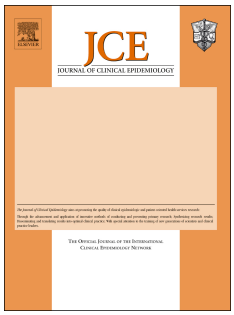
Journal Pre-proof



Non-linear effects and effect modification at the participant-level in IPD meta-analysis part 2: Methodological guidance is available

Nadine Marlin, Peter J. Godolphin, Richard Hooper, Richard Riley, Ewelina Rogozinska

PII: S0895-4356(23)00102-6

DOI: https://doi.org/10.1016/j.jclinepi.2023.04.014

Reference: JCE 11071

To appear in: *Journal of Clinical Epidemiology*

Received Date: 16 January 2023

Revised Date: 20 March 2023

Accepted Date: 26 April 2023

Please cite this article as: Marlin N, Godolphin PJ, Hooper R, Riley R, Rogozinska E, Non-linear effects and effect modification at the participant-level in IPD meta-analysis part 2: Methodological guidance is available, *Journal of Clinical Epidemiology* (2023), doi: https://doi.org/10.1016/j.jclinepi.2023.04.014.

Non-linear effects and effect modification at the participant-level in IPD meta-analysis part 2: Methodological guidance is available

Authors: Nadine Marlin[a], Peter J Godolphin[b], Richard Hooper[a], Richard Riley[c], Ewelina Rogozinska[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, London, UK
90 High Holborn, London WC1V 6LJ, UK

c) Institute of Applied Health Research, College of Medical and Dental Sciences, University of Birmingham
Birmingham, UK. B15 2TT

Corresponding author: Nadine Marlin, 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

Email: n.marlin@qmul.ac.uk, Phone: +44 (0) 20 7882 7327

Abstract

Objective: To review methodological guidance for non-linear associations (NL), and linear and non-linear 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 6466 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. Non-linearity (NL or NLEM) can be modelled using polynomials or splines to avoid categorisation.

Conclusion: Detailed methodological guidance on IPDMA of effect modification at participant-level is available. However, methodology papers for sample size and non-linearity are rarer and may not cover all scenarios. On these aspects, further guidance is needed.

Word count: 196

Key words:

- Individual participant data meta-analysis
- Methodology
- Effect modification
- interaction
- Non-linear
- Sample size

Running title: Non-linear effects and effect modification in IPD meta-analysis

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 randomised 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), non-linear covariate-outcome associations (NL), and non-linear effect modification (NLEM). Terminology varies in the literature [4] (Box 1).

IPDMA of these 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]. While 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 (for example, 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.

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 summarise 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.

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.

Non-linearity: 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 analysed in each trial separately and the estimates combined using meta-analysis techniques.

One-stage IPDMA: Data from all trials are analysed 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 variation.

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

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

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 Nov 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 Jan 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, non-linear associations and/or power calculations. We excluded methodology articles on network meta-analysis, non-frequentist 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 IPDMA 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**Error! Reference source not found.**).

