

Title: Systematic review and meta-analysis of the association between Epstein-Barr virus, Multiple Sclerosis, and other risk factors

Authors: Benjamin M Jacobs¹, Gavin Giovannoni^{1,2,3}, Jack Cuzick¹, *Ruth Dobson^{1,3}

1: Preventive Neurology Unit, Wolfson Institute of Preventive Medicine, Queen Mary University London

2: Blizard Institute, Queen Mary University London

3: Royal London Hospital, Whitechapel, London

Corresponding Author:

Dr Ruth Dobson

Preventive Neurology Unit

Wolfson Institute of Preventive Medicine

Charterhouse Square

London EC1M 6BQ

Email: ruth.dobson@qmul.ac.uk

Word count: 3217

Key Words: multiple sclerosis, clinically isolated syndrome, Epstein Barr virus, infectious mononucleosis, systematic review, meta-analysis

Funding & Disclosures: This work was funded via a grant from Barts Charity (grant ref MGU0365)

Dr Jacobs has no relevant disclosures to declare

Professor Giovannoni has no relevant disclosures to declare

Professor Cuzick has no relevant disclosures to declare

Dr Dobson has no relevant disclosures to declare

Abstract

Background: EBV infection is thought to play a central role in the development of Multiple Sclerosis (MS). If causal, it represents a target for interventions to reduce MS risk.

Objective: To examine the evidence for interaction between EBV and other risk factors, and explore mechanisms via which EBV infection may influence MS risk.

Methods: Pubmed was searched using the terms “multiple sclerosis” AND “Epstein Barr virus”, “multiple sclerosis” AND EBV, “clinically isolated syndrome” AND “Epstein Barr virus” and “clinically isolated syndrome” AND EBV. All abstracts were reviewed for possible inclusion.

Results: 262 full-text papers were reviewed. There was evidence of interaction on the additive scale between anti-EBV antibody titre and HLA genotype (AP 0.48, $p < 1 \times 10^{-4}$; RERI 3.84, $p < 5 \times 10^{-3}$; S 1.68, $p = 0.06$). Previous IM was associated with increased OR of MS in HLA-DRB1*1501 positive but not HLA-DRB1*1501 negative persons. Smoking was associated with a greater risk of MS in those with high anti-EBV antibodies (OR 2.76) but not low anti-EBV antibodies (OR 1.16). No interaction between EBV and risk factors was found on a multiplicative scale.

Conclusions: EBV appears to interact with at least some established MS risk factors. The mechanism via which EBV influences MS risk remains unknown.

Introduction

Multiple sclerosis (MS) is thought to arise as the result of acquired environmental risk in a genetically susceptible population¹⁻⁴. Environmental risk factors for MS include Epstein-Barr Virus (EBV) infection, smoking, obesity during adolescence, and low serum vitamin D². Understanding how environmental risk factors interact with each other and with genotype is crucial to developing targeted preventative strategies.

We set out to update and extend our understanding of the interaction between EBV and other MS risk factors. To our knowledge, there has been no previous attempt to integrate all data related to how EBV interacts with other MS risk factors. One meta-analysis has examined the potential interaction between EBV serostatus and HLA in MS; other previous meta-analyses have not studied risk factor interaction⁵⁻⁸.

Nested case-control studies using large health repositories^{9,10} have made a major contribution to epidemiological evidence supporting a causal relationship between EBV and MS. However, the high rate of EBV seropositivity in the general population argues against EBV seropositivity alone being a sufficient factor for causing MS⁷. The prevalence of MS in EBV-negative individuals is virtually zero when highly sensitive techniques are used to assess EBV serostatus^{11,12}. Symptomatic EBV^{5-7,13-15} infection (Infectious Mononucleosis; IM) confers a greater risk of MS than asymptomatic EBV carriage.

Population-based epidemiological studies indicate that EBV infection and other environmental risk factors may interact with genotype in the pathogenesis of MS¹⁶. To our knowledge there have been no previous attempts to systematically pool these estimates. In

this systematic review and meta-analysis, we examine all the available evidence for EBV interaction with other MS risk factors (both in terms of EBV serostatus and IM) using both multiplicative and additive models for interaction. We also examine the reported relationship between EBV and MS, and pool evidence around the relationship between active EBV turnover (as measured by PCR) and MS. Finally, we provide a narrative systematic review of the literature around MS and EBV.

Methods

Search strategy

Pubmed was searched using the terms “multiple sclerosis” AND “Epstein Barr virus”, “multiple sclerosis” AND EBV, “clinically isolated syndrome” AND “Epstein Barr virus” and “clinically isolated syndrome” AND EBV. Search dates were 1950-present. The most recent search was performed on 22nd December 2018.

All abstracts were reviewed for possible inclusion. Studies for use in the meta-analysis were screened according to the following criteria: containing both MS and control group, and using either standard techniques to establish EBV serostatus, history of IM, or PCR. Where these criteria were met, the full text was retrieved.

Following this, relevant studies were reviewed and data extracted. Where full text was not available, the authors were contacted to provide the article. Where it was judged unclear as to whether data within selected papers met the inclusion criteria (details of inclusion criteria for each analysis are given in the results section), a second co-author independently reviewed the paper, and a consensus decision was reached. The quality of data were assessed by recording the reported security of MS diagnosis (no clear criteria and/or self-reported vs. explicit

criteria used for diagnosis, the gold standard), and technique for assessing EBV (ELISA vs. immunofluorescence, the gold standard). All references of retrieved review and/or meta-analyses were reviewed for additional articles not captured during the original search.

Technical differences between study design may introduce bias and limit the validity of pooled effect estimates. Such differences included: differences in clinical criteria for MS diagnosis, differences in method of IM diagnosis (clinical, recall questionnaire, serological), differences in laboratory techniques (e.g. immunofluorescence vs ELISA), differences in HLA genotyping (molecular typing vs SNP imputation), and difference in the quantification of smoking exposure (cotinine vs questionnaire). To overcome these difficulties, we performed subgroup analyses where appropriate to stratify by these potential sources of heterogeneity.

All included full text papers were assigned to analyses covered by this review - EBV interaction with other MS risk factors, serology and MS risk, infectious mononucleosis and MS risk, EBV DNA detection and MS, papers covering possible mechanisms of EBV contribution to MS risk, and papers examining the relationship between immune response to EBV and MS-related clinical or MRI outcomes. A single paper could be assigned to any number of analyses, and each analysis/review was performed independently of all others.

Statistical methods

Meta-analyses were conducted in R v3.6.0 using the ‘meta’ package based on reported data. Odds ratios (ORs) were calculated using a Mantel-Haenszel random effects model with a continuity correction. Bias was quantified using the efficient score (a linear regression of

funnel plot asymmetry)¹⁷. For interaction studies, odds ratios were pooled using the inverse variance method. Where data were available, unadjusted odds ratios were calculated.

For interaction studies, the highest and lowest exposure groups were used - e.g. where EBNA titres were divided into quartiles, we took the lowest and the highest groups as 'EBNA lo' and 'EBNA hi' respectively. Interaction was assessed by calculating 4 measures of interaction: where the numbers of cases and controls in each risk factor group were presented, the Attributable Proportion due to interaction (AP), the relative excess risk due to interaction (RERI), the Synergy index (S), and multiplicative interaction^{18,19} were calculated. For two risk factors of interest, e.g. smoking and HLA status, if OR₁₁ indicates the Odds Ratio for MS in individuals exposed to both risk factors, OR₁₀ the OR for HLA+ non-smokers and OR₀₁ that for HLA- smokers:

$$RERI = OR_{11} - OR_{10} - OR_{01} + 1$$

$$S = OR_{11} - 1 / ((OR_{10} - 1) - (OR_{01} - 1))$$

$$AP = (OR_{11} - OR_{10} - OR_{01} + 1) / OR_{11}$$

In the absence of interaction, RERI and AP will be 0, and S will be 1. Measures of departure from additivity (AP, RERI, and S) were calculated using the indicator variable method described previously¹⁹ in R (code available on request). Multiplicative interaction was calculated by performing logistic regression with an interaction term. If OR represents the Odds Ratio for MS, x₁ one risk factor, x₂ the second risk factor, and x₁x₂ the product (interaction) term, then:

$$\text{Ln(OR)} = b_0 + b_1x_1 + b_2x_2 + b_3x_1x_2$$

The exponent of the interaction term coefficient b_3 represents the multiplicative interaction between the two risk factors. For these analyses, the regression model did not adjust for variables other than the two risk factors in question. Standard errors were calculated for measures of additive interaction using the delta method¹⁹. Standard errors for the multiplicative interaction were calculated from the output of the logistic regression model. Meta-analysis of interaction terms was performed using the inverse variance method with a random effects model.

Data availability statement

This work was performed using published data. All data sources are listed in the references and supplementary references. All R code used for the analysis are available on request.

Results

A total of 632 references were retrieved using the search terms “multiple sclerosis” AND “Epstein Barr virus”, and “multiple sclerosis” AND EBV. “Multiple sclerosis” AND EBV, “clinically isolated syndrome” retrieved 22 references, all of which had been captured in the previous search. “Clinically isolated syndrome” AND EBV retrieved a further 17 references, again all of which had been previously captured. Review of all references of meta-analyses and systematic reviews provided 6 unique new results. 370 results were discarded following review of abstracts for reasons including pre-selecting EBV positive patients only, having no control group, validation studies of new methods for EBV serology. 262 full text papers were reviewed and included as summarized in Fig. 1.

*EBV interaction with HLA-DRB1*1501 in MS*

10 papers^{20–28} were included for this analysis. All but one paper presented HLA-DRB1*1501 homo- and heterozygotes pooled into a single group (“HLA positive”), and so this grouping was used in the analysis. Where EBNA titres were divided into quartiles, we took the highest and lower quartiles to represent ‘high’ and ‘low’ titres respectively. One paper²⁶ was excluded due to overlapping participants with another paper²².

The odds ratio (OR) of MS in individuals with high anti-EBV antibody titres is increased in HLA-DRB1*1501 positive (OR 7.17, 95% CI 3.92 - 13.14) compared to HLA-DRB1*1501 negative individuals (OR 2.92, 95% CI 1.86 - 4.59, Fig. 2, Table 1). Studies differed in their method of HLA genotyping. Restricting the analysis to studies using tagging SNPs (rs3135005 or rs9271366) did not significantly alter the results (Fig. S1). Restricting the analysis to studies using PCR-based methods yielded a similar result (Fig. S1).

Individual-level data were available for five studies. We estimated the degree of interaction between HLA status and EBNA titre by calculating the AP, Synergy Index, RERI, and the degree of multiplicative interaction as described above. There was evidence of significant interaction between EBNA titre and HLA genotype on the additive scale in terms of the AP and RERI (AP 0.48, $p < 1 \times 10^{-4}$; RERI 3.84, $p < 5 \times 10^{-3}$; S 1.68, $p = 0.06$). There was no evidence of interaction on the multiplicative scale (β 1.27, $p = 0.74$) (Fig. 2, Table 1). Subgroup analyses based on method of HLA genotyping are presented in table S1.

Four papers examined the potential interaction between previous infectious mononucleosis and HLA-DRB1*1501 status and MS^{20,22,28,29}. Again, homo- and heterozygote status was pooled into “HLA positive”. A history of IM is associated with increased OR of MS in HLA-DRB1*1501 positive individuals (OR 5.11 95% CI 2.00-13.03; $p < 1 \times 10^{-3}$) but not in HLA-DRB1*1501 negative individuals (OR 1.22 95% CI 0.33-4.48; $p = 0.77$, Fig. 3, Table 2). Three studies had individual-level data available. There was no significant interaction on the additive or multiplicative scales between HLA status and IM in the meta-analysis of 3 studies with individual-level data available (figure 3, Table 2). Subgroup analysis by method of HLA genotyping did not significantly alter the results (Fig. S2).

EBV interaction with smoking in MS

5 papers studied the potential interaction between smoking status and anti-EBV antibody titre^{25,26,28,30,31}. Three studies stratified smoke exposure as ever vs never smokers, one study used second-hand smoke exposure as a variable, and one study distinguished active from inactive smoking using serum cotinine levels. Smoking is associated with a greater risk of MS in those with high anti-EBV antibodies (OR 2.76 95%CI 2.13-3.59; $p < 1 \times 10^{-5}$) but not in those with low anti-EBV antibodies (OR 1.16 95%CI 0.95-1.42; $p = 0.15$). There was no significant interaction on the multiplicative or additive scales in the meta-analysis of the four eligible studies (Fig. 4, Table 3). Exclusion of either the study using second-hand smoke as the exposure or using serum cotinine as a proxy for smoking did not significantly affect the results (Table S2).

EBV interaction with vitamin D in MS risk

Only 2 studies presented data on both EBV and vitamin D in MS^{32,33}. One of these studies looked at vitamin D levels in people with established MS³², and the other in samples taken

both prior to and following MS onset, with multiple, variable sampling points per participant³³. One study applied a correction to vitamin D levels for month of sampling³³, the other did not³². In addition, one study using a single EBNA epitope³³, whereas the other looked at specific ENA-1 domains³². Neither study demonstrated any interaction between vitamin D level and anti-EBNA titre, however for the reasons above they were not pooled.

