RESEARCH ARTICLE

Use of multiple covariates in assessing treatment-effect modifiers: A methodological review of individual participant data meta-analyses

Peter J. Godolphin[1] | Nadine Marlin[2] | Chantelle Cornett[1] |
David J. Fisher[1] | Jayne F. Tierney[1] | Ian R. White[1] |
Ewelina Rogozińska[1]

[1]MRC Clinical Trials Unit at University College London, Institute of Clinical Trials and Methodology, London, UK

[2]Pragmatic Clinical Trials Unit, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK

Correspondence
Peter J. Godolphin, 2nd Floor, 90 High Holborn, MRC Clinical Trials Unit at University College London, Institute of Clinical Trials and Methodology, London WC1V 6LJ, UK.
Email: p.godolphin@ucl.ac.uk

Abstract

Individual participant data (IPD) meta-analyses of randomised trials are considered a reliable way to assess participant-level treatment effect modifiers but may not make the best use of the available data. Traditionally, effect modifiers are explored one covariate at a time, which gives rise to the possibility that evidence of treatment-covariate interaction may be due to confounding from a different, related covariate. We aimed to evaluate current practice when estimating treatment-covariate interactions in IPD meta-analysis, specifically focusing on involvement of additional covariates in the models. We reviewed 100 IPD meta-analyses of randomised trials, published between 2015 and 2020, that assessed at least one treatment-covariate interaction. We identified four approaches to handling additional covariates: (1) Single interaction model (unadjusted): No additional covariates included (57/100 IPD meta-analyses); (2) Single interaction model (adjusted): Adjustment for the main effect of at least one additional covariate (35/100); (3) Multiple interactions model: Adjustment for at least one two-way interaction between treatment and an additional covariate (3/100); and (4) Three-way interaction model: Three-way interaction formed between treatment, the additional covariate and the potential effect modifier (5/100). IPD is not being utilised to its fullest extent. In an exemplar dataset, we demonstrate how these approaches lead to different conclusions. Researchers should adjust for additional covariates when estimating interactions in IPD meta-analysis providing they adjust their main effects, which is already widely recommended. Further, they should consider whether more complex approaches could provide better information on who might benefit most from treatments, improving patient choice and treatment policy and practice.

KEYWORDS

confounding, effect modification, individual participant data, meta-analysis, participant-level covariate, treatment-covariate interaction

1 | INTRODUCTION

Individual participant data (IPD) meta-analysis of randomised trials is recognised as the most reliable and flexible way to assess participant-level treatment effect modifiers and enables researchers to make the best use of the available data.[1] Identifying whether and how treatment effects vary across different participant groups (referred to as an interaction) is vital to informing how best to treat individual patients[2–4] and thus to improving patient outcomes.

Traditionally, meta-analysis methods have pooled results across trials for each patient subgroup and compared the subgroup meta-analyses results using a chi-square test of interaction.[5] As this approach is at risk of aggregation bias,[5–8] an alternative approach has recently been proposed, in which interactions are first estimated within each trial and then pooled using a standard meta-analysis model.[6,9] However, both approaches usually consider only one covariate at a time. Thus, estimates of interaction derived from a one-covariate at a time approach may be due, in part at least, to the interaction effect of a different covariate. This could be seen as a form of confounding. For example, if a cancer treatment is effective only in metastatic disease, and older patients are more likely to have metastatic disease, then an interaction between treatment and age would be expected. This would be a genuine interaction, because treatment would on average benefit older patients more than younger patients; but it would be a confounded interaction, because the interaction with metastatic disease provides a fuller and more clinically useful description.

Meta-analyses in clinical areas where they are many related participant-level covariates, such as tumour characteristics in prostate cancer,[10] may be at the biggest risk of such confounding occurring. In IPD meta-analysis, access to participant-level data enables additional covariates to be incorporated when estimating treatment-covariate interactions, which may alleviate the impact of potential participant-level confounding. These covariates could be included in several ways. Simple adjustment for main effects of potentially-confounding covariates is already widely recommended.[4,11] However, more complex modelling using higher-level interaction terms is also possible, and potentially allows more relevant clinical questions to be answered. Currently, it is unclear if, and how researchers utilise this opportunity, and what the impact of their chosen approach is.

The aim of this article is to describe and critique current practice when estimating treatment-covariate interactions in an IPD meta-analysis, specifically focusing on involvement of additional covariates in the models. The article is structured as follows. In Section 2, we outline the methodology of the literature review. Section 3 describes the findings of the review, detailing the snapshot of current practice in IPD meta-analysis when estimating treatment-covariate interactions. In Section 4, we present example statistical models for each identified approach, the question each model addresses and apply these approaches to an exemplar dataset for illustration purposes. We follow with a discussion and brief conclusions.

2 | REVIEW METHODS

2.1 | Protocol

Our literature review followed a prospectively registered protocol[12] (version from 2nd August 2021) that shares some methods, including the search strategy, with a review by Marlin et al.[13,14] (PROSPERO no. CRD42019126768).

2.2 | Literature search

Using a cohort of IPD meta-analysis reports published between 2015 and 2020,[13] we included IPD meta-analyses with at least two randomised trials in which at least one participant-level treatment-covariate interaction was reported. Articles were screened for eligibility by one reviewer (NM), with another reviewer (PJG, ER, and CC) independently confirming eligibility.

