ORIGINAL ARTICLE

Nonlinear effects and effect modification at the participant-level in IPD meta-analysis part 2: methodological guidance is available

Nadine Marlin[a,*], Peter J. Godolphin[b], Richard L. Hooper[a], Richard D. Riley[c], Ewelina Rogozińska[b]

[a]*Methodology Research Unit, Centre for Evaluation and Methods, Wolfson Institute of Population Health, Queen Mary University of London, 58 Turner Street, London E1 2AB, UK*
[b]*MRC Clinical Trials Unit at University College London, Institute of Clinical Trials and Methodology, 90 High Holborn, London WC1V 6LJ, UK*
[c]*Institute of Applied Health Research, College of Medical and Dental Sciences, University of Birmingham, Birmingham B15 2TT, UK*

Accepted 26 April 2023; Published online 3 May 2023

Abstract

Objectives: To review methodological guidance for nonlinear covariate-outcome associations (NL), and linear effect modification and nonlinear effect modification (LEM and NLEM) at the participant level in individual participant data meta-analyses (IPDMAs) and their power requirements.

Study Design and Setting: We searched Medline, Embase, Web of Science, Scopus, PsycINFO and the Cochrane Library to identify methodology publications on IPDMA of LEM, NL or NLEM (PROSPERO CRD42019126768).

Results: Through screening 6,466 records we identified 54 potential articles of which 23 full texts were relevant. Nine further relevant publications were published before or after the literature search and were added. Of these 32 references, 21 articles considered LEM, 6 articles NL or NLEM and 6 articles described sample size calculations. A book described all four. Sample size may be calculated through simulation or closed form. Assessments of LEM or NLEM at the participant level need to be based on within-trial information alone. Nonlinearity (NL or NLEM) can be modeled using polynomials or splines to avoid categorization.

Conclusion: Detailed methodological guidance on IPDMA of effect modification at participant-level is available. However, methodology papers for sample size and nonlinearity are rarer and may not cover all scenarios. On these aspects, further guidance is needed. © 2023 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Keywords: Individual participant data meta-analysis; Methodology; Effect modification; Interaction; Nonlinear; Sample size; Fractional polynomials; Restricted cubic splines; One- and two-stage models

* Corresponding author. Methodology Research Unit, Centre for Evaluation and Methods, Wolfson Institute of Population Health, Queen Mary University of London, 58 Turner Street, London E1 2AB, UK. Tel.: +44-0-20-7882-7327; fax: +44-0-20-7882-2552.

E-mail address: n.marlin@qmul.ac.uk (N. Marlin).

1. Introduction

Personalised medicine, also termed precision medicine, is becoming increasingly relevant in health care decision-making. It requires understanding of how treatment effects may vary depending on individual characteristics, for example, gender or age. Individual participant data meta-analysis (IPDMA) of randomized trials (RCTs) are often well suited to investigate such complex participant-level relationships, due to increased sample size over single trials and greater methodological flexibility compared to meta-analysis based on aggregated data [1−3]. This flexibility enables a reliable assessment of linear effect modification (LEM), nonlinear covariate-outcome associations (NL), and nonlinear effect modification (NLEM). Terminology varies in the literature [4] (Box 1).

What is new?

Key findings
- Methodological guidance on individual participant data meta-analysis (IPDMA) of effect modification is available, including how to separate within-trial and across-trial relationships, and how to allow for nonlinearity.

- Further research comparing various proposals for IPDMA of nonlinear covariate outcome relationships or nonlinear effect modification (NLEM) is required.

- Some guidance on a priori sample size requirements is available but not all scenarios are covered.

What this adds to what is known?
- This review provides an overview of available methodology guidance to address nonlinear covariate-outcome associations (NL) and effect modification in IPDMA.

What is the implication?
- Comparison of methodological options (e.g., splines, polynomials) for analyzing NL or NLEM is needed.

Box 1 Terminology for individual participant data meta-analysis of complex relationships

Here we present some brief explanations of commonly used terms in the literature.

LEM, NL and NLEM:

Interaction: The combined effect of two factors is different than their individual effects. During analysis a multiplicative term is included into the model in addition to the individual factors.

Effect modification: It is a type of interaction between a binary intervention indicator and a covariate called the effect modifier. The effect of an intervention differs depending on the level of the modifier characteristic. During analysis an interaction term between the intervention indicator and covariate is included in the model. If the covariate is categorical, the term is also used when the effect is estimated within subsets of data.

Subgroup effect: The effect of the intervention within a subset of patients usually defined by categorical characteristics. The term subgroup effect is used for analyses including interaction terms or analyses within subsets of data.

Nonlinearity: Estimates are not consistent across varying levels of patient characteristics, either in an effect modification or covariate-outcome relationship.

IPDMA approaches and distributional assumptions:

Two-stage IPDMA: The effect of interest is analyzed in each trial separately and the estimates combined using meta-analysis techniques.

One-stage IPDMA: Data from all trials are analyzed together while accounting for clustering by trial.

Common/Fixed effects: The true effect is assumed to be the same across trials. Differences seen in individual trial estimates are only due to sampling error.

