Estimating interactions in individual participant data meta-analysis. A comparison of methods in practice.

Ruth Walker (ruth.walker@york.ac.uk)
University of York https://orcid.org/0000-0003-2765-7363
Lesley Stewart
University of York
Mark Simmonds
University of York

Research Article

Keywords: individual patient data, meta□analysis, one□stage, two-stage, subgroups

Posted Date: March 11th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-317690/v1

Estimating interactions in individual participant data meta-analysis. A comparison of methods in practice.

Walker R[1], Stewart LA[1,] Simmonds M[1],

[1] Centre for Reviews and Dissemination, University of York, Heslington, York YO10 5DD, UK

Funding NIHR Research Methods Systematic Reviews Training Fellowship.

Medical interventions may be more effective in some types of individuals than others and identifying characteristics that modify the effectiveness of an intervention is a cornerstone of precision or stratified medicine. The opportunity for detailed examination of treatment-covariate interactions can be an important driver for undertaking an individual participant data (IPD) meta-analysis, rather than a meta-analysis using aggregate data. A number of recent modelling approaches are available. We apply these methods to the Perinatal Antiplatelet Review of International Studies (PARIS) Collaboration IPD dataset and compare estimates between them. We discuss the practical implications of applying these methods, which may be of interest to aid meta-analysists in the use of these, often complex models.

Models compared included the two-stage meta-analysis of interaction terms and one-stage models which fit multiple random effects and separate within and between trial information. Models were fitted for nine covariates and five binary outcomes and results compared.

Interaction terms produced by the methods were generally consistent. We show that where data are sparse and there is low heterogeneity in the covariate distributions across trials, the meta-analysis of interactions may produce unstable estimates and have issues with convergence. In this IPD dataset, varying assumptions by using multiple random effects in one-stage models or using only within trial information made little difference to the estimates

of treatment-covariate interaction. Method choice will depend on datasets characteristics and individual preference.

Key words: individual patient data, meta-analysis, one-stage, two-stage, subgroups

Introduction

Systematic reviews and meta-analyses are widely used within healthcare to combine relevant data from individual clinical studies. They form an integral part of evidence-based medicine and aim to provide robust evidence to inform policy and clinical practice. Systematic review with individual participant data (IPD) meta-analysis has been referred to as a 'gold standard' approach to evidence synthesis[1]. In addition to producing more reliable and robust summary effect estimates, an important reason for collecting IPD is to understand if particular types of patients benefit more or less from a treatment. This is achieved through investigating whether important clinical covariates, such as age, sex or previous medical history alter treatment effectiveness, also termed "effect modification" [2].

The ability to conduct this type of analysis is important within healthcare as it underpins more personalised approaches to clinical intervention by informing the types of individual to whom intervention is offered [3]. Tailoring healthcare to individuals also has potential to save resources, for example, restricting use to those individuals where it is most effective or cost effective or avoiding the burden of treatment in individuals who derive little or no benefit.

Methods of estimating treatment- covariate interaction

Exploring effect modification involves the estimation of treatment-covariate interaction through the inclusion of interaction parameters within meta-analytic models [4]. This can be done through the use of meta-regression based on published results/aggregate data but may

lack power and is prone to aggregation bias and confounding [4]. In IPD synthesis, treatment-covariate interaction can be estimated using either 'two-stage' or 'one-stage' meta-analysis. Two-stage methods produce estimates of interaction within each trial which are then combined using conventional meta-analysis and either fixed or random effects. One-stage methods combine data from all relevant trials in a single analysis using regression modelling. The approach does not analyse data as if they come from a single "mega-trial", but maintains the differentiation between trials and accounts for the clustering of patients within the trials [5]. Due to this, the approach is also sometimes termed a 'stratified analysis' [3].

There has been growing consideration of methods for the estimation of treatment-covariate interaction, with debate in the literature as to how they differ, and which is method best to use and when. Simmonds & Higgins (2007) first compared meta-regression, the two-stage meta-analysis of interactions (MAOI) and a one-stage method and determined the power of the methods to detect treatment-covariate interactions. Authors concluded that the power of meta-regression depends on the variation in the mean covariate values across studies and that statistical power may be lacking when studies are few. The method is also prone to aggregation bias and confounding. The power of the MAOI model depends on the variation of covariates within each study, whilst the one-stage model always has at least equal or greater statistical power compared to meta-regression and the MAOI model[4].