2.3 | Sample size

We sought to identify 100 eligible IPD meta-analyses. A random sample from the potentially eligible records was obtained by selecting each 10th record until we either reached the predefined sample size or exhausted the number of potentially eligible records. The justification for this number was that we deemed it sufficient to identify the majority of the current methods used when estimating treatment-covariate interactions in IPD meta-analysis, without being unnecessarily resource intensive.[15]

2.4 | Data collection and extraction

We used a bespoke data collection form that was piloted on five eligible IPD meta-analyses (see Appendix S1). The form was split into four sections: general information (publication year, medical area, number of trials included, IPD meta-analysis approach), non-linear effects (only relevant for the related review[13]), effect modification (number of treatment-covariate interactions assessed, number of outcomes considered for effect

modification, whether additional covariates were involved beyond the treatment-covariate interaction in any statistical models, and if so how they were selected/involved) and contact details of the IPD meta-analysis team.

Data were extracted in duplicate by at least two reviewers (two from PJG, ER, CC, and NM), with discrepancies resolved through discussion. Extracted data was collated in an Excel spreadsheet and cross-checked for accuracy against the completed data extraction forms.

We sought additional documents for eligible IPD meta-analyses (i.e., protocol, statistical analysis plan etc.) to supplement the data extraction, as appropriate. Where further clarifications were required, we contacted one of the authors of the IPD meta-analysis for more details. For pragmatic reasons, we also made several assumptions when extracting data. These included:

1. A one-stage fixed-effect approach had been used if the methods described an analysis 'stratified by trial' or 'adjusted for trial'.[16,17]
2. The software used was indicative of the analytical approach. For example, if Stata commands metan or ipdmetan[18] had been used we assumed a two-stage approach.
3. A random-effects model had been used if the methods mentioned 'a random effect for study/trial' or 'random effects model'. Conversely, a fixed-effect model was assumed if the methods described a 'random intercept'.[16]
4. We assumed that no covariates (over and above the effect modifier itself) were involved if none were mentioned in either the methods or in tables/figures displaying the results of the statistical models.

2.5 | Data synthesis and analysis

All data in this review are summarised narratively. Continuous variables are presented with mean and standard deviation or median and interquartile range. Categorical variables are described with frequency counts and percentages. All analyses were performed using Stata software (version 16.1).

3 | RESULTS OF METHODOLOGICAL REVIEW

In order to reach 100 IPD meta-analyses meeting out eligibility criteria, we had to assess a random sample of 211 from the 738 potentially eligible records (Figure 1).

The included IPD meta-analyses are evenly distributed in terms of their publication year between 2015 and 2020

(Table 1). They comprised a variety of medical fields, with 29% addressing questions in cardiovascular research. About a quarter of the IPD meta-analyses were prospectively registered on PROSPERO. Median meta-analysis size was five datasets (trials) and 2816 participants. Most of meta-analyses used a one-stage IPD method either as the sole approach (60%) or in addition to a two-stage approach (19%). A two-stage IPD method as the only approach was rare (19%). Fixed-effect models were used slightly more than a random-effects model (39% vs 30%), although a fifth of IPD meta-analyses used both fixed-effect and random-effects models, with one often employed as a sensitivity analysis.

Effect modification tended to be explored based on a small number of outcomes (median of two), although this ranged from 1 to 16 (Table 2). A median of six covariates per IPD meta-analysis were explored as potential effect modifiers. The majority of IPD meta-analyses (81%) considered at least one categorical effect modifier. A quarter of IPD meta-analyses (26%) kept continuous effect modifiers as continuous when estimating treatment-covariate interactions while in almost two-thirds (63%) continuous effect modifiers were categorised.

Only around a third of IPD meta-analyses (35%) provided sufficient detail to ascertain whether within- and across-trial information was appropriately separated out when estimating treatment-covariate interactions, either written in the main text, appendices or in supplementary documents (e.g., in protocols or statistical analysis plans) (Table 2). Of these, 15 used methods that separated out within- and across-trial information. Less than half (43/100) of the included IPD meta-analyses involved additional covariates when estimating treatment-covariate interactions.

In IPD meta-analyses that involved additional covariates, most of these covariates were selected a priori with only three IPD meta-analyses reporting use of a stepwise procedure (Table 3). The reasons for involving additional covariates varied (Table 3). They were mostly included as main effects alongside the treatment-covariate interaction of interest (35/43, 81%). In eight IPD meta-analyses, the covariates were used in a more complex model: adjustment for at least one additional two-way treatment-covariate interaction (3/43, 7%) or inclusion of a three-way interaction between two covariates and treatment (5/43, 12%).

