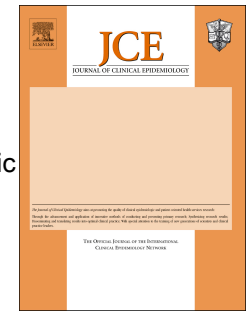
Journal Pre-proof



Methodological Guidance for Individual Participant Data Meta-Analyses: A Systematic Review

Edith Ginika Otalike, Mike Clarke, Farjana Akhter, Areti Angeliki Veroniki, Ngianga-Bakwin Kandala, Joel J. Gagnier

PII: S0895-4356(25)00422-6

DOI: https://doi.org/10.1016/j.jclinepi.2025.112089

Reference: JCE 112089

To appear in: *Journal of Clinical Epidemiology*

Received Date: 23 September 2025

Revised Date: 17 November 2025

Accepted Date: 20 November 2025

Please cite this article as: Otalike EG, Clarke M, Akhter F, Veroniki AA, Kandala N-B, Gagnier JJ, Methodological Guidance for Individual Participant Data Meta-Analyses: A Systematic Review, *Journal of Clinical Epidemiology* (2025), doi: https://doi.org/10.1016/j.jclinepi.2025.112089.

Objective: To systematically identify and synthesize methodological guidance for conducting IPD-MAs of randomized trials and observational studies, to inform the development of a critical appraisal tool for reports of IPD-MAs.

Study design and Setting: We searched nine major electronic databases and grey literature sources through June 2025 using a strategy developed with a health sciences librarian. To be eligible, articles had to report empirical, simulation-based, consensus-based, or narrative research and offer guidance on the methodology of IPD-MA. Study selection and data extraction were performed independently by two reviewers. Quality was assessed using tools tailored to study design (e.g., ADEMP, ROBIS, ACCORD, SANRA). Extracted guidance was categorized thematically and mapped to appraisal domains.

Results: After screening 14,736 records, we included 141 studies. These encompassed simulation (38%), empirical (21%), and methodological guidance (12%), among others. Key themes included IPD-MA planning, data access and harmonization, analytic strategies and other statistical issues, and reporting. While there was robust guidance for IPD-MA of randomized trials, recommendations for observational studies are sparse. Across all study types, 63% were rated high quality.

Conclusions: This review collates otherwise fragmented guidance into an integrative synthesis, highlighting best practices and critical domains for appraising IPD-MAs. These findings formed the evidence base for a Delphi consensus process to develop a dedicated IPD-MA critical appraisal tool.

Plain language summary

Meta-analyses often pool published summaries from many studies. That approach can miss important details and introduce bias. An Individual Participant Data Meta-Analysi (IPD-MA) instead re-analyses the original, participant-level data across studies. IPD-MAs are powerful but complex, and practical guidance is scattered, especially for observational studies. We wanted to bring these recommendations together in one place and identify candidate items for a tool to assess the quality of a completed IPD-MA. We systematically searched eight databases from their inception to 2025 to identify papers offering practical guidance on conducting IPD-

MAs for health interventions. We organized guidance across the full project life cycle, from planning, finding and accessing data, to preparing and checking data, analyzing results, and reporting. We highlighted where experts broadly agree and where gaps remain. We found 141 relevant papers published between 1995 and 2025. Among these, we identified 25 key topic areas and several smaller subtopics. Many papers covered more than one topic, so we allowed them to appear in multiple categories rather than forcing a single classification. From this mapping, we developed a clear set of recommendations organized around four themes: IPDMA planning; identification and access to studies and individual participant data; methods for meta-analysis; and other special considerations, such as methods for observational studies. We also identified notable gaps for observational IPD-MAs. Our synthesis offers a practical checklist for teams planning or reviewing IPD-MAs. It also lays the groundwork for a consensus-based appraisal tool to help editors, funders, and guideline developers judge quality and improve practice.

Keywords: Individual Participant Data, Meta-analysis, Randomized trials, RCT, Non-randomized studies, critical appraisal

Title: Methodological Guidance for Individual Participant Data Meta-Analyses: A Systematic Review

Word count = 5168

Methodological Guidance for Individual Participant Data Meta-Analyses: A Systematic Review

Edith Ginika Otalike[a], Mike Clarke[b], Farjana Akhter[a], Areti Angeliki Veroniki[c], Ngianga-Bakwin Kandala[a], Joel J. Gagnier [a,d]

[a] Department of Epidemiology & Biostatistics, Schulich School of Medicine & Dentistry, Western University, London, Ontario, Canada
[b] Northern Ireland Methodology Hub, Queen's University Belfast, Northern Ireland, United Kingdom
[c] Department of Health Services, Policy, and Practice, School of Public Health, Brown University, Providence, Rhode Island, USA
[d] Department of Surgery, Schulich School of Medicine & Dentistry, Western University, London, Ontario, Canada

Corresponding Author
Edith Ginika Otalike
Western Centre for Public Health and Family Medicine
1465 Richmond Street, London, ON, Canada N6A 0C1
Email: eotalike@uwo.ca
Phone: (437)299-7460

Email addresses of authors
Edith Otalike[a] – eotalike@uwo.ca
Mike Clarke[b] - m.clarke@qub.ac.uk
Farjana Akhter[a] - makhter9@uwo.ca
Areti Angeliki Veroniki[c] - argie_veroniki@brown.edu
Ngianga-Bakwin Kandala[a] - nkandala@uwo.ca
Joel J. Gagnier [a,d] - jgagnie4@uwo.ca

Abstract

Objective: To systematically identify and synthesize methodological guidance for conducting IPD-MAs of randomized trials and observational studies, to inform the development of a critical appraisal tool for reports of IPD-MAs.

