Content of the appendix

1. **Supplementary table 1**: Deviations from the protocol and the corresponding justification for each deviation.

2. **Supplementary table 2**: Categories of RCT participants who might have missing data.

3. **Supplementary statistical notes**: Data analysis.

4. **Supplementary results A:** The percentage of meta-analyses for which the 'sensitivity analysis pooled relative effect' (1) crossed the threshold of the null effect and (2) changed direction compared to the 'sensitivity analysis pooled relative effect (CCA)' when considering total possible missing data.

5. **Supplementary results B:** The percentage of meta-analyses for which the 'sensitivity analysis pooled relative effect' (1) crossed the threshold of the null effect and (2) changed direction compared to the 'original pooled relative effect' when considering definite and total possible missing data.

6. **Supplementary results C:** The percentage change in the relative effect estimate between the 'sensitivity analysis pooled effect estimate (assumption)' and the 'sensitivity analysis pooled effect estimate (CCA)', when considering participants with total possible missing data.

7. **Supplementary results D: Table 3:** Details of the percentage change in the relative effect estimate, stratified by whether the estimate is less than or greater than 1 under the complete case analysis (CCA), using either definite missing data or total possible missing data.

1. **Supplementary table 1**: Deviations from the protocol and the corresponding rationale for each deviation

Item	What was stated in protocol	The deviation	Rationale for the change
Selecting and reproducing the original meta-analysis of interest	"For each eligible meta-analysis, we will first attempt to reproduce the original analysis. When this analysis generates a different effect estimate that is not statistically significant, we will exclude the corresponding meta-analysis from this part of the study."	We included the eligible meta-analyses without reproducing the original analysis.	When we attempted to reproduce the original analysis, we found it very challenging to figure out what data the systematic reviewers used in their analysis.
Assumed effect among participants with missing data relative to effect observed among followed-up participants	"We define $RI_{NotFU/FU}$ as the relative event incidence among those not followed-up relative to the event incidence among those followed-up"	Instead of using the $RI_{NotFU/FU}$, we used the informative missing odds ratio (IMOR) method which describes the relationship between the unknown odds of the outcome among participants with missing data and the known odds among observed participants.	We decided to use the IMOR because of it easily applied in Stata (metamiss command). On the other hand, $RI_{NotFU/FU}$ has not such command available. In addition, the two methods rendered the comparable results when applied on a sample of 52 meta-analyses.

2. **Supplementary table 2**: Categories of RCT participants who might have missing data [1]

Category of participants that might have missing data	Description of the category
Explained lost to follow-up	Participants described as lost to follow-up, and trialists provided an explanation, e.g., relocated to a different country
Unexplained lost to follow-up	Participants described as lost to follow-up, and trialists did not provide an explanation
Outcome not assessable	Data of a certain outcome for a number of participants is not available because the outcome adjudicators could not assess their outcome. For example, venography could not be done for a number of participants
Data not available	Participants who are still part of the RCT, however due to incomplete or missing record, some of the outcome data of this participant are missing
Ineligible or mistakenly randomized	Participants who, subsequent to randomization, are either found not to have the condition of interest (e.g. are not pregnant in an RCT among pregnant women), or did not undergo a procedure for which the intervention is intended (e.g. did not undergo surgery in an RCT of postoperative thromboprophylaxis)
Did not receive first dose/treatment	Participants who did not receive the 'first dose' of the intervention to which they were randomized
Ineligible due to early occurrence of outcome	Participants who were eligible at baseline then developed the outcome of interest soon after enrollment. These are considered ineligible if the trialists judge that the occurrence of the outcome cannot be related to the intervention of interest
Experienced adverse events	Participants who developed adverse events but without clear indication whether or not they discontinued the RCT

Non-compliant	Participants who were non-adherent or otherwise violated the protocol
Cross-over	Participants randomized to one arm, but who received the intervention meant for another treatment arm
Withdrew consent	Participants who withdraw their consent to participate in the RCT
Discontinued due to adverse events	Participants who discontinued the RCT due to adverse events
Discontinued trial prematurely	Participants who left the RCT but for whom a reason for discontinuation was not provided
Withdrawn by investigator/clinician	Participants who left the RCT through a decision made by the investigator or clinician (e.g., due to medical necessity)
Unintended protocol violation	Participants who left the RCT due a protocol violation for which they are not responsible (e.g., unavailability of hospital beds)
Lack of efficacy	Participants who left the RCT because they perceived no benefits from the intervention they were randomized to
Protocol violation by investigator/clinician	Investigator/clinician violated the protocol (e.g., change the intended intervention) due to a medical reason
More than one category reported together	The number refers to participants belonging to two or more of the above categories
Other	Reason different than the above

1. Kahale, L.A., et al., *A guidance was developed to identify participants with missing outcome data in randomized controlled trials.* Journal of Clinical Epidemiology, 2019.

