Can statistical adjustment guided by causal inference improve the accuracy of effect estimation? A simulation and empirical research based on Meta-analyses of case-control studies

Ruohua Yan\textsuperscript{1}, Tianyi Liu\textsuperscript{1,2}, Yaguang Peng\textsuperscript{1}, Xiaoxia Peng\textsuperscript{1*}

Institutional addresses:

1. Center for Clinical Epidemiology and Evidence-based Medicine, Beijing Children’s Hospital, Capital Medical University, National Center for Children’s Health, Nanlishilu 56, Xicheng District, Beijing 100045, China

2. Evidence Generation, Medical Affairs, AstraZeneca, Level 22, International Fortune Center, Jianguomenwai Avenue 8, Chaoyang District, Beijing 100010, China

Email addresses:

yan_ruohua@163.com (Ruohua Yan)
tianyi.liu@astrazeneca.com (Tianyi Liu)
plwumi@hotmail.com (Yaguang Peng)
pengxiaoxia@bch.com.cn (Xiaoxia Peng*, assigned as the corresponding author)
Abstract

**Background:** Statistical adjustment is often considered to control confounding bias in observational studies, especially case-control studies. However, different adjustment strategies may affect the estimations of odds ratios (ORs), and in turn affect the results of their pooled analyses. Our study is aimed to investigate how to deal with the statistical adjustment in case-control studies to improve the validity of Meta-analyses.

**Methods:** We carried out a series of Monte Carlo simulation experiments based on predesigned scenarios, and assessed the accuracy of effect estimations from Meta-analyses of case-control studies by combining ORs calculated according to different adjustment strategies. The strategies included fully adjustment of all preset confounders guided by causal inference, insufficiently adjustment of less confounders, and improperly adjustment of covariates other than confounders.

**Results:** For all scenarios with different strength of causal relations, combining ORs adjusted for confounders as far as possible would get the most precise effect estimation, regardless of the sampling approaches of case-control studies and the scale of Meta-analysis. By contrast, combining ORs that were not sufficiently adjusted for confounders or improperly adjusted for mediators or colliders would easily introduce bias in causal interpretation, especially when the true effect of exposure on outcome was weak or none.

**Conclusions:** Statistical adjustment guided by causal inference are recommended for effect estimation. Therefore, when conducting Meta-analyses of case-control studies,
the causal relationship formulated by exposure, outcome, and covariates should be firstly understood through a directed acyclic graph, and then reasonable original ORs could be extracted and combined by suitable methods.

**Keywords:** simulation; confounder; causal inference; case-control study; Meta-analysis

**Background**

Meta-analysis is a well-developed statistical methodology to synthesize results of multiple original studies [1]. Since it increases the sample size for a specific research question by combining data from different independent studies, Meta-analysis enhances the accuracy of effect estimation and improves the strength of evidence [2]. The basic assumption of Meta-analysis is that each included study provides an unbiased estimator, i.e., the variability of results is only attributed to random error but not systematic error [3]. Therefore, randomized controlled trial (RCT) with low risk of bias is acknowledged as a “combinable” study type for Meta-analysis [4]. However, for certain conditions, especially in public health fields, RCTs may be unavailable in consideration of feasibility, ethics, or time, while observational studies can provide supplementary information that experimental studies cannot reflect [5, 6].

During past two decades, a growing number of Meta-analyses are conducted in observational settings [7]. Compared with RCTs, observational studies, especially case-control studies, are exposed to several potential risk of bias which may bring
systematic errors in effect estimations [8]. Besides of selection bias and information bias, confounding is a kind of important bias that may distort the association between exposure and outcome. Particularly when the association strength is weak or medium, confounding may even reverse the direction of causal inference. Therefore, when conducting Meta-analyses of case-control studies, all potential bias of original studies should be properly addressed [9, 10]. Selection bias and information bias can be evaluated and restricted by the Newcastle-Ottawa Scale in the process of Meta-analyses [11]. Confounding bias, however, is always adjusted in the analysis phase of original studies. Logistic regression model is one of the most widely used approaches to control multiple confounders simultaneously, and odds ratio (OR) is a common estimation of causal effect.