Study Design and Setting: We searched nine major electronic databases and grey literature sources through June 2025 using a strategy developed with a health sciences librarian. To be eligible, articles had to report empirical, simulation-based, consensus-based, or narrative research and offer guidance on the methodology of IPD-MA. Study selection and data extraction were performed independently by two reviewers. Quality was assessed using tools tailored to study design (e.g., ADEMP, ROBIS, ACCORD, SANRA). Extracted guidance was categorized thematically and mapped to appraisal domains.

Results: After screening 14,736 records, we included 141 studies. These encompassed simulation (38%), empirical (21%), and methodological guidance (12%), among others. Key themes included IPD-MA planning, data access and harmonization, analytical strategies, and other statistical issues, as well as reporting. While there was robust guidance for IPD-MA of randomized trials, recommendations for observational studies are sparse. Across all study types, 63% were rated high quality.

Conclusions: This review synthesizes previously fragmented guidance into an integrative synthesis, highlighting best practices and critical domains for evaluating IPD-MAs. These findings formed the evidence base for a Delphi consensus process to develop a dedicated IPD-MA critical appraisal tool.

Plain Language Summary

Meta-analyses often pool published summaries from many studies. That approach can miss important details and introduce bias. An Individual Participant Data Meta-Analysi (IPD-MA) instead re-analyses the original, participant-level data across studies. IPD-MAs are powerful but complex, and practical guidance is scattered, especially for observational studies. We wanted to

bring these recommendations together in one place and identify candidate items for a tool to assess the quality of a completed IPD-MA. We systematically searched eight databases from their inception to 2025 to identify papers offering practical guidance on conducting IPD-MAs for health interventions. We organized guidance across the full project life cycle, from planning, finding and accessing data, to preparing and checking data, analyzing results, and reporting. We highlighted where experts broadly agree and where gaps remain. We found 141 relevant papers published between 1995 and 2025. Among these, we identified 25 key topic areas and several smaller subtopics. Many papers covered more than one topic, so we allowed them to appear in multiple categories rather than forcing a single classification. From this mapping, we developed a clear set of recommendations organized around four themes: IPDMA planning; identification and access to studies and individual participant data; methods for meta-analysis; and other special considerations, such as methods for observational studies. We also identified notable gaps for observational IPD-MAs. Our synthesis offers a practical checklist for teams planning or reviewing IPD-MAs. It also lays the groundwork for a consensus-based appraisal tool to help editors, funders, and guideline developers judge quality and improve practice.

Keywords: Individual Participant Data, Meta-analysis, Randomized trials, RCT, Non-randomized studies, critical appraisal
Word count = 5168

What is new?

Key findings

- We synthesized methodological guidance for IPD meta-analyses of interventions across the full lifecycle, from planning, data access and preparation, analysis (endpoint-specific), and reporting, by reviewing 141 method papers (1995-2025). We organized guidance into four themes and 25 primary topic areas, yielding a practical set of recommendations and highlighting gaps, especially methods for IPD-MAs based on observational studies.

What this adds to what was known

- Existing guidance was fragmented and often focused on a single technique, such as interaction modelling or endpoint-specific models. Our review provides an end-to-end consolidation that integrates operational procedures with analytic guidance across outcomes and study designs. We also map guidance to conceptual appraisal domains and translate it into candidate items for a consensus-based critical appraisal tool for completed IPD-MAs.

What is the implication, and what should change now?

- The synthesized recommendations provide a consolidated methodological reference that can be used as a best-practice checklist when planning, conducting or reviewing IPD-MAs. Methods work should now target the identified gaps, particularly for IPD-MAs of observational studies to strengthen data management, design-specific analyses, and causal inference.

1. Introduction

Evidence synthesis enables the scientific community to pool together available findings from multiple studies to inform clinical practice, health policy, or an ongoing debate on a specific subject [1]. Systematic reviews and meta-analyses relevant to the effects of interventions summarize evidence from a series of separate but similar studies to evaluate these effects [2]. The commonly undertaken aggregate data meta-analysis pools summary statistics obtained from publications or authors, such as overall effect estimates (mean differences, odds ratios (ORs), hazard ratios (HRs), etc.) and their measures of variance, and sometimes the results for subgroups of participants. Limitations of aggregate data meta-analyses stem from the reliance on published summary statistics, which can be compromised by poor reporting quality, such as the reporting of p-values rather than estimates [3]. Other limitations include inconsistent definitions of outcomes, the use of different statistical methods to provide the estimates from different studies, variations in outcome measures (e.g., HR, OR, etc.) [4], lack of information regarding correlation between repeated observation [37], vulnerability to reporting bias as positive findings are more likely to be published than null or negative findings [5], and lack of power to model effect modification, which is of increasing interest in personalized medicine [6]. While aggregate data meta-analysis contributes substantially to evidence synthesis, these limitations have led to repeated calls for a transition toward individual participant data meta-analysis (IPD-MA) [7]. However, aggregate data meta-analyses have been found to produce comparable overall effect estimates to IPD-MA [8, 9, 10].

IPD-MA re-analyzes the raw data from individual participants and can circumvent some of the challenges of aggregate data meta-analysis. Accessing IPD enables independent scrutiny of data quality and integrity, reduces research waste, standardizes outcome definition and statistical analysis across studies, derives effect estimates independent of how results were reported, verifies modelling assumptions, models complexity or non-linear relationships, and summarizes subject-specific effects with greater statistical power [4, 11]. Additionally, it has reformed the design, execution, and interpretation of trials [11]. Given the numerous advantages and the robustness of evidence it produces, IPD-MA is considered the 'gold standard' in evidence synthesis of intervention studies [12, 13, 14] and has gained increasing interest across health research.

