both measured and unmeasured baseline confounding, enabling causal interpretation of an observed association.[1] Randomised controlled trials, however, are not practically or ethically suitable for many types of research questions, including rare risks of treatments. Notably, few randomised controlled trials include pregnant women. Instead, the primary source of information on the safety of treatments in pregnancy is observational studies, relying on routinely collected (secondary) data from registries and other electronic databases.[2 3] Observational studies based on secondary data allow adjustment of the analysis for the confounders measured in a particular database, whereby residual (unmeasured) confounding from poorly measured, unmeasured, or unknown characteristics might remain.

Approaches to removal or quantification of unmeasured confounding include external adjustment (correction of findings by using external information); use of instrumental variables (emulating randomisation); self-controlled designs (removal of time independent confounding); or negative and positive controls (benchmarking findings against established associations).[4–7] Another option to deal with unmeasured confounding in observational research is to use the E value, a readily calculated and easily interpreted statistical tool that assesses the minimum strength of potential unmeasured confounding needed to explain away an effect. The E value was introduced in 2017 by VanderWeele and Ding.[8] In common with many new epidemiological tools, integrating the E value into applied health research has taken time, and many researchers might still be unfamiliar with the method. In this paper, we want to raise awareness about the E value and encourage researchers to include this tool among their standard methods in sensitivity analyses. We first describe the E value method and then we show its use in a worked example of a meta-analysis combining the evidence for the association between the use of antidepressants in the first trimester of pregnancy and the risk of miscarriage.

The E value method

In a study with a specific exposure and outcome, the E value is the minimum strength of association that the unmeasured confounders should have with both the exposure and the outcome to explain away the observed association.[8] The exposure refers to any characteristic that may explain or predict the presence of a study outcome.[9] For a particular observed exposure-outcome risk ratio (RR), the E value can be calculated based on formula (1), involving only the estimate itself.[10]

$$\text{E value} = \text{RR} + \sqrt{\text{RR} \times (\text{RR} - 1)} \quad (1)$$

The same approach applies to odds ratios and hazard ratios if the outcome is rare. If the outcome is common, the odds ratio and hazard ratio must be modified (eg, by replacing RR in formula (1) with the square root of the odds ratios).[8]

A strong association has 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. In contrast, a small E value suggests that weak unmeasured confounding would be enough to explain the association. A lower E value might indicate that confounding rather than causality is a more plausible explanation than a higher E value. The E value does not have a specific range, and whether its value is considered large or small depends on the particular

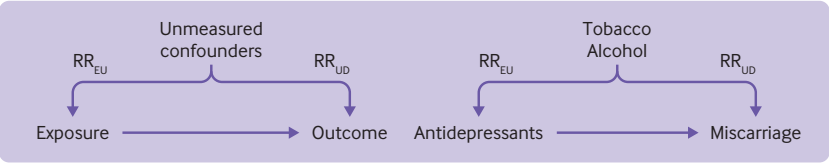

Figure 1 | Causal directed acyclic graph showing the direction of hypothesised causal effects.[23] Left: Directed acyclic graph for a generic association between unmeasured confounding, exposure, and outcome. Right: Directed acyclic graph for tobacco or alcohol as possible unmeasured confounders when examining the association between the use of antidepressants and risk of miscarriage. exRR$_{EU}$=strength of association between the unmeasured confounder and exposure; RR$_{UD}$=strength of association between the unmeasured confounder and outcome

exposure and outcome and the amount of controlled confounding.[11] For example, when examining the association between the use of glucocorticoids and the risk of suicide among patients with cancer, the incidence rate ratio was high (7.2) and the dose-response pattern showed that the highest cumulative dose was associated with a 20-fold increase in risk 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 threefold greater risk of suicide. The authors judged that such a confounder is not likely to exist given the confounding already adjusted for in the analysis.[12] Thus labelling the E value large or small depends on knowledge of the subject matter, the strength of the observed association, and amount of confounding removed.

The E value can be calculated without any assumptions regarding unmeasured confounders, such as being binary or consisting of only one unmeasured confounding factor.[10] The E value estimates the overall strength of potential unmeasured confounding rather than the effect of individual confounding factors. With this information, investigators can assess whether one or several specific unmeasured confounders could plausibly explain away the observed association in a particular study.