EBV interaction with obesity in MS risk

Only one study examined the potential interaction between EBV and obesity in risk of MS³⁴. This study demonstrated a striking potential interaction on an additive scale with an attributable proportion due to interaction of 0.8 (95%CI 0.6-1.0) in the incident study, and in the prevalent study an attributable proportion due to interaction of 0.7 (95%CI 0.5-1.0)³⁴.

EBV seropositivity and MS

56 papers were included in the final analysis for this analysis (supplementary references). Inclusion criteria were MS and control group, no pre-selection of groups based on EBV serostatus and history of IM, EBV serology measured using clearly defined methods. Reasons for exclusion included not having a control group and pre-selecting EBV positive patients. Studies were separated into those examining adult vs. paediatric MS populations given the reported differences in seroprevalence between the two groups. Following an assessment of data quality, validity analyses were performed limiting studies to those deemed to be of high quality. Seropositivity for EBV was calculated by pooling results from studies which reported seropositivity to either EBNA, VCA, or both. Where both were reported, the EBNA data were used. Studies using different EBNA1 and EBNA2 epitopes were pooled for all analysis.

EBV seropositivity was significantly more common among people with MS (adults and children) than controls (n=15554, 56 studies, OR 3.62, 95% CI 2.83-4.83, Fig. 5). There was evidence of significant heterogeneity ($Q=168.37$, $p<1\times10^{-4}$) and publication bias ($p<0.05$). Overall, 6479/7395 people with MS were EBV seropositive (87.6%) compared with 6205/8159 EBV seropositive control subjects (76.1%). EBV seropositivity was more prevalent among adults with MS compared to controls (n=13768, 46 studies, OR 3.27, 95% CI 2.45-4.35, $p<1\times10^{-4}$). There was substantial heterogeneity between studies ($Q=134.5$, $p<1\times10^{-4}$) and evidence of publication bias ($p=0.0196$), with studies demonstrating a relationship between EBV infection and MS more likely to be published. Overall, 5805/6606 adults with MS were EBV seropositive (87.9%) compared with 5702/7162 adult control subjects (79.6%). EBV seropositivity was more common among children with MS or CIS than controls (n=1786, 10 studies, OR 4.93, 95% CI 3.83-6.36, $p<1\times10^{-4}$). There was no evidence of heterogeneity ($Q=5.76$, $p=0.76$) and no evidence of publication bias ($p=0.622$). Overall, 674/789 children with MS were EBV seropositive (85.4%) compared with 485/997 control subjects (48.6%).

IgG reactivity to the Viral Capsid Antigen (VCA) was more prevalent among adults with MS (n=7688, 23 studies, OR 3.18, 95% CI 1.88-5.38, $p<1\times10^{-4}$, data not shown). There was substantial heterogeneity between studies ($Q=73.27$, $p<1\times10^{-4}$) and no evidence of publication bias ($p=0.062$). Overall, VCA seropositivity was present in 3414/3698 (92.3%) of adults with MS compared with 3595/3990 controls (90.1%). Reactivity to the EBNA antigen was again more prevalent among people with MS compared to controls (n=9537, 33 studies, OR 3.39, 95% CI 2.53-4.56, $p<1\times10^{-4}$, data not shown). There was substantial heterogeneity between studies ($Q=71.5$, $p<1\times10^{-4}$) with evidence of publication bias in these studies

($p < 0.01$). Overall, EBNA seropositivity was present in 4260/4805 (88.7%) of adults with MS compared with 3660/4732 (77.3%) controls.

The increased seroprevalence of EBV infection in people with MS/CIS remained significant when restricting included studies to those using the more sensitive technique of immunofluorescence (rather than enzyme-linked immunosorbent assay) to detect EBV antibodies (OR 4.61, 95% CI 1.86-11.40). Similarly, when restricting included studies to those which used explicit diagnostic criteria to define MS, this effect remained significant (OR 3.44, 95% CI 2.56-4.61)

Infectious mononucleosis and MS

32 full text papers were reviewed for this analysis, of which 19 met the inclusion criteria (supplementary references). Inclusion criteria were MS and control group, clearly stated methods for obtaining a previous history of IM, and no selection on the basis of reported history of IM.

Previous Infectious Mononucleosis (IM) was more common in people with MS ($n=107243$, 19 studies, OR 2.00, 95% CI 1.79-2.23, $p < 10^{-4}$, Fig. 6). There was significant heterogeneity ($Q=37.28$, $p=0.0048$) but no evidence of publication bias ($p=0.45$, Fig. 6). This effect remained significant when restricting studies to those using criteria-defined MS (OR 1.94, 95% CI 1.72-2.18, Fig. 6).

EBV DNA detectable by PCR.

31 full text papers were reviewed and 23 included in the analysis (supplementary references). 8 papers studied EBV DNA in CSF, 3 in whole blood, 7 in peripheral blood mononuclear

cells, 4 in plasma/serum and 1 in saliva. The EBNA gene was the most commonly used for EBV detection (9 studies), with BAM used in 4 studies, VCA in 3 studies, and LMP in 2 studies.

EBV DNA was detectable in whole blood/PBMC more often in people with MS versus controls ($n = 1853$, 9 studies, OR 3.48, 95% CI 1.7360-6.9659, $p < 5 \times 10^{-4}$). There was evidence of significant heterogeneity ($Q = 48.94$, $p < 1 \times 10^{-4}$) but no evidence of publication bias ($p = 0.78$). Detection of EBV DNA did not differ between MS and control serum/plasma samples ($n = 607$, OR 1.81, 95% CI 0.77-4.26; $p = 0.18$) or CSF ($n = 802$, OR 1.74, 95% CI 0.97-3.12, $p = 0.062$).

Discussion and conclusions

There is a considerable body of epidemiological evidence implicating EBV in the pathogenesis of MS. EBV infection appears to be a necessary but not sufficient requirement for developing MS, EBV seroprevalence is higher among people with MS, symptomatic EBV infection (IM) is more prevalent among people with MS, and the effect of anti-EBV antibody titre on MS risk appears contingent on the strongest genetic risk factor for MS, the HLA-DRB1*1501 genotype.

In our meta-analysis of interaction between EBV and other risk factors, we demonstrate evidence for supra-additive interaction between EBNA titre and HLA status in determining risk. The absence of strong evidence for interaction between EBV and other risk factors in our analysis demonstrates the importance of using multiple measures of interaction (AP,

RERI, Synergy Index, and multiplicative interaction) to avoid the risk of type 1 error.

Relying on the Attributable Proportion alone to make inferences about interaction may be misleading. However, the small number of studies suitable for our analysis of interaction limits the power of this meta-analysis, and therefore conclusions about interaction should be drawn cautiously from these results.

The mechanism via which EBV exerts this increased risk remains unknown, and our systematic review of the literature highlights a multitude of potential biological mechanisms that have been both demonstrated, replicated, and importantly not replicated. It seems likely that the route via which EBV exerts its effect lies in complex interactions between EBV and the host genome, the precise mechanisms of which remain to be elucidated. Despite the relatively small retrospective studies that have been performed, it also appears clear that IgG titres against EBV antigens are not able to act as a clinically useful biomarker in either treated or untreated cohorts.

Our results for the seroprevalence of EBV among people with MS are consistent with the previously published meta-analysis, which reported ORs of 4.47 (95%CI 3.26-6.11) and 4.51 (95%CI 2.84-7.16) for EBNA and VCA respectively. Our estimates of 3.39 (95%CI 2.53-4.56) and 3.18 (95%CI 1.88-5.38) are more conservative, likely reflecting new, larger studies with smaller effect sizes and our different inclusion criteria⁷. Similarly, our estimates of measures of interaction between EBNA titre, HLA status, and MS risk are similar, though not identical, to the published meta-analysis estimates³⁵. The reasons for this discrepancy are not clear.

An EBV vaccine has previously been developed, which prevents IM but not EBV seroconversion³⁶. Newer vaccines are entering phase I clinical studies. The data above indicates that IM carries an odds ratio of 5.11 in individuals with a high genetic risk - preventing IM in this population has the potential to significantly reduce MS risk. Approximately 50% of the general population are EBV seronegative at age 15³⁷; hence vaccination in early adolescence to prevent IM seems an appropriate route both reduce IM-associated morbidity and subsequent EBV-related disease.

Despite the evidence above, not all epidemiological aspects of MS can be explained by EBV infection. The relatively short latency between putative infection and subsequent MS seen in the Faroe epidemics, and the decreasing risk in migrants moving from high- to low-risk areas cannot be explained purely by EBV infection - the fact remains that MS is overwhelmingly likely to be the result of multiple environmental risk modifiers. However, evidence for EBV infection as an obligate step in MS development is increasing, and with vaccination on the horizon as a potential preventive intervention, cannot be ignored.

References

1. Belbasis, L., Bellou, V., Evangelou, E., Ioannidis, J. P. A. & Tzoulaki, I. Environmental risk factors and multiple sclerosis: an umbrella review of systematic reviews and meta-analyses. *Lancet Neurol.* **14**, 263–273 (2015).
2. Waubant, E. *et al.* Environmental and genetic risk factors for MS: an integrated review. *Ann Clin Transl Neurol* (2019). doi:10.1002/acn3.50862
3. Baranzini, S. E., Santaniello, A., Shoostari, P. & Cotsapas, C. The Multiple Sclerosis Genomic Map: Role of peripheral immune cells and resident microglia in susceptibility. *BioRxiv* (2017).
4. Amato, M. P. *et al.* Environmental modifiable risk factors for multiple sclerosis: Report from the 2016ECTRIMS focused workshop. *Mult. Scler.* 1352458516686847 (2017).
5. Thacker, E. L., Mirzaei, F. & Ascherio, A. Infectious mononucleosis and risk for multiple sclerosis: a meta-analysis. *Ann. Neurol.* **59**, 499–503 (2006).
6. Handel, A. E. *et al.* An Updated Meta-Analysis of Risk of Multiple Sclerosis following Infectious Mononucleosis. *PLoS ONE* **5**, e12496 (2010).
7. Almohmeed, Y. H., Avenell, A., Aucott, L. & Vickers, M. A. Systematic review and meta-analysis of the sero-epidemiological association between Epstein Barr virus and multiple sclerosis. *PLoS One* **8**, e61110 (2013).
8. Xiao, D. *et al.* A meta-analysis of interaction between Epstein-Barr virus and HLA-DRB1*1501 on risk of multiple sclerosis. *Sci. Rep.* **5**, 18083 (2015).
9. Ascherio, A. *et al.* Epstein-Barr Virus Antibodies and Risk of Multiple Sclerosis. *JAMA* **286**, 3083 (2001).

10. Levin, L. I. *et al.* Temporal relationship between elevation of epstein-barr virus antibody titers and initial onset of neurological symptoms in multiple sclerosis. *JAMA* **293**, 2496–2500 (2005).
11. Pakpoor, J. *et al.* The risk of developing multiple sclerosis in individuals seronegative for Epstein-Barr virus: a meta-analysis. *Mult. Scler.* **19**, 162–166 (2013).
12. Dobson, R., Kuhle, J., Middelcorp, J. & Giovannoni, G. Epstein-Barr-negative MS: a true phenomenon? *Neurol Neuroimmunol Neuroinflamm* **4**, e318 (2017).
13. Goldacre, M. J., Wotton, C. J., Seagroatt, V. & Yeates, D. Multiple sclerosis after infectious mononucleosis: record linkage study. *J. Epidemiol. Community Health* **58**, 1032–1035 (2004).
14. Marrie, R. A. *et al.* Multiple sclerosis and antecedent infections: a case-control study. *Neurology* **54**, 2307–2310 (2000).
15. Nielsen, T. R. *et al.* Multiple Sclerosis After Infectious Mononucleosis. *Archives of Neurology* **64**, 72 (2007).
16. Olsson, T., Barcellos, L. F. & Alfredsson, L. Interactions between genetic, lifestyle and environmental risk factors for multiple sclerosis. *Nat. Rev. Neurol.* **13**, 25–36 (2017).
17. Sterne, J. A. C. *et al.* Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ* **343**, d4002 (2011).
18. Hosmer, D. W. & Lemeshow, S. Confidence interval estimation of interaction. *Epidemiology* **3**, 452–456 (1992).
19. Andersson, T., Alfredsson, L., Källberg, H., Zdravkovic, S. & Ahlbom, A. Calculating measures of biological interaction. *Eur. J. Epidemiol.* **20**, 575–579 (2005).
20. De Jager, P. L. *et al.* Integrating risk factors: HLA-DRB1*1501 and Epstein-Barr virus in multiple sclerosis. *Neurology* **70**, 1113–1118 (2008).