Random effects: The true effects in each trial are assumed to follow a normal distribution allowing for between study variations.

Effects stratified by trial: The effect in each trial is independent from those in other trials.

IPDMA of such complex relationships can provide a more nuanced understanding of which patients benefit most from interventions, thereby optimising how treatments are used in practice [5]. For example, Leijten et al. showed that children with more severe conduct problems gained the most from the Incredible Years program [6]. Additionally, such analyses may also identify a need for more effective interventions in certain subgroups, for example, in pulmonary arterial hypertension patients [7]. Interpretation can be challenging, and appropriate expertise is required to properly interpret and communicate such complex analyses.

Effect modification should be considered during the design stage of IPDMAs; however this rarely occurs [8]. Planning a sufficiently powered treatment effect modification analysis requires considerably larger sample sizes than for the overall treatment effect [9,10]. Although researchers have limited impact on sample size (acquired IPD), power considerations have many benefits, such as indicating whether planned analyses have the potential to provide meaningful results. They may support decisions on which analysis to plan or even which trials to focus on for data retrieval [11].

Analysis methods for effect modification should separate within- and across-trial information, to avoid the potential for aggregation bias impacting participant-level relationships. This occurs when a between-trial relationship (e.g., trials that include a higher proportion of women find larger treatment effects) is misinterpreted as a within-trial relationship (the treatment effect is larger in women compared to men) [12−15].

Our previous review found few IPDMA studies reporting power considerations for analysis of effect modification and often inadequate methodology and reporting of LEM, NL and NLEM analyses [16]. It is unclear what guidance for these complex analyses is available.

<div style="border:1px solid">

Box 2 Data extraction

General: Date of extraction, first author, year of publication, abstract, aims, recommendations, limitations.

Analysis methods: general IPDMA approach (one-/two-stage, common/random/stratified effects), specific methods compared or described.

Approach: Literature review included yes/no, comparison of methods yes/no, methods tested on real datasets yes/no, methods tested using simulation yes/no.

Further supporting references.

</div>

In this article, we present findings of a review of current methodology for examining LEM, NL and/or NLEM at the participant-level in IPDMA, and summarize recommendations. This overview will serve anyone involved in the planning, analysis, or review of an IPDMA in exploring the range of potential approaches for their specific IPDMA project.

2. Methods

2.1. Literature review

The detailed methods including search strategy are described elsewhere [16]. In brief, we searched six databases from 01 January 2015, to 04 November 2020, without language restrictions for methods papers describing approaches for IPDMA of LEM, NL, or NLEM. The search strategy was developed in discussion with an information specialist and was sensitive and comprehensive, therefore suitable to identify research studies and methodology papers.

The search was guided by a prospectively registered protocol (CRD42019126768) and recommendations on the conduct of methodological studies [17]. Reference lists and citation indices of relevant publications were hand searched for further relevant methodological papers up to 01 November 2022. Due to the low number of publications on power calculations for LEM, NL, or NLEM in IPDMA we also included references on this topic published before 01 January 2015.

2.2. Eligibility criteria

Methodology publications were eligible if they described, reviewed, assessed, or compared methodology for IPDMA of RCTs addressing effect modification, subgroup effects, NL and/or power calculations. We excluded methodology articles on network meta-analysis, nonfrequentist methods and those dealing with summary-level data only or where the full text was not accessible.

2.3. Screening

One researcher (NM) identified potentially relevant IPD-MA methods papers by screening titles and abstracts. All potentially relevant IPDMA methodology papers underwent full-text review by one researcher (NM). If uncertain, the articles were discussed with other experienced members of the team (RR, PG).

2.4. Data extraction

Data were extracted using a prospectively developed excel spreadsheet by one researcher (NM). In addition, a narrative synthesis was developed by NM and discussed within the team.

We extracted general information, the analysis methods considered, the approach, and if available, aims, recommendations, and limitations (Box 2).

3. Results

Database searches identified 6,466 unique records including 54 potentially eligible methodology articles published between 2015 and 2020 (Fig. 1). They were considered in full text. Of these, 23 were relevant and included in the narrative synthesis together with a further seven articles published after November 2020 and identified up to 01 November 2022, and two articles published before January 2015. These 32 relevant articles mainly focused on the analysis of subgroup effects and effect modification and are considered below. References of excluded articles are listed in the Appendix.

3.1. IPDMA approaches for subgroup effects and linear effect modification

Table 1 presents 21 methodological papers and 1 book chapter considering subgroup effects and effect modification at the participant-level published since 2015. Many of the methodologies presented draw on work by earlier authors, of which most are referenced in the reviews by Riley, Fisher, Gao, Hua, and Simmonds [2,13,18,19,33,35].

3.1.1. Comparison of one- vs. two-stage and common vs. random effects models for subgroup effects and interaction terms

IPDMA of effect modification can either be performed in two stages, where analyses are performed within each trial and the summary measures combined, or in one stage, where individual level data from all trials is analyzed together while accounting for clustering by trial [23] (Box 1).