In addition to the E value, Ding and VanderWeele introduced the joint bounding factor, B (formula 2).[10]

$$B = \frac{RR_{UD} \times RR_{EU}}{(RR_{EU} + RR_{UD} - 1)} \quad (2)$$

RR$_{EU}$ denotes the strength of association between the unmeasured confounder and the exposure. B does not require assumptions about the structure of the unmeasured confounding.[10] RR$_{UD}$ denotes the strength of association between the unmeasured confounder and the outcome, as illustrated in a directed acyclic graph representing a simplified form of the confounding structure (figure 1).

The joint bounding factor, B, could take an infinite number of values depending on RR$_{UD}$ and RR$_{EU}$. If the joint bounding factor B is set to equal the observed risk ratio, the joint bounding factor describes the different combinations of RR$_{UD}$ and RR$_{EU}$ that would have the joint minimum strength to explain away the association. The E value=B when RR$_{UD}$=RR$_{EU}$, which is one possible combination of RR$_{UD}$ and RR$_{EU}$. A range of different possible combinations exist, which are often illustrative for the setting studied (figure 2).

For example, if RR$_{EU}$ is known, the joint bounding factor can be used to estimate the minimum strength of association between the unmeasured confounder and the outcome (RR$_{UD}$) needed to explain away the association.[8] In other settings, information might be available on the magnitude of the association between the strongest unobserved confounder and the outcome (RR$_{UD}$). The joint bounding factor can similarly be used to estimate RR$_{EU}$.

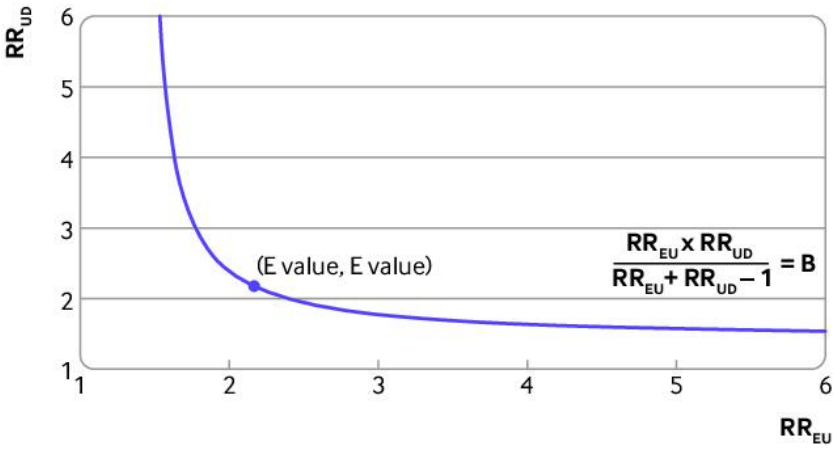

Figure 2 | Illustration of different combinations of RR$_{UD}$ and RR$_{EU}$ for a possible joint bounding factor, B, of 1.41. RR$_{EU}$=strength of association between the unmeasured confounder and exposure; RR$_{UD}$=strength of association between the unmeasured confounder and outcome

Example: meta-analysis of the association between use of antidepressants in pregnancy and risk of miscarriage

One in five women have depression during pregnancy, and about 13% of pregnant women take antidepressants.[13 14] Because both untreated depression and drug treatments might adversely affect pregnancy, treatment of depression in pregnant women involves balancing the benefits of treated depression against the potential treatment related risks to the mother and unborn child. Information about the safety of the use of antidepressants is therefore important for the healthcare professional when making these decisions.

We conducted a meta-analysis combining the evidence for the association between the use of antidepressants in the first trimester of pregnancy and the risk of miscarriage. Eligible studies were identified by a search in the PubMed and Embase databases from 2000 to February 2021, based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (online supplemental appendix).[15] The combined effect