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

Box 2: Data extraction

3) Results

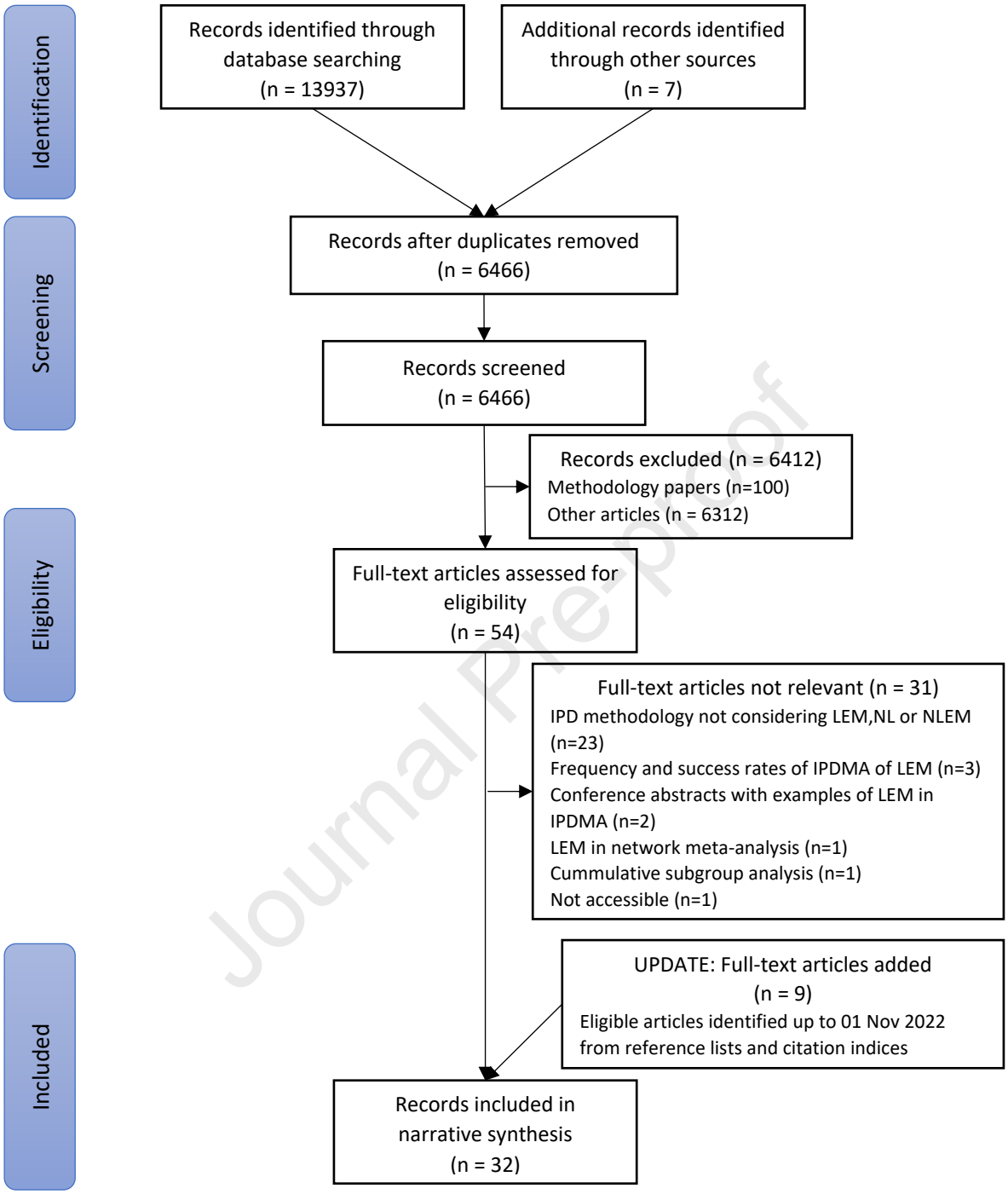


Figure 1: Flow diagram

Database searches identified 6466 unique records including 54 potentially eligible methodology articles published between 2015 and 2020 (Figure 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 Nov 2020 and identified up to 01 Nov 2022 and two articles published before Jan 2015. These 32 relevant articles mainly focussed 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, 18-21]

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 analysed together while accounting for clustering by trial [22] (Box 1).

Riley and colleagues provide comprehensive guidelines on analysis of effect modification in one- or two-stage settings [2, 18]. Both publications highlight the problems with categorisation 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, 23, 24]. 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 [22]. Two further articles by da Costa and Hua compared one-stage approaches with both emphasizing the need to separate within- and across-trial variation [20, 25]. 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 [26]. Convergence and stability issues may dictate the choice of method and pre-specification 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 pre-specified 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) [27]. In the presence of heterogeneity an overall treatment effect can still be clinically relevant if it is standardised to a reference population of interest.

Two articles by Fokkema and Mistry describe the exploration of larger numbers of potential subgroup effects using tree-based methods [28, 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 employing two-step approach that first links study-specific vectors of parameters and then estimates hyperparameters using a multivariate random-effects meta-analysis model [30].

The META-STEPP approach estimates subpopulation treatment effects based on overlapping patient subpopulations [31]. Treatment effects are analysed 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 specialised issues when analysing effect modification: use of pseudo IPD [32], analysis of repeated measures data [33], measures of heterogeneity [1] and multivariate meta-analysis of multiple outcomes [34].

3.1.3. Reporting

Fisher reviewed the methods used to analyse effect modification in IPDMA research studies published between 2011-2014 [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 [19].

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 [35]. Effect modification analysis and inclusion of non-linear terms is possible.