In our previous study, we have made a secondary data analysis based on all Meta-analyses of passive smoking and breast cancer in non-smoking women published from 1966 to 2016, as well as all original studies included in these Meta-analyses [12]. We found an apparent inconsistency in statistical methodology among Meta-analyses of case-control studies, including the selection of crude or adjusted OR for the calculation of pooled OR, and the number of covariates adjusted in original case-control studies. These inconsistencies might introduce heterogeneity of original studies and challenge the validity of Meta-analysis. Although we detected these phenomena from a single case study, it is hard to draw conclusions and extrapolate to other Meta-analyses of case-control studies. Furthermore, the empirical
research cannot tell the true effect based on counterfactual hypothesis, and thus cannot judge which adjustment strategy has the best precision in estimating the true effect.

Therefore, we designed this simulation study to assess the accuracy of effect estimations from Meta-analyses by combining ORs of original case-control studies calculated according to different adjustment strategies. The strategies included fully adjustment of all preset confounders guided by causal inference, insufficiently adjustment of less confounders, and improperly adjustment of covariates other than confounders such as mediators or colliders. We set several scenarios and compared the performances of pooled ORs, and thereby provided recommendations on how to choose original ORs for Meta-analyses under different circumstances. Then we used the data from an empirical review to give illustrations.

**Methods**

**Simulation study**

We carried out a series of Monte Carlo simulation experiments to create original case-control studies and their Meta-analyses. The design of the simulation study is displayed in Supplementary Figure S1. We first simulated a target population with pre-determined exposure, outcome, and covariates (Supplementary Table S1) [13]. Then we randomly selected cases and controls from the population, and generated a number of case-control studies according to predesigned scenarios (Supplementary Table S2). We calculated series of ORs for each case-control study by adjusting for
different covariates. Then we conducted Meta-analyses to pool these ORs [14]. The above-mentioned process was repeated for 1000 times to obtain the empirical distribution of pooled OR [15]. The simulation ensured that all generated case-control studies were homogeneous in study population and research methodology, so that the effect estimations were free from selection bias and information bias.

**Generation of target population**

Suppose that we are interested in the causal effect of a dichotomous exposure variable $A$ (1: exposed, 0: unexposed) on a dichotomous outcome variable $Y$ (1: case, 0: control). The causation from $A$ to $Y$ can be reached in four ways, i.e., the direct path $A \rightarrow Y$, the indirect path through mediators $A \rightarrow M \rightarrow Y$ ($M$ denotes a set of mediators of $A$ and $Y$), the backdoor path through common causes $A \leftarrow L \rightarrow Y$ or $A \leftarrow L \rightarrow R \rightarrow Y$ ($L$ denotes a set of confounders of $A$ and $Y$, and $R$ denotes a set of risk factors of $Y$ that have no causations with $A$), and the front-door path by conditioning on common effects $A \rightarrow C \leftarrow Y$ ($C$ denotes a set of colliders of $A$ and $Y$). The simplified causal directed acyclic graph (DAG) between $A$ and $Y$ is shown in Figure 1, and the interpretation of the DAG is provided in Supplementary Method S1.

Without loss of generality, we assumed a vector of 6 dichotomous confounders $L = [L_1, L_2, \ldots, L_6]$ was sufficient to block all backdoor paths from $A$ to $Y$. Only a dichotomous risk factor $R$, a dichotomous mediator $M$, a dichotomous collider $C$ existed in the $A-Y$ causal pathway. $R$ was affected by $L_1, L_2, L_3$, and $L_4$. By specifying
the positive probabilities of all variables and the association parameters, a target population with certain number of observations would be generated. Detailed method is shown in Supplementary Method S2.