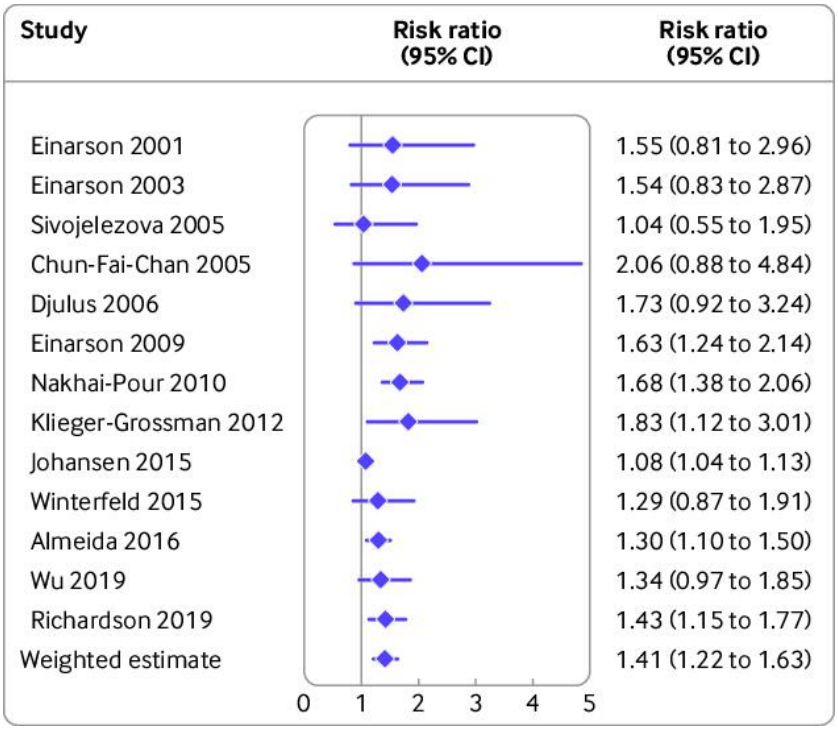

Figure 3 | Forest plot showing risk ratio and 95% confidence interval (CI) for each study included in the meta-analysis. A random effects model was used. Computing measures of heterogeneity (τ^2=0.039 and Q=223.28) showed low heterogeneity apart from the study by Johansen, but removing this study did not change the overall estimate of risk ratio substantially

was estimated with the random effects model and Episheet software (figure 3).[16]

The weighted estimate for the risk ratio of miscarriage in those who received treatment with antidepressants was 1.41 (95% confidence interval 1.22 to 1.63). To evaluate if this estimate is potentially biased by unmeasured confounding, we applied the E value, which we estimated at 2.17. Thus the unmeasured confounders must be associated with both use of antidepressants and risk of miscarriage by a risk ratio of at least 2.17 to fully explain away the observed risk ratio of 1.41. Tobacco and alcohol are prevalent and plausible confounders associated with both miscarriage and depression and hence with the use of antidepressants (figure 1).[17–20] We used these confounders in our application of the E value method. Petersen et al estimated that the risk of substantial (>35 units/week) alcohol use and use of antidepressants was RR_{EU}=10.25.[18] Sundermann et al found that alcohol use (>35 units/week) increased the risk of miscarriage (RR_{UD}=3.1).[17] Because both estimates are stronger than the calculated E value, substantial alcohol use could explain the observed association between the use of antidepressants and risk of miscarriage.

Based on data from Johansen et al, we estimated that the association between smoking and use of antidepressants was RR_{EU}=2.06.[21] Because the risk ratio was <2.17 (the E value), smoking could not fully explain the association when applying the E value method. By using the joint bounding factor, however, smoking could still explain the observed association if the association between smoking and miscarriage is sufficient. Figure 4 shows the different combinations of RR_{EU} and RR_{UD} that jointly would have the minimum strength required to explain the association.

Smoking had an RR_{EU} value of 2.06, which requires RR_{UD}=2.3 to fully explain the association. Pineles et al found an increased risk of miscarriage associated with smoking of 1.32.[19] Because this value is below the required RR_{UD} of 2.3, smoking is unlikely to explain the observed risk ratio of 1.41 on its own. The E value can similarly be applied to the results of any individual study.