Table 1: Methodological articles focusing on effect modification

Reference	IPDMA approach*	Focus*	Aggregation bias considered	Recommendation*
All outcomes				
Gao 2021 [19]	One- and two-stage	IPDMA of EM in cancer studies	Yes	Pre-specify and fully report methods and results of subgroup analyses
Riley 2021 [18]	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 [30]	Two-stage	Confidence Distributions based mapping method	Yes	Approach for analysing multiple related covariates across studies
Belias 2019 [14]	One- and two-stage	Comparison of one- and two-stage models	Yes	Centred one-stage model recommended for binary outcomes

		for binary effect modifiers		
Vo 2019 [27]	Two-stage	Case-mix heterogeneity	Yes	Address case-mix heterogeneity when subgroups are not of interest
Mistry 2018 [29]	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 [22]	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 [35]	Two-stage	Stata command IPDMETAN	Yes	Convenient way to model two-stage IPDMA
Riley 2015 [34]	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 [32]	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 [25]	One-stage, meta-regression	Methods comparison for MA of subgroup effects	Yes	Allow for the between-trial variation in interaction effects
Noma 2019 [33]	Two-stage	IPDMA of EM for longitudinal data	Yes	Two-stage mixed effects model approach for main and interaction effects
Fokkema 2018 [28]	One-stage	Tree-based approach, Categorical effect modifiers only	No	Approach for exploring large numbers of effect modifiers

Morris 2018 [24]	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 [23]	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 [26]	One- and two-stage	Case study comparison of one- and two-stage models	Yes	Pre-specify methods, more real-world explorations are needed
Hua 2017 [20]	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 [31]	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 categorisation.

*IPD… individual participant data, MA… meta-analysis, EM… Effect modification

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

We found no published reviews of IPDMA methods for non-linear associations. We identified six methodological papers and one book chapter that described methods for either non-linear effect modification or non-linear relationships between covariates and outcomes (Table 2).

Splines and fractional polynomials can be used to model non-linear covariate-outcome relationships and effect modification in two-stage models [2, 36-38]. Best fitting non-linear effect (modification) is identified in the first stage and then combined in the second stage pointwise (metacurve, [39]) or using multivariate meta-analysis (mvmeta, [40]). The former allows for study-specific polynomial functions, the latter only for common functions. White also show the advantages of allowing for non-linear covariate-outcome relationships over the commonly used categorization approach [37].

Riley and colleagues suggest using restricted cubic splines for their increased flexibility compared to fractional polynomials and combining them using multivariate meta-analysis [2, 18]. 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 analysing odds ratios but not when analysing risk ratios [41]. 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 non-linear effects and effect modification [42]. While the choice of spline had little impact, some differences were found for the pooling methods. A one-stage approach using generalised additive mixed effects models (GAMMs) handled splines from differing data ranges and sample sizes better than pointwise meta-analysis or multivariate meta-analysis. However, modelling GAMMs is complex and requires great care.

Belias description of the use of GAMMs is the only guideline on modelling non-linear effect modification in a one-stage setting we identified. Some other possible approaches and their challenges have been discussed by Riley and colleagues [2, 18].

DeJong describes how non-linear terms and interactions can be used to model baseline hazard functions and non-proportional hazards in survival analysis [43]. For details on the modelling, they refer to other articles [37, 44, 45]. Instead of using non-linear terms the authors suggest achieving proportionality of non-proportional hazards by modelling on a different scale and describe the example of a log-logistic model. If non-linear 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 modelling the intervention effect as random to account for heterogeneity.

Table 2: Publications on methods for non-linear associations and non-linear effect modification

Reference	Type of outcome	IPDMA approach	Focus*	Recommendations*
Non-linear effect modification				
Belias 2022 [42]	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 modelling.
Sauerbrei 2022 [36]	Any	Two-stage	MFPI and pointwise MA ("metaTEF")	Report analysis using the MethProf-MA profile
Riley 2021 [18]	Any	One- and two-stage	Restricted cubic splines and multivariate MA	Non-linear 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 categorisation and allows for non-linearity in effect modification analyses
Non-linear covariate-outcome relationships				
White 2019 [37]	Any	Two-stage	FP for non-linear associations	Modelling non-linear effects is superior to dichotomization and subgroup analysis
Non-linear associations in baseline risk				

DeJong 2020 [43]	Time to event	One- and two-stage	Modelling baseline hazard and non-PH	Model non-PH Cox models by rescaling instead of non-linear or interaction terms.

