**Statistical Analysis Protocol (March 11th, 2021)**

1. **Prepare data for analysis**
* Include cohort and case-control studies with any type of effect estimate, i.e., maximally-adjusted hazard ratio (HR)/risk ratio (RR)/odds ratio (OR)/mean difference (MD).
* Exclude studies that reported only crude effect estimates.
* To facilitate intuitive interpretation of results [1,2], we will convert effect estimates on HR/OR scale to RR scale. When the outcome is relatively rare ($<$ 15% by the end of follow-up), we approximately convert HR/OR to RR using [3]:

$$RR≈HR or OR$$

when the outcome is common ($\geq $ 15%) [2]:

$$RR≈\frac{1-0.5^{\sqrt{HR}}}{1-0.5^{\sqrt{\frac{1}{HR}}}}or \sqrt{OR}$$

* Finally, for binary outcomes, log RRs and corresponding sampling variances will be calculated for subsequent analyses; for continuous outcomes measured on MD scale, the effect estimates with corresponding sampling variances will be used for analyses.
1. **Robust random-effects meta-analysis**

We will apply a random-effects meta-analysis model throughout this systematic review. The random-effects model assumes that there is a distribution of population effect sizes across studies [4]. Since studies included in this systematic review were collected from published literature, the random-effects model is a more plausible match to the underlying effect distribution compared to the fixed-effect model [5]. Of note, the summary effect under the random-effects model is an estimate of themeanof a distribution of population effects [5].

We will use the robust variance estimation (RVE) method with random-effects weights for meta-analysis [6,7]. Compared with the conventional random-effects model with a normality assumption [4], the RVE method has the following advantages:

* It imposes no distributional assumptions on the population effects.
* With small-sample adjustments [8], RVE method can be effectively used to make inferences even in meta-analyses with a small number of studies.
* RVE method can accommodate dependence among effect estimates (e.g., multiple estimates based on the same individuals at different time points, multiple estimates computed using a common control group), without requiring knowledge of the dependence structure.
1. **Predictive distribution**

In presence of effect heterogeneity, the random-effects summary estimate and its confidence interval are usually insufficient to summarize the whole body of evidence, because they only represent an estimate of the mean effect and its precision [4]. Thepredictive distribution can describe how the population effect sizes are distributed around the mean effect size [4]. We plan to characterize this effect distribution by using three recently proposed metrics (see below), in addition to the heterogeneity estimate. These new metrics use “calibrated” estimates [9] without making any distributional assumptions. Based on simulation results [10], they are recommended for use in meta-analyses of $\geq $ 10 studies.

* + **Tau (**$T$**)**––the estimate of the standard deviation of population effect sizes [11]––it quantifies the absolute amount of heterogeneity.
	+ **95% prediction interval**––estimation of the middle 95% area of the effect distribution [9]––it predicts with 95% confidence the population effect in a new study that is similar to the studies in the meta-analysis.
	+ $\hat{P}\left(θ<q\right)$––estimation of the lower tail of the effect distribution [10,12]––it estimates the proportion of population effects ($θ$) below a threshold ($q$) of scientific importance. $q$ is defined as follows: for outcome measured as MD, $q$ = 0; for RR, $q$ = 0.9 or 1.0.
	+ $\hat{P}\left(θ>q^{\*}\right)$––estimation of the upper tail of the effect distribution [10,12]––it estimates the proportion of population effects ($θ$) above a threshold ($q^{\*}$) of scientific importance. $q^{\*}$ is defined as follows: for outcome measured as MD, $q^{\*}$ = 0; for RR, $q^{\*}$ = 1.0 or 1.1.
1. **Small-study effects**

Small-study effects describe the tendency for smaller studies in a meta-analysis to show more pronounced effects than larger studies [13]. Possible causes include publication bias, true heterogeneity, data irregularities, etc [14]. We plan to use a random-effects Egger’s regression to examine whether there is indication of small-study effects [14]. Egger’s regression is recommended for use in meta-analyses of $\geq $ 10 studies [15].

* Indication of small-study effects [14]: “$P<$ 0.10(two-sided) of Egger’s regression” plus “the random-effects summary estimate being larger than the point estimate of the most precise study (the study with the smallest standard error) in the meta-analysis”.
1. **Publication bias**

Publication bias describes the “iceberg phenomenon” where the studies included in a systematic review are systematicallyunrepresentative of all studies that have been conducted on a topic [16,17]. It may lead to an exaggerated or wholly distorted conclusion of the actual body of evidence. In this systematic review, we will use one modelling method, S-value [18] (see below), to evaluate the robustness of meta-analysis results to potential publication bias. This method assumes a one-tailed selection process [19], where the publication process selects studies with both point estimates in the direction of summary estimate and small *P*-values (e.g., $<$ 0.05). This assumption is justified by empirical findings on how applied researchers interpret *P*-values [20].