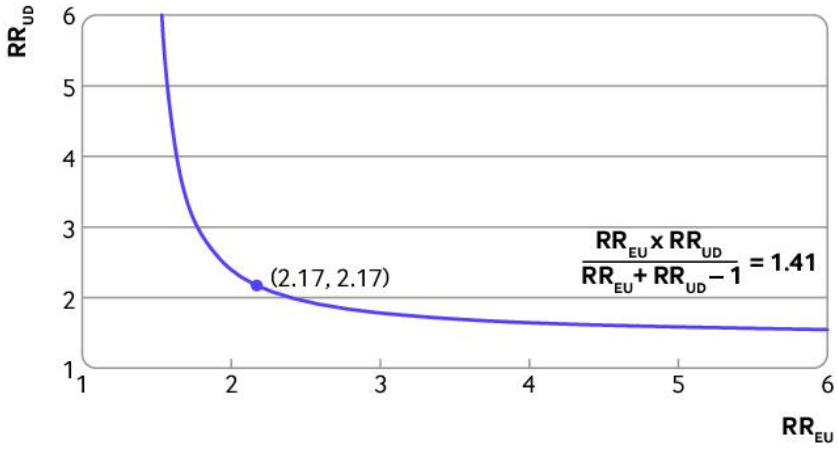

Figure 4 | Different combinations of RR_{EU} and RR_{UD} that jointly would have the minimum strength required to explain the observed association in the meta-analysis. RR_{EU}=strength of association between the unmeasured confounder and exposure; RR_{UD}=strength of association between the unmeasured confounder and outcome

Discussion

Confounded estimates could make safe treatments seem unsafe and might potentially limit treatment options for pregnant women. Equally problematic is that treatments might seem safe when they are not, hindering the use of safer alternatives. In our example, the weighted estimate from the meta-analysis suggested that the use of antidepressants in pregnancy increased the risk of miscarriage. Hence the treatment would be considered unsafe to use during pregnancy. The use of the E value method, however, suggested that the association could be explained away by unobserved confounding, plausibly attributable to alcohol consumption. In other words, an analysis incorporating adequately measured alcohol consumption could give a null finding. The E value method also showed that confounding by smoking could not explain the observed association. Tools helping to understand the potential role of confounding are

therefore important when analysing observational studies.

The ultimate goal of causal inference is removal of confounding. In a particular dataset, total confounding might come from a few or many confounders. Unmeasured confounding is a combination of unknown and residual confounding.[22] The E value is a relatively new addition to the strategies for assessing unmeasured confounding in observational studies. Although the E value does not allow partitioning of confounding sources, it provides an estimate of the joint unmeasured confounding, needed for causal inference. Another advantage of the E value is that no assumptions are needed about the structure of the unmeasured confounding, and it is simple to calculate from the risk ratio. The E value is therefore easy to implement as a sensitivity analysis in observational studies. Knowledge of the subject matter should help identify whether any specific unmeasured confounders can plausibly cause the potential confounding.

The E value has limitations. A low E value does not always mean that an unmeasured confounder (eg, the smoking example) could fully explain the association but possibly only to some extent. Similarly, a high E value does not always rule out unmeasured confounding. For example, confounding by indication might strongly confound an association even when the E value is high. Some unmeasured confounders (eg, those with a low prevalence) that fulfil the requirements of the E value might not explain the observed effect. Thus the prevalence of a particular confounder should be considered carefully. Moreover, the E value deals with confounding that inflates the magnitude of an association and cannot help evaluate confounding that masks a true association.[22] Another important caveat, applicable to any setting, is that any estimate of RR_{EU} and RR_{UD} used in computing the E value is itself subject to an assumption of no bias. Finally, confounding is not the only source of bias in epidemiological studies. Therefore, the results of an E value analysis should be part of a series of sensitivity analyses addressing all threats to validity.

Twitter Christine Marie Eggertsen @Chris_Eggertsen, Henrik Støvring @hstovring, Vera Ehrenstein @EhrensteinVera and Irene Petersen @i_petersen

Contributors CME and TG conducted the literature search, carried out the statistical analysis, and jointly wrote this paper as part of their bachelor thesis. HS, VE, and IP supervised the project and the writing. HS is a prominent methodologist and senior statistician. VE and IP are experts in reproductive pharmacoepidemiology. IP is an expert in methodology, has studied the effects of alcohol use and the association with antidepressant use in pregnancy, and has previously contributed to a similar BMJ paper on study design. VE is the guarantor. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial, or not-for-profit sectors.

Competing interests All authors have completed the ICMJE uniform disclosure form at www.icmje.org/disclosure-of-interest/ and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data sharing not applicable as no datasets generated and/or analysed for this study.