* MA… meta-analysis, GAMM…, MFPI… Multivariable fractional polynomial interaction approach, non-PH… non-proportional hazards, NL… non-linearity, FP… fractional polynomial(s)

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 modelling of effect modification and specification of non-linear 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 [49]. 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, 48].

Riley and colleagues present closed form approaches for continuous and binary outcomes addressing these issues [2, 47]. 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 [18, 47].

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*
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 [48]	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 [49]	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 [47]	Two-stage	Closed form	Yes	Improved approximation of variances based on existing aggregate data. Stata and R code are provided
Kovalchik 2012 [48]	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

* EM… Effect modification

4) Discussion

4.1. Main findings

In this article we present a review of methodology publications for IPDMA of linear effect modification, non-linear covariate-outcome relationships, and non-linear effect modification 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 non-linear associations. Simulation approaches could be adapted in these cases.

Guidelines on avoiding categorisation by analysing non-linear covariate-outcome relationships and non-linear effect modification 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 Nov 2022 non-systematically. 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]. Easier 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 modelling assumptions are matching including how each parameter (treatment effect, covariate effects, intercept, residual variances etc) is modelled: 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 modelled 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 [20].

3. Avoid categorisation of continuous covariates

Analysing continuous covariates instead of categorizing them (1) increases power to detect effect modification if it exists and (2) allows investigation of non-linearity. If data is shared as continuous then categorisation should only be used for exploration but not for primary analysis [10, 37].

4. Consider non-linearity for effect modification of continuous covariates

Non-linearity in effect modification should be considering when analysing effect modification by a continuous covariate [2, 18, 49].

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 modelled 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 utilised more widely. This will provide higher quality evidence to better support clinical decision making.

6) References

1. Chen, B. and A. Benedetti, *Quantifying heterogeneity in individual participant data meta-analysis with binary outcomes.* Systematic Reviews, 2017. **6**(1).
2. Riley, R.D., et al., *Individual participant data meta-analysis to examine interactions between treatment effect and participant-level covariates: Statistical recommendations for conduct and planning.* Statistics in Medicine, 2020. **39**(15): p. 2115-2137.
3. Tierney JF, S.L., Clarke M, *Chapter 26: Individual participant data.*, in *Cochrane Handbook for Systematic Reviews of Interventions version 6.3 (updated February 2022)*, T.J. Higgins JPT, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, Editor. 2022, Cochrane.
4. Schandelmaier, S., 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): p. E901-E906.
5. Marlin, N. and J. Allotey, *The difference between effect modification and covariate confounding.* BJOG, 2021. **128**(10): p. 1574.
6. Leijten, P., 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.* Journal of the American Academy of Child and Adolescent Psychiatry, 2020. **59**(8): p. 933-943.
7. Rhee, R.L., et al., *Comparison of treatment response in idiopathic and connective tissue disease-associated pulmonary arterial hypertension.* American Journal of Respiratory and Critical Care Medicine, 2015. **192**(9): p. 1111-1117.
8. Schmidt, A.F., et al., *Tailoring treatments using treatment effect modification.* Pharmacoepidemiol Drug Saf, 2016. **25**(4): p. 355-62.
9. Brookes, S.T., et al., *Subgroup analyses in randomized trials: risks of subgroup-specific analyses; power and sample size for the interaction test.* J Clin Epidemiol, 2004. **57**(3): p. 229-36.
10. Ensor, J., et al., *Simulation-based power calculations for planning a two-stage individual participant data meta-analysis.* BMC medical research methodology, 2018. **18**(1): p. 41.
11. Riley, R.D. and D.J. Fisher, *Chapter 12: Power Calculations for Planning an IPD Meta-Analysis*. Individual Participant Data Meta-Analysis: A Handbook for Healthcare Research, ed. R.D. Riley, J.F. Tierney, and L.A. Stewart. 2021: Wiley
12. Fisher, D.J., et al., *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**(9): p. 949-67.
13. Fisher, D.J., et al., *Meta-analytical methods to identify who benefits most from treatments: Daft, deluded, or deft approach?* BMJ (Online), 2017. **356**.
14. Belias, M., et al., *Statistical approaches to identify subgroups in meta-analysis of individual participant data: a simulation study.* BMC medical research methodology, 2019. **19**(1): p. 183.
15. Godolphin, P.J., et al., *Estimating interactions and subgroup-specific treatment effects in meta-analysis without aggregation bias: A within-trial framework.* Res Synth Methods, 2022.
16. Marlin, N., et al., *Examining Non-Linear Effects and Effect Modification at the Participant-Level in IPD Meta-Analysis Part 1: Analysis Methods are Often Substandard*. preprint 2023. Available at https://dx.doi.org/10.2139/ssrn.4333137.
17. Mbuagbaw, L., et al., *A tutorial on methodological studies: the what, when, how and why.* BMC Med Res Methodol, 2020. **20**(1): p. 226.
18. Riley, R.D. and D.J. Fisher, *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, ed. R.D. Riley, J.F. Tierney, and L.A. Stewart. 2021: Wiley