4 | APPROACHES TO ESTIMATING TREATMENT-COVARIATE INTERACTIONS IN THE CONTEXT OF INCLUSION OF ADDITIONAL COVARIATES

In our review we identified four different approaches to handling additional covariates when estimating treatment-
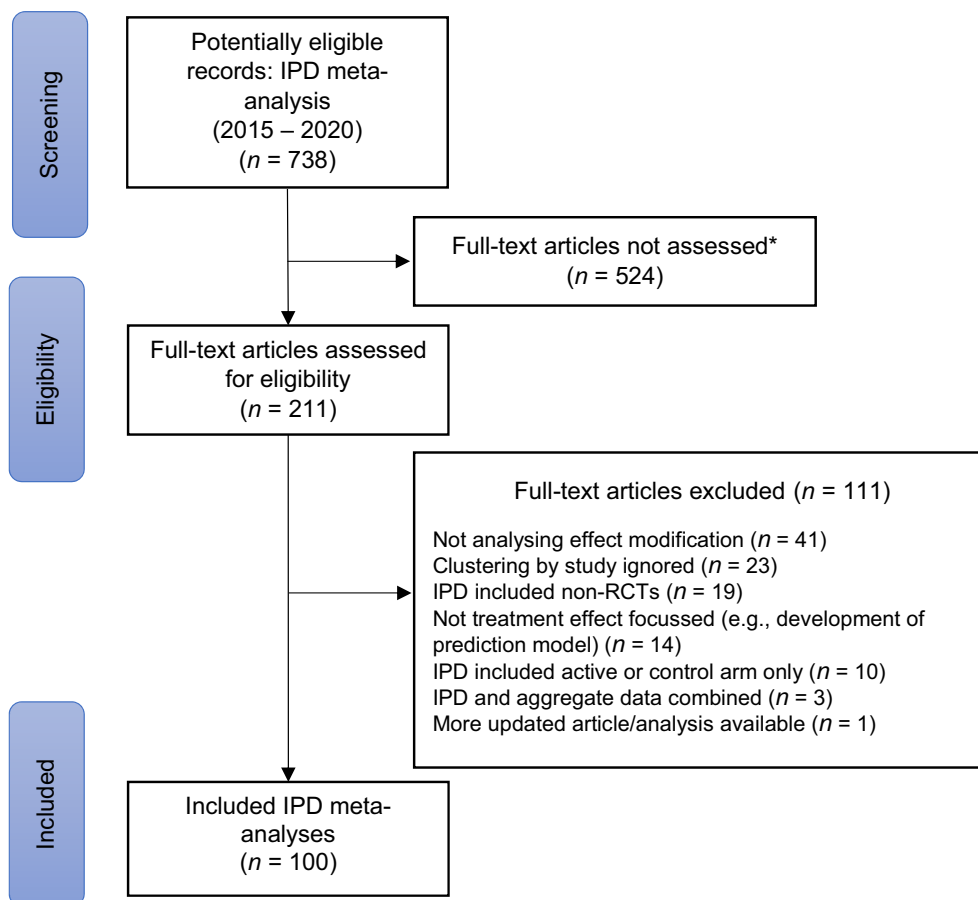
FIGURE 1 Study selection flow diagram. *Full-text articles were not assessed for eligibility once 100 eligible IPD meta-analyses had been identified IPD refers to individual participant data, RCT refers to randomised controlled trial. [Colour figure can be viewed at wileyonlinelibrary.com]

covariate interactions, with most of the IPD meta-analyses estimating interactions without including additional covariates. The identified approaches are as follows:

Approach 1: *Single interaction model (unadjusted).* When estimating the treatment-covariate interaction the analysis does not involve any additional covariates and is unadjusted.

Approach 2: *Single interaction model (adjusted).* When estimating the treatment-covariate interaction the analysis is adjusted for the main effect of at least one additional covariate.

Approach 3: *Multiple interactions model.* When estimating the treatment-covariate interaction the analysis adjusts for at least one two-way interaction between an additional covariate (often a further potential effect modifier) and treatment.

Approach 4: *Three-way interaction model.* When estimating the treatment-covariate interaction the analysis includes a three-way interaction term with treatment, the potential effect modifier, and an additional covariate (often a further potential effect modifier).

Table 4 contains example statistical models for the four approaches for both one-stage and two-stage meta-

analysis, with these models presented for a continuous outcome and correctly separating out within- and across-trial information.[4,8] The approaches 2–4 (Table 4) assume only involvement of one additional covariate, although approaches 2 and 3 can easily be extended by making w_{ij}, β_{2i} and γ_2 vectors. Approach 4 can potentially be extended to include higher-order interaction terms as appropriate. Note, we do not make a distinction here between fixed-effect and random-effect models, as this is not relevant when specifying the covariates to be included in a model. Further, it is important to note that in these models it is inappropriate to draw inference from the coefficient of the average treatment effect, θ. Through inclusion of treatment-covariate interaction terms, these models all assume that there is heterogeneity in this treatment effect. Therefore, θ is not interpretable on its own and instead now depends on the values of other coefficients in the model.

The four identified approaches address three different questions (Table 4). Approaches 1 and 2 use different methods to address the same question in the single interaction model: this validly estimates an interaction effect, and this information could be appropriately used to inform treatment decisions. However, if there is at least one true additional effect modifier, then incorporating

TABLE 1 Characteristics of the included individual participant data meta-analyses.

	Total ($n = 100$)
Year published	
2015	14 (14%)
2016	8 (8%)
2017	19 (19%)
2018	19 (19%)
2019	21 (21%)
2020	19 (19%)
Medical field	
Cardiovascular	29 (29%)
Neurology	13 (13%)
Cancer	12 (12%)
Women's Health	7 (7%)
Mental Health	7 (7%)
Critical Care	5 (5%)
Public Health	4 (4%)
Infectious Diseases	3 (3%)
Neonatal Health	3 (3%)
Other[a]	17 (17%)
PROSPERO registration	
Yes	22 (22%)
Not reported	78 (78%)
Number of trials[b]	
Median [25th, 75th]	5 [3, 11]
Min, Max	2, 34
Number of participants included[c]	
Median [25th, 75th]	2816 [1094, 4754]
Min, Max	73, 174,000
IPD synthesis method	
One-stage	60 (60%)
Two-stage	19 (19%)
Both	19 (19%)
Unclear	2 (2%)
Statistical model(s) used	
Fixed-effect only	39 (39%)
Random-effects only	30 (30%)
Both	20 (20%)
Unclear	11 (11%)

Note: All data are frequency (%) unless stated. IPD refers to individual participant data.

[a]Other category includes a variety of medical fields that are represented by two or less IPD meta-analyses (see Appendix S2).

[b]Excludes five IPD meta-analyses that included a varying number of datasets.

[c]Excludes nine IPD meta-analyses where multiple numbers of participants were reported (n = 7) or no information was reported ($n = 2$).

TABLE 2 Description of investigation into effect modification in the included sample of individual participant data meta-analyses.