**Generation of case-control studies**

From the target population, a series of case-control studies would be generated using random sampling method. For each case-control study, 11 original ORs were calculated by different adjustment strategies. One kind of strategy was insufficient adjustment, i.e., not all confounders were controlled in the logistic regression model. The effect of A on Y was estimated by adjusting for 0 to 6 measured confounders, respectively. Another kind of strategy was improper adjustment, i.e., covariates other than confounders were controlled in the logistic regression model. The effect of A on Y was estimated by adjusting for $L_1$ to $L_6$ in addition to $R, M, C$, and all, respectively.

**Generation of Meta-analyses**

Meta-analyses were generated by combining several case-control studies, and 11 pooled ORs were estimated under each adjustment strategy. For simplicity, we assumed all included case-control studies were from a same population and shared a same sampling process to ensure homogeneity, and thus fixed-effect model would be suitable for Meta-analyses. Further, we considered a more realistic situation that original case-control studies were sampled from different sources, and the method
used for pooling ORs would be either fixed- or random-effect model. We compared
the 11 pooled ORs with the true effect specified in the target population, and thereby
evaluated the performances of 11 adjustment strategies.

**Scenario settings**

Several factors may affect the performances of adjustment strategies. The first is
the causation between exposure $A$ and outcome $Y$ in target population, including the
total effect of $A$ on $Y$ (OR$_{AY}$), the independent associations of covariates $U = [L, R, M, C]$ with $A$ (OR$_{UA}$) and $Y$ (OR$_{UY}$), and the correlations among different variables of $U$ (r$_{UU}$). Suppose that the positive probabilities of $A$, $Y$, and $U$ were 20% in subjects
unexposed to any parent variables, and the associations of $U$ with $A$ or $Y$ were equal.
We specified (1) OR$_{AY}$, (2) OR$_{UA}$, and (3) OR$_{UY}$ as 0.2, 0.5, 0.8, 1, 1.25, 2, or 5 in
different scenarios to represent strong, medium, weak, and no associations with
opposite directions. We also specified (4) r$_{UU}$ as 0, 0.2, 0.5, or 0.8, with r$_{UU} \neq 0$
indicating the nonindependence of covariates.

The second factor that may affect the performances of adjustment strategies is
the sample size of original case-control studies, which involves the number of cases
and the matching approach (matching ratio). To reflect various scales of original
studies, we specified (5) the number of cases as 20, 100, or 500, and (6) the matching
approach as frequency matching or individual matching (base on $L_6$; 1:1, 1:2, or 1:4).
Case-control studies with individual matching design should be analyzed using
conditional logistic regression models.

The third factor that may affect the performances of adjustment strategies is the number of original studies included and the pooling method used in Meta-analyses. We specified (7) the number of original studies as 5, 20, or 50, referring to real Meta-analyses extracted by our previous research [12]. We also specified (8) the pooling method as fixed-effect model, random-effect model, or either depended on the result of heterogeneity test (if $I^2 \leq 50\%$, then fixed-effect model, else random-effect model). Detailed scenario settings are presented in Supplementary Table S2.

**Performance measures**

A total of 32 scenarios were designed. In each scenario, 1000 Meta-analyses of case-control studies were generated, and 1000 pooled ORs of exposure $A$ on outcome $Y$ were estimated for certain adjustment strategies. The parameter $\beta = \ln(\text{OR}_{AY})$ was of interest.

The repetition times ($n = 1000$) was decided by the equation $n = \left(\frac{Z_{\alpha/2}\sigma}{\delta}\right)^2$ [15], where $Z_{\alpha/2}$ was the $1 - \alpha/2$ quantile of the standard normal distribution ($\alpha = 0.05$), $\sigma$ was the standard deviation for $\beta$ ($\sigma = 0.16$ referring to real Meta-analyses extracted by our previous research) [12], and $\delta$ was the permissible difference from the true value of $\beta$ (1000 repetitions could at least ensure the accuracy of estimated $\hat{\beta}$ achieve $\delta = 0.01$, i.e., the accuracy of estimated $\hat{\text{OR}}$ archive 1%).
The Monte Carlo means of pooled ORs were calculated by \( \exp(\bar{\beta}) = \exp\left[(1/n) \sum_{i=1}^{n} \hat{\beta}_i\right] \), while the confidence intervals (CIs) were calculated by \( \exp\left[\bar{\beta} \pm Z_{\alpha/2} \times \sqrt{(1/n - 1) \sum_{i=1}^{n} (\hat{\beta}_i - \bar{\beta})^2}\right] \). The performances of different adjustment strategies were evaluated by the following 6 measures based on \( \hat{\beta} \) [15, 16]:

\[
\text{bias} = \frac{1}{n} \sum_{i=1}^{n} (\hat{\beta}_i - \beta)
\]

\[
\text{relative bias} = \left[\frac{1}{n} \sum_{i=1}^{n} (\hat{\beta}_i - \beta) / \beta\right] \times 100\%, \text{ when } \beta \neq 0
\]

\[
\text{mean square error (MSE)} = \frac{1}{n} \sum_{i=1}^{n} (\hat{\beta}_i - \beta)^2
\]

\[
\text{width of CI} = \frac{1}{n} \sum_{i=1}^{n} (\hat{\beta}_{\text{upp},i} - \hat{\beta}_{\text{low},i})
\]

\[
\text{coverage} = \left[\frac{1}{n} \sum_{i=1}^{n} 1(\hat{\beta}_{\text{low},i} \leq \beta \leq \hat{\beta}_{\text{upp},i})\right] \times 100\%
\]

\[
\text{power} = \left[\frac{1}{n} \sum_{i=1}^{n} 1(\hat{\beta}_{\text{low},i} > 0 \text{ or } \hat{\beta}_{\text{upp},i} < 0)\right] \times 100\%, \text{ when } \beta \neq 0
\]

where \( \hat{\beta}_{\text{upp}} \) and \( \hat{\beta}_{\text{low}} \) represented the upper and lower limits of the CI of \( \hat{\beta} \), respectively. Specially, when \( \text{OR}_{\text{AY}} = 1 (\beta = 0) \) in scenario 1-4, coverage was equal to the probability of not making type I error, while in other scenarios that \( \text{OR}_{\text{AY}} \neq 1 (\beta \neq 0) \), power was equal to the probability of not making type II error. All simulation processes and statistical analyses were conducted by SAS 9.4. The main SAS code is presented in the Appendix.
Empirical research

We chose an empirical Meta-analysis focused on passive smoking and breast cancer in nonsmoking women to illustrate the replicability of simulation experiments [17]. Similar to the process of simulation study, we firstly investigate the causal relationship among passive smoking, breast cancer, and potential confounders through a DAG. The DAG was determined on both literature evidence and subject-matter knowledge, i.e., the nodes of DAG were identified by variables adjusted in each original case-control study, and the direction of arrow between every two nodes was judged by the author and was further approved by clinical experts.

Then we selected original ORs that were calculated by the most appropriate adjustment strategy based on the causal diagram. Fixed- or random-effect model was used to pool ORs according to the size of heterogeneity (decided by $I^2$). Publication bias was assessed by funnel plots. Moreover, sensitivity analyses were conducted to combine original ORs that seemed to underestimate and overestimate the true effect, respectively, through the guidance of causal inference. All Meta-analyses were performed with Review Manager 5.3.

Results

Effect estimations in Meta-analyses of case-control studies

Among all scenarios defined in the simulation study, set scenario Ref be the primary analysis. Figure 2 presents the Monte Carlo pooled ORs of Meta-analyses in
When no covariates were adjusted in original case-control studies, the average effect estimation of Meta-analyses was 2.82 (95% CI, 2.46-3.22), which significantly overestimated the true effect of exposure A on outcome Y (OR_{AY} = 2). The overestimation gradually decreased with the adjustment of more confounders. Combining original ORs that adjusted for all 6 confounders had a mean pooled OR of 2.01 (95% CI, 1.72-2.32), which was the closest estimation to OR_{AY}. Further adjusted for risk factor did not substantially change the estimation (OR, 2.05; 95% CI, 1.74-2.36). However, further adjusted for mediator or collider in addition to confounders did underestimate the true effect. The underestimation was similar for mediator and collider, if they had an equal association strength with exposure and outcome. A more particular interpretation of the results is shown in Supplementary Result S1 and Supplementary Figure S2.