3. **Supplementary statistical notes:** Data analysis

We opted to select fixed IMOR with variance of 0 since we are dealing with 100 meta-analyses of different topics. It would have been so challenging (almost impossible) to obtain a range of uncertainty value for each topic covered in the pool of 100 meta-analyses. In addition, this approach allows us to stabilize the uncertainty value across all 100 meta-analyses, to make sure that the observed change is purely due to the assumption applied.

For the calculation of the change in the relative effect estimate, we initially attempted to compare the 'sensitivity analysis pooled relative effect' to the 'original pooled relative effect'. However for the following two reasons, this was not feasible. First, for a significant number of systematic reviews, we could not reproduce the original meta-analysis as it was not clear how the systematic review authors dealt with missing data. Indeed, when we compared the 'best-case scenario pooled relative effect' to the 'original pooled relative effect', 10% of the meta-analyses shifted closer to the null value of one which contradicts the nature of this assumption (i.e., best-case scenario shifts the effect estimate away from the null value of one). Whereas, when we compared the 'best-case scenario pooled relative effect' to the 'complete case analysis pooled relative effect', all meta-analysis shifted away from the null value of one. A very likely explanation is that in first scenario (comparing the 'best-case scenario pooled relative effect' to the 'original pooled relative effect'), missing data were identified and handled differently in the 'original pooled relative effect' by the systematic review authors than how we identified and handled missing data when calculating the 'best-case scenario pooled relative effect'. Second, our approach complies with the GRADE guidance that recommends conducting complete case analysis in the primary analysis and some form of sensitivity analysis, in order to assess the risk of bias associated with missing data [2].

According to Gamble and Hollis [3] and other methods papers [4], sensitivity analysis based on the worst-case scenario is intended to not favor the intervention. So, irrespective of the rate of missing data per arm, worst-case scenario for an outcome with effect estimate >1, will always assume events in the control arm and zero events in the intervention arm. Similarly, for outcomes with effect estimate <1, the worst-case scenario will always assume

events in the intervention arm and zero events in the control arm. Hence, challenging the robustness, by shifting the original effect estimate closer to the null value of one.

Under the best-case scenario, Stata imputes missing data as ones in the intervention group and zeroes in the control group. Under the worst-case scenario, Stata imputes missing data as zeroes in the intervention group and ones in the control group. However, the best-case scenario is intended to shift the original effect estimate away from the null value of one, whereas the worst-case scenario is intended to shift the original effect estimate closer to the null value of one. Thus, when applying the worst-case scenario for an outcome with an effect estimate less than 1, we imputed missing data as zeroes in the intervention group and ones in the control group.

When the statistical method of the original meta-analysis of interest was not reported, we used Mantel-Haenszel for the implausible but common assumptions and the inverse-variance method when using IMOR (via a two-stage approach in metamiss command).

1. Kahale LA, Guyatt GH, Agoritsas T, et al. A guidance was developed to identify participants with missing outcome data in randomized controlled trials. *Journal of Clinical Epidemiology* 2019
2. Guyatt GH, Ebrahim S, Alonso-Coello P, et al. GRADE guidelines 17: Assessing the Risk of Bias Associated with Missing Participant Outcome Data in a body of evidence. *J Clin Epidemiol* 2017 doi: 10.1016/j.jclinepi.2017.05.005
3. Gamble C, Hollis S. Uncertainty method improved on best-worst case analysis in a binary meta-analysis. *J Clin Epidemiol* 2005;58(6):579-88. doi: 10.1016/j.jclinepi.2004.09.013
4. Mavridis D, Chaimani A, Efthimiou O, et al. Addressing missing outcome data in meta-analysis. *Evid Based Ment Health* 2014;17(3):85-9. doi: 10.1136/eb-2014-101900