19. Gao, Y., 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**: p. 156-167.
20. Hua, H., et al., *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.* Statistics in Medicine, 2017. **36**(5): p. 772-789.
21. Simmonds, M., G. Stewart, and L. Stewart, *A decade of individual participant data meta-analyses: A review of current practice.* Contemporary Clinical Trials, 2015. **45**: p. 76-83.
22. Burke, D.L., J. Ensor, and R.D. Riley, *Meta-analysis using individual participant data: one-stage and two-stage approaches, and why they may differ.* Statistics in Medicine, 2017. **36**(5): p. 855-875.
23. Kontopantelis, E., *A comparison of one-stage vs two-stage individual patient data meta-analysis methods: A simulation study.* Research synthesis methods, 2018. **9**(3): p. 417-430.
24. Morris, T.P., et al., *Meta-analysis of Gaussian individual patient data: Two-stage or not two-stage?* Statistics in Medicine, 2018. **37**(9): p. 1419-1438.
25. da Costa, B.R. and A.J. Sutton, *A comparison of the statistical performance of different meta-analysis models for the synthesis of subgroup effects from randomized clinical trials.* BMC medical research methodology, 2019. **19**(1): p. 198.
26. Walker, R., L. Stewart, and M. Simmonds, *Estimating interactions in individual participant data meta-analysis: a comparison of methods in practice.* Syst Rev, 2022. **11**(1): p. 211.
27. Vo, T.T., et al., *A novel approach for identifying and addressing case-mix heterogeneity in individual participant data meta-analysis.* Research synthesis methods, 2019. **10**(4): p. 582-596.
28. Fokkema, M., et al., *Detecting treatment-subgroup interactions in clustered data with generalized linear mixed-effects model trees.* Behavior research methods, 2018. **50**(5): p. 2016-2034.
29. Mistry, D., N. Stallard, and M. Underwood, *A recursive partitioning approach for subgroup identification in individual patient data meta-analysis.* Statistics in Medicine, 2018. **37**(9): p. 1550-1561.
30. Jiao, Y., et al., *A CD-based mapping method for combining multiple related parameters from heterogeneous intervention trials.* Statistics and its Interface, 2020. **13**(4): p. 533-549.
31. Wang, X.V., et al., *Meta-STEPP: subpopulation treatment effect pattern plot for individual patient data meta-analysis.* Statistics in Medicine, 2016. **35**(21): p. 3704-3716.
32. Papadimitropoulou, K., et al., *Meta-analysis of continuous outcomes: using pseudo IPD created from aggregate data to adjust for baseline imbalance and assess treatment-by-baseline modification.* Research synthesis methods, 2020.
33. Noma, H., et al., *Efficient two-step multivariate random effects meta-analysis of individual participant data for longitudinal clinical trials using mixed effects models.* BMC medical research methodology, 2019. **19**(1): p. 33.
34. Riley, R.D., et al., *Multivariate meta-analysis using individual participant data.* Research synthesis methods, 2015. **6**(2): p. 157-174.
35. Fisher, D.J., *Two-stage individual participant data meta-analysis and generalized forest plots.* Stata Journal, 2015. **15**(2): p. 369-396.
36. Sauerbrei, W. and P. Royston, *Investigating treatment-effect modification by a continuous covariate in IPD meta-analysis: an approach using fractional polynomials.* BMC Med Res Methodol, 2022. **22**(1): p. 98.
37. White, I.R., et al., *Meta-analysis of non-linear exposure-outcome relationships using individual participant data: A comparison of two methods.* Statistics in Medicine, 2019. **38**(3): p. 326-338.