21. Sundström, P., Nyström, L., Jidell, E. & Hallmans, G. EBNA-1 reactivity and HLA DRB1*1501 as statistically independent risk factors for multiple sclerosis: a case-control study. *Mult. Scler.* **14**, 1120–1122 (2008).
22. Sundqvist, E. *et al.* Epstein-Barr virus and multiple sclerosis: interaction with HLA. *Genes Immun.* **13**, 14–20 (2012).
23. Pandit, L., Malli, C., D’Cunha, A., Shetty, R. & Singhal, B. Association of Epstein-Barr virus infection with multiple sclerosis in India. *J. Neurol. Sci.* **325**, 86–89 (2013).
24. Lucas, R. M. *et al.* Current and past Epstein-Barr virus infection in risk of initial CNS demyelination. *Neurology* **77**, 371–379 (2011).
25. Simon, K. C. *et al.* Combined effects of smoking, anti-EBNA antibodies, and HLA-DRB1*1501 on multiple sclerosis risk. *Neurology* **74**, 1365–1371 (2010).
26. Sundqvist, E. *et al.* Lack of replication of interaction between EBNA1 IgG and smoking in risk for multiple sclerosis. *Neurology* **79**, 1363–1368 (2012).
27. van der Mei, I. A. F. *et al.* Human leukocyte antigen-DR15, low infant sibling exposure and multiple sclerosis: gene-environment interaction. *Ann. Neurol.* **67**, 261–265 (2010).
28. van der Mei, I. *et al.* Population attributable fractions and joint effects of key risk factors for multiple sclerosis. *Mult. Scler.* **22**, 461–469 (2016).
29. Nielsen, T. R. *et al.* Effects of infectious mononucleosis and HLA-DRB1*15 in multiple sclerosis. *Mult. Scler.* **15**, 431–436 (2009).
30. Lavery, A. M. *et al.* The contribution of secondhand tobacco smoke exposure to pediatric multiple sclerosis risk. *Mult. Scler.* **25**, 515–522 (2019).
31. Salzer, J., Stenlund, H. & Sundström, P. The interaction between smoking and Epstein-Barr virus as multiple sclerosis risk factors may depend on age. *Mult. Scler.* **20**, 747–750 (2014).
32. Salzer, J. *et al.* Epstein-Barr virus antibodies and vitamin D in prospective multiple sclerosis biobank samples. *Mult. Scler.* **19**, 1587–1591 (2013).

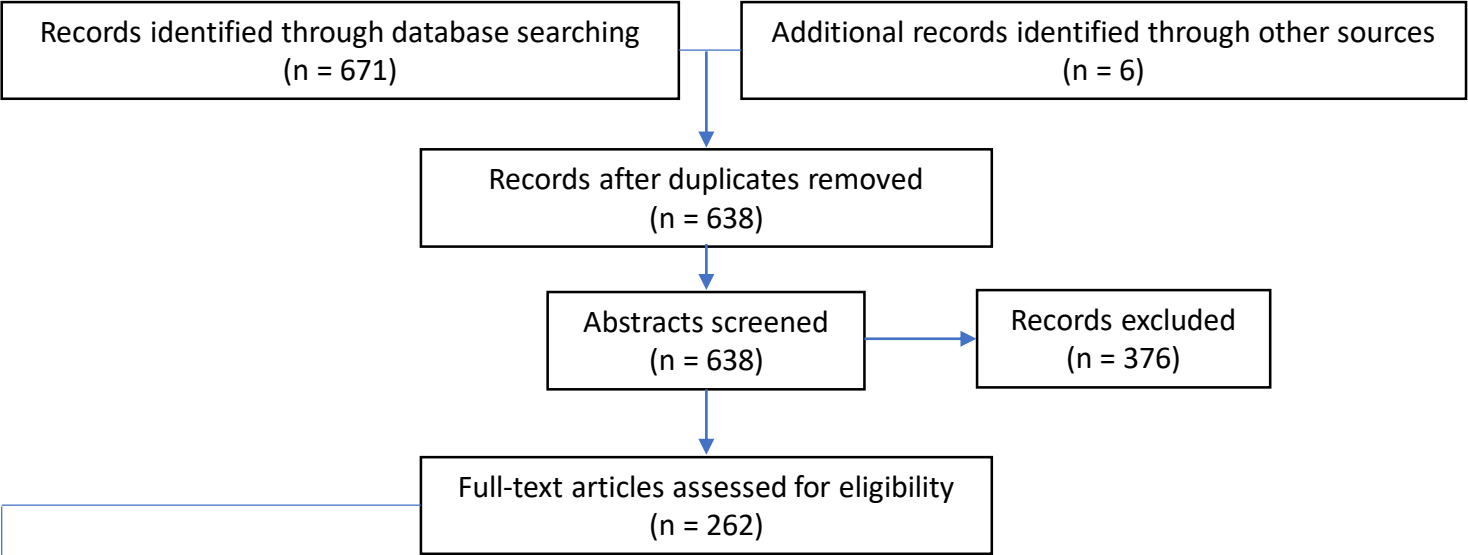
33. Décard, B. F. *et al.* Low vitamin D and elevated immunoreactivity against Epstein-Barr virus before first clinical manifestation of multiple sclerosis. *J. Neurol. Neurosurg. Psychiatry* **83**, 1170–1173 (2012).
34. Hedström, A. K., Lima Bomfim, I., Hillert, J., Olsson, T. & Alfredsson, L. Obesity interacts with infectious mononucleosis in risk of multiple sclerosis. *Eur. J. Neurol.* **22**, 578–e38 (2015).
35. Xiao, D. *et al.* A meta-analysis of interaction between Epstein-Barr virus and HLA-DRB1*1501 on risk of multiple sclerosis. *Sci. Rep.* **5**, 18083 (2015).
36. Sokal, E. M. *et al.* Recombinant gp350 vaccine for infectious mononucleosis: a phase 2, randomized, double-blind, placebo-controlled trial to evaluate the safety, immunogenicity, and efficacy of an Epstein-Barr virus vaccine in healthy young adults. *J. Infect. Dis.* **196**, 1749–1753 (2007).
37. Morris, M. C. *et al.* Sero-epidemiological patterns of Epstein-Barr and herpes simplex (HSV-1 and HSV-2) viruses in England and Wales. *J. Med. Virol.* **67**, 522–527 (2002).

Identification

Screening

Eligibility

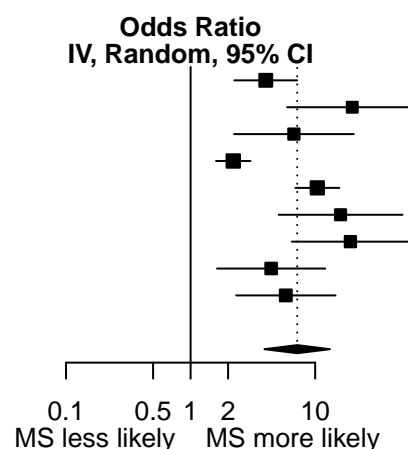
Included



Meta-analysis	Full-text articles assessed for eligibility	Studies included in quantitative synthesis	Full-text articles excluded
Seropositivity & MS	127	56	71
IM & MS	32	19	13
EBV DNA & MS	31	23	8
HLA, EBV, & MS	10	10	0
IM, EBV & MS	4	4	0
Smoking, EBV, & MS	5	5	0

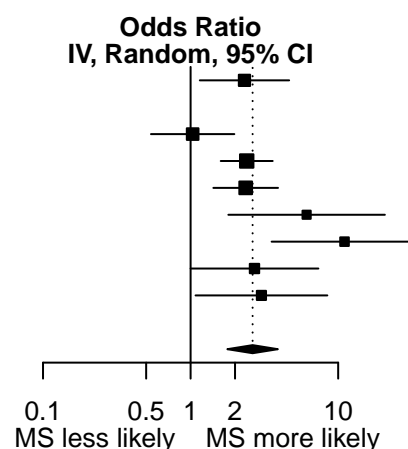
Study	ln(OR)	SE(ln(OR))	Weight	OR [95% CI]
De Jager 2008	1.39	0.2965	12.8%	4.01 [2.24; 7.17]
Lucas 2011	2.99	0.6144	9.2%	19.84 [5.95; 66.15]
Pandit 2013	1.91	0.5641	9.8%	6.75 [2.23; 20.38]
Simon 2010	0.79	0.1625	13.9%	2.20 [1.60; 3.02]
Sundqvist 2011	2.34	0.2068	13.6%	10.40 [6.93; 15.59]
Sundstrom 2008	2.77	0.5837	9.5%	15.98 [5.09; 50.17]
van der Mei 2010	2.95	0.5526	9.9%	19.17 [6.49; 56.62]
van der Mei 2010	1.49	0.5101	10.4%	4.43 [1.63; 12.04]
Van der Mei 2016	1.76	0.4684	10.9%	5.81 [2.32; 14.55]

Total (95% CI) **100.0%** **7.17 [3.92; 13.14]**
Heterogeneity: $\tau^2 = 0.6583$; $\chi^2 = 53.80$, $df = 8$ ($P < 0.01$); $I^2 = 85\%$



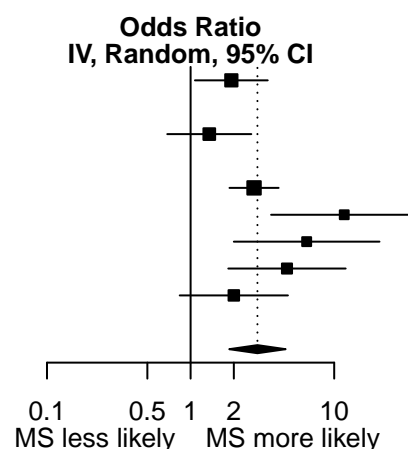
Study	ln(OR)	SE(ln(OR))	Weight	OR [95% CI]
De Jager 2008	0.84	0.3557	13.9%	2.32 [1.15; 4.65]
Lucas 2011	.	.	0.0%	
Pandit 2013	0.03	0.3319	14.8%	1.03 [0.54; 1.98]
Simon 2010	0.88	0.2069	19.6%	2.40 [1.60; 3.60]
Sundqvist 2011	0.86	0.2574	17.6%	2.36 [1.43; 3.91]
Sundstrom 2008	1.81	0.6231	7.3%	6.11 [1.80; 20.72]
van der Mei 2010	2.40	0.5803	8.1%	11.04 [3.54; 34.43]
van der Mei 2010	1.00	0.5086	9.6%	2.71 [1.00; 7.34]
Van der Mei 2016	1.11	0.5246	9.2%	3.02 [1.08; 8.44]

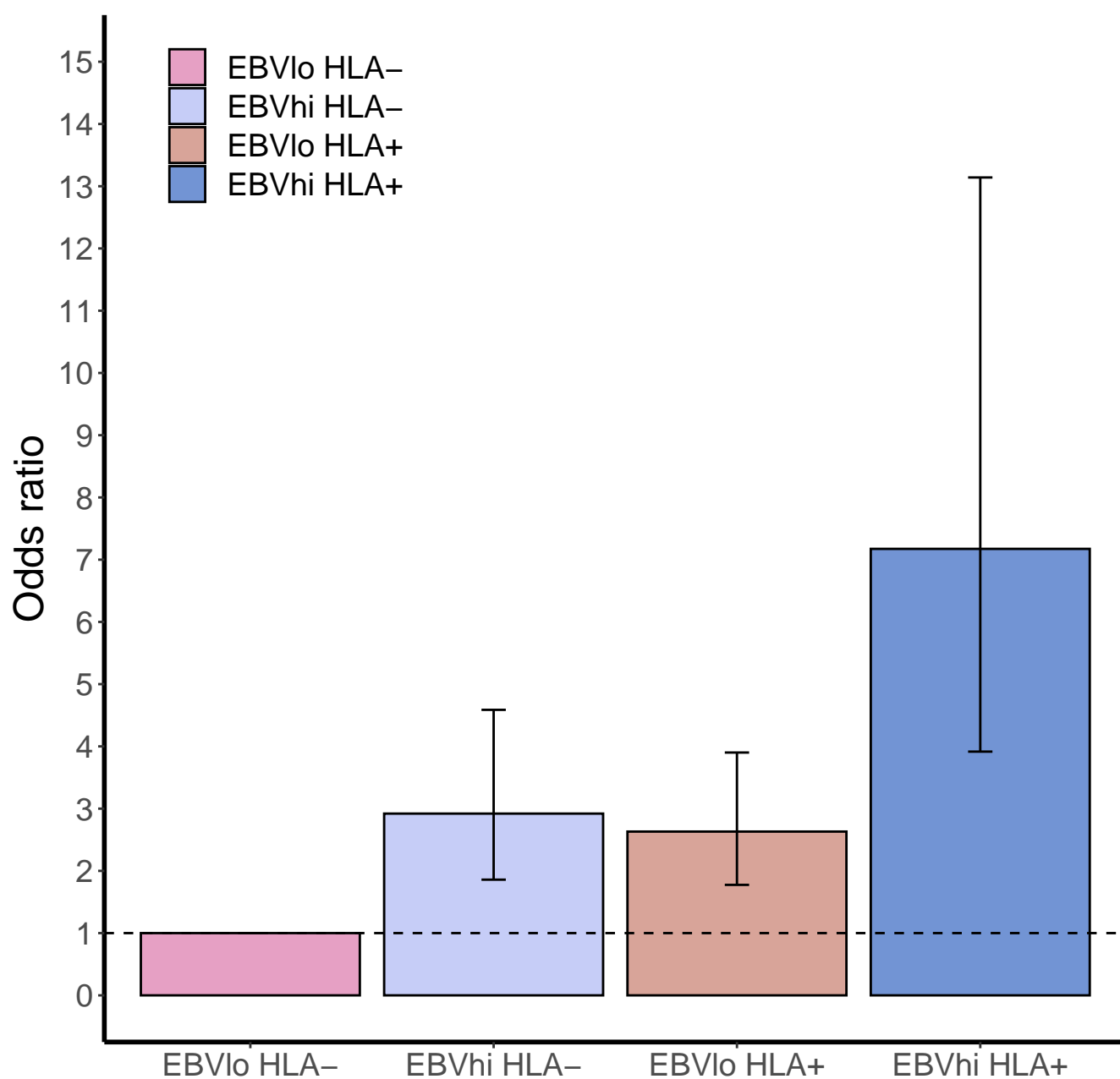
Total (95% CI) 100.0% **2.63 [1.78; 3.90]**
Heterogeneity: $\tau^2 = 0.1632$; $\chi^2 = 15.89$, $df = 7$ ($P = 0.03$); $I^2 = 56\%$