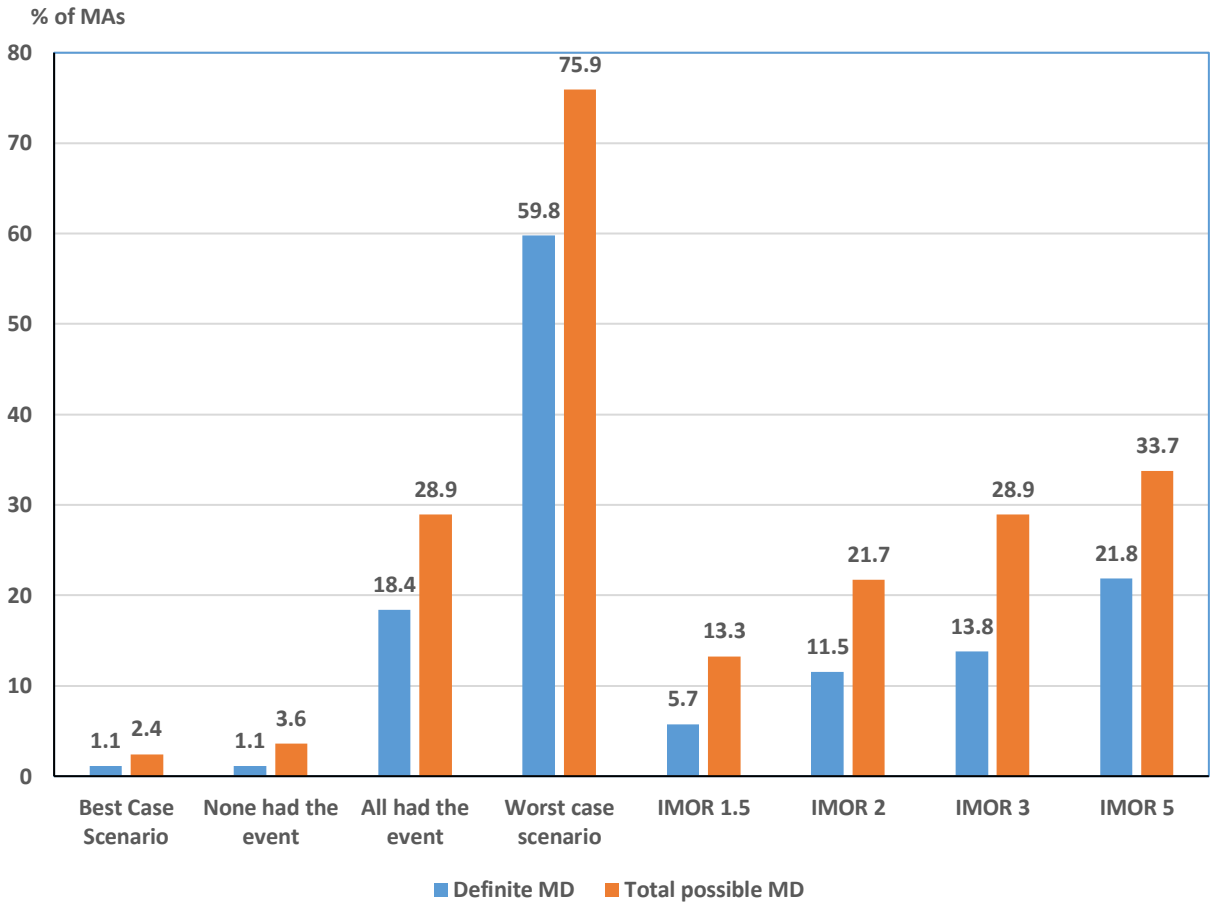
4. **Supplementary results A:** The percentage of meta-analyses for which the 'sensitivity analysis pooled relative effect' (1) crossed the threshold of the null effect and (2) changed direction compared to the 'sensitivity analysis pooled relative effect (CCA)' when considering total possible missing data.

The below figures show the results for the comparison of the 'sensitivity analysis pooled relative effect' to the 'sensitivity analysis pooled relative effect (CCA)' for each method. Specifically, they show the numbers of meta-analyses that (1) crossed the threshold of the null effect and (2) changed direction respectively, when considering participants with definite and total possible missing data.

For the four implausible but commonly used assumptions, the percentage of meta-analyses that crossed the threshold of the null effect varied from 2% (best case scenario) to 4% (none of the participants with missing data had the event) to 30% (all participants with missing data had the event) to 76% (worst case scenario). For the plausible assumptions based on IMOR, the percentage of meta-analyses that crossed the threshold of the null effect varied from 5% (least stringent assumption IMOR 1.5) to 34% (most stringent assumption IMOR 5).