Since then, methods have advanced and it has become possible to fit more complex one-stage models, including those with multiple random effects. There is now suggestion these may be better than models that use a single random effect [6]. Furthermore, as one-stage models do not automatically avoid aggregation bias when estimating treatment-covariate interactions, there has been consideration of methods that separate within-trial and across-trial information. It has been recommended that one-stage IPD-MA consider only within-trial estimates to avoid

biased results driven by aggregation bias [7-9].

Simulation studies have compared these methods for estimating treatment-covariate interaction. Da Costa and Sutton (2019) used simulated IPD to compare six one-stage models, which varied in whether or not they accounted for variation in between-trial interaction effects. Those accounting for this variation, were less prone to bias and had more accurate standard errors [10]. These results are generalisable to trials using continuous outcomes. In this context, the research highlights the importance of separating out with and between trial effects results. Another simulation study conducted by Kontopantelis (2018) used continuous outcome data to compare one-stage models that used fixed or random treatment effects and the two-stage MAOI. The study found that one-stage models consistently outperformed the two-stage model when estimating interactions, but considered only one-stage models that combined within and across-study information[11], and so did not account for aggregation bias.

Belias et al. (2019) used simulated binary IPD to compare four models of estimating treatment-covariate interaction. These were meta-regression, by-trial subgroup analysis, the MAOI model, a one stage model (referred to as a 'naïve one-stage IPD-MA') that used only a random treatment effect and a one-stage model that centered effect modifiers by their mean in each trial to account for potential aggregation bias. Both the one-stage models were found to have greater power than other models, were unaffected by heterogeneity levels and showed increased power in scenarios with aggregation bias. The MAOI model had less power but was unaffected by heterogeneity levels and showed increased power in scenarios with aggregation bias. Conversely, the by-trial subgroup analysis lost power in scenarios where between-study heterogeneity was increased. Meta-regression showed poor power in all

scenarios [12].

Using simulated data with often-extreme scenarios, to compare methods, may not accurately reflect how the methods will perform in practice, and rather demonstrate how they are expected to perform in theory. Comparison of methods for estimating treatment-covariate interaction using real world datasets is limited. Stewart et al. (2012) used obstetric data, collected by the Perinatal Antiplatelet Review of International Studies (PARIS) Collaboration to compare one-stage models with fixed vs. random treatment effects and interaction effects and one model with separated within and across-trial treatment-covariate interactions. One-stage and two-stage models were used to estimate treatment-covariate interactions for pre-eclampsia (the main outcome). Authors concluded that models produced similar results, but discussed advantages of the one-stage model, over the two-stage model, including greater flexibility to explore model structure [3].

In this paper we use the same PARIS dataset to consider and compare additional one-stage and two-stage models to estimate interactions, building on the work by Stewart et al. (2012). We also consider more recent modelling approaches which fit multiple random effects and separate within and between trial information. We analyse five main outcomes and nine covariates in the PARIS dataset and examine whether different methods produced different results. We also explore and discuss the practical implications of applying these methods to a large IPD dataset, which will be of interest to statisticians and using the methods in practice.

Methods

The PARIS dataset contains IPD from 31 randomised controlled trials with health outcomes for 32,217 women and their 32,819 babies. It explores the efficacy of antiplatelet therapy in

the prevention of pre-eclampsia and its complications. Women were randomised to receive an antiplatelet therapy or a placebo [13].

We used PARIS data for five main outcomes: pre-eclampsia, pre-term birth prior to 34 weeks, small for gestational age infant, fetal or neonatal death and the composite outcome, pregnancy with a serious adverse outcome (SAO). Treatment-covariate analysis for these outcomes considered all available binary covariates in the PARIS dataset.

- 1st pregnancy– Family history of hypertensive disorder of pregnancy (HPD)
- 2nd pregnancy – Previous history of HPD
- 1st pregnancy – any high risk factor*
- 2nd pregnancy – any high risk factor*
- Pre-existing renal disease
- Pre-existing hypertension
- Pre-existing diabetes
- Previous infant SGA
- Multiple pregnancy

*A 'high risk' pregnancy was defined as a current pregnancy with any of the following: maternal autoimmune disease, renal disease, diabetes or chronic hypertension, or with abnormal uterine artery Doppler flow, multiple pregnancy, family history of HDP, or an unspecified risk factor as defined within the trial. Otherwise, a previous pregnancy with a history of any of the following: gestational hypertension, pre-eclampsia, eclampsia, foetal or neonatal death each of which were collected and included in the dataset as individual variables [13].