38. Kasenda, B., et al., *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): p. e011148.
39. Sauerbrei, W. and P. Royston, *A new strategy for meta-analysis of continuous covariates in observational studies.* Stat Med, 2011. **30**(28): p. 3341-60.
40. White, I.R., *Multivariate random-effects meta-analysis.* Stata Journal, 2009. **9**(1): p. 40-56.
41. Shrier, I. and M. Pang, *Confounding, effect modification, and the odds ratio: common misinterpretations.* J Clin Epidemiol, 2015. **68**(4): p. 470-4.
42. Belias, M., et al., *Predicting personalised absolute treatment effects in individual participant data meta-analysis: An introduction to splines.* Res Synth Methods, 2022.
43. de Jong, V.M.T., et al., *Individual participant data meta-analysis of intervention studies with time-to-event outcomes: A review of the methodology and an applied example.* Research synthesis methods, 2020. **11**(2): p. 148-168.
44. Hess, K.R., *Assessing time-by-covariate interactions in proportional hazards regression models using cubic spline functions.* Stat Med, 1994. **13**(10): p. 1045-62.
45. Giorgi, R., et al., *A relative survival regression model using B-spline functions to model non-proportional hazards.* Stat Med, 2003. **22**(17): p. 2767-84.
46. Kontopantelis, E., et al., *Simulation-based power calculations for mixed effects modeling: Ipdpower in stata.* Journal of Statistical Software, 2016. **74**.
47. Riley, R.D., et al., *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**(24): p. 4822-4837.
48. Kovalchik, S.A. and W.G. Cumberland, *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): p. 370-84.
49. Simmonds, M.C. and J.P. Higgins, *Covariate heterogeneity in meta-analysis: criteria for deciding between meta-regression and individual patient data.* Stat Med, 2007. **26**(15): p. 2982-99.
50. Kahan, B.C., et al., *A comparison of methods to adjust for continuous covariates in the analysis of randomised trials.* BMC Med Res Methodol, 2016. **16**: p. 42.

Appendix excluded references

1. Chen, D. G., et al. (2020). "Relative efficiency of using summary versus individual data in random-effects meta-analysis." Biometrics.
2. Hemming, K., et al. (2020). "Extending the I-squared statistic to describe treatment effect heterogeneity in cluster, multi-centre randomized trials and individual patient data meta-analysis." Statistical Methods in Medical Research.
3. Riley, R. D., et al. (2020). "One-stage individual participant data meta-analysis models for continuous and binary outcomes: Comparison of treatment coding options and estimation methods." Statistics in Medicine 39(19): 2536-2555.
4. Belhechmi, S., et al. (2019). "An alternative trial-level measure for evaluating failure-time surrogate endpoints based on prediction error." Contemporary Clinical Trials Communications 15.
5. Fanshawe, T. R. and R. Perera (2019). "Conducting one-stage IPD meta-analysis: Which approach should i choose?" BMJ Evidence-Based Medicine 24(5): 190.
6. Papadimitropoulou, K., et al. (2019). "One-stage random effects meta-analysis using linear mixed models for aggregate continuous outcome data." Research synthesis methods 10(3): 360-375.
7. Schuit, E., et al. (2019). "How often can meta-analyses of individual-level data individualize treatment? A meta-epidemiologic study." International Journal of Epidemiology 48(2): 596-608.
8. Sofeu, C. L., et al. (2019). "One-step validation method for surrogate endpoints using data from multiple randomized cancer clinical trials with failure-time endpoints." Statistics in Medicine.
9. Vo, T., et al. (2019). "Rethinking meta-analysis: Addressing problems of non-transportability when combining treatment effects across patient populations." Revue d'Epidemiologie et de Sante Publique 67: S121-S122.
10. Freeman, S. C., et al. (2018). "A framework for identifying treatment-covariate interactions in individual participant data network meta-analysis." Research synthesis methods 9(3): 393-407.
11. Legha, A., et al. (2018). "Individual participant data meta-analysis of continuous outcomes: A comparison of approaches for specifying and estimating one-stage models." Statistics in Medicine 37(29): 4404-4420.
12. Snell, K. I. E., et al. (2018). "Meta-analysis of prediction model performance across multiple studies: Which scale helps ensure between-study normality for the C-statistic and calibration measures?" Statistical Methods in Medical Research 27(11): 3505-3522.
13. Kunkel, D. and E. E. Kaizar (2017). "A comparison of existing methods for multiple imputation in individual participant data meta-analysis." Statistics in Medicine 36(22): 3507-3532.
14. Landau, S., et al. (2017). "Assessing treatment effect moderation in trials of psychological interventions: A case for individual participant data meta-analysis of pooled trials." Trials 18.
15. Thomas, D., et al. (2017). "A comparison of analytic approaches for individual patient data meta-analyses with binary outcomes." BMC medical research methodology 17(1): 28.
16. Egger, M., et al. (2016). "GetReal: from efficacy in clinical trials to relative effectiveness in the real world." Research synthesis methods 7(3): 278-281.
17. Huang, Y. (2016). The ability of aggregate data meta-analysis in predicting individual patient data meta-analysis, ProQuest Information & Learning. 76.
18. Huang, Y., et al. (2016). "Comparing the Overall Result and Interaction in Aggregate Data Meta-Analysis and Individual Patient Data Meta-Analysis." Medicine (United States) 95(14).