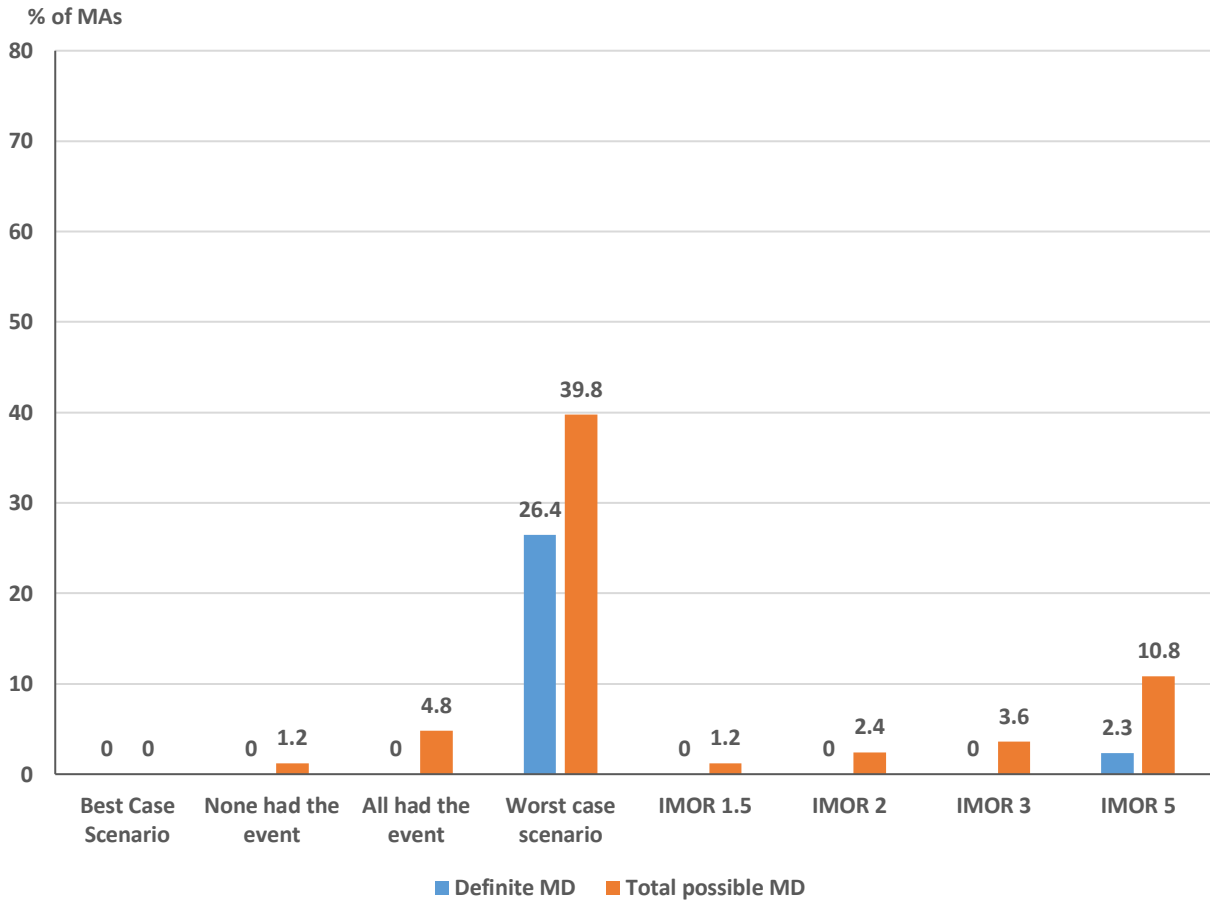
The percentage of meta-analyses that changed direction varied from 0% (best case scenario), to 1% (none of the participants with missing data had the event), to 5% (all participants with missing data had the event), to 40% (worst case scenario). As for the five plausible assumptions, the percentage of meta-analyses that changed direction varied from 1.2% (least stringent assumption IMOR 1.5) to 11% (most stringent assumption IMOR 5).

Supplementary results A figure 1: Results of meta-analyses that crossed the threshold of the null effect when considering participants with definite (in blue) and total possible missing data (in orange) and comparing the 'sensitivity analysis pooled relative effect (assumption)' to the 'sensitivity analysis pooled relative effect (CCA)' (n=87 systematic reviews that did not cross the threshold of the null effect under the CCA)



***Abbreviations**: IMOR: informative missing odds ratio; MAs: meta-analyses; MD: missing data*

Supplementary results A figure 2: Results of meta-analyses that changed direction when considering participants with definite (in blue) and total possible missing data (in orange) and comparing the 'sensitivity analysis pooled relative effect (assumption)' to the 'sensitivity analysis pooled relative effect (CCA)' (n=87 systematic reviews that did not cross the threshold of the null effect under the CCA)



***Abbreviations**: IMOR: informative missing odds ratio; MAs: meta-analyses; MD: missing data*

5. **Supplementary results B:** The percentage of meta-analyses for which the 'sensitivity analysis pooled relative effect' (1) crossed the threshold of the null effect and (2) changed direction compared to the 'original pooled relative effect' when considering definite and total possible missing data.

The below figures show the results for the comparison of the 'sensitivity analysis pooled relative effect' to the 'original pooled relative effect' for each method. Specifically, they show the numbers of meta-analyses that (1) crossed the threshold of the null effect and (2) changed direction respectively, when considering participants with definite and total possible missing data.

Using definite missing data: Under CCA, the results of 87% of meta-analyses did not cross the threshold of the null effect. For the four implausible but commonly used assumptions, the percentage of meta-analyses that crossed the threshold of the null effect varied from 3% (best case scenario) to 12% (none of the participants with missing data had the event) to 27% (all participants with missing data had the event) to 65% (worst case scenario). For the plausible assumptions based on IMOR, the percentage of meta-analyses that crossed the threshold of the null effect varied from 18% (least stringent assumption IMOR 1.5) to 32% (most stringent assumption IMOR 5).