Performances of statistical adjustment strategies

Figure 3 displays the performances of statistical adjustment strategies in different scenarios. MSE, which is a comprehensive indicator for variance and bias (MSE = Var(\hat{\beta}) + bias^2), was closest to 0 when combining original ORs that fully adjusted for 6 confounders but not needlessly adjusted for other covariates. With more insufficient or improper adjustment of covariates in original studies, the estimated parameter \hat{\beta} was more away from the true value. Detailed data are in Supplementary Table S3-S10.
Coverage showed similar tendencies with MSE, i.e., CI estimations based on full adjustment strategy had the highest coverage rate to the true effect (Supplementary Figure S3).

Power was large in most scenarios where the total effect of exposure A on outcome Y (OR_{AY}) was specified as 2. However, for scenarios 1-3 and 1-5 with weaker effects (OR_{AY} = 0.8 and 1.25, respectively), power became insufficient and type II error rate exceeded 20% under inappropriate adjustment strategies. Moreover, for scenario 1-4 with null effect (OR_{AY} = 1), type I error rate exceeded 5% if confounders were not correctly adjusted. Error rates were acceptable only for full adjustment strategy when OR_{AY} was around 1 (Figure 4).

Although the accuracy of effect estimations under each adjustment strategy was sensitive to OR_{AY} in target population, it was rarely affected by OR_{UA}, OR_{UY}, or r_{UU}, except for some extreme situations. Characteristics of case-control studies and Meta-analyses also had little impact on the precision of pooled ORs for a same adjustment strategy. A more particular interpretation of the results is shown in Supplementary Result S2 and Supplementary Figure S4.

**Empirical illustrations**

From the simulation experiments we noticed that, in order to get accurate estimations of causal effect in Meta-analyses, ORs calculated by appropriate adjustment strategies in original case-control studies should be extracted and
combined. However, how to apply our simulation results in practice is still unclear. We now use a Meta-analysis conducted by Lee and Hamling in 2016 to give illustrations [17].

The Meta-analysis was focused on passive smoking and risk of breast cancer in nonsmoking women. Since the research question could not be fulfilled by experimental studies due to ethical reasons, all 47 original studies involved in the Meta-analysis were observational studies. Among them, 30 were case-control studies, 15 were prospective studies, and 2 were case-control studies nested within prospective studies. In the principal analysis, 29 case-control studies and 16 prospective studies were included, with the pooled effect estimations as 1.26 (95% CI, 1.13-1.41) and 1.02 (95% CI, 0.97-1.08), respectively. A clear difference has been found between study types (P<0.001). We supposed that the result of prospective studies might be more credible, since prospective studies generally exposed to less bias and provided relatively higher quality of evidence. However, without a background knowledge of the true effect, we could not definitely conclude whether the association existed or not. Therefore, we tried to re-analyze the data from the case-control studies and give a more decisive causal inference.

First of all, the causal relationship between passive smoking and breast cancer should be understood. We summarized the information of 29 original case-control studies in Supplementary Table S11, where the adjusted covariates were potential confounders identified by each study. However, most studies controlled variables that
showed significance in baseline comparisons or univariate analyses, without distinguishing confounders with risk factors, mediators, or colliders. We should draw a DAG to make detailed differentiation (Supplementary Result S3).