ORCID iDs

Henrik Støvring http://orcid.org/0000-0002-5821-2351
Vera Ehrenstein http://orcid.org/0000-0002-3415-3254
Irene Petersen http://orcid.org/0000-0002-0037-7524

REFERENCES

1 Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. Am J Epidemiol 2016;183:758–64. 10.1093/aje/kwv254
2 Thurin NH, Pajouheshnia R, Roberto G, *et al*. From inception to conception: genesis of a network to support better monitoring and communication of medication safety during pregnancy and breastfeeding. Clin Pharmacol Ther 2022;111:321–31. 10.1002/cpt.2476
3 Huybrechts KF, Bateman BT, Hernández-Díaz S. Use of real-world evidence from healthcare utilization data to evaluate drug safety during pregnancy. Pharmacoepidemiol Drug Saf 2019;28:906–22. 10.1002/pds.4789
4 Schneeweiss S. Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics. Pharmacoepidemiol Drug Saf 2006;15:291–303. 10.1002/pds.1200
5 Dekkers OM. On causation in therapeutic research: observational studies, randomised experiments and instrumental variable analysis. Prev Med 2011;53:239–41. 10.1016/j.ypmed.2011.08.003
6 Lipsitch M, Tchetgen Tchetgen E, Cohen T. Negative controls: a tool for detecting confounding and bias in observational studies. Epidemiology 2010;21:383–8. 10.1097/EDE.0b013e3181d61eeb
7 Petersen I, Douglas I, Whitaker H. Self controlled case series methods: an alternative to standard epidemiological study designs. BMJ 2016;354:i4515. 10.1136/bmj.i4515
8 VanderWeele TJ, Ding P. Sensitivity analysis in observational research: introducing the E-value. Ann Intern Med 2017;167:268–74. 10.7326/M16-2607
9 Kestenbaum B. Epidemiology and biostatistics. 2019. 10.1007/978-3-319-97433-0_2 Available: http://link.springer.com/10.1007/978-3-319-97433-0_2
10 Ding P, VanderWeele TJ. Sensitivity analysis without assumptions. Epidemiology 2016;27:368–77. 10.1097/EDE.0000000000000457
11 VanderWeele TJ, Mathur MB. Commentary: developing best-practice guidelines for the reporting of E-values. Int J Epidemiol 2020;49:1495–7. 10.1093/ije/dyaa094
12 Laugesen K, Farkas DK, Vestergaard M, *et al*. Glucocorticoid use and risk of suicide: a Danish population-based case-control study. World Psychiatry 2021;20:142–3. 10.1002/wps.20831
13 Ajinkya S, Jadhav PR, Srivastava NN. Depression during pregnancy: prevalence and obstetric risk factors among pregnant women attending a tertiary care hospital in Navi Mumbai. Ind Psychiatry J 2013;22:37–40. 10.4103/0972-6748.123615

14 Cooper WO, Willy ME, Pont SJ, *et al*. Increasing use of antidepressants in pregnancy. Am J Obstet Gynecol 2007;196:544. 10.1016/j.ajog.2007.01.033
15 Page MJ, McKenzie JE, Bossuyt PM, *et al*. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. 10.1136/bmj.n71
16 Rothman K. Episheet. 2015. Available: https://www.rtihs.org/sites/default/files/2022-04/Episheet.xls
17 Sundermann AC, Zhao S, Young CL, *et al*. Alcohol use in pregnancy and miscarriage: a systematic review and meta-analysis. Alcohol Clin Exp Res 2019;43:1606–16. 10.1111/acer.14124
18 Petersen I, Evans SJ, Gilbert R, *et al*. Selective serotonin reuptake inhibitors and congenital heart anomalies: comparative cohort studies of women treated before and during pregnancy and their children. J Clin Psychiatry 2016;77:e36–42. 10.4088/JCP.14m09241
19 Pineles BL, Park E, Samet JM. Systematic review and meta-analysis of miscarriage and maternal exposure to tobacco smoke during pregnancy. Am J Epidemiol 2014;179:807–23. 10.1093/aje/kwt334
20 Zuckerman B, Amaro H, Bauchner H, *et al*. Depressive symptoms during pregnancy: relationship to poor health behaviors. Am J Obstet Gynecol 1989;160(5 Pt 1):1107–11. 10.1016/0002-9378(89)90170-1
21 Johansen RLR, Mortensen LH, Andersen A-MN, *et al*. Maternal use of selective serotonin reuptake inhibitors and risk of miscarriage-assessing potential biases. Paediatr Perinat Epidemiol 2015;29:72–81. 10.1111/ppe.12160
22 Freemantle N, Marston L, Walters K, *et al*. Making inferences on treatment effects from real world data: propensity scores, confounding by indication, and other perils for the unwary in observational research. BMJ 2013;347:bmj.f6409. 10.1136/bmj.f6409
23 Lipsky AM, Greenland S. Causal directed acyclic graphs. JAMA 2022;327:1083–4. 10.1001/jama.2022.1816

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/bmjmed-2022-000366).