Riley and colleagues provide comprehensive guidelines on analysis of effect modification in one- or two-stage settings [2,19]. Both publications highlight the problems with

categorization of continuous covariates, the challenges of one-stage approaches when it comes to separating within- and across trial variation and the need to power IPDMAs for analysis of subgroup effects.

We identified three articles comparing common-effect and random-effects and one-stage and two-stage models through simulation [14,30,31]. Belias and Kontopantelis advocate a one-stage approach although Kontopantelis' simulation studies merged across and within-trial relationships, and are therefore prone to aggregation bias. Morris warned that one-stage models are far easier to specify incorrectly but found little difference between two- and one-stage approaches if the models are correctly specified. This is in line with the theoretical comparison performed by Burke [23]. Two further articles by da Costa and Hua compared one-stage approaches with both emphasizing the need to separate within- and across-trial variation [27,33]. Walker used an IPD dataset to compare two- and one-stage approaches with varying assumptions of random effects and found the effect modification estimates to be similar [32]. Convergence and stability issues may dictate the choice of method and prespecification of methods is vital to avoid data dredging.

3.1.2. Other approaches to effect modification

The articles described above consider subgroup analysis or inclusion of prespecified interaction terms into the meta-analysis model. The following four articles describe alternatives when dealing with effect modification.

Vo suggests separation of "case-mix heterogeneity" (i.e., effect modifiers) and "beyond case-mix heterogeneity" (i.e., other differences between studies such as design) [21]. In the presence of heterogeneity an overall treatment effect can still be clinically relevant if it is standardised to a reference population of interest.
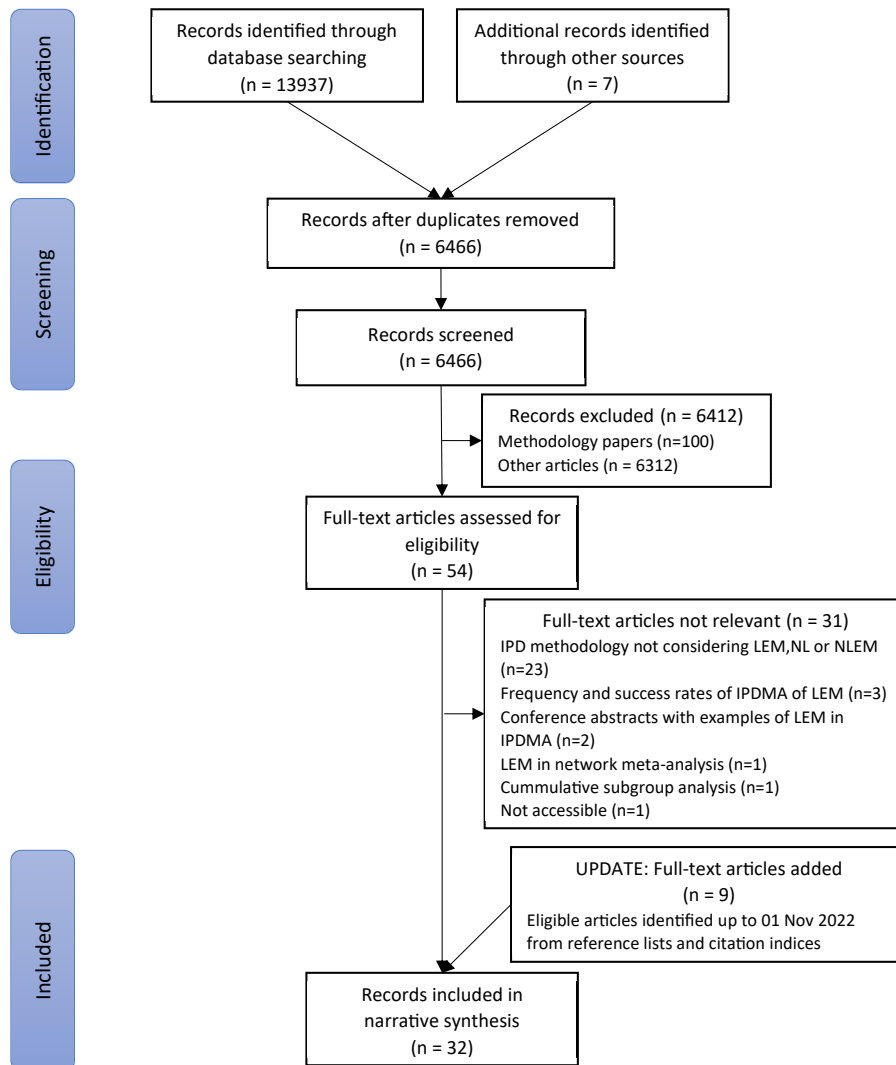

Fig. 1. Flow diagram.