Based on the DAG in Supplementary Figure S5, we evaluated the accuracy of original ORs and made stratification analysis. Among 29 case-control studies, 8 (27.6%) gave reasonable effect estimations and were included in the primary analysis. Meanwhile, 12 (41.4%) were underestimated due to not adjusting for negative confounders of family history (2/12), adjusting for mediators of benign breast disease (9/12), or adjusting for colliders of cardiovascular disease (1/12); 9 (31.0%) were overestimated due to not adjusting for positive confounders of age or body mass index (Supplementary Table S11). None of the original studies were subject to the risk of overfitting. From the forest plot in Figure 5, we detected a weak but significant association between passive smoking and breast cancer in primary analysis (OR, 1.18; 95% CI, 1.01-1.39). Underestimated results slightly shrank the effect and gave a false negative estimation (OR, 1.15; 95% CI, 0.99-1.33). Overestimated results substantially amplified the effect (OR, 1.62; 95% CI, 1.17-2.25). The fixed-effect OR of primary analysis (1.18; 95% CI, 1.07-1.29) was same to the random-effect OR. The funnel plot in Figure 6 further showed that, compared with underestimated or overestimated results that might expose to publication bias (studies with few cases tended to report positive associations), original ORs in primary analysis were symmetrically scattered on both sides of 1.18 with relatively small standard errors.
Therefore, we believed there is a causal relationship between passive smoking and breast cancer in non-smoking women. The conclusion was consistent with the main finding of Lee and Hamling’s review, that the relative risk from all 45 observational studies was 1.15 (95% CI, 1.07-1.23) [17].

**Discussion**

Our study used simulation technique and found that statistical adjustment strategy guided by causal inference would improve the accuracy of effect estimation from Meta-analyses of case-control studies. For all scenarios with different strength of causal relations, combining original ORs that adjusted for confounders as far as possible would get the most precise estimation of pooled effect, regardless of the sampling approaches of case-control studies and the scale of Meta-analysis. By contrast, combining original ORs that were not sufficiently adjusted for confounders or improperly adjusted for mediators or colliders would easily introduce bias in causal interpretation, especially when the true effect of exposure on outcome was weak or none. The pooling method of Meta-analysis could be selected based on the result of heterogeneity test. However, if the causal effect was constant among populations and confounders were sufficiently controlled in all case-control studies, both fixed- and random-effect model would get valid effect estimations.

The findings of our simulation study were further verified by an empirical research, that is, pooled OR calculated by appropriate adjustment strategy yielded an
unbiased estimation of the causal effect. By constructing a DAG with the help of adjusted variables identified by each original study in a systematic review, we could judge which study gave credible results, and combined the results together for pooled effect estimation. Other underestimated or overestimated results could be considered in sensitivity analysis to support the causal interpretation.

Nowadays, simulation study is becoming a powerful supplement to empirical research, especially when the outcomes cannot be derived from mathematical formulae or experimental replications [18]. The thought of Monte Carlo simulation has been widely applied in the field of epidemiology [19, 20]. Regarding our study, because the true causal effect could not be obtained from real Meta-analyses, and the possible adjustment strategies could not be exhausted by real case-control studies, empirical researches were not able to give confirmative conclusions, while statistical simulation was a good solution.

On the other hand, empirical research is also an important illustration of simulation results. Compared with the methodological representation, the practical application of the findings is always of value. By analyzing a motivating example, the reason for conducting the simulation study would be clarified, and the parameter settings of the scenarios would be justified. Moreover, a proper instance would help the technical paper popularize to non-technical audiences.

To our knowledge, this is the first simulation study to evaluate how adjustment strategies of original case-control studies impact the pooled effect estimations of
Meta-analyses. From our previous case study on Meta-analyses of passive smoking and breast cancer [12], we detected an inconsistency in adjustment strategies used for calculating original ORs, which might not eliminate confounding bias but introduce new bias during Meta-analyses. However, the previous study could not determine the best strategy without knowing the true effect, and could not extrapolate the best strategy to other situations as well. In the present study, by specifying the true OR of exposure on outcome and setting various parameters in multiple scenarios, we compared the accuracy of adjustment strategies under different circumstances and gave methodological recommendations correspondingly. Our findings are expected to help improve the validity of Meta-analyses of case-control studies, and provide high-quality evidence for medical decision-making.