19. Kast, J., et al. (2016). "Assessment of covariate effect based on individual patient data vs. Model-based meta-analysis of aggregate data for DPP-4 inhibitors." Clinical Pharmacology and Therapeutics 99: S105.
20. Kaufmann, E., et al. (2016). "Avoiding methodological biases in meta-analysis." Zeitschrift fur Psychologie / Journal of Psychology 224(3): 157-167.
21. Lueza, B., et al. (2016). "Bias and precision of methods for estimating the difference in restricted mean survival time from an individual patient data meta-analysis." BMC medical research methodology 16: 37.
22. Richter, A., et al. (2016). "Simple pooling of data from different studies is increasingly used but not in line with methodological recommendations: A systematic review of methods applied in the field of rheumatoid arthritis." Annals of the Rheumatic Diseases 75: 108.
23. Smith, C. T., et al. (2016). "Individual participant data meta-analyses compared with meta-analyses based on aggregate data." Cochrane Database of Systematic Reviews(9): 56.
24. Song, F. and M. O. Bachmann (2016). "Cumulative subgroup analysis to reduce waste in clinical research for individualised medicine." BMC Medicine 14(1).
25. Waldron, L. and M. Riester (2016). Meta-analysis in gene expression studies. Methods in Molecular Biology, Humana Press Inc. 1418: 161-176.
26. Debray, T. P. A., et al. (2015). "Individual Participant Data (IPD) Meta-analyses of Diagnostic and Prognostic Modeling Studies: Guidance on Their Use." PLoS Medicine 12(10).
27. Debray, T. P., et al. (2015). "Get real in individual participant data (IPD) meta-analysis: a review of the methodology." Research synthesis methods 6(4): 293-309.
28. Riley, R. D., et al. (2015). "Meta-analysis of test accuracy studies: An exploratory method for investigating the impact of missing thresholds." Systematic Reviews 4(1).
29. Riley, R. D., et al. (2015). "Multivariate meta-analysis of prognostic factor studies with multiple cut-points and/or methods of measurement." Statistics in Medicine 34(17): 2481-2496.
30. Riley, R. D., et al. (2015). "Summarising and validating test accuracy results across multiple studies for use in clinical practice." Statistics in Medicine 34(13): 2081-2103.
31. Simmonds, M., et al. (2015). "A decade of individual participant data meta-analyses: A review of current practice." Contemporary Clinical Trials 45: 76-83.

Declaration of interests

☐ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

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

Nadine Marlin reports financial support was provided by National Institute for Health and Care Research.

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 non-linearity.
- Further research comparing various proposals for IPDMA of non-linear covariate outcome relationships or non-linear effect modification 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 non-linear associations and effect modification in IPDMA.

What is the implication?

- Comparison of methodological options (e.g splines, polynomials) for analysing non-linear associations or non-linear effect modification is needed.

Nadine Marlin: Conceptualization; Data curation; Formal analysis; Investigation; Roles/Writing - original draft; Supervision

Peter J Godolphin: Data curation; Investigation; Roles/Writing - review & editing; Supervision

Richard Riley: Conceptualization; Roles/Writing - review & editing

Richard Hooper: Conceptualization; Roles/Writing - review & editing

Ewelina Rogozińska: Data curation; Investigation; Roles/Writing - review & editing; Supervision