Study	ln(OR)	SE(ln(OR))	Weight	OR [95% CI]
De Jager 2008	0.65	0.2977	18.0%	1.92 [1.07; 3.44]
Lucas 2011	.	.	0.0%	
Pandit 2013	0.30	0.3438	16.3%	1.35 [0.69; 2.65]
Simon 2010	.	.	0.0%	
Sundqvist 2011	1.02	0.2010	21.5%	2.77 [1.87; 4.10]
Sundstrom 2008	2.46	0.5985	9.4%	11.75 [3.64; 37.98]
van der Mei 2010	1.86	0.5958	9.5%	6.43 [2.00; 20.67]
van der Mei 2010	1.55	0.4802	12.1%	4.69 [1.83; 12.02]
Van der Mei 2016	0.69	0.4426	13.2%	2.00 [0.84; 4.76]

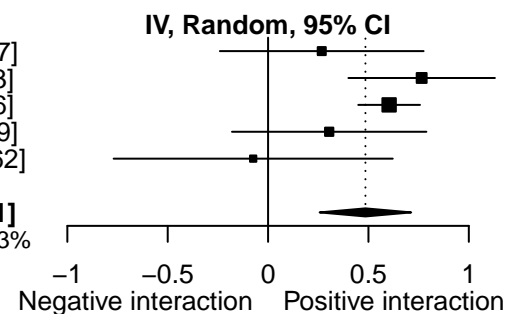
Total (95% CI) **100.0%** **2.92 [1.86; 4.59]**
Heterogeneity: $\tau^2 = 0.2068$; $\chi^2 = 15.28$, $df = 6$ ($P = 0.02$); $I^2 = 61\%$





Study	AP	SE	Weight	AP [95% CI]
De Jager 2008	0.27	0.2589	14.3%	0.27 [-0.24; 0.77]
Pandit 2013	0.77	0.1862	21.7%	0.77 [0.40; 1.13]
Sundqvist 2011	0.60	0.0784	39.9%	0.60 [0.45; 0.76]
Sundstrom 2008	0.30	0.2470	15.3%	0.30 [-0.18; 0.79]
van der Mei 2010	-0.07	0.3547	8.8%	-0.07 [-0.77; 0.62]

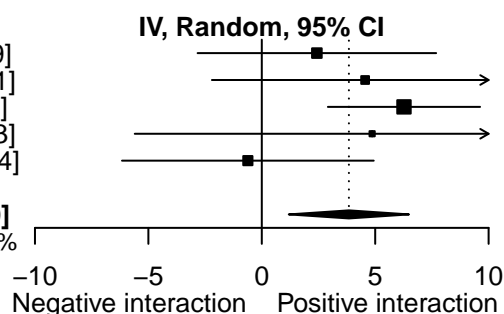
Total (95% CI) **100.0%** **0.48 [0.26; 0.71]**
Heterogeneity: $\tau^2 = 0.0278$; $\chi^2 = 7.06$, $df = 4$ ($P = 0.13$); $I^2 = 43\%$

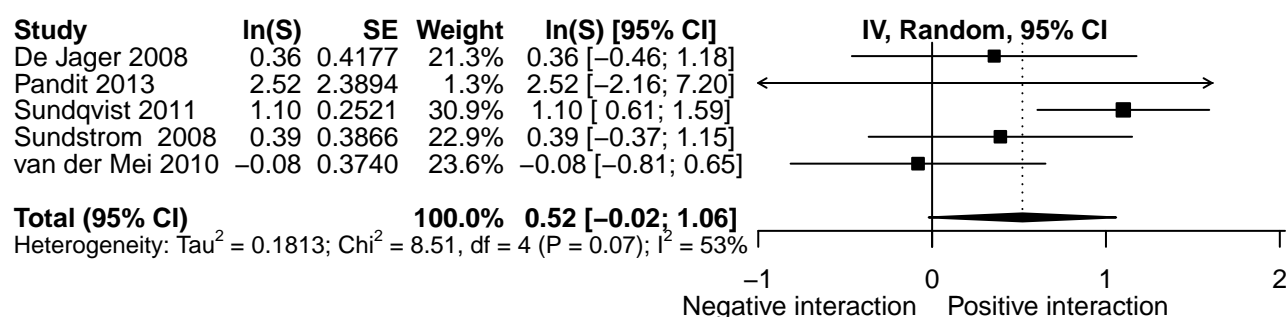


Study	RERI	SE	Weight	RERI [95% CI]
De Jager 2008	2.43	2.6830	20.8%	2.43 [-2.83; 7.69]
Pandit 2013	4.56	3.4463	13.6%	4.56 [-2.20; 11.31]
Sundqvist 2011	6.27	1.7115	40.5%	6.27 [2.91; 9.62]
Sundstrom 2008	4.86	5.3396	6.1%	4.86 [-5.60; 15.33]
van der Mei 2010	-0.61	2.8314	19.0%	-0.61 [-6.16; 4.94]

Total (95% CI) **100.0%** **3.84 [1.19; 6.49]**

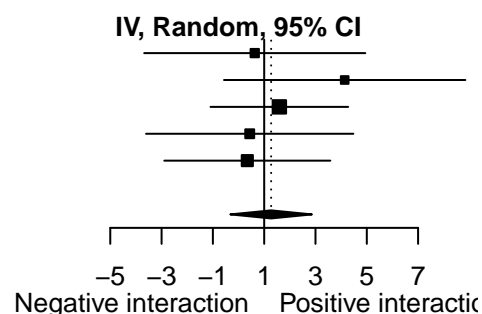
Heterogeneity: $\tau^2 = 1.5954$; $\chi^2 = 4.80$, $df = 4$ ($P = 0.31$); $I^2 = 17\%$





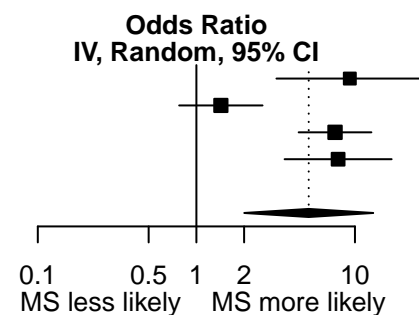
Study	Interaction term	SE	Weight	MIT [95% CI]
De Jager 2008	0.63	2.1994	13.6%	0.63 [-3.68; 4.94]
Pandit 2013	4.14	2.4002	11.5%	4.14 [-0.57; 8.84]
Sundqvist 2011	1.59	1.3691	35.2%	1.59 [-1.09; 4.27]
Sundstrom 2008	0.44	2.0614	15.5%	0.44 [-3.60; 4.48]
van der Mei 2010	0.34	1.6515	24.2%	0.34 [-2.89; 3.58]
Total (95% CI)			100.0%	1.27 [-0.32; 2.86]

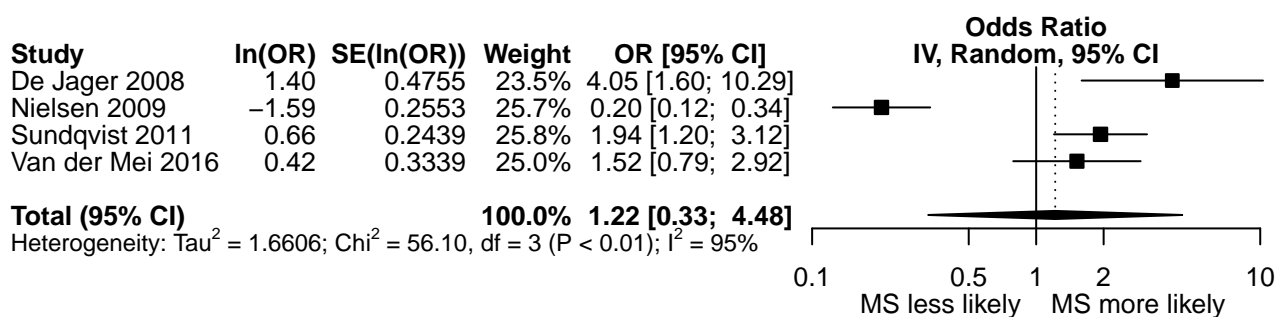
Heterogeneity: $\tau^2 = 0$; $\chi^2 = 2.05$, $df = 4$ ($P = 0.73$); $I^2 = 0\%$

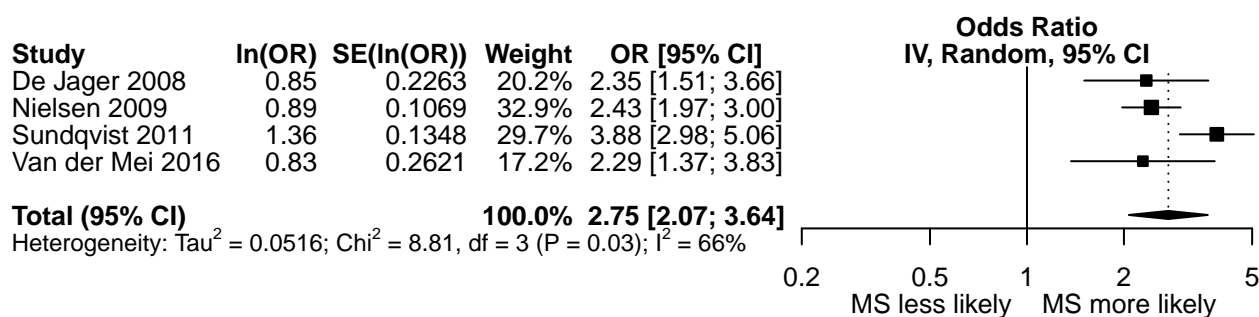


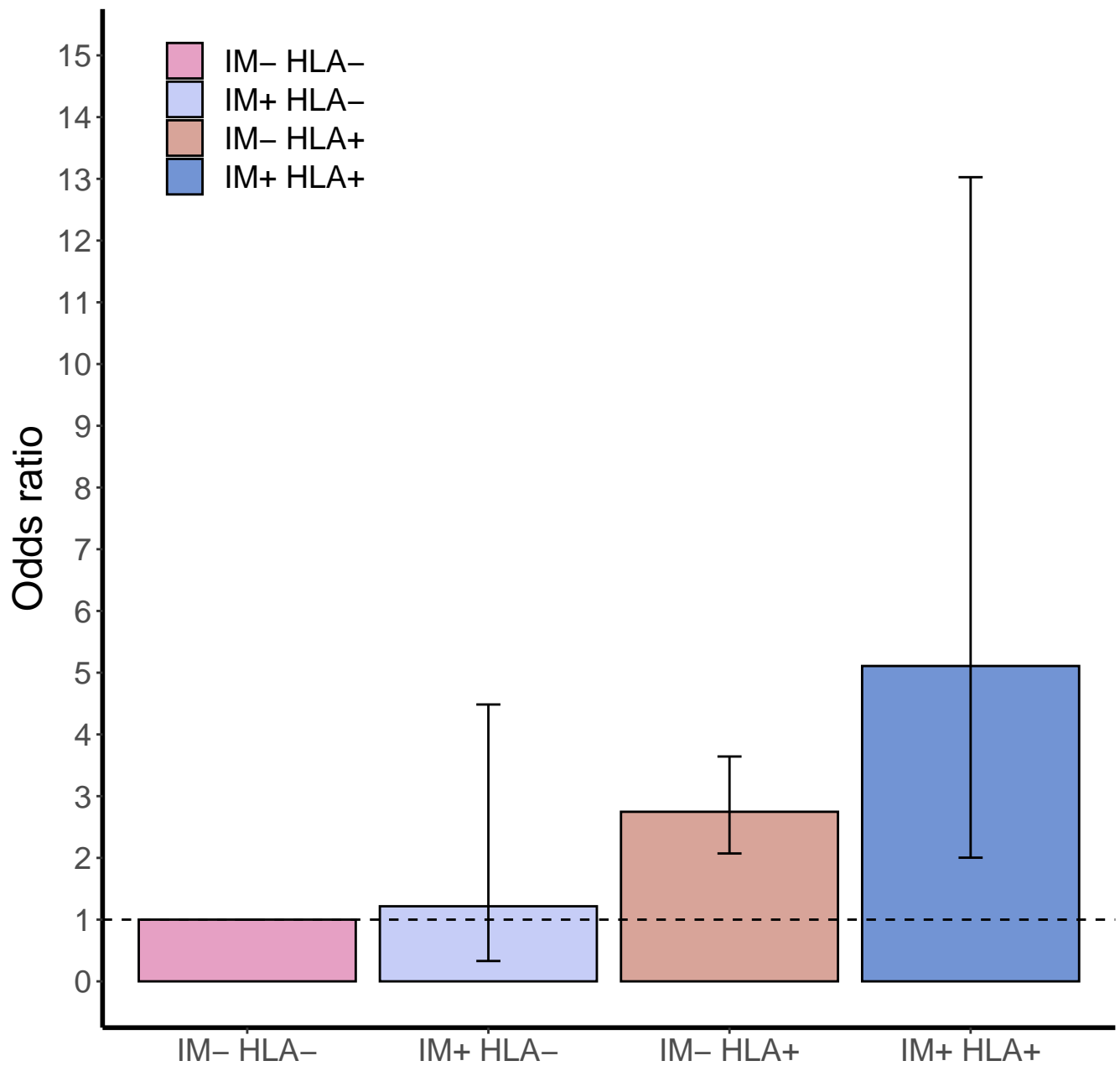
Study	ln(OR)	SE(ln(OR))	Weight	OR [95% CI]
De Jager 2008	2.23	0.5436	21.5%	9.28 [3.20; 26.94]
Nielsen 2009	0.36	0.3088	26.5%	1.43 [0.78; 2.61]
Sundqvist 2011	2.01	0.2695	27.2%	7.49 [4.42; 12.70]
Van der Mei 2016	2.06	0.3964	24.7%	7.83 [3.60; 17.03]
Total (95% CI)			100.0%	5.11 [2.00; 13.03]