Despite its strengths, IPD-MA does not resolve all the limitations inherent in evidence synthesis. It is resource-intensive, demanding substantial time, effort, funding, resources, and coordination, as well as collaboration with the original investigators, careful attention to ethical considerations around data sharing, and advanced statistical expertise. Despite these investments, IPD-MA can be vulnerable to selection and publication bias [15] and to the inherent biases and methodological flaws of the original studies, such as flawed randomization. Recent evidence from a quality assessment of 323 IPD-MA studies revealed that they performed poorly in general and in IPD-MA-identified specific methodological items, indicating a gap in the literature for a consensus-based tool to assess the methodological quality of IPD-MA [16]. Quality assessment has relied on guidance and reporting tools, such as the Preferred Reporting Items for a Systematic Review and Meta-analysis extension for IPD-MA (PRISMA-IPD) [17], IPD-MA handbook for healthcare research (CheckMap) [18], and Cochrane guidance on the use of IPD-MA of randomized trials [19]. Therefore, given the increased interest in IPD-MA, there is a need for a rigorous critical appraisal tool for their reports.

To inform the development of a consensus-based critical appraisal tool, we undertook a systematic review to identify and collate existing guidance relevant to IPD-MA. While several reviews of IPD-MA methodology have been conducted, they have addressed selected aspects of IPD-MA, such as statistical techniques [14], strategies for estimating interaction effects [20, 21], or guidance tailored

to specific outcome data (e.g., binary outcomes) [22]. Reviews explicitly tailored to methods for IPD-MA of observational studies are limited. A recent review focused on pooled longitudinal IPD-MA but does not cover outcome-specific modelling or design-based analysis strategies (e.g., complex survey designs) [23]. Therefore, there is a need for a review that provides an integrated synthesis of guidance spanning the full spectrum of IPD-MA methodology for estimating the effects of interventions, covering IPD-MA planning, study identification, data access, analysis, and dissemination.

1.1 Review Objective

This systematic review aims to identify and synthesize existing methodological guidance relevant to the conduct of IPD-MA of randomized trials and observational studies, and to categorize this content into conceptual domains aligned with the proposed checklist structure that informs the selection of candidate items for the subsequent Delphi consensus process.

2. Methods

This review was conducted following guidance for systematic reviews [24] and reported following PRISMA 2020 guidance [25]

2.1 Protocol registration

A protocol for this review, produced in accordance with the PRISMA extension for Protocols (PRISMA-P) [26], is available on the Open Science Framework (OSF, https://osf.io/v5n9r/).

2.2 Eligibility criteria

Studies were eligible for inclusion if they were published in English, available in full-text format, and had been peer-reviewed or disseminated as preprints. Eligible studies were required to focus on methodological issues related to IPD-MA of randomized trials, observational studies of the effects of interventions, or both. Specifically, articles were eligible if they provided recommendations or guidance, addressed statistical considerations, discussed software implementation, or presented simulation-based or empirical comparisons designed to inform the conduct of IPD-MA. We focused on studies that contribute to the development, evaluation, or refinement of IPD-MA methods, such as those exploring model selection, effect estimation strategies, data harmonization, handling of missing data, and the integration of individual-level and aggregate data in intervention studies. Tutorials that present new or consolidated methodological recommendations for IPD-MA were eligible. We excluded studies if they used IPD-MA solely to answer a specific research question, if they lacked a clear focus on methodological guidance, or if the full text of the article was not accessible.

2.3 Information sources

On June 9-10, 2024, our team (E.O. and F.A.), in collaboration with the Western University Librarian (R.I), conducted a comprehensive literature search across the following databases: Cochrane Database of Systematic Reviews, MEDLINE, Embase, CINAHL, Scopus, Web of Science, Health Technology Assessment Review Database, and Journal of Research Synthesis Methods. The search in Embase and MEDLINE was updated quarterly, while the searches in the other databases were updated once, using the same search strategy, up to June 11, 2025. We also searched reference lists and grey literature. Details of these searches are provided in Table 1.

Table 1: Literature databases, coverage years, and update information for the search strategy

S/N	Database	Coverage years	Update description
1	Cochrane Library (CDSR)	1996 to 2025	Updated quarterly; last update on June 11, 2025
2	MEDLINE (via Ovid)	1946 to 2025	Updated quarterly; last update on June 11, 2025
3	EMBASE (via Ovid)	1947 to 2025	Updated quarterly; last update on June 11, 2025
4	CINAHL (via EBSCOhost)	1981 to 2025	Initially searched June 9-10, 2024; updated June 11, 2025
5	Scopus	2004 to 2025	Initially searched June 9-10, 2024; updated June 11, 2025
6	Web of Science (Core Collection)	1900 to 2025	Initially searched June 9-10, 2024; updated June 11, 2025
7	Health Technology Assessment Database (HTA)	1996 to 2025	Initially searched June 9-10, 2024; updated June 11, 2025
8	Research Synthesis Methods (manual search)	2010 to 2025	Initially searched June 9-10, 2024; updated June 11, 2025

2.4 Search Strategy

We searched the databases using keywords and Medical Subject Headings (MeSH) terms related to IPD and methodology to identify relevant articles, with no restrictions on publication year. Our search approach was based on strategies previously published in related methodology reviews [14, 16, 27]. We developed our initial set of search terms by examining the titles, abstracts, and index terms or keywords of these earlier publications and adapted their strategies for the current review. The final search strategy underwent peer review by librarians at Western University, who provided suggestions for syntax, structure, and spelling corrections. However, we did not formally apply the PRESS checklist [28]. Full details of the search strategy are provided in Supplementary File 1.