Although our findings provide important implications on how to choose original ORs of case-control studies for Meta-analyses, caution is needed when the situation is more complex. First, we assumed all generated case-control studies were homogenous in study population and research methodology, so that the original studies were free from selection bias and information bias. However, in practice, it is difficult to ensure the methodological consistency in case-control studies included in a specific Meta-analysis. Therefore, the quality of original studies should be carefully evaluated [11], and the heterogeneity among original studies should be controlled, either by statistical approaches such as random-effect model or Meta-regression, or by subgroup analysis to pool all “combinable” results together [21]. Otherwise, if the
methodologies of original studies are far from each other, qualitative systemic reviews rather than quantitative pooled analyses are recommended [8, 22].

Second, we made all statistical adjustment in case-control studies based on logistic regression models. But in real cases, other adjustment methods such as propensity score and instrument variable are also used [23]. How to combine the results calculated by different adjustment methods need to be further investigated.

Third, we focused on Meta-analyses of case-control studies, while in most actual reviews, observational studies including case-control studies and cohort studies are combined together to calculate pooled ORs. Compared with case-control studies, cohort studies are exposed to more uncertain factors, such as different follow-up durations, different rates of loss to follow-up, etc. Whether the results of the present study are still valid for Meta-analyses of cohort studies, is also an important question to be answered in our future research.

Conclusions

Statistical adjustment strategy guided by causal inference are recommended for effect estimations. Thus, when conducting Meta-analyses of case-control studies, the causal relationship between exposure and outcome should be firstly understood through a DAG, and then reasonable original ORs should be extracted and combined by suitable methods to get accurate pooled ORs.
List of abbreviations

RCT, randomized controlled trial
OR, odds ratio
DAG, directed acyclic graph
CI, confidence interval
MSE, mean square error

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Availability of data and material
The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests
The authors declare that they have no competing interests.
Funding

Not applicable.

Authors’ contributions

XP and RY designed the study. RY analyzed the data. RY, XP, and YP interpreted the results. RY and TL wrote the first draft of the manuscript. XP and YP revised the manuscript for important intellectual content. All authors read and approved the final manuscript.

Acknowledgements

Not applicable.

References

1. Glass GV. Primary, secondary, and meta-analysis of research. Educational Researcher. 1976;5:3-8.

2. Jadad AR, Cook DJ, Jones A, Klassen TP, Tugwell P, Moher M, et al. Methodology and reports of systematic reviews and meta-analyses: a comparison of Cochrane reviews with articles published in paper-based journals. JAMA. 1998;280(3):278-80.

3. Egger M, Smith GD, Phillips AN. Meta-analysis: principles and procedures. BMJ. 1997;315(7121):1533-7.
4. Whitehead A WJ. A general parametric approach to the meta-analysis of randomised clinical trials. Stat Med. 1991;10(11):1665-77.

5. Black N. Why we need observational studies to evaluate the effectiveness of health care. BMJ. 1996;312(7040):1215-8.

6. Yang W, Zilov A, Soewondo P, Bech OM, Sekkal F, Home PD. Observational studies: going beyond the boundaries of randomized controlled trials. Diabetes Res Clin Pract. 2010;88 Suppl 1:S3-9.

7. Shrier I, Boivin JF, Steele RJ, Platt RW, Furlan A, Kakuma R, et al. Should meta-analyses of interventions include observational studies in addition to randomized controlled trials? A critical examination of underlying principles. Am J Epidemiol. 2007;166(10):1203-9.

8. Blettner M, Sauerbrei W, Schlehofer B, Scheuchenpflug T, Friedenreich C. Traditional reviews, meta-analyses and pooled analyses in epidemiology. Int J Epidemiol. 1999;28(1):1-9.

9. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. JAMA. 2000;283(15):2008-12.

10. Valentine JC, Thompson SG. Issues relating to confounding and meta-analysis when including non-randomized studies in systematic reviews on the effects of interventions. Res Synth Methods. 2013;4(1):26-35.
11. Stang A. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. Eur J Epidemiol. 2010;25(9):603-5.