Heterogeneity: $\tau^2 = 0.7649$; $\chi^2 = 21.10$, $df = 3$ ($P < 0.01$); $I^2 = 86\%$

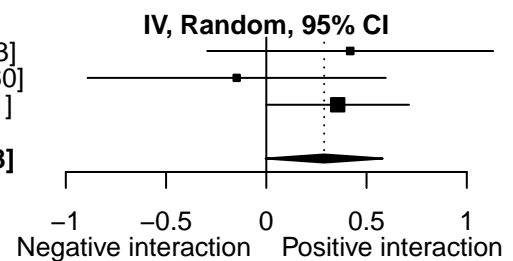


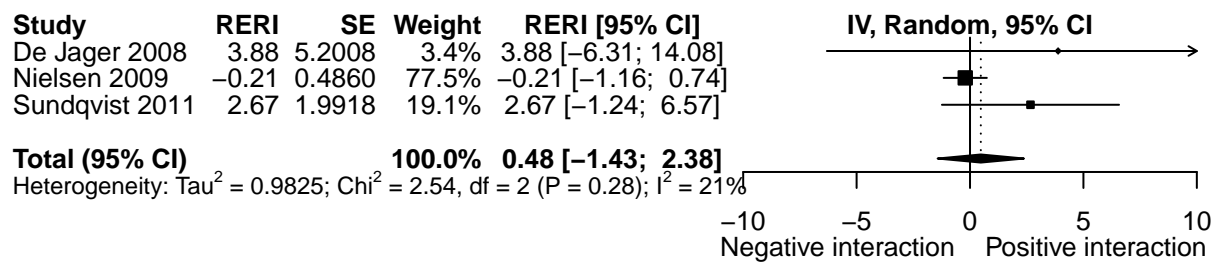


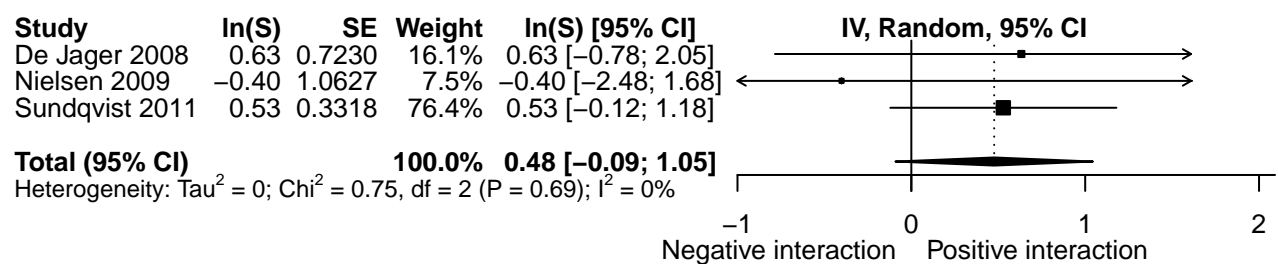




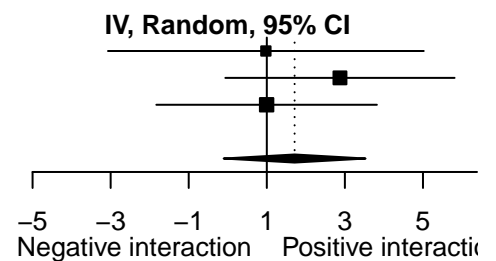
Study	AP	SE	Weight	AP [95% CI]
De Jager 2008	0.42	0.3637	16.8%	0.42 [-0.29; 1.13]
Nielsen 2009	-0.15	0.3793	15.5%	-0.15 [-0.89; 0.60]
Sundqvist 2011	0.36	0.1813	67.7%	0.36 [0.00; 0.71]
Total (95% CI)			100.0%	0.29 [0.00; 0.58]
Heterogeneity: $\tau^2 = 0$; $\chi^2 = 1.59$, $df = 2$ ($P = 0.45$); $I^2 = 0\%$				





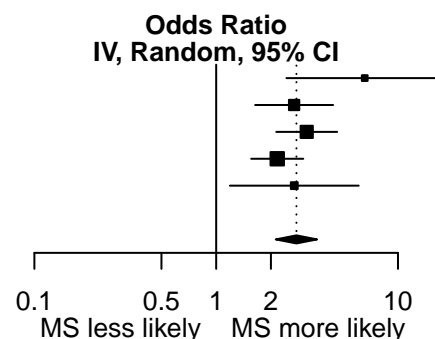


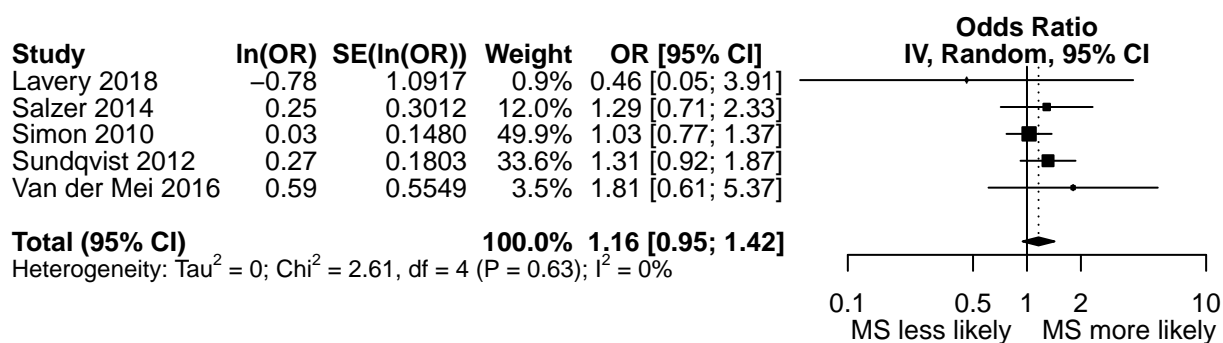
Study	Interaction term	SE	Weight	MIT [95% CI]
De Jager 2008	0.98	2.0634	20.2%	0.98 [-3.07; 5.02]
Nielsen 2009	2.87	1.4980	38.4%	2.87 [-0.06; 5.81]
Sundqvist 2011	1.00	1.4413	41.4%	1.00 [-1.83; 3.82]
Total (95% CI)			100.0%	1.71 [-0.11; 3.53]
Heterogeneity: $\tau^2 = 0$; $\chi^2 = 0.98$, $df = 2$ ($P = 0.61$); $I^2 = 0\%$				



Study	ln(OR)	SE(ln(OR))	Weight	OR [95% CI]
Lavery 2018	1.88	0.5059	6.4%	6.56 [2.43; 17.69]
Salzer 2014	0.99	0.2519	20.7%	2.68 [1.64; 4.40]
Simon 2010	1.15	0.1968	29.0%	3.15 [2.14; 4.63]
Sundqvist 2012	0.77	0.1683	34.9%	2.17 [1.56; 3.01]
Van der Mei 2016	0.99	0.4161	9.1%	2.69 [1.19; 6.08]

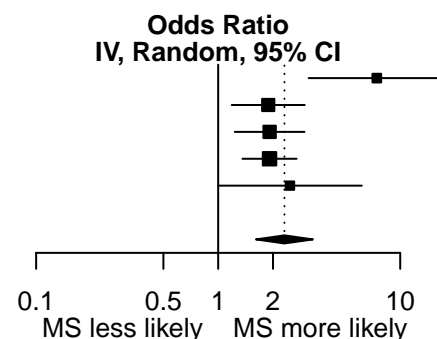
Total (95% CI) **100.0%** **2.76 [2.13; 3.59]**
Heterogeneity: $\tau^2 = 0.0229$; $\chi^2 = 5.40$, $df = 4$ ($P = 0.25$); $I^2 = 26\%$