2.5 Selection process

Following the completion of the database search, all identified references were imported into Covidence systematic review software [29], and duplicates were removed. Two reviewers (E.O. and S.A.) independently screened titles and abstracts for relevance according to the predefined eligibility criteria. Articles deemed potentially eligible proceeded to full-text screening, which was also conducted independently by two reviewers (E.O. and F. A.). Discrepancies at either stage were resolved through discussion between E.O. and S.A., and E.O. and F.A., respectively.

2.6 Data Extraction

We conducted data extraction using a pre-designed Microsoft Excel form, developed to address the specific objectives of this systematic review of methodology. The form captured general study characteristics (e.g., publication type, study design, geographical setting), as well as recommendations about methodology and quality elements relevant to the development of a critical appraisal tool for IPD-MA. Extracted elements included the authors' conclusions on the proposed appraisal domains, preferred statistical models for various outcome data types, the choice of modelling, estimation methods, and any recommendations regarding software or analytical strategies. The proposed domains were informed by existing methodology checklists and frameworks identified through the literature [17, 18, 30].

Before data extraction, the form underwent a pilot test on a randomly selected 5% sample of the included studies. Feedback obtained during this preliminary phase was used to refine and calibrate the extraction template, thereby ensuring enhanced clarity and consistency. One reviewer (E. O.) extracted data from all eligible studies into the revised form, while a second reviewer (F. A.) extracted data from a random 10% sample of the included studies. Any discrepancies were addressed through discussion between E.O. and F.A. Furthermore, a third reviewer (J. G.) cross-verified all extracted data.

2.7 Risk of Bias Assessment

We assessed the methodological quality of the included studies using tools appropriate to their respective designs. For simulation studies, we used the framework proposed by Morris et al., following the ADEMP (Aims, Data generating mechanism, Estimands, Methods, and Performance measures) framework [31]. We adapted this framework to assess empirical studies. For narrative and critical reviews, we employed the Scale for the Assessment of Narrative Review Articles (SANRA) [32]. For systematic reviews, we utilized the Risk of Bias in Systematic Reviews (ROBIS) [33]. For consensus studies, such as Delphi processes and guideline development, we employed the ACCORD (Appraisal of Guidelines for Research & Evaluation using Delphi) checklist [34]. We refrained from conducting formal risk of bias assessments for expert opinions, tutorials, book chapters, or methodological guides. One reviewer (E.O.) did the risk of bias assessments for all eligible studies. A second reviewer independently assessed a randomly selected 10% subset. We incorporated the study quality into the study characteristic table (Supplementary Table 1). We also present a graphical summary of the overall quality rating of the included studies.

2.8 Data Analysis and Presentation

2.8.1 Descriptive Thematic Synthesis

We conducted a descriptive synthesis of the included studies, summarizing key characteristics such as study design, research context, publication year, funding sources, recommendations for the methods of IPD-MA, and analytic approaches employed for various outcome data types. We also conducted a thematic categorization based on the primary methodological focus of each study, adapting the approach described by Thomas and Harden (2008) for use in qualitative synthesis [35]. The study served as the unit of analysis, given the multifaceted nature of methodological guidance in IPD-MA papers; each included study often addressed multiple aspects. For instance, a single paper could simultaneously provide recommendations on estimation methods, modelling choices, and analytical approaches for different types of outcome data. We first identified the dominant methodological theme of each paper, which was then mapped to one of four overarching thematic domains developed iteratively during data extraction and refined through team discussion. Additional elements beyond the primary focus were recorded as subcategories to preserve the breadth of content and to account for overlaps, while ensuring each study was represented by its core contribution to this review. The thematic coding was carried out by one reviewer (E.O.) and independently verified by a second reviewer (F.A.).

2.8.2 Mapping to Appraisal Domains

In parallel with the thematic synthesis, we organized the extracted recurring item-level guidance (e.g., protocols, data quality checks, analytic methods) into a conceptual appraisal domain aligned with existing appraisal frameworks [17, 18, 30], as described elsewhere [36]. Two reviewers (E.O. and J.G.) independently reviewed and classified individual signalling questions and methodological statements, which were then refined through discussion with the broader team (E.O., J.G., M.C., and

K.N.). This domain mapping generated a structured set of candidate items that directly informed the subsequent Delphi survey.

3. Results

3.1 Study selection

The initial search in June 2024 yielded 13,589 records. After updating the search in June 2025, the total number of records increased to 14,736. After removing 5,194 duplicates (54 manually, and 5,140 by Covidence), we screened 9,543 unique records. This led to the review of 351 full texts, and 210 were excluded (see Figure 1). Relevant references from the excluded reviews were assessed for additional eligible studies. Ultimately, 141 studies met the inclusion criteria (Figure 1).