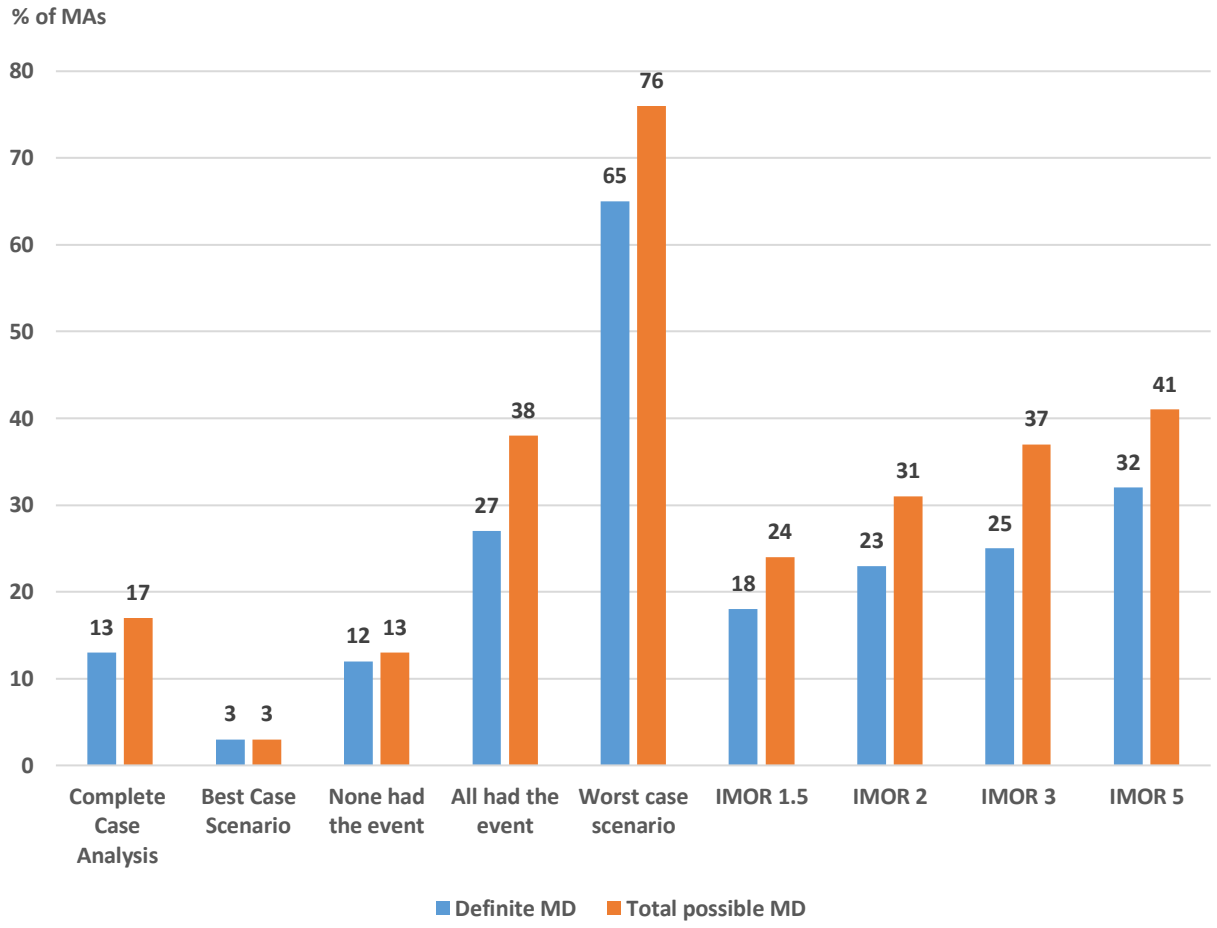
The percentage of meta-analyses that changed direction was 3% under CCA. It varied from 1% (best case scenario and none of the participants with missing data had the event) to 4% (all participants with missing data had the event) to 33% (worst case scenario). As for the five plausible assumptions, the percentage of meta-analyses that changed direction varied from 3% (least stringent assumption IMOR 1.5) to 6% (most stringent assumption IMOR 5).

Using total possible missing data: Under CCA, the results of 83 meta-analyses did not cross the threshold of the null effect. For the four implausible but commonly used assumptions, the percentage of meta-analyses that crossed the threshold of the null effect varied from 3% (best case scenario) to 13% (none of the participants with missing data had the event)

to 38% (all participants with missing data had the event) to 76% (worst case scenario). For the plausible assumptions based on IMOR, the percentage of meta-analyses that crossed the threshold of the null effect varied from 24% (least stringent assumption IMOR 1.5) to 41% (most stringent assumption IMOR 5).