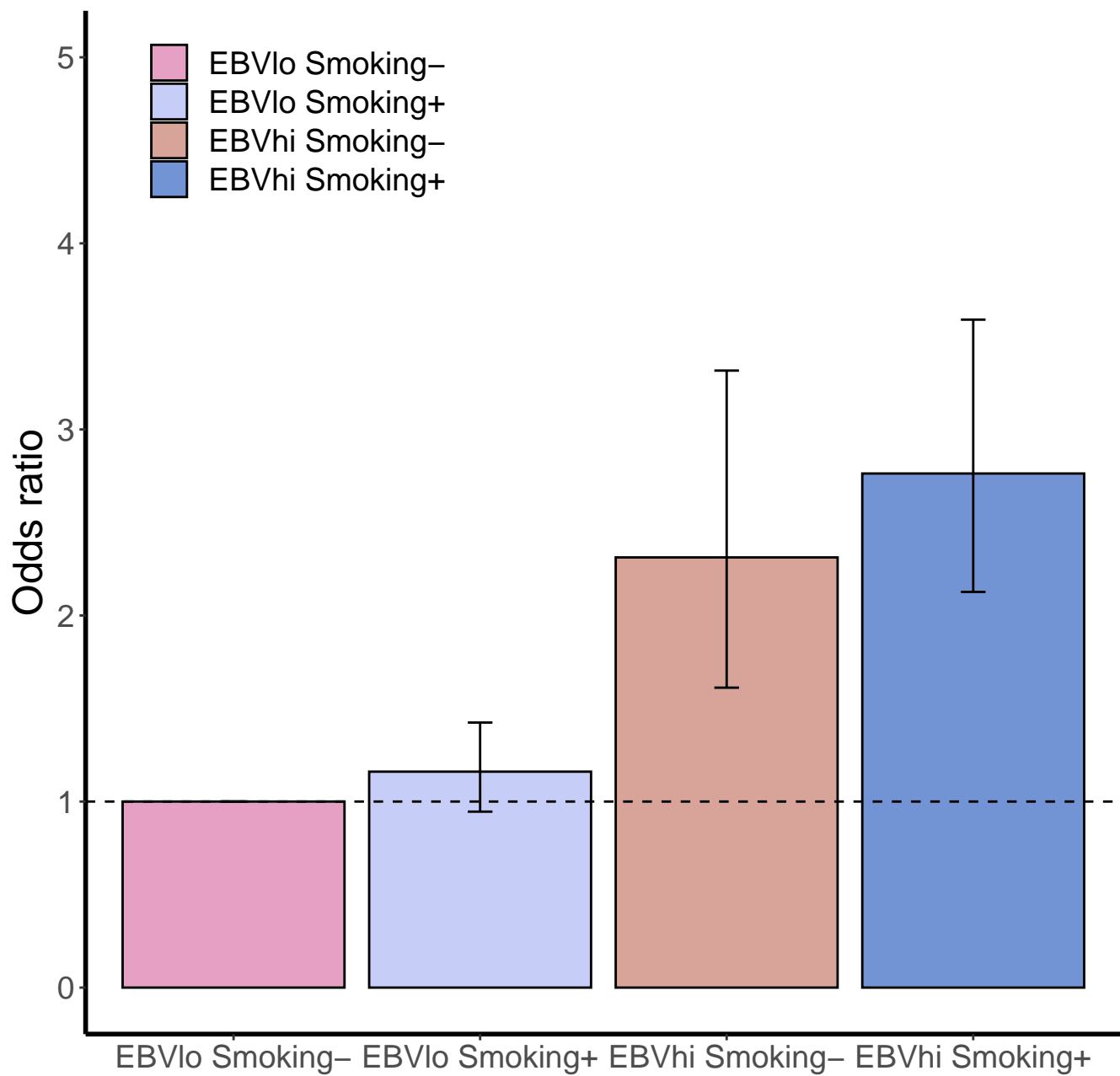


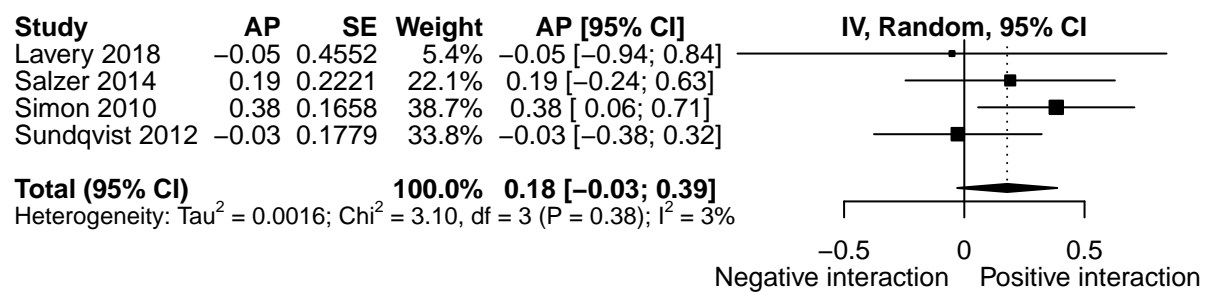


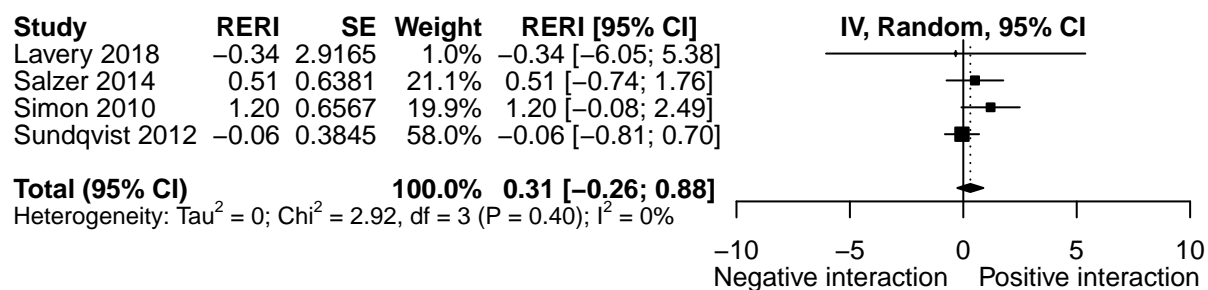
Study	ln(OR)	SE(ln(OR))	Weight	OR [95% CI]
Lavery 2018	2.01	0.4401	12.0%	7.44 [3.14; 17.62]
Salzer 2014	0.63	0.2355	23.7%	1.88 [1.19; 2.99]
Simon 2010	0.65	0.2251	24.5%	1.92 [1.23; 2.98]
Sundqvist 2012	0.65	0.1748	28.6%	1.91 [1.36; 2.70]
Van der Mei 2016	0.91	0.4634	11.2%	2.48 [1.00; 6.15]

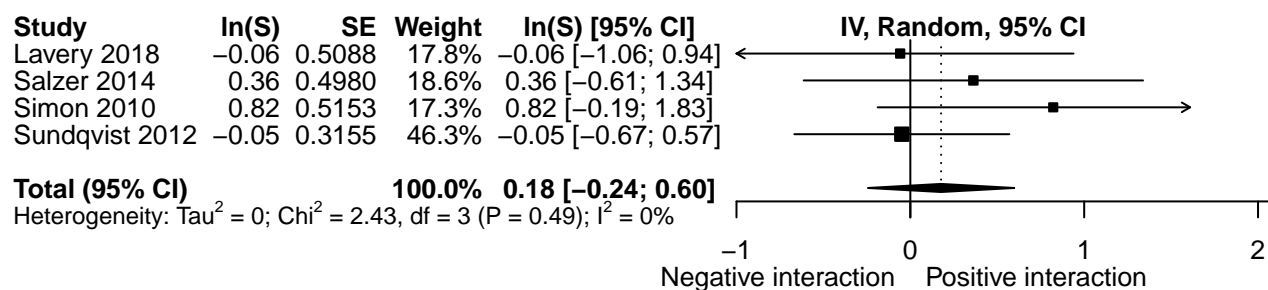
Total (95% CI) **100.0%** **2.31 [1.61; 3.32]**
Heterogeneity: $\tau^2 = 0.0877$; $\chi^2 = 9.04$, $df = 4$ ($P = 0.06$); $I^2 = 56\%$





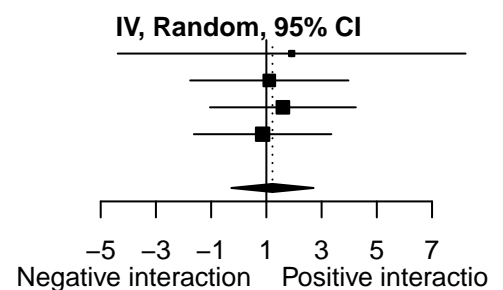


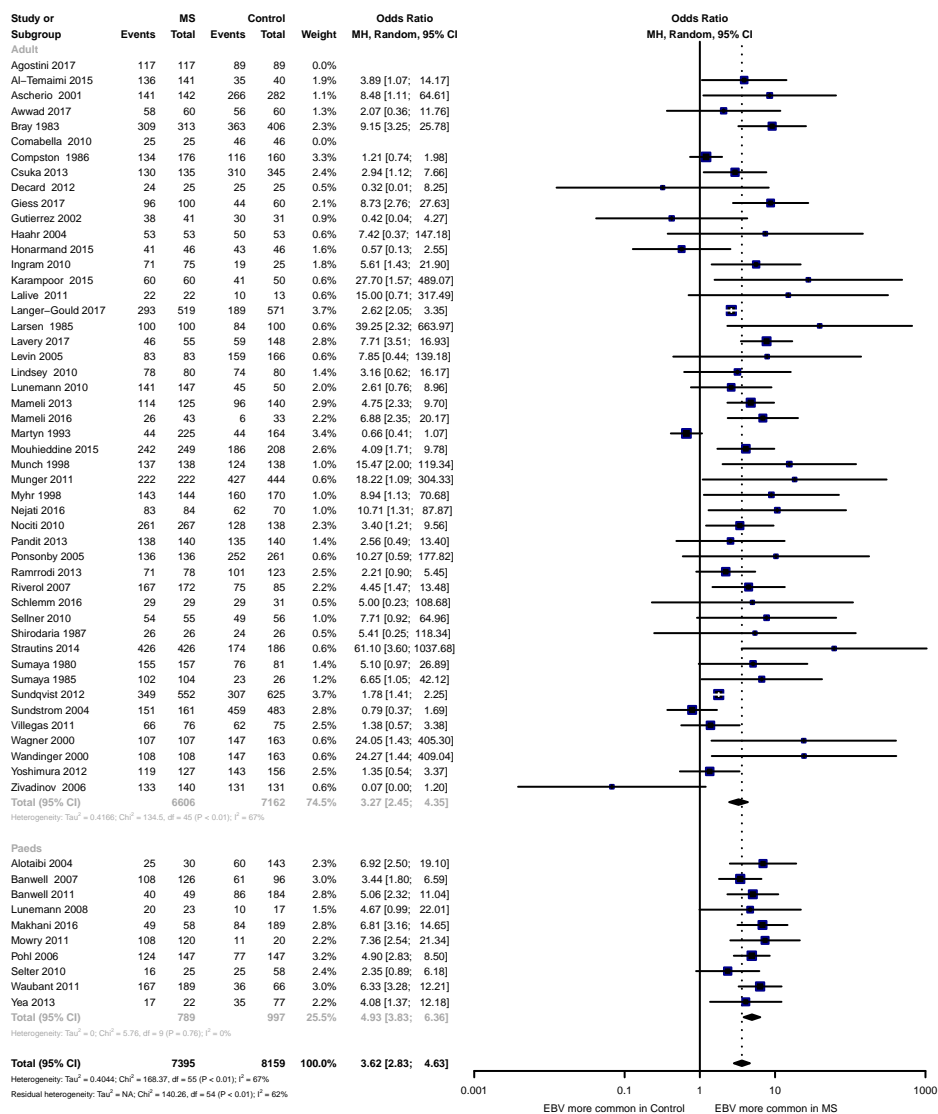


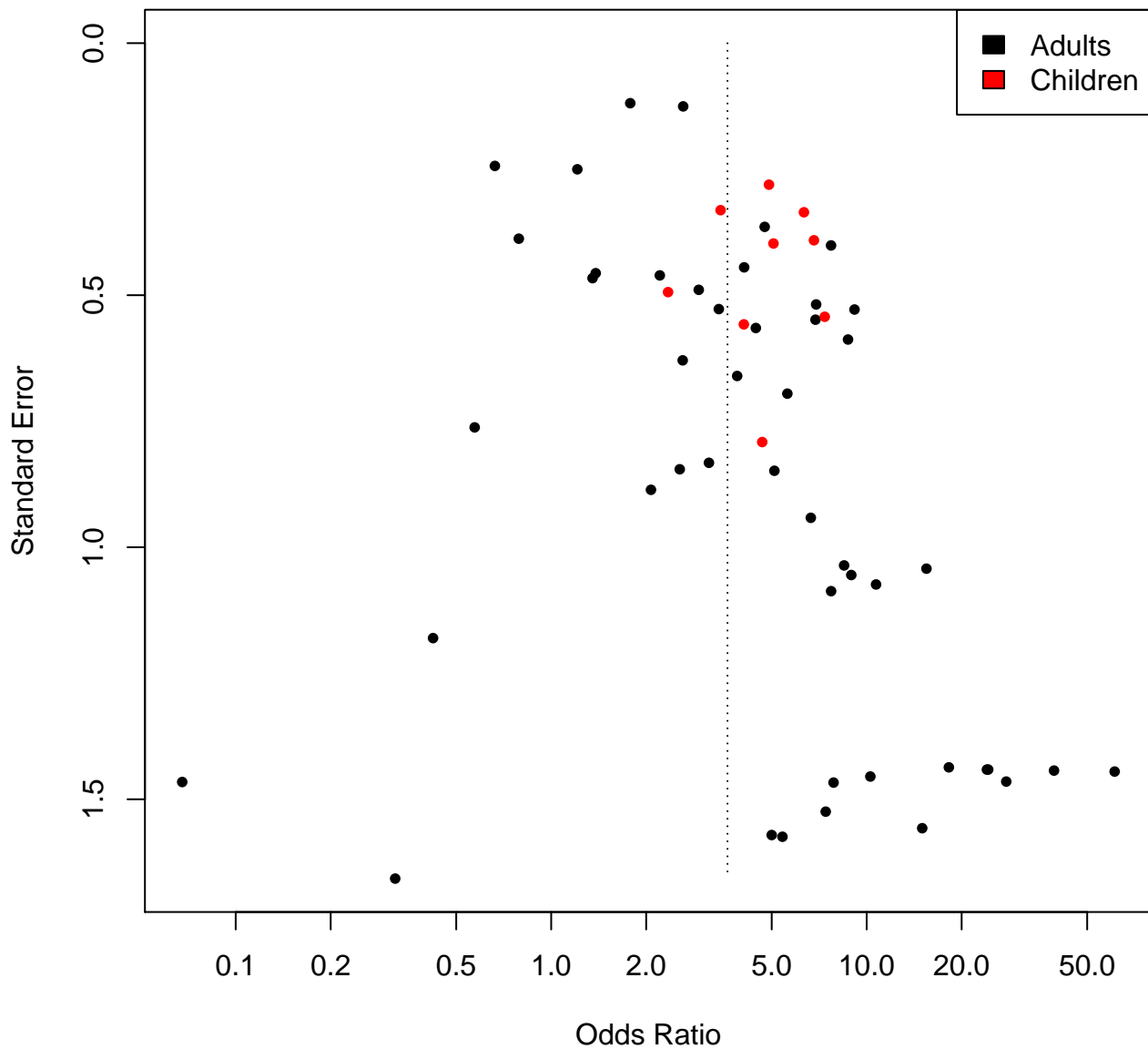


Study	Interaction term	SE	Weight	MIT [95% CI]
Lavery 2018	1.92	3.2137	5.6%	1.92 [-4.38; 8.21]
Salzer 2014	1.11	1.4622	27.0%	1.11 [-1.76; 3.97]
Simon 2010	1.60	1.3479	31.7%	1.60 [-1.04; 4.24]
Sundqvist 2012	0.86	1.2703	35.7%	0.86 [-1.63; 3.35]
Total (95% CI)			100.0%	1.22 [-0.27; 2.71]