Analysis originally conducted by Askie et al. (2007) used participant-level subgroup analysis[13], an approach commonly in early IPD meta-analyses which produces subgroup-level effect estimates within study before pooling these estimates in meta-analyses using conventional techniques. Pooled estimates are compared using a test for interaction such as the Cochrans' Q test.

We investigated the use of more recent methods and compared the estimates of treatment-covariate interaction coefficients produced by the two-stage MAOI model and five one-stage models including those with common or random interaction effects[6] and one model that uses only within-trial information on the treatment-covariate interaction. The interaction estimates indicate the extent to which one subgroup is likely to benefit more or less from a treatment. The methods are described in detail below.

Two-stage model: Meta-analysis of interactions

In the first stage, the MAOI method uses a maximum likelihood regression model within each trial (Simmonds & Higgins 2007[4]), including a treatment effect and a treatment-covariate interaction term. Namely:

$$g(y_{ij}) = \Phi_i + \theta_i x_{ij} + \mu_i z_{ij} + \gamma_i x_{ij} z_{ij} \qquad (1)$$

Where, i indicates the trial (1 to k), and j participants within each trial (1 to n_i) y_{ij} is the participant outcome with an identity for continuous outcomes or a logit link (odds ratios) or log link (risk ratio) for dichotomous outcomes; x_{ij} usually takes the value one for treatment group and zero for control group; z_{ij} is value of the covariate for each participant. Hence Φ_i is the intercept term, θ_i is the treatment effect, μ_i the covariate effect, and γ_i is the treatment-

covariate interaction (the parameter of interest). These interaction terms estimate the impact of the covariate on treatment in each trial. In the second stage the interaction effect estimates from each trial ($\hat{\gamma}_i$) are combined using conventional meta-analysis techniques (in this case, the inverse-variance meta-analysis using the DerSimonian-Laird random effect method), producing a summary treatment-covariate interaction estimate.

One-stage models

A one-stage maximum likelihood regression model includes both a treatment effect and a treatment-covariate interaction term, with data from all studies in the same model. The common effect version of the model is as for equation (1) except now the parameters are assumed common across all studies:

$$g(y_{ij}) = \Phi_i + \theta x_{ij} + \mu z_{ij} + \gamma x_{ij} z_{ij} \qquad (2)$$

We note that separate Φ_i parameters are used to retain the distinction between studies, so this is not a "mega-trial" model.

Common interaction effect: model 1: *(One stage model: with a random effect on the treatment parameter)*

In this model, we extended the above to included random effects on the treatment parameter (Tuner *et al.* 2000 [14]), allowing the treatment effect to differ between studies but assuming that the effect of the covariate and the treatment-covariate interaction are common to all trials.

$$\begin{gathered} g(y_{ij}) = \Phi_i + (\theta + u_i)\, x_{ij} + \mu z_{ij} + \gamma x_{ij} z_{ij} \\ u_i \sim N(0, \tau^2) \end{gathered} \qquad (3)$$

Common interaction effect: model 2: *(One-stage model that includes correlated trial and treatment random effects [6]).*

Here we assume random effects for the trial and random effects for the treatment are correlated (Jackson et al 2018[6]).

$$g(y_{ij}) = (\Phi + v_i) + (\theta + u_i)\, x_{ij} + \mu z_{ij} + \gamma x_{ij} z_{ij}$$
$$\begin{pmatrix} u_i \\ v_i \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_\theta^2 & \lambda \\ \lambda & \tau_\phi^2 \end{pmatrix} \right) \qquad (4)$$

Common interaction effect: model 3: *(One-stage model that includes uncorrelated trial and treatment random effects[6]).*

Here we assume the random effects for the trial and random effects for the treatment are uncorrelated (Jackson et al. 2018[6]), so $\lambda = 0$ in Model (4).

Random interaction: *(One-stage model that includes a random effect on the interaction terms.)*

This model is further extended to incorporate random-effects for both the treatment effect and the treatment-covariate interaction.

$$g(y_{ij}) = (\Phi + v_i) + (\theta + u_i)\, x_{ij} + \mu z_{ij} + (\gamma + w_i) x_{ij} z_{ij}$$
$$\begin{pmatrix} u_i \\ v_i \\ w_i \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_\theta^2 & 0 & 0 \\ 0 & \tau_\phi^2 & 0 \\ 0 & 0 & \tau_\gamma^2 \end{pmatrix} \right) \qquad (5)$$

Within study model: *One-stage model that avoids the potential for aggregation bias, by splitting within and between trial data.*