Table 1. Methodological articles focusing on effect modification

Reference	IPDMA approach[a]	Focus[a]	Aggregation bias considered	Recommendation[a]
All outcomes				
Gao 2021 [18]	One- and two-stage	IPDMA of EM in cancer studies	Yes	Prespecify and fully report methods and results of subgroup analyses
Riley 2021 [19]	One- and two-stage	Guideline on analysis of effect modification	Yes	Avoid aggregation bias and categorization of continuous covariates. Presence of effect modification may depend on scale.
Schandelmaier 2020 [4]	One- and two-stage	Credibility of EM analyses	Yes	Tool for judging EM analyses
Riley 2020 [2]	One- and two-stage	Guideline on analysis of effect modification	Yes	Aggregation bias in one-stage analysis can be dealt with by centering or stratification of nuisance parameters
Jiao 2020 [20]	Two-stage	Confidence Distributions based mapping method	Yes	Approach for analyzing multiple related covariates across studies
Belias 2019 [14]	One- and two-stage	Comparison of one- and two-stage models for binary effect modifiers	Yes	Centred one-stage model recommended for binary outcomes
Vo 2019 [21]	Two-stage	Case-mix heterogeneity	Yes	Address case-mix heterogeneity when subgroups are not of interest
Mistry 2018 [22]	One-stage	Tree-based approach, Categorical effect modifiers only	No	Approach for exploring large numbers of effect modifiers, Performs well with large between trial variation
Burke 2017 [23]	One- and two-stage	Differences between one- and two-stage models	Yes	Correctly specified one- and two-stage models perform equally well unless most studies are sparse
Fisher 2017 [13]	One- and two-stage, meta-regression	Validity and reporting of MA of EM	Yes	Meta-analyse interactions, not subgroup effects
Fisher 2015 [24]	Two-stage	Stata command IPDMETAN	Yes	Convenient way to model two-stage IPDMA
Riley 2015 [25]	Two-stage	Multivariate MA of multiple outcomes	Yes	Estimation of within-study correlations in a joint linear regression using Bayesian and frequentist methods
Continuous outcomes				
Papadimitropoulou 2020 [26]	One- and two-stage, meta-regression	Pseudo IPD reconstructed from published aggregate data	Yes	Use of Pseudo IPD is valid if IPD is unavailable and suitable aggregate data about baseline and follow-up is available
da Costa 2019 [27]	One-stage, meta-regression	Methods comparison for MA of subgroup effects	Yes	Allow for the between-trial variation in interaction effects
Noma 2019 [28]	Two-stage	IPDMA of EM for longitudinal data	Yes	Two-stage mixed effects model approach for main and interaction effects
Fokkema 2018 [29]	One-stage	Tree-based approach, Categorical effect modifiers only	No	Approach for exploring large numbers of effect modifiers
Morris 2018 [30]	One- and two-stage	Comparison of one- and two-stage models	Yes	One- and two-stage models perform equally well if correctly specified
Kontopantelis 2018 [31]	One- and two-stage	Comparison of one- and two-stage models	Yes	Use fully specified 1 stage model
Binary outcomes including time-to-event analyses				
Walker 2022 [32]	One- and two-stage	Case study comparison of one- and two-stage models	Yes	Prespecify methods, more real-world explorations are needed
Hua 2017 [33]	One- and two-stage	Addressing aggregation bias	Yes	Separate within-and across-trial variation
Chen 2017 [1]	One- and two-stage	Quantifying heterogeneity	No	Performance of measurements depend on model
Wang 2016 [34]	Two- stage	Visual exploration of continuous effect modifiers. Univariate common effects model only	Yes	Meta-STEPP: Method to identify and model complex EM patterns avoiding categorization.

[a] IPD, individual participant data; MA, meta-analysis; EM, effect modification.

Table 2. Publications on methods for nonlinear covariate-outcome associations and nonlinear effect modification

Reference	Type of outcome	IPDMA approach	Focus[a]	Recommendations[a]
Nonlinear effect modification				
Belias 2022 [36]	Binary	One- and two-stage	4 spline approaches and pointwise MA, multivariate MA, GAMMs	Presence of effect modification may depend on scale. GAMMs are powerful but require careful modeling.
Sauerbrei 2022 [37]	Any	Two-stage	MFPI and pointwise MA ("metaTEF")	Report analysis using the MethProf-MA profile
Riley 2021 [19]	Any	One- and two-stage	Restricted cubic splines and multivariate MA	Nonlinear treatment-covariate interactions should be investigated. Two-stage multivariate IPDMA of restricted cubic spline functions. Results may depend on the scale.
Riley 2020 [2]	Any	Two-stage	Multivariate MA of splines for NL	Separate within/across trial variation and allow for NL.
Kasenda 2016 [38]	Any	Two-stage	MFPI and pointwise MA	MFPI avoids categorization and allows for nonlinearity in effect modification analyses
Nonlinear covariate-outcome relationships				
White 2019 [39]	Any	Two-stage	FP for nonlinear associations	Modeling nonlinear effects is superior to dichotomization and subgroup analysis
Nonlinear associations in baseline risk				
DeJong 2020 [40]	Time to event	One- and two-stage	Modeling baseline hazard and non-PH	Model non-PH Cox models by rescaling instead of nonlinear or interaction terms.

[a] MA, meta-analysis; GAMM, generalized additive mixed effects model; MFPI, multivariable fractional polynomial interaction approach; non-PH, nonproportional hazards; NL, nonlinearity; FP, fractional polynomial(s).