	Total ($n = 100$)
Number of outcomes investigated for effect modification	
Median [25th, 75th]	2 [1,3]
Min, Max	1, 16
Number of effect modifiers considered	
Median [25th, 75th]	6 [2, 9]
Min, Max	1, 28
Covariate type of effect modifiers in each study[a]	
Categorical	81 (81%)
Continuous	26 (26%)
Categorised continuous	63 (63%)
Handling of within- and across-trial information in analysis of effect modification	
Separated out[b]	15 (15%)
Conflated[b]	20 (20%)
Unclear	66 (66%)
Involvement of additional covariates when assessing effect modification	
No	57 (57%)
Yes	43 (43%)

Note: All data are frequency (%) unless stated.

[a]Data are not mutually exclusive.

[b]One IPD meta-analysis carried out one analysis that correctly separated out within- and across-trial information and another that conflated within- and across-trial information, so this IPD meta-analysis appears in both categories.

this in the multiple interactions model (approach 3) could lead to more nuanced treatment recommendations. With multiple potential effect modifiers, the three-way interaction model (approach 4) could be used to test whether the multiple interactions model is appropriate, or whether there is an additional relationship between the effect modifiers that needs to be accounted for. In this situation, the question addressed could be thought of as "Does the multiple interactions model describe the data adequately?" Note that for the three-way interaction model, $\gamma_1 - \gamma_3$ could all be coefficients of interest, and, as the covariate-specific treatment effects are all linear combinations of $\gamma_1 - \gamma_3$, in this case the question addressed is 'How does the average treatment effect vary between individuals with higher and lower levels of w and z?'

These models can be used to test hypotheses about interactions (e.g., is the single interaction model

TABLE 3 Description of involvement of additional covariates when assessing effect modification in the 43 individual participant data meta-analyses where this occurred.

Item	Total (n = 43)
How were additional covariates selected	
Determined a priori	28 (65%)
Stepwise procedure	3 (7%)
Not mentioned/unclear	12 (28%)
How were covariate(s) involved when assessing effect modification	
Main effect of covariate(s) adjusted for	35 (81%)
Two-way interaction with covariate adjusted for	3 (7%)
Three-way interaction formed with treatment, effect modifier and covariate[a]	5 (12%)
Reasons given for involving additional covariate(s)	
Confounding explicitly mentioned	5 (12%)
Confounding not explicitly mentioned	28 (65%)
Known prognostic factors	13 (30%)
To account for baseline covariates/baseline imbalance	10 (23%)
Based on previous research	3 (7%)
Stepwise selection	2 (5%)
No reasons given	10 (23%)

Note: All data are frequency (%).

[a]One study stratified by a covariate and then assessed effect modification within these strata, which is indirectly forming a three-way interaction with the covariate, the effect modifier and treatment and is included in this category.

(adjusted) adequate or do we need a more complex approach, such as the multiple interactions model), and hence to select the most appropriate model. Further, conditional on the selected model, they can be used to predict treatment effects for patients.

4.1 | Application to exemplar dataset

To demonstrate how these approaches may provide different estimates of interaction effects, we used an exemplar dataset that was simulated to resemble in its characteristics an average IPD meta-analysis from the sample of 100 included in our review. This exemplar dataset comprises five 'trials' of varying sizes (range from 200 to 1500), with a total of 3200 participants and a continuous outcome (score from 0 to 50), where greater scores are assumed to indicate benefit. Analysis was undertaken in Stata software (version 16.1).

Under a fixed-effect inverse-variance weighted meta-analysis using a two-stage model and unadjusted for covariates, this example shows a beneficial main effect of treatment and limited statistical heterogeneity (Figure 2). A fixed-effect model was used in this example for convenience.

To demonstrate the methods in action, this dataset contained two binary covariates (referred to as *Covariate 1* and *Covariate 2*, with these covariates indicating some baseline characteristic such as a co-morbidity) that had a degree of negative correlation (Pearson's correlation coefficient, $\rho = -0.31$). Using a two-stage framework, an interaction model was initially fitted within each trial in the first stage and then interaction coefficients (differences in mean differences due to the continuous nature of the outcome) were pooled in the second stage in a fixed-effect inverse-variance weighted meta-analysis, that correctly separated out across- and within trial information.[9] This was repeated for both covariates and for three different interaction models in the first stage: (1) single interaction model (unadjusted); (2) single interaction model (adjusted); (3) multiple interactions model. Equations for the five models fitted (the multiple interactions model is fitted once with both covariates) are displayed in Appendix S3.

Initially, both covariates appeared to modify the effect of treatment on their own (using the single interaction model (unadjusted), see Appendix S3, models 1 and 2). We then fitted the single interaction model (adjusted), adjusting for the other covariate (see Appendix S3, models 3 and 4) and then finally fitted the multiple interactions model (see Appendix S3, model 5). The pooled differences in mean differences and interaction *p*-values for each approach and each covariate are presented in Figure 3.

In our illustrative example, there is consistent evidence that Covariate 1 has a modifying effect on treatment, indicating that those with the Covariate (e.g., co-morbidity) have greater benefit from treatment compared to those without the Covariate (Figure 3). However, evidence of Covariate 2's modifying effect reduces as the main effect of Covariate 1 is adjusted for (Approach 2) and disappears when the two-way interaction of treatment and Covariate 1 is adjusted for (Approach 3). This indicates that the Covariate 1 by treatment interaction explains the Covariate 2 by treatment interaction. If either of the two single interaction models had been the main approach in this example, then it may have been concluded that both covariates are important when making treatment decisions, whereas the multiple interactions model demonstrates that Covariate 1 is driving this modifying effect. We found no evidence of a three-way interaction between both covariates and treatment, suggesting that the multiple interactions model describes the data adequately.