Here we separate the within-trial information on the treatment-covariate interaction from the between-trials information, producing an estimate that includes only within-trial information[7].

$$\mathrm{g}(y_{ij}) = (\Phi + v_i) + (\theta + u_i)\, x_{ij} + \mu z_{ij} + \xi x_{ij}(z_{ij} - \bar{z}_i) + \eta \bar{z}_i$$
$$\begin{pmatrix} u_i \\ v_i \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_\theta^2 & 0 \\ 0 & \tau_\phi^2 \end{pmatrix}\right) \quad (6)$$

where $\bar{z}_i$ is the average covariate value in trial i, so ξ is the parameter for the within-trial interaction.

One-stage models may also be extended to include multiple covariates, which would correct for any correlation between covariates. Models incorporating all covariates were not considered here, as in the PARIS dataset, different covariates were reported in different trials.

Also, not considered are meta-regression and by-trial subgroup analyses, as these analyses are only useful for study level covariates or characteristics common across the trial, and therefore, hold no substantive benefit above conducting a convention meta-analysis using aggregate data[12, 15]. The two-stage participant-level subgroup analysis is also not considered here as the method does not produce a treatment-covariate interaction coefficient. Furthermore, comparisons using this methods are made between subgroups within studies and so the method sometimes lacks the statistical power to be able to detect treatment-covariate interactions [3, 4].

We note that our analyses used a 1/0 coding for the treatment variable. Other authors have noted that a +0.5/-0.5 coding may be preferable in some circumstances, by ensuring a common variance for treatment and control groups and improving maximum likelihood estimation [6, 16], particularly in random effect models where trials are few and the estimation of correlation between two random effects is problematic [14]. Another option is "study-specific centering" (coding 1/0 minus the study-specific proportion of participants in the treatment group) that may reduce the downward bias of between-study variance when using maximum likelihood estimation [16].

Results

Fifty-four outcome-covariate combinations were considered equating to 270 analyses (figures 1-2, supporting information figure 1-3, tables 1-5). Figures 1-2 contain estimates of treatment-covariate interaction coefficients for the main outcome pre-eclampsia and a rarer outcome, fetal or neonatal death, produced by the MAOI model and the five one-stage models. These demonstrate that in general, the methods produced consistent estimates throughout.

Overall, few analyses showed any clear evidence of treatment-covariate interaction, generally having wide confidence intervals. The only exceptions were history of hypertensive disorder for several outcomes including fetal or neonatal death (figures 2, supplementary material figure 1-2) and for multifetal pregnancy, small for gestational age infant and pre-term birth <34 weeks (supplementary material figures 2-3). Here interaction estimates were consistent with confidence intervals generally excluding the line of no effect.

In some cases, the choice of method altered the statistical significance of the interaction estimate, for example, for history of HPD-pregnancy with covariate pregnancy with a serious adverse outcome, history of HPD-pregnancy with covariate pre-term birth <34 weeks and multifetal pregnancy with covariate SGA (supplementary material figures 1-3). However, in all cases the maximum difference in point estimate between methods was generally small ≤0.08 and the confidence intervals were close to the line of no effect.

There were occasions where interaction estimates between methods differed more substantially, for example, for the outcome fetal or neonatal death and covariate renal disease (figure 2). For this analysis, the greatest difference in the treatment-covariate interaction coefficient between methods was 1.06, between the common interaction effect: model two (OR 1.1 (95%CI 0.40-3.00)) and the aggregation bias: within study model (OR 2.16 (95%CI 0.48-7.78)). Estimates for the latter model were more uncertain (figure 2). Differences in interaction estimates produced by the random interaction effect model were also noted for several outcomes with the covariate multifetal pregnancy (figures 1-2, supplementary material figures 1 and 3).

The aggregation bias: within-study model and the MAOI model more commonly produced estimates that differed from the one-stage models, which "amalgamate" within trial and across-trial information. For example, those produced for the covariate diabetes (figure 1-2). These differences, however, were not always consistent between the MAOI and aggregation bias: within study model (figure 1-2) and estimates generally had wider confidence intervals than those produce by the one-stage common interaction effect models.

When producing estimates using the MAOI, within-study regression models would not converge within trials where there were low to zero events, or where all participants within a trial all had the same event status for a given covariate. These trials dropped out of the analysis and therefore, in some cases, the number of trials contributing to the meta-analysis of interactions and to one-stage models differed. This was sometimes coupled with observed differences in estimates, for example, with the outcome baby death and covariate renal disease. In eight instances, the one-stage models produced estimates of treatment-covariate interaction, whereas the meta-analysis of interaction failed to converge across any trial and produced no estimate (supplementary material tables 1-3).