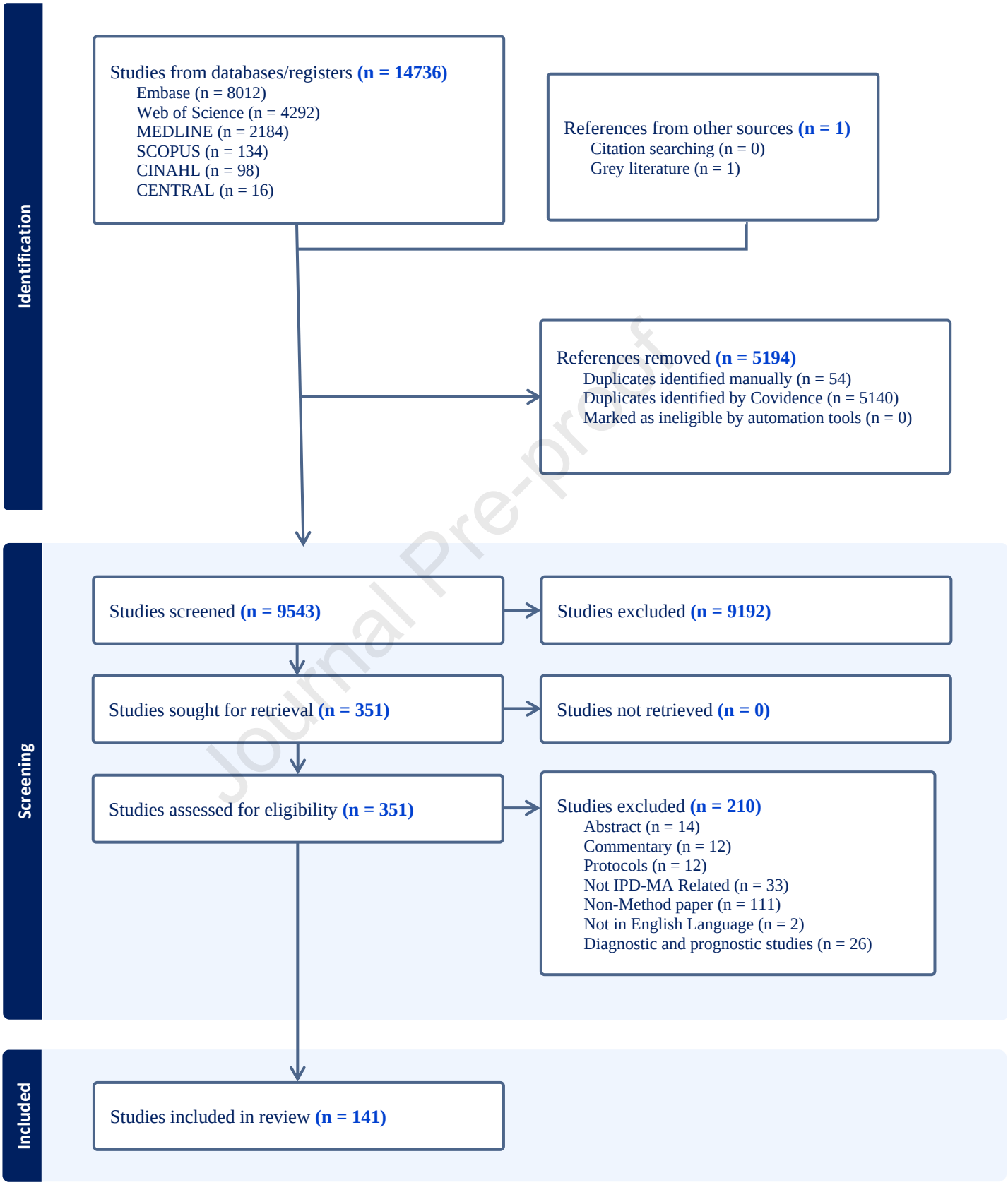


Figure 1: PRISMA Flow chat diagram

3.2 Study characteristics

The 141 included papers were published between 1995 and June 2025, with 83% (n = 117) published in or after 2010. Thirty-nine percent of the studies were simulation studies (n = 54), 21% were empirical applications demonstrating methods using real datasets (n = 29), and 12% were formal methodology guidance articles (n = 17). The remaining 41 papers (36%) were narrative reviews, systematic methodology reviews, consensus statements, chapters from handbooks, etc., that nonetheless contained extractable guidance on the methods for IPD MA. Primary (corresponding) authors were affiliated with institutions in 14 countries: United Kingdom (n = 71), United States (n = 18), Netherlands (n = 16), Canada (n = 9), France (n = 7), Japan and Germany (4 each), Australia (3), China and Switzerland (2 each), and Belgium, Denmark, Italy, Norway, and Ireland (1 each). Regarding methodological focus, the three most frequently addressed topics were treatment-covariate interaction modelling (n = 26, 14%), choice of analytical approach (one-stage vs two-stage; n = 21, 16%), and specialized statistical techniques for varied outcome types (n = 15, 12%). Funding information was reported for 69.5% (n=98) of the papers, with over half acknowledging public or governmental support 50.3% (n=71), 12.8% (n=18) citing institutional, mixed, industry, and Non-Governmental Organization (NGO) sources of funding, and 5.7% (n=8) explicitly reported no funding. Funding was not reported in 31.2% (n = 44) (see Table 2).

Table 2: Descriptive characteristics of included resources that reported recommendations for IPD-MA (N = 141)