TABLE 4 Summary of the identified approaches to involving additional covariates when estimating treatment-covariate interactions in the sample of 100 individual participant data meta-analyses.

Approach	Description	Example one-stage linear mixed model	Example two-stage linear model	Question addressed by the coefficient in bold
1	Single interaction model (unadjusted)	$Y_{ij} = \alpha_i + \beta_{1i}z_{ij} + \theta x_{ij} + \boldsymbol{\gamma} x_{ij}(z_{ij} - \bar{z}_i)$	Stage 1: $Y_{ij} = \alpha_i + \beta_{1i}z_{ij} + \theta_i x_{ij} + \gamma_i x_{ij}z_{ij}$ Stage 2: Pool γ_i in a meta-analysis to estimate $\boldsymbol{\gamma}$	How does the average treatment effect vary between individuals with higher and lower levels of z?
2	Single interaction model (adjusted)	$Y_{ij} = \alpha_i + \beta_{1i}z_{ij} + \beta_{2i}w_{ij} + \theta x_{ij} + \boldsymbol{\gamma} x_{ij}(z_{ij} - \bar{z}_i)$	Stage 1: $Y_{ij} = \alpha_i + \beta_{1i}z_{ij} + \beta_{2i}w_{ij} + \theta x_{ij} + \gamma_i x_{ij}z_{ij}$ Stage 2: Pool γ_i in a meta-analysis to estimate $\boldsymbol{\gamma}$	Same as 1, but with possibly greater power due to covariate adjustment
3	Multiple interactions model	$Y_{ij} = \alpha_i + \beta_{1i}z_{ij} + \beta_{2i}w_{ij} + \theta x_{ij} + \boldsymbol{\gamma_1} x_{ij}(z_{ij} - \bar{z}_i) + \gamma_2 x_{ij}(w_{ij} - \bar{w}_i)$	Stage 1: $Y_{ij} = \alpha_i + \beta_{1i}z_{ij} + \beta_{2i}w_{ij} + \theta_i x_{ij} + \gamma_{i1}x_{ij}z_{ij} + \gamma_{i2}x_{ij}w_{ij}$ Stage 2: Pool γ_{i1} in a meta-analysis to estimate $\boldsymbol{\gamma_1}$	Among individuals with the same level of w, how does the average treatment effect vary between individuals with higher and lower levels of z?
4	Three-way interaction model	$Y_{ij} = \alpha_i + \beta_{1i}z_{ij} + \beta_{2i}w_{ij} + \beta_{3i}z_{ij}w_{ij} + \theta x_{ij} + \gamma_1 x_{ij}(z_{ij} - \bar{z}_i) + \gamma_2 x_{ij}(w_{ij} - \bar{w}_i) + \boldsymbol{\gamma_3} x_{ij}(z_{ij} - \bar{z}_i)(w_{ij} - \bar{w}_i)$	Stage 1: $Y_{ij} = \alpha_i + \beta_{1i}z_{ij} + \beta_{2i}w_{ij} + \beta_{3i}z_{ij}w_{ij} + \theta x_{ij} + \gamma_{i1}x_{ij}z_{ij} + \gamma_{i2}x_{ij}w_{ij} + \gamma_{i3}x_{ij}z_{ij}w_{ij}$ Stage 2: Pool γ_{i3} in a meta-analysis to estimate $\boldsymbol{\gamma_3}$	Does the way the average treatment effect varies between individuals with higher and lower levels of w itself vary with the level of z?

Note: Y_{ij} is the outcome, which is continuous in these example models. α_i is intercept for each trial. z_{ij} is the potential effect modifier. w_{ij} is potential participant-level confounding covariate (or other potential effect modifier). $\bar{z}_i$ bar is study specific mean of this covariate. θ is the treatment effect, with x_{ij} an indicator variable for treatment (Yes, No). γ is the coefficient for the within-trial interaction. The coefficient of interest is bolded for each approach (e.g., γ). Note that the model numbers correspond to the four approaches described in Section 4.

5 | DISCUSSION

Most IPD meta-analyses of randomised trials included in our sample did not involve additional covariates when assessing patient-level, treatment-covariate interactions. When they did, it was often to adjust main effects, rather than to adjust for additional treatment-covariate interactions or to investigate higher-order interactions between treatment and multiple covariates. Very few IPD meta-analyses included multiple treatment-covariate interactions in the same model, yet such models have potential to be most adept at accounting for participant-level confounding and produce the most appropriate estimates of treatment-covariate interactions.

The interactions discussed here are not confounded in the same way that exposure effects are confounded in observational studies. For example, using our exemplar dataset, we found evidence that the Covariate 2-by-treatment interaction is confounded by the Covariate 1-by-treatment interaction. This means the best way to predict treatment effects is by using Covariate 1 (as shown by the multiple interactions model); but it remains true that treating based on Covariate 2 is not invalid if Covariate 1 is not available, even if this is driven, in part, by the Covariate 1-by-treatment interaction.

Whilst the availability of IPD from trials creates the possibility to fit more complex models, these can present practical problems. Incorporating multiple treatment-covariate interactions may be challenging due to model convergence issues, which could arise both for one-stage models[19] and during the first stage of a two-stage approach if data is sparse in at least one trial. Inclusion of more covariates increases the probability of missing covariate data for some participants and/or completely missing covariates in one or more trials. This could result in a loss of power through fitting multiple treatment-covariate interactions in the same model and/or modelling more complex higher-order interactions. Therefore, such models may be of most use when there is sufficient sample size to provide satisfactory power to test for and estimate these more complex interactions, where such assessments can be considered a priori[14] and where the covariates selected are known to be well collected and mostly complete. Potential solutions to these challenges may exist, such as mean imputation of missing data in an interacting covariate at the trial level, conditional on all other interacting covariates.[20,21] However, further evaluation of the implications that these methodological challenges pose for the results of treatment-covariate interaction analyses is warranted.