Two articles by Fokkema and Mistry describe the exploration of larger numbers of potential subgroup effects using tree-based methods [22,29]. Amalgamation of within- and across-trial variation is not addressed in these articles. Jiao presents a mapping approach for investigating multiple covariates across datasets using two-step approach that first links study-specific vectors of parameters and then estimates hyperparameters using a multivariate random-effects meta-analysis model [20].

The META-STEPP approach estimates subpopulation treatment effects based on overlapping patient subpopulations [34]. Treatment effects are analyzed by standard common-effects meta-analysis methodology. This approach may be useful for larger numbers of effect modifiers and complex effect modification but does not separate within- and across trial variation.

Four further papers address specialized issues when analyzing effect modification: use of pseudo IPD [26], analysis of repeated measures data [28], measures of heterogeneity [1] and multivariate meta-analysis of multiple outcomes [25].

3.1.3. Reporting
Fisher reviewed the methods used to analyse effect modification in IPDMA research studies published between 2011 and 2014 [13]. Of those few with sufficient description, most did not separate within- and across-trial variation correctly and were at risk of aggregation bias. Two-stage IPDMAs of interaction terms inherently address this issue, whilst one-stage approaches require more care in model specification. A review of cancer IPD studies by Gao found a similar lack in clear reporting and appropriate analysis methods used, with all IPDMAs that included continuous covariates categorizing them when assessing effect modification [18].

Schandelmaier developed the ICEMAN tool to score the credibility of effect modification analyses [4]. Credibility is gained on factors including the use of random-effects models, the separation of within- and between-study effects and avoiding categorizing continuous covariates.

3.1.4. Statistical software
Fisher published the Stata command (IPDMETAN), which performs both stages of a two-stage IPD meta-analysis [24]. Effect modification analysis and inclusion of nonlinear terms is possible.

3.2. IPDMA approaches for nonlinear covariate-outcome relationships and nonlinear effect modification

We found no published reviews of IPDMA methods for NL. We identified six methodological papers and one book

Table 3. Publications on methods for sample size calculation of LEM, NL or NLEM in IPDMA

Reference	IPDMA approach	Calculation approach	Aggregation bias considered	Recommendation[a]
All outcomes				
Riley 2021 [11]	One- and two-stage	Simulation-based approach, Closed form	Yes	Extension to allow for heterogeneity
Ensor 2018 [10]	Two-stage	Simulation-based approach	Yes	When planning an IPDMA assess power for main effect and effect modification
Kontopantelis 2016 [46]	One-stage	Simulation-based approach	No	Stata command IPDPOWER, but does not separate out within and across-trial relationships, so power will be inflated
Continuous outcomes				
Riley 2020 [2]	Two-stage	Closed form	Yes	Assume no between-study heterogeneity in size of EM
Kovalchik 2012 [47]	One-stage, meta-regression	Closed form	No	Estimate power of IPDMA of effect modification using aggregate data. Power estimates are prone to error due to approximations and amalgamation of within and across-trial information
Simmonds 2007 [48]	One- and two-stage, meta-regression	Closed form	Yes (two-stage), No (one-stage)	Power of each method depends on covariate distribution and sample size, Q statistics measures covariate heterogeneity
Binary outcomes				
Riley 2022 [49]	Two-stage	Closed form	Yes	Improved approximation of variances based on existing aggregate data. Stata and R code are provided
Kovalchik 2012 [47]	One-stage, meta-regression	Closed form	No	Estimate power of IPDMA of effect modification using aggregate data. Power estimates are prone to error due to approximations and amalgamation of within and across-trial information

[a] EM, effect modification.

chapter that described methods for either NLEM or nonlinear relationships between covariates and outcomes (Table 2).

Splines and fractional polynomials can be used to model nonlinear covariate-outcome relationships and effect modification in two-stage models [2,37–39]. Best fitting nonlinear effect (modification) is identified in the first stage and then combined in the second stage pointwise (meta-curve [41]) or using multivariate meta-analysis (mvmeta [42]). The former allows for study-specific polynomial functions, the latter only for common functions. White also show the advantages of allowing for nonlinear covariate-outcome relationships over the commonly used categorization approach [39].

Riley and colleagues suggest using restricted cubic splines for their increased flexibility compared to fractional polynomials and combining them using multivariate meta-analysis [2,19]. If a one-stage approach is desired this can be done by stratifying the trial parameters outside the interaction term. They highlight that effect modification may depend on the scale of the analysis and refer to a theoretical

example by Shrier and Pang who found a statistically significant interaction when analyzing odds ratios but not when analyzing risk ratios [43]. This is due to differences in baseline risk, and can therefore also be seen, for example, in survival analysis of time-to-event outcomes.

Belias compares four types of splines and three pooling methods for nonlinear effects and effect modification [36]. Although the choice of spline had little impact, some differences were found for the pooling methods. A one-stage approach using generalized additive mixed effects models (GAMMs) handled splines from differing data ranges and sample sizes better than pointwise meta-analysis or multivariate meta-analysis. However, modeling GAMMs is complex and requires great care.

Belias description of the use of GAMMs is the only guideline on modeling NLEM in a one-stage setting we identified. Some other possible approaches and their challenges have been discussed by Riley and colleagues [2,19].