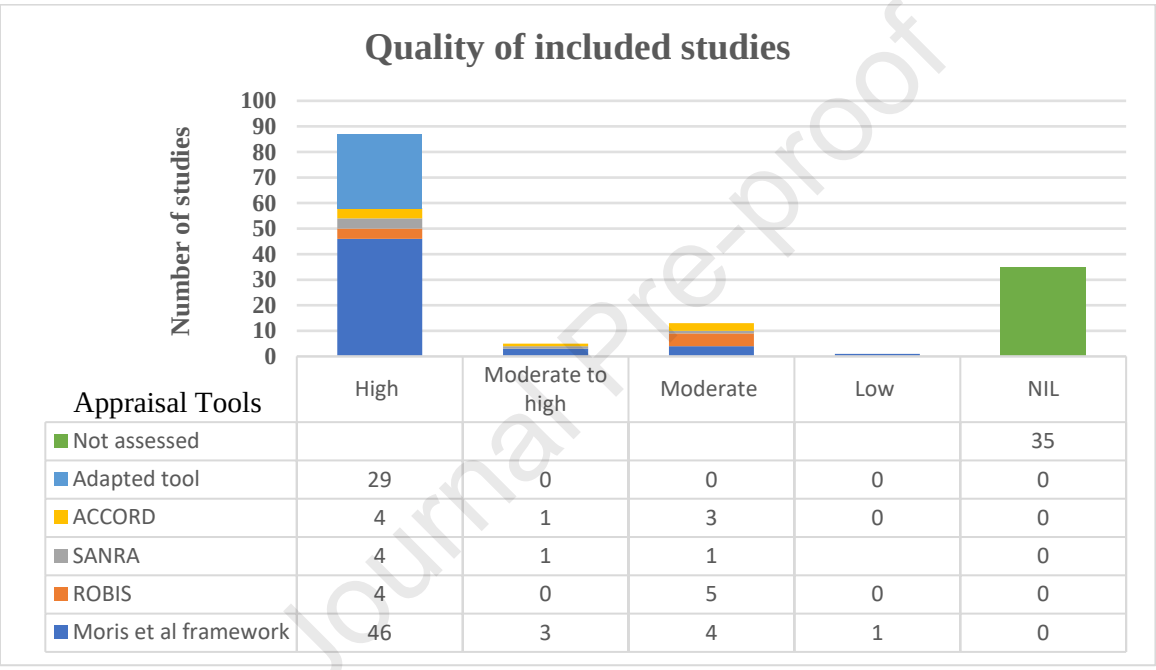
Category	Characteristic	Number of studies (n,%)
Year of publication	1995 to 1999	2 (1.4)
	2000 to 2004	8 (5.7)
	2005 to 2009	14 (9.9)
	2010 to 2014	26 (18.4)
	2015 to 2019	43 (30.5)
	2020 to 2025	48 (34.1)
Type of study	Simulation	54 (38.3)
	Empirical	29 (20.6)
	Methodological guide	17 (12.1)
	Statistical guide	10 (7.1)
	Systematic Methodology Review	9 (6.4)
	Consensus	8 (5.6)
	Critical/Narrative Review	6 (4.3)
	Handbook chapters	3 (2.1)
	Tutorial	2 (1.4)
	Empirical meta research	2 (1.4)
	Scoping review	1 (0.7)
Research setting	Europe	105 (74.5)
	North America	27 (19.1)

	Asia	6 (4.3)
	Oceanic	3 (2.1)
Type of funding	Public	71 (50.3)
	Not reported	44 (31.2)
	None	8 (5.7)
	Institutional/Internal	8 (5.7)
	Mixed	8 (5.7)
	Non-Governmental Organization (NGO)	1 (0.7)
	Industry	1 (0.7)
Methodological focus	Treatment-covariate interaction	26 (18.4)
	Analytical approach (e.g., one-stage vs two-stage)	20 (14.2)
	Comparing/combining IPD and AD	13 (9.3)
	Other statistical methods (e.g., Network meta-analysis, baseline imbalance, repeated measures, structural equation modelling)	10 (7.1)
	Missing data	10 (7.1)
	General guidance (e.g., timepoints, clustering, practical challenges, planning and conduct)	10 (7.1)
	Data processing	10 (7.1)
	Effect estimation models	9 (6.4)
	Heterogeneity	6 (4.3)
	Identification of subgroups	5 (3.6)
	Special consideration for observational studies	4 (2.8)
	Methodological quality assessment	4 (2.8)
	Standards and best practices	4 (2.8)
	Power calculation	4 (2.8)
	Bias assessment in IPDMA	3 (2.1)
	Reporting quality	3 (2.1)

3.3 Quality assessment

The quality of the included studies is summarised in Supplementary Table S1. Among the 141 records, 105 (74%) could be paired with an existing assessment tool. Simulation studies formed the largest block, comprising 54 papers that were appraised using the ADEMP framework [29], of which 46 were rated as high quality, 4 as moderate to high, 3 as moderate, and one as low quality. Strengths included clear objectives, well-specified data-generating mechanisms, target parameters, and performance metrics. The most consistent weakness was code unavailability, which may limit reproducibility. Flaws were primarily related to the limited description of elements, including data generation (e.g., unclear repetition schemes) and performance metrics [37-44], as well as limited explanation of the estimands [39, 41]. Nine systematic methodology reviews were appraised using ROBIS. Four were rated as high, while five were rated as moderate quality, mostly due to insufficient search strategies or detailed bias assessments [21, 22, 23, 45, 46]. Narrative or concept papers (n = 6) scored well on SANRA for scientific reasoning and referencing, but several lacked

searchable methods or structured data presentations [47, 48]. All eight consensus or guideline papers met the ACCORD criteria for stakeholder inclusiveness and reporting transparency; however, none had prospectively registered protocols, an omission that raises the possibility of selective outcome or method reporting. Additional limitations include a lack of formal validation or tool piloting, mostly in older consensus papers [17, 49-51]. An additional 29 empirical papers were examined using a checklist adapted from the ADEMP framework, which showed predominantly high scores. Across tools, 87 studies (63%) were rated high quality, 5 (4%) were moderate-to-high, 12 (9%) were moderate, and one was low; 35 papers (method, statistical guide, tutorials, and chapters from a handbook) were not assessed because no validated instrument exists for those designs (Figure 2).