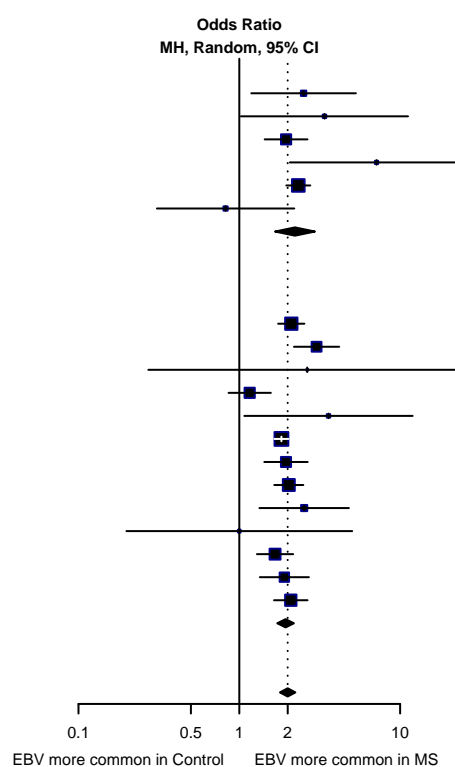
Heterogeneity: $\tau^2 = 0$; $\chi^2 = 0.21$, $df = 3$ ($P = 0.98$); $I^2 = 0\%$

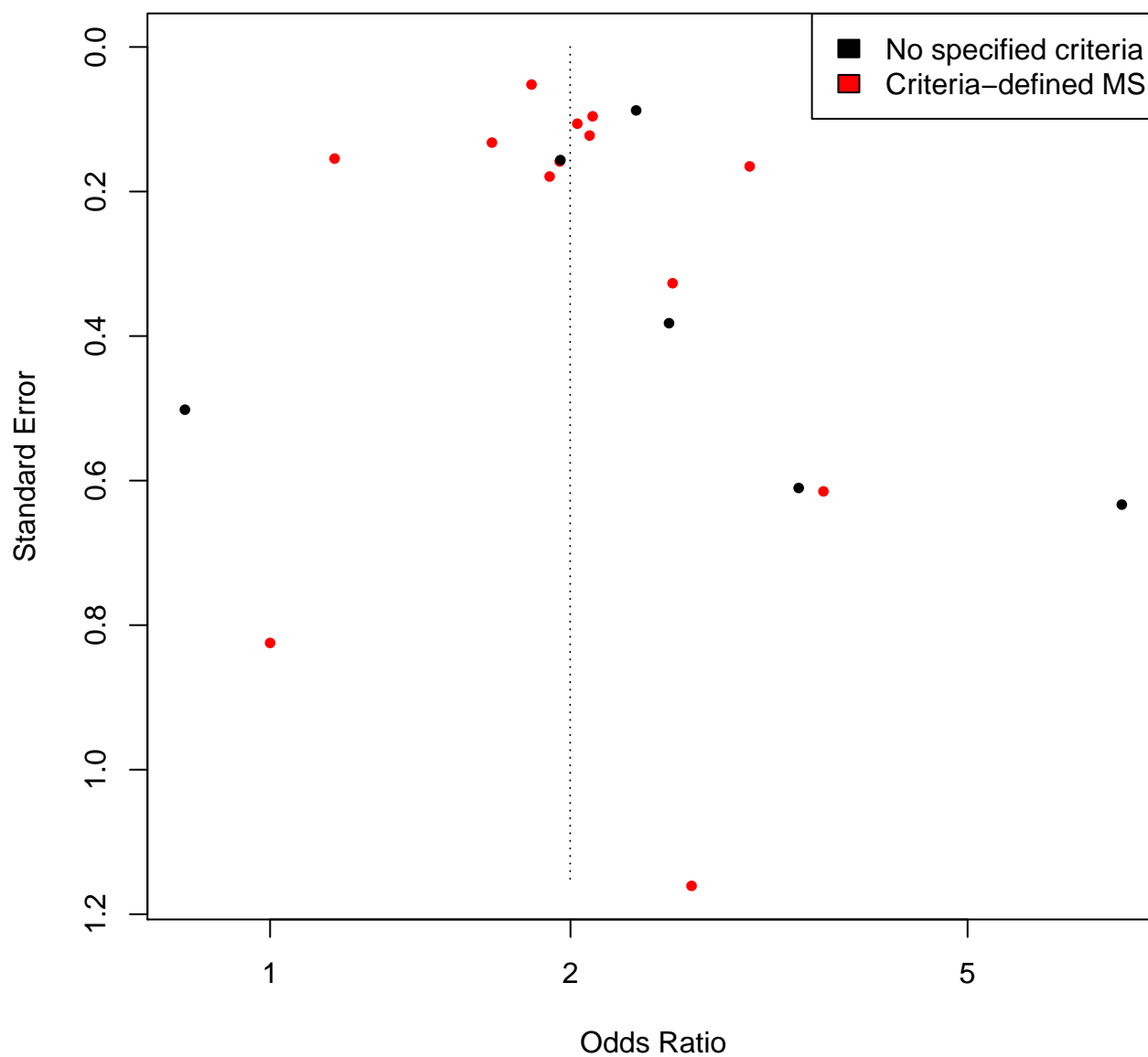






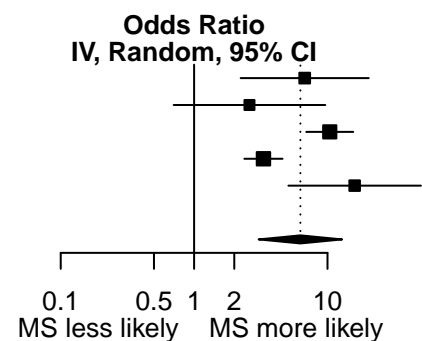
Study or Subgroup	Events	MS Total	Control Events	Control Total	Weight	Odds Ratio MH, Random, 95% CI
Specific_criteria = No						
Haahr 1995	16	28	6837	19711	1.9%	2.51 [1.19; 5.31]
Marrie 2000	5	225	6	900	0.8%	3.39 [1.02; 11.20]
Nielsen 2007	104	172	11079	25234	6.8%	1.95 [1.44; 2.66]
Operskalsi 1989	19	145	3	145	0.7%	7.14 [2.06; 24.69]
Ramagopalan 2009	699	14362	165	7671	10.5%	2.33 [1.96; 2.76]
Zorzon 2003	8	140	9	131	1.2%	0.82 [0.31; 2.20]
Total (95% CI)		15072		53792	21.9%	2.22 [1.67; 2.95]
Heterogeneity: $\tau^2 = 0.0443$; $\chi^2 = 8.86$, $df = 5$ ($P = 0.11$); $I^2 = 44\%$						
Specific_criteria = Yes						
Bjornevik 2016	242	1904	239	3694	10.0%	2.10 [1.74; 2.54]
Claire Simon 2015	315	1065	50	410	6.4%	3.02 [2.19; 4.18]
Gusev 1994	3	155	1	135	0.2%	2.64 [0.27; 25.73]
Gustavsen 2014	84	756	106	1090	6.9%	1.16 [0.86; 1.57]
Haahr 2004	12	53	4	53	0.8%	3.59 [1.07; 11.97]
Hedstrom 2014	988	6282	734	7924	12.6%	1.83 [1.65; 2.02]
Hernan 2001	70	301	178	1324	6.7%	1.95 [1.43; 2.66]
Lossius 2014	199	1462	198	2751	9.4%	2.03 [1.65; 2.50]
Martyn 1993	43	225	14	164	2.5%	2.53 [1.33; 4.80]
Souberbielle 1990	3	153	3	153	0.5%	1.00 [0.20; 5.03]
Sundqvist 2011	114	678	161	1490	7.9%	1.67 [1.29; 2.16]
van der Mei 2016	74	258	92	528	5.8%	1.91 [1.34; 2.71]
Zaadstra 2008	226	2821	102	2550	8.5%	2.09 [1.64; 2.66]
Total (95% CI)		16113		22266	78.1%	1.94 [1.72; 2.18]
Heterogeneity: $\tau^2 = 0.0188$; $\chi^2 = 24.33$, $df = 12$ ($P = 0.02$); $I^2 = 51\%$						
Total (95% CI)	31185		76058		100.0%	2.00 [1.79; 2.23]
Heterogeneity: $\tau^2 = 0.0224$; $\chi^2 = 37.28$, $df = 18$ ($P < 0.01$); $I^2 = 52\%$						
Residual heterogeneity: $\tau^2 = NA$; $\chi^2 = 33.19$, $df = 17$ ($P = 0.01$); $I^2 = 49\%$						





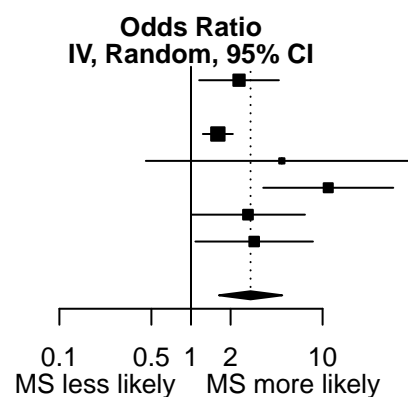
Study	ln(OR)	SE(ln(OR))	Weight	OR [95% CI]
Pandit 2013	1.91	0.5641	16.9%	6.75 [2.23; 20.38]
Simon_Swedish 2010	0.96	0.6675	14.6%	2.60 [0.70; 9.62]
Sundqvist 2011	2.34	0.2068	25.7%	10.40 [6.93; 15.59]
Sundqvist 2012	1.20	0.1683	26.4%	3.31 [2.38; 4.60]
Sundstrom 2008	2.77	0.5837	16.4%	15.98 [5.09; 50.17]
Total (95% CI)			100.0%	6.26 [3.05; 12.88]

Heterogeneity: $\tau^2 = 0.4837$; $\chi^2 = 23.30$, $df = 4$ ($P < 0.01$); $I^2 = 83\%$

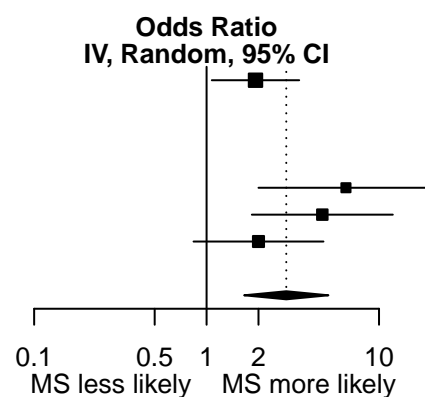


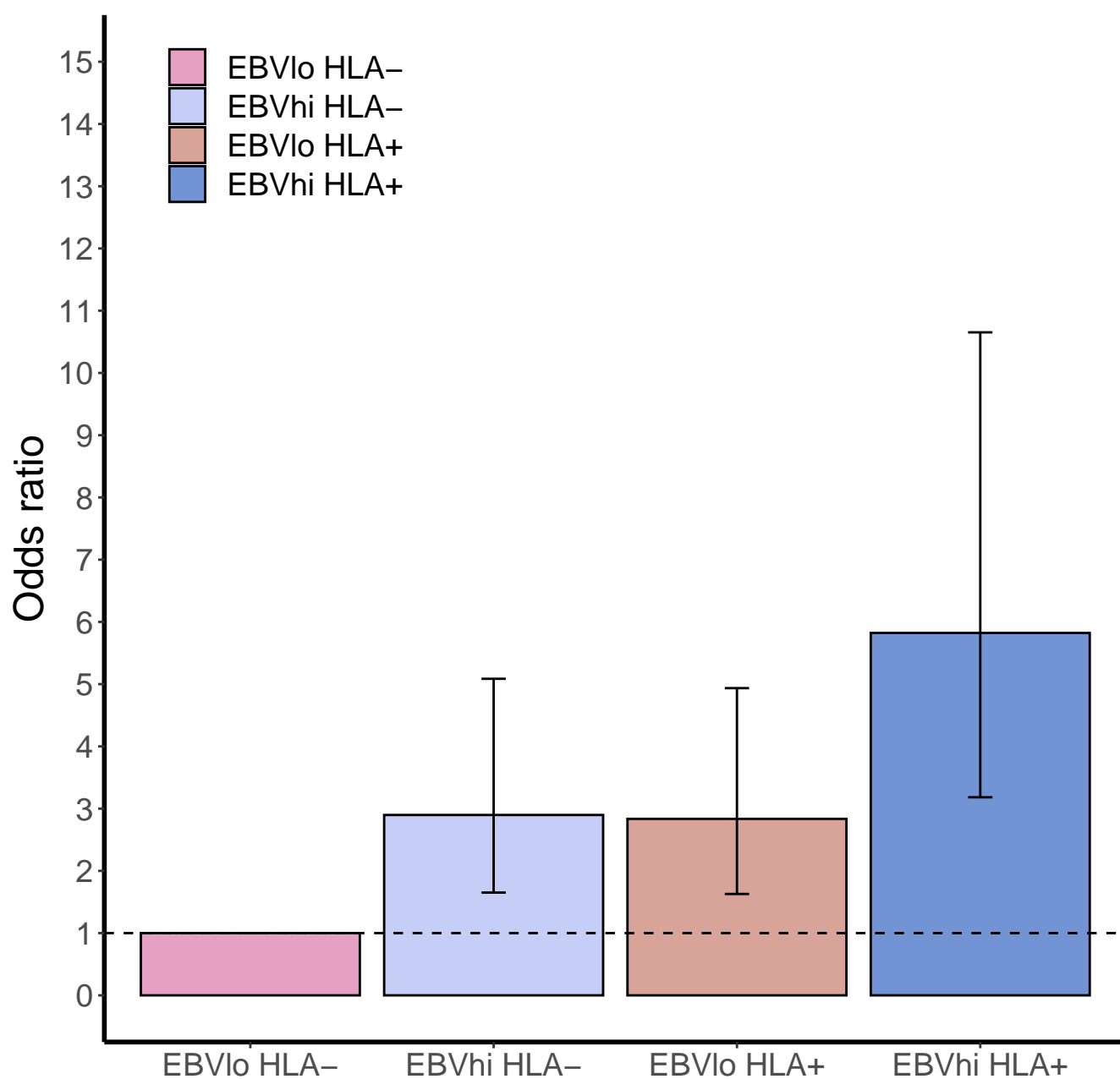
Study	ln(OR)	SE(ln(OR))	Weight	OR [95% CI]
De Jager 2008	0.84	0.3557	21.2%	2.32 [1.15; 4.65]
Lucas 2011	.	.	0.0%	
Simon_NHS 2010	0.47	0.1339	29.7%	1.60 [1.23; 2.08]
Simon_Tas 2010	1.59	1.2141	4.6%	4.90 [0.45; 52.92]
van der Mei 2010	2.40	0.5803	13.6%	11.04 [3.54; 34.43]
van der Mei 2010	1.00	0.5086	15.7%	2.71 [1.00; 7.34]
Van der Mei 2016	1.11	0.5246	15.2%	3.02 [1.08; 8.44]
Total (95% CI)			100.0%	2.84 [1.63; 4.94]