DeJong describes how nonlinear terms and interactions can be used to model baseline hazard functions and

nonproportional hazards in survival analysis [40]. For details on the modeling, they refer to other articles [39,44,45]. Instead of using nonlinear terms the authors suggest achieving proportionality of nonproportional hazards by modeling on a different scale and describe the example of a log-logistic model. If nonlinear terms are used, interpretation can be challenging and the article describes two potentially more clinically meaningful effect measurements, restricted mean survival time difference and percentile ratio. DeJong suggests for one-stage approaches of sufficient sample size, stratification of all parameters is the safest choice and modeling the intervention effect as random to account for heterogeneity.

3.3. Sample size calculation for complex relationships in IPDMA

We identified six articles and one book chapter discussing sample size calculation for IPDMAs (Table 3). Three describe simulation-based approaches that allow for modeling of effect modification and specification of nonlinear terms [10,11,46]. Closed form approaches are used in five references [2,11,47–49].

Simmonds first compared the power of three methods to model effect modification: two-stage or one-stage meta-analysis of interaction terms and meta-regression [48]. One-stage models were found to have the largest power but only under a common effects model and ignoring aggregation bias. These are strong assumptions which may not hold. The one-stage approaches presented by Kovalchik and Kontopantelis also do not account for aggregation bias, and can therefore result in too small sample size estimations [46,47].

Riley and colleagues present closed form approaches for continuous and binary outcomes addressing these issues [2,49]. One of the challenges is to estimate the amount of heterogeneity in the size of the interaction in advance and initially the authors suggest assuming an ideal case where no such heterogeneity exists. However, extensions to allow for between-trial heterogeneity are discussed in their book and publication [19,49].

4. Discussion

4.1. Main findings

In this article we present a review of methodology publications for IPDMA of LEM, nonlinear covariate-outcome relationships, and NLEM including their sample size calculations. Our preceding review of IPD research studies showed that such analyses are common in IPD but rarely implemented correctly or powered for [16]. Easy to follow guidance is needed to support researchers in producing unbiased results that underpin clinical decision making.

We have identified numerous publications describing how to correctly model effect modification at the participant level in a one-or two-stage setting. Many of these have been published in the years considered (2015–2020) although earlier authors (such as Fisher [12]) indicate the challenges in a one-stage setting. It is, therefore, perhaps not surprising that most of the IPDMA research studies published during this time did not implement unbiased procedures although this may be an issue of reporting rather than methodology [16].

Only a few methodology publications on sample size considerations were found and they may not cover all scenarios especially around one-stage approaches and NL. Simulation approaches could be adapted in these cases.

Guidelines on avoiding categorization by analyzing nonlinear covariate-outcome relationships and NLEM are so far focussed on two-stage approaches with some extension for one-stage models.

4.2. Limitations

The literature search covered the years 2015 to 2020 and was then updated in November 2022 nonsystematically. It is, therefore, possible that relevant publications during 2021 and 2022 may have been missed. However, we did perform extensive searches through reference lists and citation indices and discussed with experts in the field, thus identifying the most relevant publications.

Additionally, we found little variation in authorship. Most of the articles, including the current review, are authored by a small number of established teams. However, we used a sensitive search strategy, and our exclusion criteria did not discriminate against references by less established authors in the field, for example by favouring high impact journals. We believe this is a comprehensive overview of the currently available methodology guidance.

4.3. Best practice recommendations

Based on this review and the preceding review of research studies we make the following recommendations for planning, analysis, and reporting of complex associations in IPDMA.

1. Consider the power for effect modification a priori

Power calculations for assessing effect modification in IPDMA are currently not part of PRISMA-IPD reporting guidelines but help reveal if the IPDMA is worth the time and cost especially if effect modification is part of the main research question. This can be done before IPD collection, based on summary aggregate data from published trials, and under assumptions about true interaction effect sizes [11].

Well defined closed form solutions may not be available for all scenarios, but a simulation-based approach should work in such cases [10]. Easy-to-follow guidance for all scenarios is needed.

2. Choose an appropriate analysis model a priori and consider assumptions and implications

One- and two-stage methods produce similar results if modeling assumptions are matching including how each parameter (treatment effect, covariate effects, intercept, residual variances etc.) is modeled: common, random, or stratified effect. None of the IPDMAs in the preceding review described all these assumptions [16]. However, this choice can strongly impact results and their interpretations [2].

Assessment of effect modification at the participant-level needs to be based on within-trial information alone to avoid the potential for aggregation bias. In cases without any heterogeneity in the estimated effect this is automatically the case. In a two-stage approach this is also automatically done as interaction terms are modeled within each study first and then combined. In the one-stage model within-trial and across-trial information need to be actively separated out, by (1) stratifying all parameters outside the interaction by trial or (2) centering the effect modifier by its trial-specific mean [33].

3. Avoid categorization of continuous covariates

Analyzing continuous covariates instead of categorizing them (1) increases power to detect effect modification if it exists and (2) allows investigation of nonlinearity. If data is shared as continuous then categorization should only be used for exploration but not for primary analysis [10,39].