12. Liu T, Nie X, Wu Z, Zhang Y, Feng G, Cai S, et al. Can statistic adjustment of OR minimize the potential confounding bias for meta-analysis of case-control study? A secondary data analysis. Bmc Med Res Methodol. 2017;17(1):179.

13. Wicklin R. Simulating Data with SAS. Cary, NC: SAS Institute Inc; 2013.

14. Senn S, Weir J, Hua TSA, Berlin C, Branson M, Glimm E. Creating a suite of macros for meta-analysis in SAS: A case study in collaboration. Stat Probabil Lett. 2011;81(7):842-51.

15. Burton A, Altman DG, Royston P, Holder RL. The design of simulation studies in medical statistics. Stat Med. 2006;25(24):4279-92.

16. Morris TP, White IR, Crowther MJ. Using simulation studies to evaluate statistical methods. Stat Med. 2019;38(11):2074-102.

17. Lee PN, Hamling JS. Environmental tobacco smoke exposure and risk of breast cancer in nonsmoking women. An updated review and meta-analysis. Inhal Toxicol. 2016;28(10):431-54.

18. Tadano K, Isobe T, Sato E, Takei H, Kobayashi D, Mori Y, et al. Basics of Monte-Carlo Simulation: Focusing on Dose-to-medium and Dose-to-water. Igaku Butsuri. 2016;36(3):148-50.

19. Li H, Yuan Z, Su P, Wang T, Yu Y, Sun X, et al. A simulation study on matched
case-control designs in the perspective of causal diagrams. Bmc Med Res Methodol. 2016;16(1):102.

20. Kontopantelis E, Reeves D. Performance of statistical methods for meta-analysis when true study effects are non-normally distributed: a comparison between DerSimonian-Laird and restricted maximum likelihood. Stat Methods Med Res. 2012;21(6):657-9.

21. Egger M, Schneider M, Davey Smith G. Spurious precision? Meta-analysis of observational studies. BMJ. 1998;316(7125):140-4.

22. Shapiro S. Is meta-analysis a valid approach to the evaluation of small effects in observational studies? J Clin Epidemiol. 1997;50(3):223-9.

23. Cnossen MC, van Essen TA, Ceyisakar IE, Polinder S, Andriessen TM, van der Naalt J, et al. Adjusting for confounding by indication in observational studies: a case study in traumatic brain injury. Clin Epidemiol. 2018;10:841-52.

**Figure legends**

**Figure 1. Directed acyclic graph in target population**

A, exposure; Y, outcome; L, confounder; R, risk factor; M, mediator; C, collider.

**Figure 2. Pooled ORs of Meta-analyses in scenario Ref (OR\(_{AY} = 2\))**

OR, odds ratio; CI, confidence interval; A, exposure; Y, outcome.
Figure 3. Mean square error of effect estimations under different adjustment strategies

Mean square error was presented according to different (A) OR\(_{AY}\), (B) OR\(_{UA}\), (C) OR\(_{UY}\), and (D) \(r_{UU}\) in target population; (E) number of cases and (F) matching approach in original case-control studies; and (G) number of studies and (H) pooling method in Meta-analyses.

OR, odds ratio; A, exposure; Y, outcome; U, covariate.

Figure 4. Error rate of effect estimations under different adjustment strategies

Error rate was presented according to different OR\(_{AY}\) in target population. Solid symbol represents type II error (where OR\(_{AY}\) ≠ 1), and hollow symbol represents type I error (where OR\(_{AY}\) = 1).

OR, odds ratio; A, exposure; Y, outcome.

Figure 5. Forest plot of an empirical Meta-analysis on passive smoking and breast cancer [17]

SE, standard error; IV, inverse variance; CI, confidence interval.

Figure 6. Funnel plot of an empirical Meta-analysis on passive smoking and breast cancer [17]

SE, standard error; OR, odds ratio.