Methods for meta-analysis

1. Study eligibility criteria

To get an overview of which studies should be included in the meta-analysis we defined the inclusion criteria by defining population, exposure, comparator, and outcome (PECO) [1]. Non-original research e.g., reviews, replies and meta-analyses were excluded.

In the *population* each pregnancy must only be included once in the meta-analyses to avoid bias of the weighted estimate and therefore the population in each study must be carefully considered. Especially information regarding the country and the time of which the population are recruited must be obtained and compared between the studies to avoid overlapping cohorts.

The *exposure* included any antidepressants in the first trimester. The risk of spontaneous abortion differs greatly within gestational age and therefore it is necessary to define the exposure window when examining studies of pregnant women [2]. Spontaneous abortion is defined as loss before the 22nd week of pregnancy, and in most cases it occurs before the 12th week [3]. Consequently the etiologically relevant window in the included studies must be between the conception and the 22nd week of pregnancy in case the pregnancy should be considered exposed to antidepressants. We therefore included studies that obtained information on antidepressant exposure in the 1st trimester.

The *comparator* group must be pregnancies that are not exposed to the antidepressant during the etiologically defined window. There are several options when defining this group. It could consist of unexposed pregnancy (regardless of depression diagnosis), pregnancies exposed to nonteratogens or pregnancies exposed to other antidepressants than the antidepressant concerned, a so-called active comparator group. Pregnant women with a diagnosis of untreated depression could be used to reduce confounding by indication and thereby increase the internal validity [2]. All comparison groups were considered relevant.

The *outcome* was spontaneous abortion defined as a record of pregnancy loss before the 22nd week of gestation.

2. Systematic literature search

Based on the above principles, we conducted a literature review applying the inclusion and exclusion criteria as follows. We searched the PubMed database and Embase in February 2021 using the terms shown in Table 1:

PubMed (all fields): (antidepressant OR TCA OR SNRI OR SSRI) AND (spontaneous abortion OR miscarriage OR pregnancy loss) Filters: Humans, Danish, English, from 2000 - 2021 (141 hits) Embase (all fields): ('antidepressant agent' OR TCA OR 'serotonin noradrenalin reuptake inhibitor' OR 'serotonin uptake inhibitor') AND ('spontaneous abortion' OR miscarriage OR 'pregnancy loss') AND [2000-2021]/py (382 hits)

Table 1 - Search fields used for the systematic literature search in PubMed and Embase

References listed in relevant articles were examined to make sure suitable articles were not overlooked. Two additional articles were found by this strategy (Einarson 2003[4], Sivojelozova 2005[5]). The filter “humans" was applied in the PubMed search since the aim was to look at human species. Original research articles in English or Danish language published between 2000 to 2021 were included. The inclusion criteria made it possible to look for suitable articles among those obtained by the specific search. This was done by reading the title and abstract of the articles and evaluating the relevance for the research question i.e. the association between antidepressant-use and spontaneous abortion. Two independent reviewers (Christine and Tobias) screened the articles making it less likely to overlook relevant articles. Articles not satisfying the inclusion criteria were excluded. Out of the 471 initially identified articles, 16 were selected for full-text review. Of those, 13 articles met the inclusion criteria and were included the meta-analysis *(figure 1).*