4. Consider nonlinearity for effect modification of continuous covariates

Nonlinearity in effect modification should be considering when analyzing effect modification by a continuous covariate [2,19,48].

Two main approaches have been suggested using either splines or fractional polynomials. In a single trial setting, little difference has been found between the two methods [50] although they have not been formal compared in an IPD setting. Both approaches are easily modeled in a two-stage IPDMA but challenges arise in a one-stage setting.

5. Adhere to PRISMA-IPD reporting guidelines and include statistical code/formal model specification in publications

When reporting IPDMAs, researchers should adhere to guideline such as PRISMA-IPD and if possible, publish software code or write out the formal model specification to improve understanding and reproducibility especially of one-stage models.

5. Conclusion

Guidance on correct IPDMA of complex relationships using one- or two-stage approaches is available and should be used more widely. This will provide higher quality evidence to better support clinical decision making.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jclinepi.2023.04.014.

References

[1] Chen B, Benedetti A. Quantifying heterogeneity in individual participant data meta-analysis with binary outcomes. Syst Rev 2017;6(1): 243.

[2] 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.

[3] Tierney JF, Stewart LA, Clarke M. Chapter 26: individual participant data. In: Higgins JPT, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editors. Cochrane handbook for systematic reviews of interventions version 6.3 (updated February 2022). Cochrane; 2022. Available at www.training.cochrane.org/handbook. Accessed June 12, 2023.

[4] Schandelmaier S, Briel M, Varadhan R, Schmid CH, Devasenapathy N, Hayward RA, et al. Development of the instrument to assess the credibility of effect modification analyses (ICEMAN) in randomized controlled trials and meta-analyses. CMAJ 2020; 192(32):E901−6.

[5] Marlin N, Allotey J. The difference between effect modification and covariate confounding. BJOG 2021;128:1574.

[6] Leijten P, Scott S, Landau S, Harris V, Mann J, Hutchings J, et al. Individual participant data meta-analysis: impact of conduct problem severity, comorbid attention-deficit/hyperactivity disorder and emotional problems, and maternal depression on parenting program effects. J Am Acad Child Adolesc Psychiatry 2020;59(8):933−43.

[7] Rhee RL, Gabler NB, Sangani S, Praestgaard A, Merkel PA, Kawut SM. Comparison of treatment response in idiopathic and connective tissue disease-associated pulmonary arterial hypertension. Am J Respir Crit Care Med 2015;192:1111−7.

[8] Schmidt AF, Klungel OH, Nielen M, de Boer A, Groenwold RH, Hoes AW. Tailoring treatments using treatment effect modification. Pharmacoepidemiol Drug Saf 2016;25(4):355−62.

[9] Brookes ST, Whitely E, Egger M, Smith GD, Mulheran PA, Peters TJ. Subgroup analyses in randomized trials: risks of subgroup-specific analyses; power and sample size for the interaction test. J Clin Epidemiol 2004;57:229−36.

[10] Ensor J, Burke DL, Snell KIE, Hemming K, Riley RD. Simulation-based power calculations for planning a two-stage individual participant data meta-analysis. BMC Med Res Methodol 2018;18:41.

[11] Riley RD, Fisher DJ. In: Riley RD, Tierney JF, Stewart LA, editors. Chapter 12: power calculations for planning an IPD meta-analysis. Individual participant data meta-analysis: a handbook for healthcare research. Hoboken, NJ: Wiley; 2021. https://doi.org/10.1002/9781119333784.

[12] Fisher DJ, Copas AJ, Tierney JF, Parmar MK. A critical review of methods for the assessment of patient-level interactions in individual participant data meta-analysis of randomized trials, and guidance for practitioners. J Clin Epidemiol 2011;64:949—67.

[13] 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.

[14] 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.

[15] 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 2022;14:68—78.

[16] Marlin N, Godolphin PJ, Hooper R, Riley R, Rogozinska E. Examining non-linear effects and effect modification at the participant-level in IPD meta-analysis part 1: analysis methods are often substandard. 2023. Available at https://dx.doi.org/10.2139/ssrn.4333137. Accessed March 19, 2023.

[17] 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.

[18] Gao Y, Liu M, Shi S, Niu M, Li J, Zhang J, et al. Prespecification of subgroup analyses and examination of treatment-subgroup interactions in cancer individual participant data meta-analyses are suboptimal. J Clin Epidemiol 2021;138:156—67.

[19] Riley RD, Fisher DJ. In: Riley RD, Tierney JF, Stewart LA, editors. Chapter 7: using IPD meta-analysis to examine interactions between treatment effect and participant-level covariates. Individual participant data meta-analysis: a handbook for healthcare research. Hoboken, NJ: Wiley; 2021. https://doi.org/10.1002/9781119333784.

[20] Jiao Y, Mun EY, Trikalinos TA, Xie M. A CD-based mapping method for combining multiple related parameters from heterogeneous intervention trials. Stat Interface 2020;13(4):533—49.