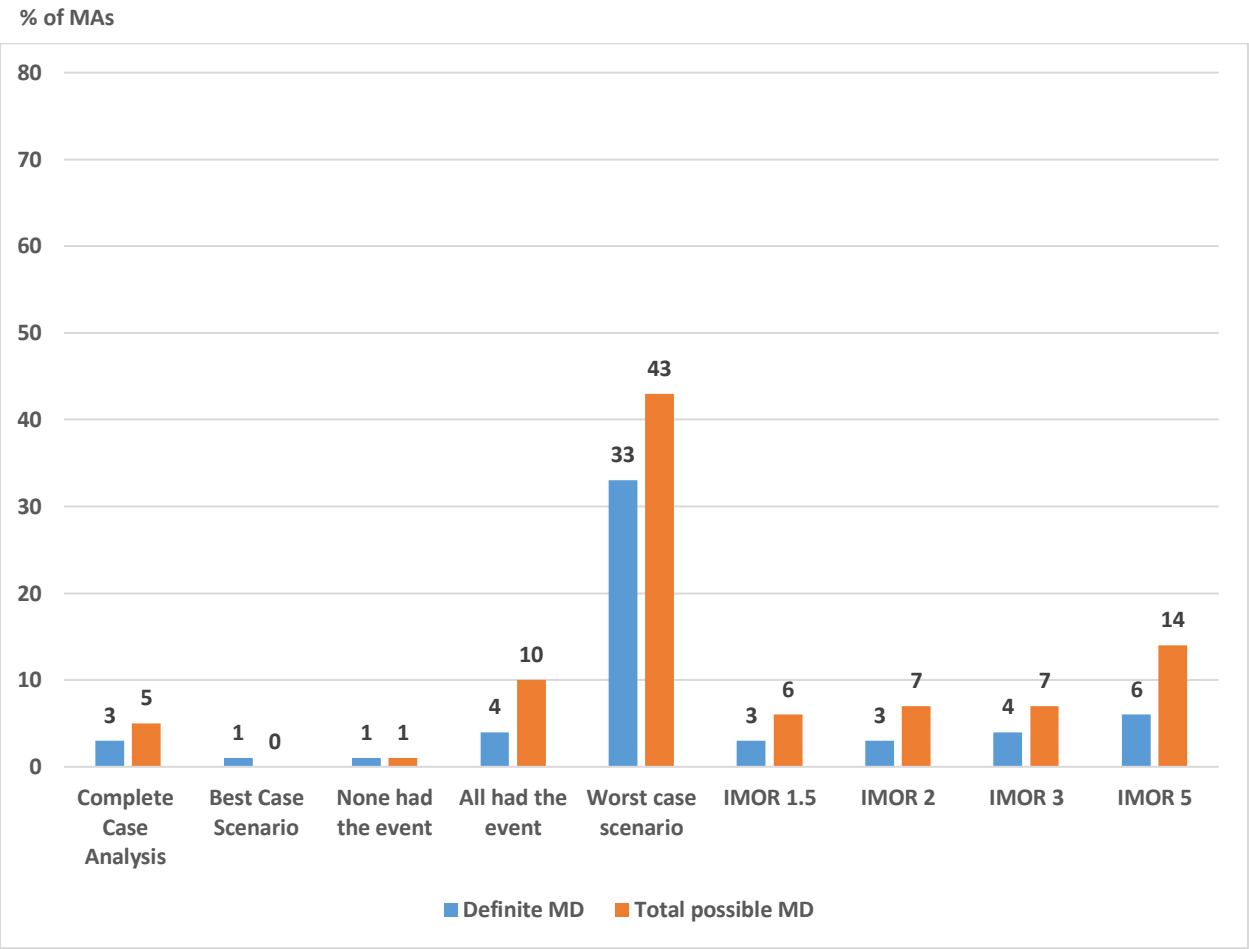
The percentage of meta-analyses that changed direction was 5% under CCA. It varied from 0% (best case scenario), to 1% (none of the participants with missing data had the event), to 10% (all participants with missing data had the event), and to 43% (worst case scenario). As for the five plausible assumptions, the percentage of meta-analyses that changed direction varied from 6% (least stringent assumption IMOR 1.5) to 14% (most stringent assumption IMOR 5).