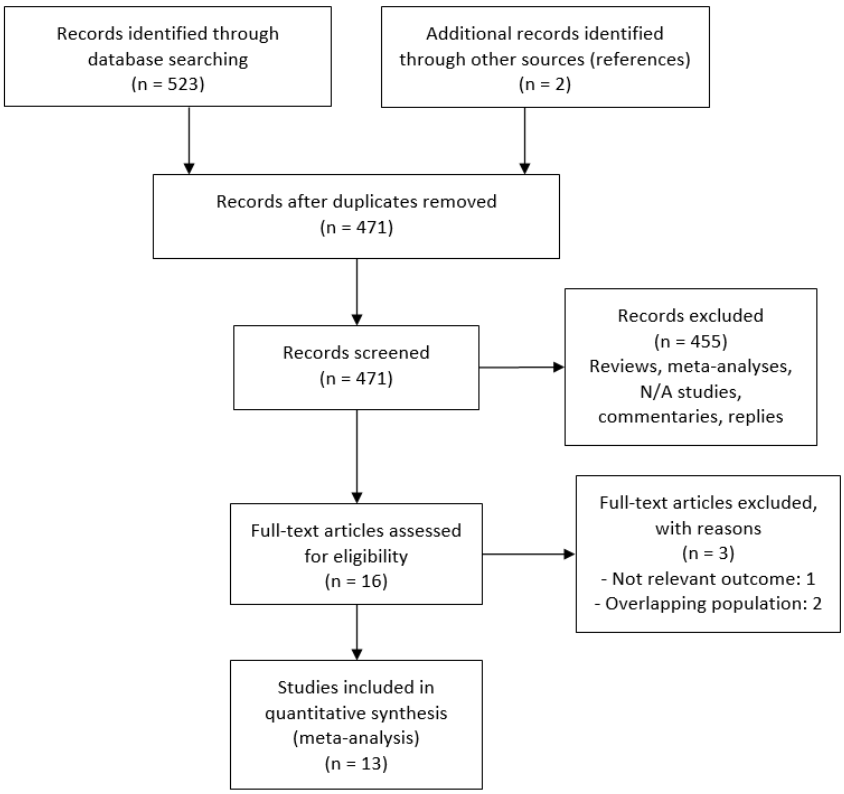

Figure 1 – Flowchart of included and excluded articles

Three of the identified records – all Danish registry-based studies – had overlapping populations (Andersen et al. Kjaersgaard et al. and Johansen et al [6-8]). All three articles adjusted for maternal age, income, and education [6-8]. Both Johansen et al. and Andersen et al. also adjusted for previous spontaneous abortion [6 8]. Johansen et al. further adjusted for self-reported mental illness, self-perceived health, use of alcohol, smoking, exercise, BMI, and if the pregnancy was planned [8]. We choose to include the study by Johansen et al. for the meta-analysis because it accounted for several confounders that the others did not.

3. Statistical methods

In the main meta-analysis, the aim was to estimate the effect of any antidepressant use in pregnancy and the risk of spontaneous abortion. This entailed different decisions on how the data should be analyzed. Firstly, some of the studies included in the meta-analysis had a second reference group of antidepressant users besides the non-teratogenic comparison-group. As this group used different antidepressant agents than the exposed group it was decided that data from the exposed group and the antidepressant exposed reference group could be combined. When this was the case the prevalence ratio, standard error and 95%-confidence interval were

calculated from the reported data by using Epibasic [9]. The prevalence ratios calculated were all from cross-sectional studies and are thus estimates of relative risk in cohort studies. Secondly, Almeida et al. had two comparison groups: women with and without the diagnosis of depression [10]. As the other studies did not differentiate their comparison group in depressed and non-depressed the two comparison groups were combined to one group. Thirdly, spontaneous abortion is a common outcome, with a prevalence of 15-20% [3]. This is important as the studies included in the meta-analysis used different measures of effect. Usually when it is a common outcome odds ratio and risk ratio cannot be assumed to be the same. However, there are exceptions to this. A nested control study included in our study used density sampling. The odds ratio can in this case be seen as the rate ratio without a rare disease assumption [11]. Furthermore, with the follow up time being short (pregnancy lasts about 9 months) the difference in person-time risk between cases and controls will be limited why the rate ratio can be seen as the relative risk. Finally, to estimate the combined effect both the random effects model and the fixed effect model were considered. Looking at similarities and heterogeneity, some of the studies included in the meta-analysis were alike in setup as several used the same teratology information service and way of collecting data. On the other hand, several studies were different regarding methods and even study design. Additionally, the composition of the antidepressant exposed group was different in the studies, which is why variation in the true effect for each study could be expected. As the random-effects model estimates the mean of the distribution of true effects, this model was chosen to estimate the combined effect [12]. Based on these decisions and calculations the combined effect was estimated using 'Episheet' software [13].