Discussion

Previous comparison of one-stage and two-stage methods of meta-analysis for estimating treatment-covariate interaction have used simulated data [10 11]. We have compared these methods using IPD to understand how the methods perform in practice. We considered five main outcomes and nine covariates in the PARIS dataset. For the majority of analyses, the MAOI model and five one-stage model produced very similar estimates of treatment-covariate interaction coefficients, aligning with findings of previous research [4 3].

Generally, the choice of analytic method had very limited impact on the estimate of treatment-covariate interaction and very rarely altered the statistical significance of an estimate. Where this did occur, point estimates did not vary greatly and were close to the null, making it unlikely that these differences in statistical significance would lead to differing conclusions.

However, in some analyses, the MAOI model and the within-study model produced point estimates that differed substantially from those produced by other methods (figure 2). These models synthesise only the within-study information to examine treatment-covariate relationship and avoid making inference about individual relationships within trials, based on the observed across-study relationships (trials may differ in ways other than the covariate under examination). As such, both of the methods avoid aggregation bias.

Most one-stage models do not automatically avoid aggregation bias when estimating treatment-covariate interactions. As such, comparing of the results from one-stage models with those from the MAOI model and the within study model, might reveal erroneous estimates produced by the one-stage models methods that aggregate both within-trial and across-trial information. However, in our analyses the MAOI model and the within study model did not produce differences with the one-stage models consistently, suggesting it may not be a real effect driving these differences. As differences were more common for outcomes where events were rare, it is more likely that the observed differences in results are attributable to limited data where the meta-analysis of interactions model lacks power[4].

The MAOI model and the within-study model, should, in theory, be the most unbiased methods as they use only the within-trial information when estimating the treatment-covariate interaction[7]. The approach has previously been recommended because the inclusion of across-study information may lead to erroneous inference, should differences in mean covariate values exist across studies [8 7, 9]. In analyses like ours, where trial populations are sufficiently similar to one another that aggregation bias does not pose a great issue, there may be a case to include across-trial information, which would improve the power of the interaction estimate compared to within-trial information alone[8].

Pragmatically, issues arose when we implemented the MAOI method, with regression models within study failing to converge. Applying this method to large IPD datasets where outcome events are rare and there are few participants with particular clinical covariates, may be difficult in practice. This is because participants with and without both the covariate and the outcome are needed to produce a within-study estimate using the MAOI method. Zero cell counts for binary outcomes can be overcome by using continuity correction, where 0.5 is added to cells in the available 2x2 table. However, this approach should be applied with caution as it has previously been shown to influence the magnitude of the effect estimates and their variances [5].

For one-stage models methods that aggregate both within-trial and across-trial information, our analyses generally showed little difference between estimates. Previous work has suggested that using models with multiple random effects produce more accurate estimates than those with single random effects [6]. Differences in the treatment-covariate interaction estimate produced by the random interaction effect model were noted for analyses that included the covariate multifetal pregnancy. As numbers of multifetal pregnancies were generally low across trials <50, few trials had enough data to estimate the treatment-covariate interaction. Adding a random effect on the interaction needs some within-trial interaction data to estimate the heterogeneity and so in these analyses, the models became unstable and the estimates uncertain. More generally, when applied to the PARIS data, the model produced very similar results to all three common interaction effect models that had only a single random effect.

One-stage and two-stage models examining treatment-covariate interactions are each associated with advantages and disadvantages. For example, the meta-analysis of interactions method comes with the ability to produce forest plots enabling easier visualisation of the contribution that each study makes to the summary effect estimate. This may be useful during

preliminary investigations. Software environments such as R with increasingly available pre-written code means that the level of statistical expertise required to implement the methods is now similar between two-stage and one-stage models, despite some previous suggestion that one-stage models are computationally more complex [3].

When considering which method to use to examine potential covariate treatment interactions, knowledge of 'covariate heterogeneity' or the heterogeneity in the covariate distributions across studies, is needed. We recommended that the number of events occurring in covariate groups for outcomes of interest be checked prior to data-analyses, by cross tabulating the data. If covariate heterogeneity is low, then meta-analysis of interactions models may fail to converge, and estimates produced by such models may be unstable. In this case, a one-stage model would be the preferable choice. Should covariate heterogeneity be high, then the meta-analysis of interactions will likely be comparable to the one-stage model [4] and wider factors may be considered when choosing between methods.