Supplementary results B figure 1: Results of meta-analyses that crossed the threshold of the null effect when considering participants with definite (in blue) and total possible missing data (in orange) when comparing the 'sensitivity analysis pooled relative effect' to the 'original pooled relative effect' (N=100)



Abbreviations: *IMOR: informative missing odds ratio; MAs: meta-analyses; MD: missing data*

Supplementary results B figure 2: Results of meta-analyses that changed direction when considering participants with definite (in blue) and total possible missing data (in orange) when comparing the 'sensitivity analysis pooled relative effect' to the 'original pooled relative effect' (n=100)



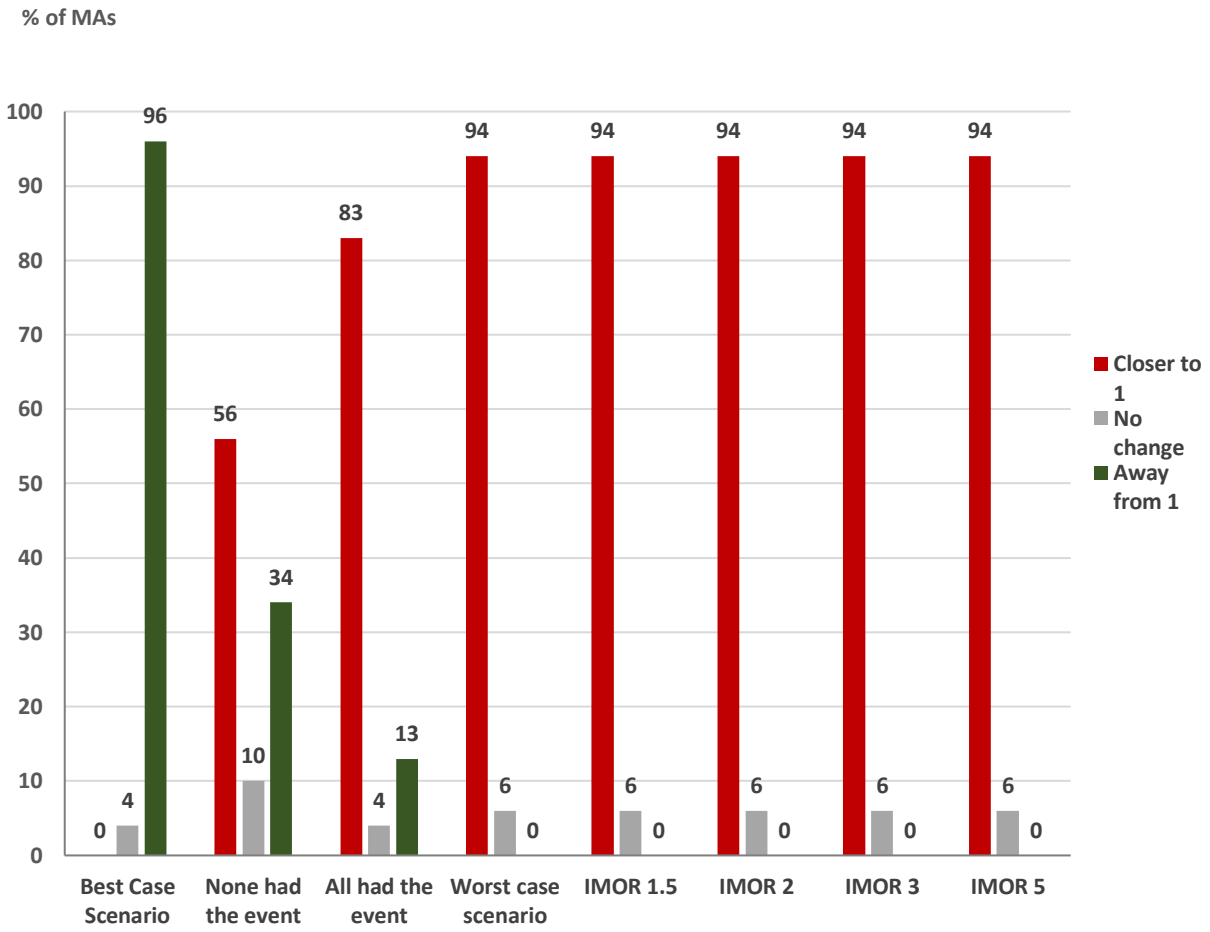
Abbreviations: *IMOR: informative missing odds ratio; MAs: meta-analyses; MD: missing data*

6. **Supplementary results C:** The percentage change in the relative effect estimate between the 'sensitivity analysis pooled effect estimate (assumption)' and the 'sensitivity analysis pooled effect estimate (CCA)', when considering participants with total possible missing data.

For the four implausible but commonly used assumptions, the percentage of meta-analyses with increased relative effect estimate (shifted away from the null value of 1) was 96% for 'best case scenario' assumption, 34% with 'none of the participants with missing data had the event' assumption, and 13% with 'all participants with missing data had the event' assumption. The median increase in the relative effect estimate varied from 0% for the' worst case scenario' assumption to 25.2% (IQR 11.7%-47.3%) for the 'best case scenario' assumption. The percentage of meta-analyses with reduced relative effect estimate (shifted closer towards the null value of 1) was 94% for the 'worst case scenario' assumption, 56% for 'none of the participants with missing data had the event' assumption, and 83% for 'all participants with missing data had the event' assumption. The median reduction in the relative effect estimate varied from 0% for the 'best case scenario' assumption to 52.8% (IQR 21.5%-94.2%) for the 'worst case scenario' assumption (please see below figure).