* **S-value** [18]––It can accommodate dependence among effect estimates, non-normal population effects, and small meta-analyses. The one-tailed selection process will be modeled using a single two-sided *P*-value cutoff of 0.05, such that publication selects “affirmative” results (i.e., statistically significant point estimates in the direction of summary estimate) over “non-affirmative” results (i.e., significant point estimates but in the opposite direction, or non-significant ones). It can calculate the following metrics: a. a summary estimate corrected for worst-case publication bias; b. the severity of publication bias (i.e., the ratio, $η$, by which affirmative studies are more likely to be published than non-affirmative studies) that would be required to shift the pooled estimate or its confidence interval limit to a chosen threshold of scientific importance (i.e., the null or a non-null value $q$). The threshold is defined as follows: for outcome measured as MD, null = 0; for RR, null = 1.0, $q$ = 0.9 or 1.1 ($q$ is the value in the same direction of the pooled estimate). A large $η$ would indicate that the meta-analysis is relatively robust to publication bias, whereas a small $η$ would indicate that the meta-analysis is relatively sensitive to publication bias. Informed by the empirical benchmarks for plausible values of $η$ in medicine, a $η$ of $\geq $ 4 would represent implausibly severe or extreme publication bias.
1. **Sensitivity analysis for residual confounding**

In this systematic review, we will use the E-value (see below) [3,21], which assesses how strong residual confounding would have to be to “explain away” an observed exposure-outcome association, to represent the robustness of meta-analysis results to potential residual confounding.

* **E-value** [3,21]––The minimum strength of association, on RR scale, that residual confounding would need to have with both the exposure and outcome, conditional on the measured covariates and on average across studies, to shift the pooled estimate or its confidence interval limit to a chosen threshold of scientific importance (i.e., the null or a non-null value $q$). The threshold is defined as follows: null = 1.0, $q$ = 0.9 or 1.1 ($q$ is the value in the same direction of the pooled estimate). A large E-value would indicate that the meta-analysis mean estimate is relatively robust to residual confounding, whereas a small E-value would indicate that the meta-analysis mean estimate is relatively sensitive to residual confounding.

**eReferences**

1. Grant RL. Converting an odds ratio to a range of plausible relative risks for better communication of research findings. BMJ. 2014;348:f7450.

2. VanderWeele TJ. Optimal approximate conversions of odds ratios and hazard ratios to risk ratios. Biometrics. 2020;76(3):746-752.

3. VanderWeele TJ, Ding P. Sensitivity analysis in observational research: introducing the E-value. Ann Intern Med. 2017;167(4):268-274.

4. Higgins JP, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. J R Stat Soc Ser A Stat Soc. 2009;172(1):137-159.

5. Borenstein M, Hedges LV, Higgins JP, Rothstein HR. A basic introduction to fixed-effect and random-effects models for meta-analysis. Res Synth Methods. 2010;1(2):97-111.

6. Hedges LV, Tipton E, Johnson MC. Robust variance estimation in meta-regression with dependent effect size estimates. Res Synth Methods. 2010;1(1):39-65.

7. Hedges LV, Tipton E, Johnson MC. Erratum: robust variance estimation in meta-regression with dependent effect size estimates. Res Synth Methods. 2010;1(2):164-165.

8. Tipton E. Small sample adjustments for robust variance estimation with meta-regression. Psychol Methods. 2015;20(3):375-393.

9. Wang CC, Lee WC. A simple method to estimate prediction intervals and predictive distributions: summarizing meta-analyses beyond means and confidence intervals. Res Synth Methods. 2019;10(2):255-266.

10. Mathur MB, VanderWeele TJ. Robust metrics and sensitivity analyses for meta-analyses of heterogeneous effects. Epidemiology. 2020;31(3):356-358.

11. Borenstein M, Higgins JP, Hedges LV, Rothstein HR. Basics of meta-analysis: I2 is not an absolute measure of heterogeneity. Res Synth Methods. 2017;8(1):5-18.

12. Mathur MB, VanderWeele TJ. New metrics for meta-analyses of heterogeneous effects. Stat Med. 2019;38(8):1336-1342.

13. Sterne JA, Gavaghan D, Egger M. Publication and related bias in meta-analysis: power of statistical tests and prevalence in the literature. J Clin Epidemiol. 2000;53(11):1119-1129.

14. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997;315(7109):629-634.

15. Sterne JA, Sutton AJ, Ioannidis JP, Terrin N, Jones DR, Lau J, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. BMJ. 2011;343:d4002.

16. Boissel JP, Haugh MC. The iceberg phenomenon and publication bias: the editors' fault? Clin Trials Metaanal. 1993;28(6):309-315.

17. Song F, Parekh S, Hooper L, Loke YK, Ryder J, Sutton AJ, et al. Dissemination and publication of research findings: an updated review of related biases. Health Technol Assess. 2010;14(8):iii, ix-xi, 1-193.

18. Mathur MB, VanderWeele TJ. Sensitivity analysis for publication bias in meta-analyses. J R Stat Soc Ser C Appl Stat. 2020;69(5):1091-1119.

19. Vevea JL, Hedges LV. A general linear model for estimating effect size in the presence of publication bias. Psychometrika. 1995;60(3):419-435.

20. McShane BB, Gal D. Statistical significance and the dichotomization of evidence. J Am Stat Assoc. 2017;112(519):885-895.

21. Mathur MB, VanderWeele TJ. Sensitivity analysis for unmeasured confounding in meta-analyses. J Am Stat Assoc. 2020;115(529):163-172.