If such models are fitted, then attention turns to which covariates to include and how to select covariates. In the 43 IPD meta-analyses from our sample that included additional covariates, the majority selected these a priori, with only three using a stepwise procedure.
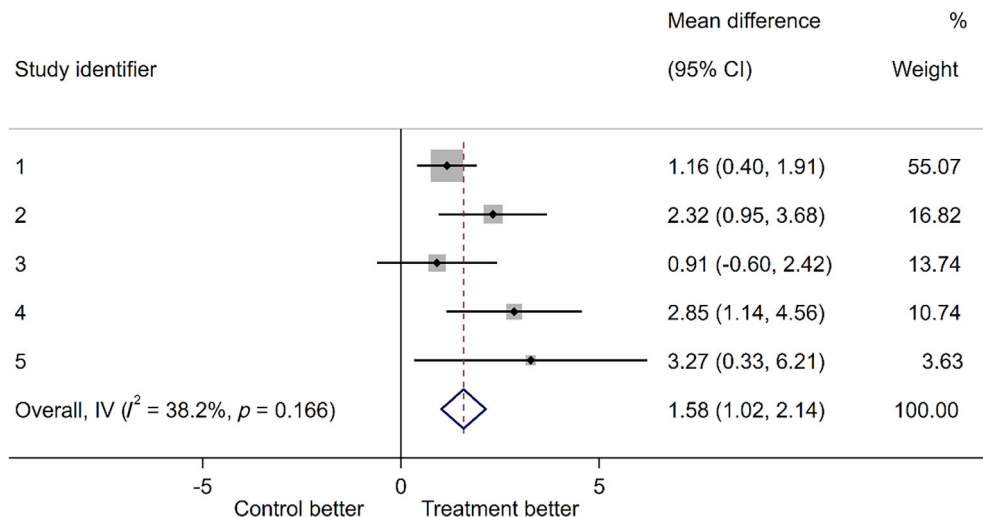
Study identifier	Mean difference (95% CI)	% Weight
1	1.16 (0.40, 1.91)	55.07
2	2.32 (0.95, 3.68)	16.82
3	0.91 (-0.60, 2.42)	13.74
4	2.85 (1.14, 4.56)	10.74
5	3.27 (0.33, 6.21)	3.63
Overall, IV (I^2 = 38.2%, p = 0.166)	1.58 (1.02, 2.14)	100.00

-5   0   5
Control better   Treatment better

FIGURE 2 Fixed-effect meta-analysis of five trials for a continuous outcome using exemplar dataset. [Colour figure can be viewed at wileyonlinelibrary.com]
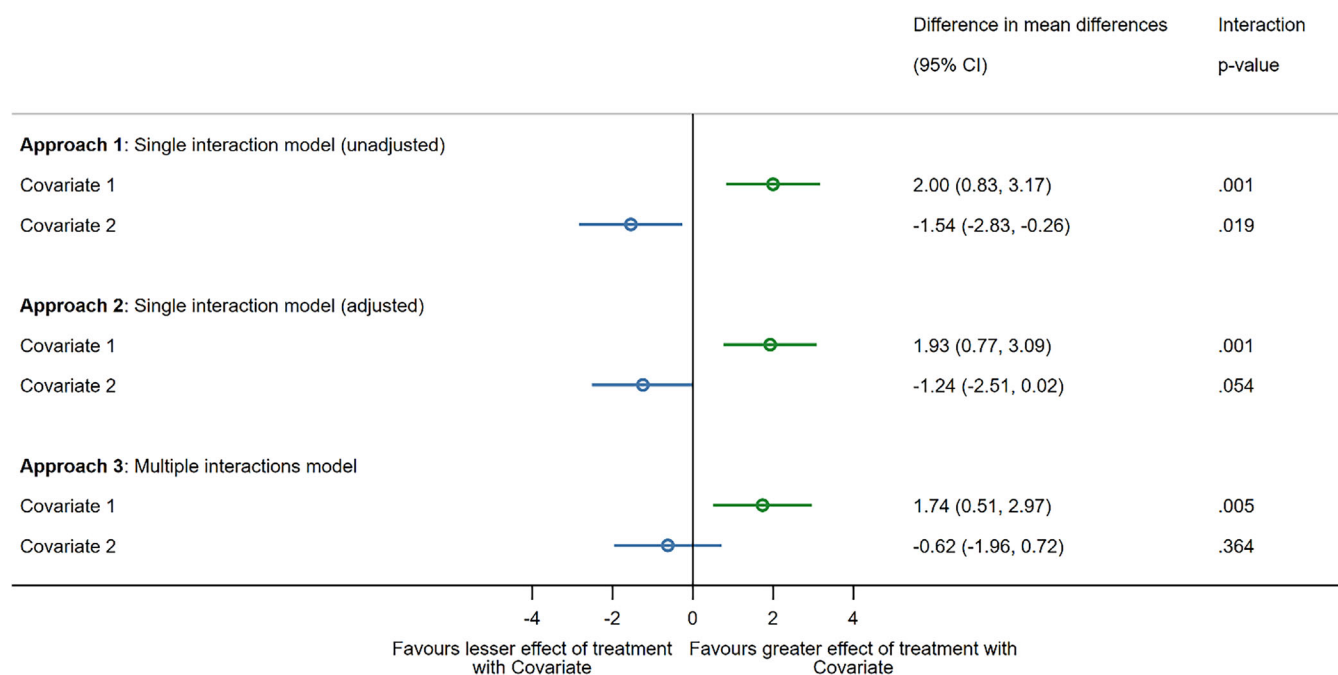
	Difference in mean differences (95% CI)	Interaction p-value
Approach 1: Single interaction model (unadjusted)		
Covariate 1	2.00 (0.83, 3.17)	.001
Covariate 2	-1.54 (-2.83, -0.26)	.019
Approach 2: Single interaction model (adjusted)		
Covariate 1	1.93 (0.77, 3.09)	.001
Covariate 2	-1.24 (-2.51, 0.02)	.054
Approach 3: Multiple interactions model		
Covariate 1	1.74 (0.51, 2.97)	.005
Covariate 2	-0.62 (-1.96, 0.72)	.364