Conclusion

In this empirical example, varying assumptions within the one-stage model made little difference when estimating treatment-covariate interaction. As trial populations were sufficiently similar, aggregation bias did not pose a great issue, and as such, applying models that separate within and between study information did not hold substantive value. A wider range of real-world applications, including those which consider continuous covariates are needed to understand which methods perform best in practice. As the methods are capable of producing differing results in some circumstances, it is important to pre-specify method choice in a study protocol, to avoid post-hoc testing which attempts to achieve statistical significance.

Abbreviations

HPD- hypertensive disorder of pregnancy

IPD- Individual participant data

MA- Meta-analysis

MAOI- Meta-analysis of interactions

PARIS- Perinatal Antiplatelet Review of International Studies

SAO- serious adverse outcome.

Declarations

Ethics approval and consent to participate

This IPD-meta-analysis does not require ethical approval.

Consent for publication

Not applicable.

Availability of data and materials

Data sharing is not applicable to this article as no new data were created or analysed in this study. Access to the PARIS dataset was granted by the PARIS collaboration who hold the dataset.

Competing interests

The authors have no competing interests.

Funding

This research project was funded by a National Institute for Health Research (NIHR), (Research Methods Systematic Reviews Training Fellowship). This publication presents independent research funded by the National Institute for Health Research (NIHR). The

views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

Authors' contributions

RW, MS and LS conceived and designed the study. RW and MS conducted the analysis and interpreted the results. RW, MS and LS wrote the manuscript.

Acknowledgements

We acknowledge and thank the PARIS collaboration and the PARIS trial investigators for making this work possible.

References:

1. Stewart L.A., J.F. T. TO IPD OR NOT TO IPD? Advantages and Disadvantages of Systematic Reviews Using Individual Patient Data. *EVALUATION & THE HEALTH PROFESSIONS* 2002;25:76-97.

2. Simmonds MC, Higgins JP, Stewart L.A., Tierney J.F., Clarke M.I., S.G. T. Meta-analysis of individual patient data from randomized trials: a review of methods used in practice. *Clincial trials* 2005;2:209-17.

3. Stewart GB, Altman DG, Askie LM, Duley L, Simmonds MC, Stewart LA. Statistical analysis of individual participant data meta-analyses: a comparison of methods and recommendations for practice. *PLoS One* 2012;7:e46042.

4. Simmonds MC, Higgins JP. Covariate heterogeneity in meta-analysis: criteria for deciding between meta-regression and individual patient data. *Stat Med* 2007;26:2982-99.

5. Burke DL, Ensor J, Riley RD. Meta-analysis using individual participant data: one-stage and two-stage approaches, and why they may differ. *Stat Med* 2017;36:855-75.

6. Jackson D, Law M, Stijnen T, Viechtbauer W, White IR. A comparison of seven random-effects models for meta-analyses that estimate the summary odds ratio. *Stat Med* 2018;37:1059-85.

7. Hua H, Burke DL, Crowther MJ, Ensor J, Tudur Smith C, Riley RD. One-stage individual participant data meta-analysis models: estimation of treatment-covariate interactions must avoid ecological bias by separating out within-trial and across-trial information. *Stat Med* 2017;36:772-89.

8. Fisher DJ, Carpenter JR, Morris TP, Freeman SC, Tierney JF. Meta-analytical methods to identify who benefits most from treatments: daft, deluded, or deft approach? *BMJ* 2017;356:j573.

9. Riley RD, Debray TPA, Fisher D, Hattle M, Marlin N, Hoogland J, et al. Individual participant data meta-analysis to examine interactions between treatment effect and participant-level covariates: Statistical recommendations for conduct and planning. *Stat Med* 2020;39:2115-37.

10. da Costa BR, Sutton AJ. A comparison of the statistical performance of different meta-analysis models for the synthesis of subgroup effects from randomized clinical trials. *BMC Med Res Methodol* 2019;19:198.

11. Kontopantelis E. A comparison of one-stage vs two-stage individual patient data meta-analysis methods: A simulation study. *Res Synth Methods* 2018;9:417-30.

12. Belias M, Rovers MM, Reitsma JB, Debray TPA, IntHout J. Statistical approaches to identify subgroups in meta-analysis of individual participant data: a simulation study. *BMC Med Res Methodol* 2019;19:183.