Appraisal Tools	High	Moderate to high	Moderate	Low	NIL
Not assessed					35
Adapted tool	29	0	0	0	0
ACCORD	4	1	3	0	0
SANRA	4	1	1		0
ROBIS	4	0	5	0	0
Moris et al framework	46	3	4	1	0

Figure 2: Quality of included studies

3.4 Results of individual studies

Each included study was mapped to the specific methodological guidance it contributed, reflecting its stated focus, contribution type and thematic relevance to IPD-MA. The specific methodological guidance from each study ranges from broad overviews on how to conduct IPD-MA to more targeted recommendations. Each study's guidance topic, captured in the authors' own terms, is presented in Supplementary Table S1.

3.5 Results of syntheses

Our thematic synthesis organized the 141 studies into four overarching themes, each subdivided into 25 primary methodological foci and several granular subcategories, reflecting the dominant focus of each paper while also capturing secondary guidance areas, thereby retaining the full breadth of content. Because many articles spanned more than one aspect of IPD-MA, overlaps are shown within the sub-categories rather than forcing each study into a single box. We distilled a curated set of recommendations aligned to the four domains (Table 3). The complete mapping (theme,

subcategory, recommendation text, and contributing sources) is provided in Supplementary Table S2.

3.5.1 IPD-MA Planning

Among the studies addressing foundational topics of concern in IPD-MA, several provided guidance spanning the entire research process, from study conception to dissemination. Across these 20 papers, a recurring recommendation was that IPD-MA should only be undertaken when it offers clear methodological advantages over aggregate data meta-analyses, such as the ability to explore participant-level interactions, standardize outcome definitions, or address reporting biases. These benefits are typically articulated during the protocol development phase, which should incorporate a structured feasibility assessment, engagement with knowledge users, ethical review, and an operational coordination plan [4, 12, 17, 19, 52]. The use of standardized resources, such as field-tested toolkits, standard operating procedures (SOPs), and harmonization templates, was advocated to support complex multi-study workflows across the lifecycle of an IPD-MA [12, 52].

Key protocol elements should include a clear rationale, predefined objectives, eligibility criteria, and a procedure for protocol amendment (with decision rule and documentation). Detailed method should specify procedures for IPD acquisition, risk of bias assessment and mitigation strategies, definitions of key variables, prespecified statistical analyses (including estimands and power considerations for treatment-covariate interactions), and strategies for integrating IPD and aggregate data. All methodological decisions should be documented a priori and made publicly accessible (e.g., PROSPERO/OSF/publication) [4, 12, 17, 18, 30, 33, 53]. In alignment with best practices, IPD meta-analyses conducted within systematic reviews are expected to follow established principles for these reviews [12]. They should also incorporate centralized data integrity checks and statistical models that account for within-trial clustering, thereby minimizing bias and misinterpretation [12, 15, 53, 54]. The intention-to-treat principle was identified as the preferred analytical approach for the primary estimand [12].

Interpretation of findings should explicitly consider the impact of missing IPD, selection bias, publication bias, small-study effects [12, 15], and other design-related sources of bias [23, 54 - 55]. Moreover, an appraisal of the quality of the completed study was deemed essential [18]. For reporting, the use of established guidelines was recommended for IPD-MA based on randomized trials [17] and non-randomized studies [23]. Overall, the reviewed evidence emphasized the importance of structured planning, comprehensive documentation, and rigorous conduct throughout all phases of an IPD-MA. However, the findings also highlight implementation gaps, particularly in the consistent use of reporting checklists and formal bias assessment tools [4, 16, 17, 46, 56]. A summary of subcategories and corresponding references is provided in Table 3.

3.5.2 Study Selection, Data Access, and Preparation

Guidance converged on a staged, collaborative workflow that begins before data acquisition and extends through to the creation of a harmonized, analysis-ready master dataset. This workflow encompasses study selection, data governance, ethical and legal frameworks, data processing, and quality assurance [19, 51, 52, 56]. Although often peripheral in aggregate data meta-analyses, these processes are central to the conduct of IPD syntheses and were the subject of several papers providing detailed procedural guidance [17, 19, 47, 52, 56]. Initial steps should include the development of a transparent study selection dossier, documenting all trial invitations, the verification of eligibility criteria, and the rationale for non-participation or data access refusals [17, 19, 56]. Once participation is confirmed, governance procedures become paramount;

Recommended practices include standardized data-sharing agreements, authorship/publication policies, and requisite ethical/institutional approvals to facilitate timely data transfer while protecting participant confidentiality [19, 52, 53, 56]. The data transfer and verification phases were repeatedly highlighted as the critical data quality and integrity gate. Consistent with established guidance, systematic IPD "input audits," including validation against trial protocols or case report forms, cross-variable logic checks (e.g., age versus date of birth), completeness assessments, range and outlier detection, and adherence to predefined codebook [12, 19, 51, 57, 58], particularly for unpublished data [59]. Because results can differ materially when such checks are undertaken, best practice requires building quality assurance workflows (ideally with independent verification for critical steps) and maintaining a version-controlled audit trail linking raw data uploads to the final analysis dataset [52, 56, 60]. The audit trail should include embedded metadata documenting all data transformations, exclusions, imputations, and decisions made throughout the process, preserving all unmodified raw datasets to support reproducibility and enable error tracking during synthesis [17, 51, 52, 56, 57, 60]. In contexts where data sponsors restrict access to secure remote data environments, pre-specify fallback strategies and plan sensitivity analyses to address discrepancies that may arise from remote access limitations [12, 19, 47, 56]. For harmonization, prescriptive guidance recommends a dynamic data dictionary that maps variable names, units, and derived values across studies; consistent pre-model processing where appropriate (centring or rescaling of covariates for interaction modelling); and systematic logging of all transformations in an auditable pipeline [19, 54, 60, 61, 62]. Furthermore, plan sensitivity analyses that examine the impact of data-processing decisions themselves (e.g., comparing results before versus after harmonization, or excluding datasets that fail to meet the predefined completeness criteria [54, 59].