-4   -2   0   2   4
Favours lesser effect of treatment with Covariate   Favours greater effect of treatment with Covariate

FIGURE 3 Differences in mean differences and interaction *p*-values for the investigation of effect modifiers Covariate 1 and Covariate 2 in the exemplar dataset, dependent on model choice. [Colour figure can be viewed at wileyonlinelibrary.com]

We strongly recommend that the covariates included are determined *a priori*, that selection is restricted to only a small number of covariates with a strong clinical rationale that have a clear stated hypothesis, and that p-values are not used to determine selection.[22]

Of note, in our review, none of the IPD meta-analyses used a risk-based approach as advocated by the PATH statement[23] or shrinkage approaches that incorporate multiple treatment-covariate interactions at once.[24] Both methods have been popularised only in recent years, and our sample does not reflect this newer methodology, which could be due to ending our search in 2020, or from us not reviewing the entire sample of IPD meta-analyses,

or from these methods not yet being widely adopted. Further work is ongoing that aims to establish what barriers may exist to using such methodology or the more complex approaches identified in our review, and to identify why these approaches have not had large take-up yet from the evidence synthesis community.

Within our sample, reporting on whether within- and across-trial information was separated was limited. Appropriate methodology to separate out across- and within-trial information for both one-stage and two-stage approaches exists and has been available for several years.[6,7] Reassuringly, a greater number of IPD meta-analyses in our review used an appropriate analysis that did not introduce

aggregation bias (15%) compared to a prior review by Fisher et al.[6] (2%). However, the number of meta-analyses undertaking appropriate analyses and/or reporting clearly is still severely lacking and methodological and reporting improvements are warranted. Research should be guided by the Preferred Reporting Items for a Systematic Review and Meta-analysis of IPD[25]; which preceded many of the IPD meta-analyses included in this review.

Our study was guided by a prospectively developed protocol however is not without limitations. First, we included a pre-determined sample of 100 IPD meta-analyses, and we did not assess an additional 200–300 IPD meta-analyses of randomised trials between 2015 and 2020. However, we do believe that this is a representative sample as our sample is comparable to IPD meta-analyses identified in a systematic review[26] for a number of characteristics such as clinical area, number of included studies and number of participants included. Also, some IPD meta-analyses only investigated one treatment-covariate interaction and others may not have found evidence of multiple effect modifiers. Both factors may have impacted whether IPD meta-analyses would have considered involving additional covariates at all, potentially reducing the number that utilised more complex approaches. However, we reviewed statistical analysis plans and protocols (where available) and rarely found consideration of involving additional covariates contained in this additional documentation. Further, whilst including multiple treatment-covariate interactions or forming three-way interactions may not have been appropriate for these cases, simple covariate would have been possible. Finally, we often had to make assumptions when extracting data, but we did contact authors for clarification on key information, such as reasons for involving additional covariates, where required.

Future methodological studies should establish a general strategy for estimating treatment-covariate interactions whilst taking account of other covariates. Empirical data should be analysed to understand in which circumstances the various possible approaches will lead to different conclusions and based on these findings, advocate that the most appropriate approaches should be used in future IPD meta-analyses of randomised trials.

6 | CONCLUSIONS

Researchers should make the most of the IPD they collect, and at a minimum adjust for prognostic factors when estimating treatment-covariate interactions, provided they adjust their main effects, which is already widely recommended. Where multiple effect modifiers are explored, consideration of additional methods that alleviate any potential participant-level confounding, such as the multiple interactions model, may provide

better information on which participants are most likely to benefit from treatments, leading to more informed treatment policy and practice.

AUTHOR CONTRIBUTIONS

Peter J. Godolphin: Conceptualization; formal analysis; investigation; supervision; writing – original draft. **Nadine Marlin:** Investigation; resources; supervision; writing – review and editing. **Chantelle Cornett:** Investigation; writing – review and editing. **David J. Fisher:** Writing – review and editing. **Jayne F. Tierney:** Writing – review and editing. **Ian R. White:** Writing – review and editing. **Ewelina Rogozinska:** Conceptualization; investigation; supervision; writing – review and editing.

CONFLICT OF INTEREST STATEMENT
The authors declare no conflict of interests.

DATA AVAILABILITY STATEMENT
All data that was extracted as part of the literature review and used in the exemplar dataset is available on the Open Science Framework (see https://osf.io/zmu75/).