[21] Vo TT, Porcher R, Chaimani A, Vansteelandt S. A novel approach for identifying and addressing case-mix heterogeneity in individual participant data meta-analysis. Res Synth Methods 2019;10(4):582—96.

[22] Mistry D, Stallard N, Underwood M. A recursive partitioning approach for subgroup identification in individual patient data meta-analysis. Stat Med 2018;37:1550—61.

[23] 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.

[24] Fisher DJ. Two-stage individual participant data meta-analysis and generalized forest plots. Stata J 2015;15(2):369—96.

[25] Riley RD, Price MJ, Jackson D, Wardle M, Gueyffier F, Wang J, et al. Multivariate meta-analysis using individual participant data. Res Synth Methods 2015;6(2):157—74.

[26] Papadimitropoulou K, Stijnen T, Riley RD, Dekkers OM, Cessie SL. Meta-analysis of continuous outcomes: using pseudo IPD created from aggregate data to adjust for baseline imbalance and assess treatment-by-baseline modification. Res Synth Methods 2020;11:780—94.

[27] 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.

[28] Noma H, Maruo K, Gosho M, Levine SZ, Goldberg Y, Leucht S, et al. Efficient two-step multivariate random effects meta-analysis of individual participant data for longitudinal clinical trials using mixed effects models. BMC Med Res Methodol 2019;19:33.

[29] Fokkema M, Smits N, Zeileis A, Hothorn T, Kelderman H. Detecting treatment-subgroup interactions in clustered data with generalized linear mixed-effects model trees. Behav Res Methods 2018;50(5):2016—34.

[30] 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—38.

[31] Kontopantelis E. A comparison of one-stage vs two-stage individual patient data meta-analysis methods: a simulation study. Res Synth Methods 2018;9(3):417—30.

[32] Walker R, Stewart L, Simmonds M. Estimating interactions in individual participant data meta-analysis: a comparison of methods in practice. Syst Rev 2022;11(1):211.

[33] 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.

[34] Wang XV, Cole B, Bonetti M, Gelber RD. Meta-STEPP: subpopulation treatment effect pattern plot for individual patient data meta-analysis. Stat Med 2016;35:3704—16.

[35] Simmonds M, Stewart G, Stewart L. A decade of individual participant data meta-analyses: a review of current practice. Contemp Clin Trials 2015;45:76—83.

[36] Belias M, Rovers MM, Hoogland J, Reitsma JB, Debray TPA, IntHout J. Predicting personalised absolute treatment effects in individual participant data meta-analysis: an introduction to splines. Res Synth Methods 2022;13:255—83.

[37] Sauerbrei W, Royston P. Investigating treatment-effect modification by a continuous covariate in IPD meta-analysis: an approach using fractional polynomials. BMC Med Res Methodol 2022;22:98.

[38] Kasenda B, Sauerbrei W, Royston P, Mercat A, Slutsky AS, Cook D. Multivariable fractional polynomial interaction to investigate continuous effect modifiers in a meta-analysis on higher versus lower PEEP for patients with ARDS. BMJ Open 2016;6(9):e011148.

[39] White IR, Kaptoge S, Royston P, Sauerbrei W. Meta-analysis of non-linear exposure-outcome relationships using individual ;participant data: a comparison of two methods. Stat Med 2019;38:326—38.

[40] de Jong VMT, Moons KGM, Riley RD, Tudur Smith C, Marson AG, Eijkemans MJC, et al. Individual participant data meta-analysis of intervention studies with time-to-event outcomes: a review of the methodology and an applied example. Res Synth Methods 2020;11(2):148—68.

[41] Sauerbrei W, Royston P. A new strategy for meta-analysis of continuous covariates in observational studies. Stat Med 2011;30:3341—60.

[42] White IR. Multivariate random-effects meta-analysis. Stata J 2009;9(1):40—56.

[43] Shrier I, Pang M. Confounding, effect modification, and the odds ratio: common misinterpretations. J Clin Epidemiol 2015;68:470—4.

[44] Hess KR. Assessing time-by-covariate interactions in proportional hazards regression models using cubic spline functions. Stat Med 1994;13:1045—62.

[45] Giorgi R, Abrahamowicz M, Quantin C, Bolard P, Esteve J, Gouvernet J, et al. A relative survival regression model using B-spline functions to model non-proportional hazards. Stat Med 2003;22:2767—84.

[46] Kontopantelis E, Springate DA, Parisi R, Reeves D. Simulation-based power calculations for mixed effects modeling: ipdpower in stata. J Stat Softw 2016;74:1—25.

[47] Kovalchik SA, Cumberland WG. Using aggregate data to estimate the standard error of a treatment-covariate interaction in an individual patient data meta-analysis. Biom J 2012;54(3):370—84.

[48] 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.

[49] Riley RD, Hattle M, Collins GS, Whittle R, Ensor J. Calculating the power to examine treatment-covariate interactions when planning an individual participant data meta-analysis of randomized trials with a binary outcome. Stat Med 2022;41:4822−37.

[50] Kahan BC, Rushton H, Morris TP, Daniel RM. A comparison of methods to adjust for continuous covariates in the analysis of randomised trials. BMC Med Res Methodol 2016;16:42.