Heterogeneity: $\tau^2 = 0.2513$; $\chi^2 = 13.02$, $df = 5$ ($P = 0.02$); $I^2 = 62\%$



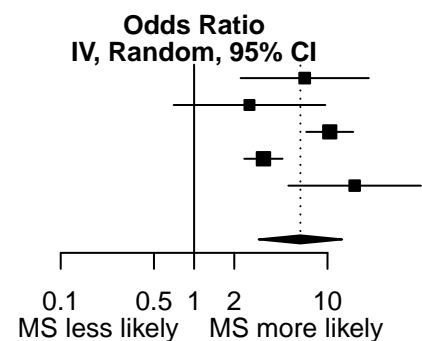
Study	ln(OR)	SE(ln(OR))	Weight	OR [95% CI]
De Jager 2008	0.65	0.2977	36.3%	1.92 [1.07; 3.44]
Lucas 2011	.	.	0.0%	
Simon_NHS 2010	.	.	0.0%	
Simon_Tas 2010	.	.	0.0%	
van der Mei 2010	1.86	0.5958	16.7%	6.43 [2.00; 20.67]
van der Mei 2010	1.55	0.4802	22.3%	4.69 [1.83; 12.02]
Van der Mei 2016	0.69	0.4426	24.7%	2.00 [0.84; 4.76]
Total (95% CI)			100.0%	2.90 [1.65; 5.09]
Heterogeneity: $\tau^2 = 0.1382$; $\chi^2 = 5.20$, $df = 3$ ($P = 0.16$); $I^2 = 42\%$				





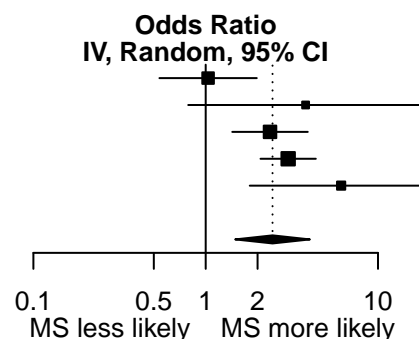
Study	ln(OR)	SE(ln(OR))	Weight	OR [95% CI]
Pandit 2013	1.91	0.5641	16.9%	6.75 [2.23; 20.38]
Simon_Swedish 2010	0.96	0.6675	14.6%	2.60 [0.70; 9.62]
Sundqvist 2011	2.34	0.2068	25.7%	10.40 [6.93; 15.59]
Sundqvist 2012	1.20	0.1683	26.4%	3.31 [2.38; 4.60]
Sundstrom 2008	2.77	0.5837	16.4%	15.98 [5.09; 50.17]
Total (95% CI)			100.0%	6.26 [3.05; 12.88]

Heterogeneity: $\tau^2 = 0.4837$; $\chi^2 = 23.30$, $df = 4$ ($P < 0.01$); $I^2 = 83\%$



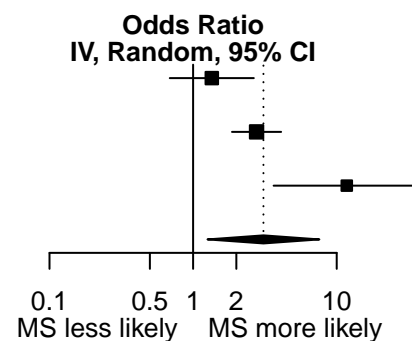
Study	ln(OR)	SE(ln(OR))	Weight	OR [95% CI]
Pandit 2013	0.03	0.3319	22.8%	1.03 [0.54; 1.98]
Simon_Swedish 2010	1.34	0.8003	8.0%	3.80 [0.79; 18.24]
Sundqvist 2011	0.86	0.2574	26.9%	2.36 [1.43; 3.91]
Sundqvist 2012	1.10	0.1886	30.8%	3.01 [2.08; 4.36]
Sundstrom 2008	1.81	0.6231	11.5%	6.11 [1.80; 20.72]
Total (95% CI)			100.0%	2.44 [1.48; 4.03]

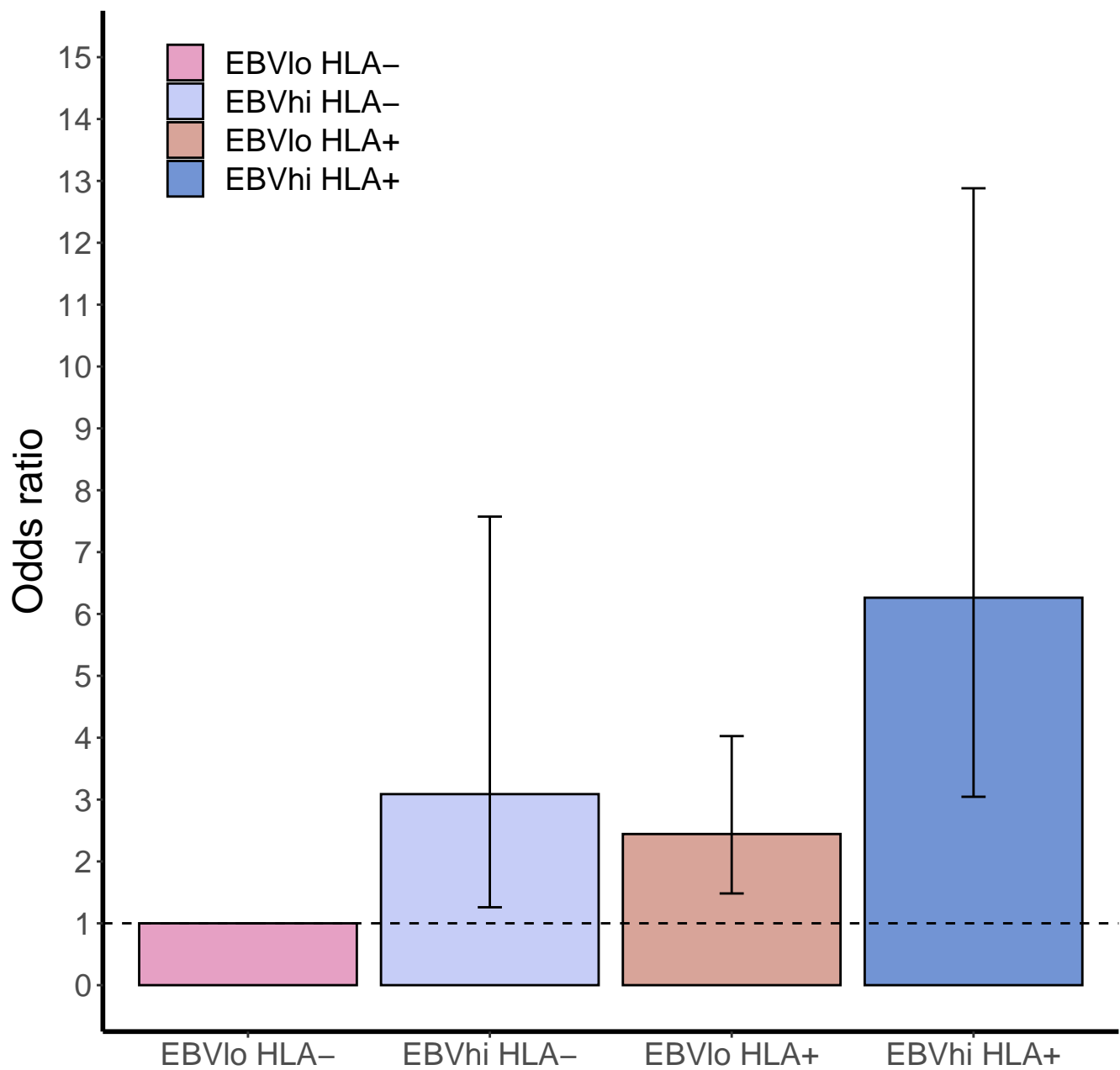
Heterogeneity: $\tau^2 = 0.1752$; $\chi^2 = 10.44$, $df = 4$ ($P = 0.03$); $I^2 = 62\%$



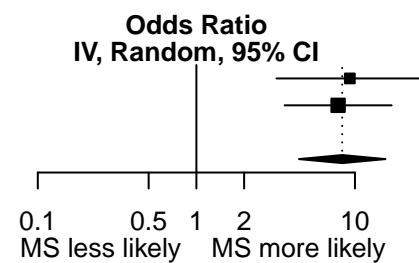
Study	ln(OR)	SE(ln(OR))	Weight	OR [95% CI]
Pandit 2013	0.30	0.3438	34.9%	1.35 [0.69; 2.65]
Simon_Swedish 2010	.	.	0.0%	
Sundqvist 2011	1.02	0.2010	40.1%	2.77 [1.87; 4.10]
Sundqvist 2012	.	.	0.0%	
Sundstrom 2008	2.46	0.5985	24.9%	11.75 [3.64; 37.98]
Total (95% CI)			100.0%	3.09 [1.26; 7.57]

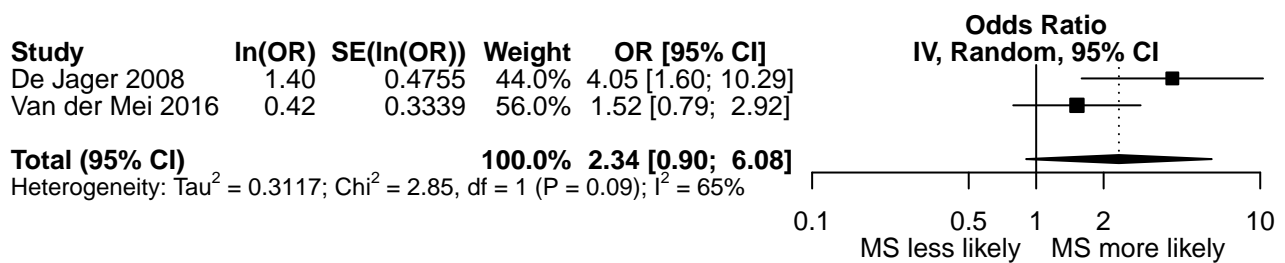
Heterogeneity: $\tau^2 = 0.4817$; $\chi^2 = 10.09$, $df = 2$ ($P < 0.01$); $I^2 = 80\%$





Study	ln(OR)	SE(ln(OR))	Weight	OR [95% CI]
De Jager 2008	2.23	0.5436	34.7%	9.28 [3.20; 26.94]
Van der Mei 2016	2.06	0.3964	65.3%	7.83 [3.60; 17.03]
Total (95% CI)			100.0%	8.31 [4.43; 15.56]
Heterogeneity: $\tau^2 = 0$; $\chi^2 = 0.06$, $df = 1$ ($P = 0.80$); $I^2 = 0\%$				





Study	ln(OR)	SE(ln(OR))	Weight	OR [95% CI]
De Jager 2008	0.85	0.2263	57.3%	2.35 [1.51; 3.66]
Van der Mei 2016	0.83	0.2621	42.7%	2.29 [1.37; 3.83]
Total (95% CI)			100.0%	2.32 [1.66; 3.25]
Heterogeneity: $\tau^2 = 0$; $\chi^2 = 0.00$, $df = 1$ ($P = 0.94$); $I^2 = 0\%$				