ORCID
Peter J. Godolphin https://orcid.org/0000-0003-0648-0992
Nadine Marlin https://orcid.org/0000-0002-2841-4825
Chantelle Cornett https://orcid.org/0000-0003-3789-0361
David J. Fisher https://orcid.org/0000-0002-2512-2296
Jayne F. Tierney https://orcid.org/0000-0002-4734-3014
Ian R. White https://orcid.org/0000-0002-6718-7661

Ewelina Rogozińska https://orcid.org/0000-0003-3455-0644

REFERENCES

1. Stewart LA, Tierney JF. To IPD or not to IPD? Advantages and disadvantages of systematic reviews using individual patient data. *Eval Health Prof.* 2002;25:76-97.
2. Tierney JF, Vale CL, Riley R, et al. Individual participant data (IPD) meta-analyses of randomised controlled trials: guidance on their use. *PLoS Med.* 2015;12:e1001855. doi:10.1371/journal.pmed.1001855
3. Tierney JF, Riley RD, Tudur-Smith C, et al. Rationale for embarking on an IPD meta-analysis project. In: Riley RD, Tierney JF, Stewart LA, eds. *Individual Participant Data Meta-Analysis: A Handbook for Healthcare Research.* Wiley; 2021:9-19.
4. Riley RD, Fisher DJ. Using IPD meta-analysis to examine interactions between treatment effect and participant-level covariates. In: Riley RD, Tierney JF, Stewart LA, eds. *Individual Participant Data Meta-Analysis: A Handbook for Healthcare Research.* Wiley; 2021:163-198.
5. Fisher DJ, Copas AJ, Tierney JF, Parmar MKB. A critical review of methods for the assessment of patient-level interactions in individual patient data (IPD) meta-analysis of randomised trials, and guidance for practitioners. *J Clin Epidemiol.* 2011;64:949-967.
6. 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. doi:10.1136/bmj.j573
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;36:772-789. doi:10.1002/sim.7171
8. Riley RD, Debray TPA, Fisher D, 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;39:2115-2137. doi:10.1002/sim.8516
9. Godolphin PJ, White IR, Tierney JF, Fisher DJ. Estimating interactions and subgroup-specific treatment effects in meta-analysis without aggregation bias: a within-trial framework. *Res Synth Methods.* 2023;14:68-78. doi:10.1002/jrsm.1590
10. Vale CL, Fisher DJ, Godolphin PJ, et al. Which patients with metastatic hormone-sensitive prostate cancer benefit from docetaxel: a systematic review and meta-analysis of individual participant data from randomised trials. *Lancet Oncol.* 2023;24:783-797. doi:10.1016/S1470-2045(23)00230-9
11. VanderWeele TJ, Knol MJ. Interpretation of subgroup analyses in randomized trials: heterogeneity versus secondary interventions. *Ann Intern Med.* 2011;154:680-683. doi:10.7326/0003-4819-154-10-201105170-00008
12. Rogozińska E, Cornett C, Marlin N, et al. Is Confounding Considered when Estimating Treatment-Covariate Interactions in Individual Participant Data Meta-Analyses of Randomised Trials? Protocol for a methodological study. OSF Preprints. 2021.
13. Marlin N, Godolphin PJ, Hooper RL, Riley RD, Rogozińska E. Nonlinear effects and effect modification at the participant-level in IPD meta-analysis part 1: analysis methods are often substandard. *J Clin Epidemiol.* 2023;159:309-318. doi:10.1016/j.jclinepi.2023.04.013
14. Marlin N, Godolphin PJ, Hooper RL, Riley RD, Rogozińska E. Non-linear effects and effect modification at the participant-level in IPD meta-analysis part 2: methodological guidance is available. *J Clin Epidemiol.* 2023;159:319-329. doi:10.1016/j.jclinepi.2023.04.014
15. Mbuagbaw L, Lawson DO, Puljak L, Allison DB, Thabane L. A tutorial on methodological studies: the what, when, how and why. *BMC Med Res Methodol.* 2020;20:226. doi:10.1186/s12874-020-01107-7
16. 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;36:855-875. doi:10.1002/sim.7141
17. Morris TP, Fisher DJ, Kenward MG, Carpenter JR. Meta-analysis of Gaussian individual patient data: two-stage or not two-stage? *Stat Med.* 2018;37:1419-1438. doi:10.1002/sim.7589
18. Fisher DJ. Two-stage individual participant data meta-analysis and generalized forest plots. *Stata J.* 2015;15:369-396.
19. 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. doi:10.1371/journal.pone.0046042
20. Sullivan TR, White IR, Salter AB, Ryan P, Lee KJ. Should multiple imputation be the method of choice for handling missing data in randomized trials? *Stat Methods Med Res.* 2018;27:2610-2626. doi:10.1177/0962280216683570
21. White IR, Thompson SG. Adjusting for partially missing baseline measurements in randomized trials. *Stat Med.* 2005;24:993-1007. doi:10.1002/sim.1981
22. Shrier I, Redelmeier A, Schnitzer ME, Steele RJ. Challenges in interpreting results from 'multiple regression' when there is interaction between covariates. *BMJ Evidence-Based Med.* 2021;26:53-56. doi:10.1136/bmjebm-2019-111225
23. Kent DM, Paulus JK, van Klaveren D, et al. The predictive approaches to treatment effect heterogeneity (PATH) statement. *Ann Intern Med.* 2020;172:35-45. doi:10.7326/m18-3667
24. Seo M, White IR, Furukawa TA, et al. Comparing methods for estimating patient-specific treatment effects in individual patient data meta-analysis. *Stat Med.* 2021;40:1553-1573. doi:10.1002/sim.8859
25. Stewart LA, Clarke M, Rovers M, et al. Preferred reporting items for a systematic review and meta-analysis of individual participant data: the PRISMA-IPD statement. *JAMA.* 2015;313:1657-1665. doi:10.1001/jama.2015.3656
26. Nevitt SJ, Marson AG, Davie B, Reynolds S, Williams L, Smith CT. Exploring changes over time and characteristics associated with data retrieval across individual participant data meta-analyses: systematic review. *BMJ.* 2017;357:j1390. doi:10.1136/bmj.j1390

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.