13. Askie LM, Duley L, Henderson-Smart DJ, Stewart LA. Antiplatelet agents for prevention of pre-eclampsia: a meta-analysis of individual patient data. *The Lancet* 2007;369:1791-98.

14. Turner RM, Omar, R.Z., Yang, M., Goldstein, H. and Thompson, S.G. A multilevel model framework for meta-analysis of clinical trials with binary outcomes. *Statist. Med* 2000;19:3417-32.

15. Debray TP, Moons KG, van Valkenhoef G, Efthimiou O, Hummel N, Groenwold RH, et al. Get real in individual participant data (IPD) meta-analysis: a review of the methodology. *Res Synth Methods* 2015;6:293-309.

16. Riley RD, Legha A, Jackson D, Morris TP, Ensor J, Snell KIE, et al. One-stage individual participant data meta-analysis models for continuous and binary outcomes: Comparison of treatment coding options and estimation methods. *Stat Med* 2020;39:2536-55.

Figures

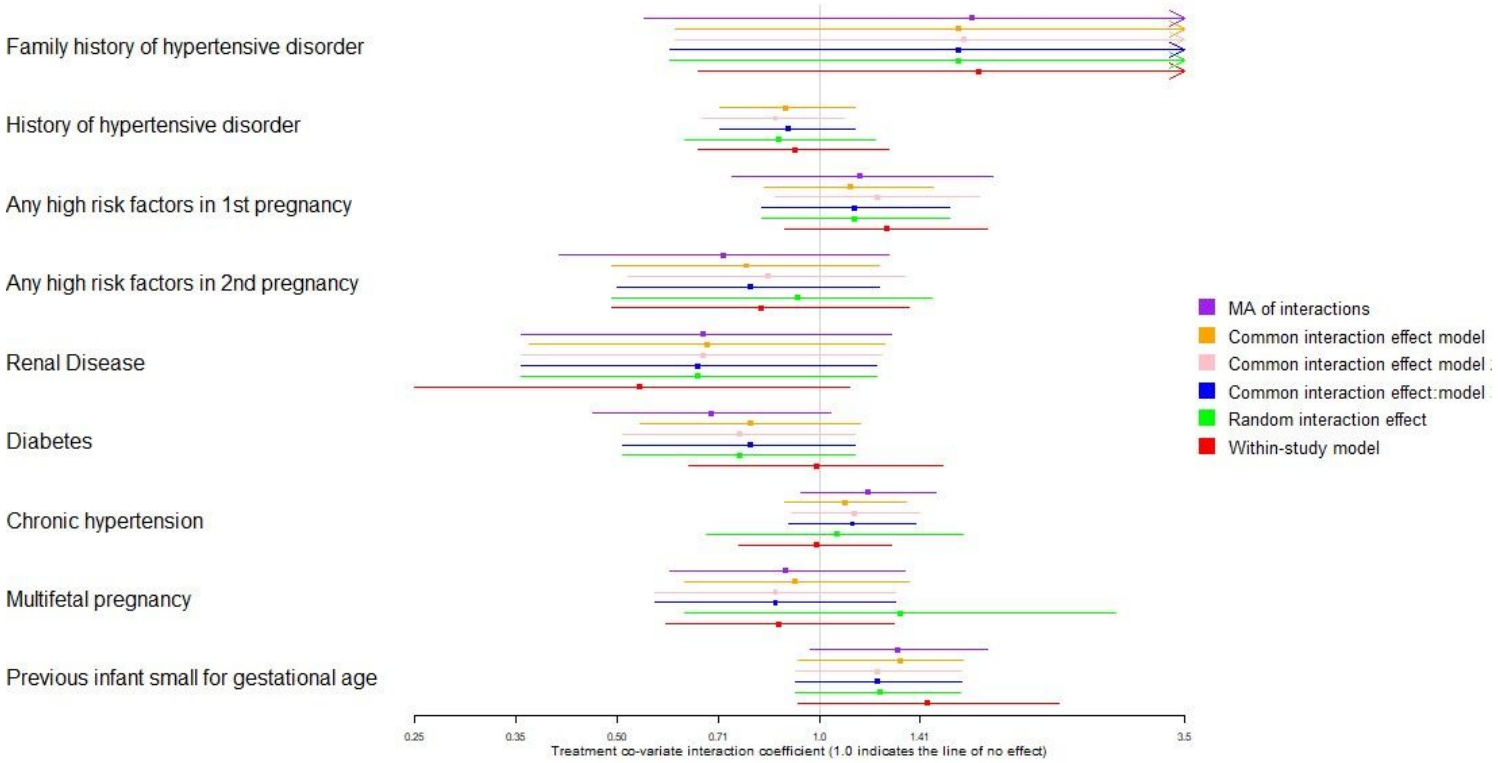

Figure 1

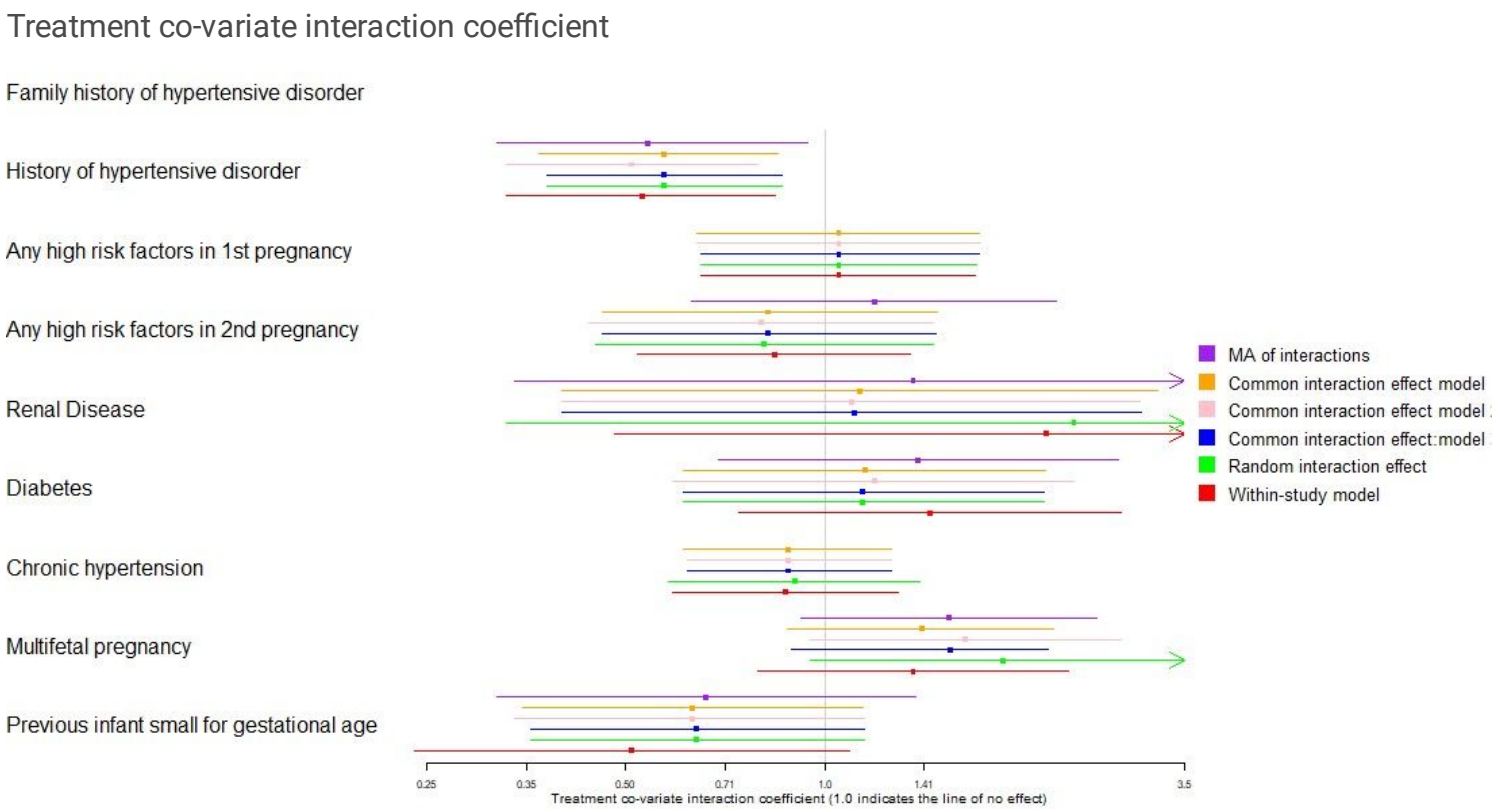

Figure 2

Treatment co-variate interaction coefficient

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Rcomputercode.docx
- Supplementaryinformationfigure1.jpg
- Supplementaryinformationfigure2.jpg
- Supplementaryinformationfigure3.jpg
- Supportinginformationtables15.docx