For the plausible assumptions based on the IMOR, the percentage of meta-analyses with increased relative effect estimate was 94% of across all stringent assumptions. The median reduction in relative effect estimate varied from 2.1% (IQR 0.9%-4.7%) for IMOR 1.5 assumption to 11.6% (IQR 5.0%-22.5%) for IMOR 5 assumption.

Supplementary results C figure 1: Change of relative effect estimate (by direction) between the 'sensitivity analysis pooled effect estimate (assumption) and the 'sensitivity analysis pooled effect estimate (CCA)' when considering participants with total possible missing data. Bars in the upper part of the figure represent the percentage of meta-analyses with change of relative effect estimate (by direction). The numerical values in the bottom part represent the median (IQR) for, respectively, the increase and decrease in relative effect estimate (N=100)



Abbreviations: *IMOR: informative missing odds ratio; IQR: interquartile range; MAs: meta-analyses; MD: missing data*

7. **Supplementary results D: Table 3**: Details of the percentage change in the relative effect estimate, stratified by whether the estimate is less than or greater than 1 under the complete case analysis (CCA), using either definite missing data or total possible missing data

Definite missing data

	Best Case Scenario	None had the event	IMOR 1.5	IMOR 2	IMOR 3	IMOR 5	All had the event	Worst case scenario
Effect estimate < 1 under the complete case analysis								
Closer to 1 n (%)	0	24	59	62	62	62	53	63
Median (IQR)	-	4.3 (1.8 – 7.8)	1.3 (0.6 – 2.7)	2.2 (1.0 – 4.8)	4.0 (1.7 – 8.6)	6.6 (2.7 – 14.5)	14.8 (2.8 – 38.2)	30.3 (11.4 – 89.2)
No change n (%)	8	32	13	10	10	10	7	9
Away from 1 n (%)	64	16	0	0	0	0	12	0
Median (IQR)	16.7 (6.1 – 31.5)	1.5 (0.8 – 3.3)	-	-	-	-	0.9 (0.3 – 2.1)	-
Effect estimate > 1 under the complete case analysis								
Closer to 1 n (%)	0	14	26	26	26	26	22	27
Median (IQR)	-	1.3 (0.7 – 4.6)	2.1 (0.7 – 5.0)	3.9 (1.3 – 8.3)	6.5 (2.1 – 12.9)	9.7 (3.2 – 20.6)	10.8 (3.5 – 41.3)	32.6 (10.3 – 65.3)
No change n (%)	1	5	2	2	2	2	1	1
Away from 1 n (%)	27	9	0	0	0	0	5	0
Median (IQR)	22.2 (9.3 – 117.2)	2.5 (1.0 – 9.9)	-	-	-	-	1.9 (0.3 – 3.1)	-

Total possible missing data

	Best Case Scenario	None had the event	IMOR 1.5	IMOR 2	IMOR 3	IMOR 5	All had the event	Worst case scenario
Effect estimate < 1 under the complete case analysis								
Closer to 1 n (%)	0	36	68	68	68	68	62	67
Median (IQR)	-	1.9 (0.7 – 7.8)	1.8 (0.9 – 4.1)	3.6 (1.7 – 7.4)	6.6 (3.2 – 13.9)	9.9 (5.2 – 21.9)	14.7 (4.5 – 39.9)	66.4 (28.9 – 125.8)
No change n (%)	3	9	4	4	4	4	3	5
Away from 1 n (%)	69	27	0	0	0	0	7	0
Median (IQR)	24.6 (11.6 – 39.2)	2.2 (0.7 – 3.7)	-	-	-	-	1.2 (0.9 – 4.7)	-
Effect estimate > 1 under the complete case analysis								
Closer to 1 n (%)	0	20	26	26	26	26	21	27
Median (IQR)	-	2.5 (0.9 – 9.0)	3.4 (1.0 – 5.4)	6.1 (1.7 – 9.7)	9.4 (2.6 – 15.2)	13.6 (3.7 – 22.1)	11.2 (4.4 – 44.6)	35.1 (15.2 – 70.3)
No change n (%)	1	1	2	2	2	2	1	1
Away from 1 n (%)	27	7	0	0	0	0	6	0
Median (IQR)	31.7 (12.4 – 134.6)	4.6 (1.1 – 9.5)	-	-	-	-	2.4 (1.0 – 5.4)	-

Abbreviations: *IMOR: informative missing odds ratio; IQR: interquartile range*