In addition to the main meta-analysis it was of interest and possible based on the studies included to make two separate sub-analyses. The first sub-analysis examined the association between SSRI use during pregnancy and spontaneous abortion. This sub-analysis was of interest as it allows to address the specificity of association, and because SSRIs are the most prevalent antidepressants used by Danish women of reproductive age (*Figure 2*). The second sub-analysis compared risk of spontaneous abortion in antidepressant-exposed pregnancies with pregnancies among women with untreated depression. This was done to reduce the risk of confounding by indication. Data for the two sub-analyses were extracted from the studies in which relevant information was directly reported or could be inferred from the publication.

The influence of a study result on a publication decision is called publication bias. As a result, published studies do not represent the real population of studies done. Large studies that show an association are more likely to be published than small or "null" studies [1]. To visualize this, a funnel plot is often used, with standard error on the vertical axis and natural logarithm to the effect size on the horizontal axis. This means larger studies generally will be grouped at the top of the funnel plot due to smaller standard error and the smaller studies at the bottom. An asymmetrical funnel plot is indicative of potential publication bias.

1. Dekkers OM, Vandenbroucke JP, Cevallos M, et al. COSMOS-E: Guidance on conducting systematic reviews and meta-analyses of observational studies of etiology. *PLoS Med* 2019;16(2):e1002742. doi: 10.1371/journal.pmed.1002742 [published Online First: 2019/02/23]
2. Huybrechts KF, Bateman BT, Hernández-Díaz S. Use of real-world evidence from healthcare utilization data to evaluate drug safety during pregnancy. *Pharmacoepidemiol Drug Saf* 2019;28(7):906-22. doi: 10.1002/pds.4789 [published Online First: 2019/05/11]
3. Trolle D. Spontan abort: Sundhed.dk; 2018 [Available from: https://www.sundhed.dk/borger/patienthaandbogen/graviditet/sygdomme/komplikationer-i-graviditeten/spontan-abort/ accessed 14-04 2021.
4. Einarson A, Bonari L, Voyer-Lavigne S, et al. A multicentre prospective controlled study to determine the safety of trazodone and nefazodone use during pregnancy. *Can J Psychiatry* 2003;48(2):106-10. doi: 10.1177/070674370304800207 [published Online First: 2003/03/27]
5. Sivojelezova A, Shuhaiber S, Sarkissian L, et al. Citalopram use in pregnancy: prospective comparative evaluation of pregnancy and fetal outcome. *Am J Obstet Gynecol* 2005;193(6):2004-9. doi: 10.1016/j.ajog.2005.05.012 [published Online First: 2005/12/06]
6. Andersen JT, Andersen NL, Horwitz H, et al. Exposure to selective serotonin reuptake inhibitors in early pregnancy and the risk of miscarriage. *Obstet Gynecol* 2014;124(4):655-61. doi: 10.1097/aog.0000000000000447 [published Online First: 2014/09/10]
7. Kjaersgaard MI, Parner ET, Vestergaard M, et al. Prenatal antidepressant exposure and risk of spontaneous abortion - a population-based study. *PLoS One* 2013;8(8):e72095. doi: 10.1371/journal.pone.0072095 [published Online First: 2013/09/10]
8. Johansen RL, Mortensen LH, Andersen AM, et al. Maternal use of selective serotonin reuptake inhibitors and risk of miscarriage - assessing potential biases. *Paediatr Perinat Epidemiol* 2015;29(1):72-81. doi: 10.1111/ppe.12160 [published Online First: 2014/11/11]
9. Epibasic v4.3 [program], 2019.
10. Almeida ND, Basso O, Abrahamowicz M, et al. Risk of Miscarriage in Women Receiving Antidepressants in Early Pregnancy, Correcting for Induced Abortions. *Epidemiology* 2016;27(4):538-46. doi: 10.1097/ede.0000000000000484 [published Online First: 2016/04/01]
11. Pearce N. What does the odds ratio estimate in a case-control study? *Int J Epidemiol* 1993;22(6):1189-92. doi: 10.1093/ije/22.6.1189 [published Online First: 1993/12/01]
12. Borenstein M, Hedges LV, Higgins JP, et al. A basic introduction to fixed-effect and random-effects models for meta-analysis. *Res Synth Methods* 2010;1(2):97-111. doi: 10.1002/jrsm.12 [published Online First: 2010/04/01]
13. Episheet [program], 2015.