3.5.3 Meta-analysis

Across the included studies, the analysis stage is consistently portrayed as the analytical core of an IPD-MA. Authors emphasized the need for statistical methods that (i) preserve within-trial randomization, (ii) quantify between-trial heterogeneity, and (iii) maintain robustness in the presence of missing data or partially available aggregate data. Of the 141 reviewed papers, 107 (76%) provided explicit guidance on analytical strategy, estimation method, and model choices, across a range of outcome types, with the greatest focus on treatment-covariate interaction, analytical strategy (one-stage vs two-stage), most often for binary and time-to-event endpoints, as well as integrating aggregate data with IPD (Table 4). A recurring recommendation in the included literature is the use of one-stage mixed-effects models for most outcome types, given their ability to analyse all participants within a single likelihood while accounting for trial-level clustering; two-stage pooling remains a valid alternative often preferred for feasibility, especially when model convergence fails or only summary statistics are available [22, 40, 63-66], but tends to agree with one-stage when sample sizes and event rates are sufficient, and modelling assumptions and specification align [64, 67-70].

Outcome-specific guidance is highly granular. For binary outcomes, prefer a one-stage generalized linear mixed model with study-specific random intercepts, by Maximum likelihood (Laplace or adaptive Gauss-Hermite quadrature) [63, 66, 68, 71]. Penalized quasi-likelihood (PQL) can be acceptable in large datasets but may misestimate variance components and perform poorly with sparse data; prefer ML where feasible [71]. In small meta-analyses or with rare events, one-stage models generally yield less biased estimates [71]. When a one-stage analysis is infeasible or some studies only contribute aggregate data, a two-stage analysis using specific log-odds ratios (fixed or random effects) is a viable alternative. Mantel-Haenszel is reasonable in sparse settings, whereas

inverse variance random effects (e.g., with a DL) are common [22, 68, 71]. Under similar assumptions, with sufficient sample size or events, the two approaches agree [68, 71]. For continuous outcomes, one-stage linear mixed models with random intercepts (and, when heterogeneity is substantial, random slopes) were preferred; Centering treatment or covariate within trials helps disaggregate within- and across-trial information [63, 72-73]. For time-to-event endpoints, hazard-ratio-based analyses implemented as either one-stage (e.g., shared-frailty Cox or flexible parametric survival models) or two-stage pooling of study log-HRs under random effects, yield comparable results under similar model specification [64, 75]. In two-stage synthesis, meta-analyses should use inverse-variance pooling of log-HRs (or log-rank derived summaries, O-E and V); approaches that rely solely on study-level log-rank statistics are less efficient and may be biased [40]. When fixed-effect assumptions are appropriate, a one-stage stratified Cox model and a two-stage fixed-effect inverse-variance synthesis of log-HRs give nearly equivalent results [77]. Piecewise-exponential (Poisson) models are useful alternatives and can accommodate non-proportional hazards via time-varying effects [76]. If proportional hazards are doubtful or a time-bounded estimand is preferred, consider the restricted mean survival time (RMST) difference; simulation work supports a two-stage pooled Kaplan-Meier approach with random effects for RMST meta-analysis [78]. Specialized models exist for ordinal data [79], zero-inflated count outcomes [80], and single-case experimental designs [81]. IPD-MA involving observational data or survey designs should apply causal inference frameworks that incorporate sampling weights, clustering, and marginal structural modelling when identifiability assumptions permit [23, 55, 82]. Handling heterogeneity and effect modification were recurring focuses. The recommended strategy involves pre-specifying effect modifiers, estimating within- and across-trial interaction effects simultaneously in a mixed model framework [21, 45, 48, 54, 83]. Address non-linearity with restricted cubic splines or fractional polynomials, and visualized treatment-covariate interactions through model-based marginal effects plots [45, 54, 83]. Exploratory tools, such as model-based recursive partitioning, RECPAM, and rough-set classifiers, may be a secondary follow-up when prespecified interaction models fail to explain heterogeneity. This approach should be paired with cross-validation or bootstrapping to prevent overfitting and spurious findings [83-85]. Recommendations for handling missing data are explicit. Address sporadic missingness with single-level multiple imputation by chained equations (MICE) within each trial [86], whereas systematically missing covariates require multilevel imputation methods that borrow strength across studies [74, 87]. For longitudinal or repeated-measures data, the preferred approach is to model the full trajectory using multilevel mixed models. If computational feasibility becomes a barrier, extracting visit-specific contrasts and synthesizing in a multivariate two-stage meta-analysis [88].

Finally, a cluster of additional methods widens the toolbox: baseline-imbalance adjustments for continuous outcome via within-trial ANCOVA [89]; IPD meta-analytic structural-equation modelling for latent mediators and pathways [90]; and simulation-based power planning, implemented via Stata's *ipdpower* for one-stage mixed effects designs (including interaction terms and endpoints specific targets) to assess feasibility of detecting effect modification [54, 91-92] with a simulation framework when planning a two-stage IPD-MA [93]. (Table 3).

3.5.4 IPD Network Meta-analysis (IPD-NMA)

IPD-based network meta-analysis is justified when indirect comparisons are necessary for decision-making and when treatment-covariate interactions may compromise the transitivity assumptions underlying standard aggregate data NMA. Pre-specifying whether IPD will be

collected for all studies, a subset, and which part of the network, as this affects network coverage, model selection, and power [44, 94, 99]. Prioritize obtaining IPD where it adds the most value (e.g., sparse links, effect-modifier imbalance) [94, 99]. Data harmonization is central to IPD-NMA and should include mapping treatment variables to a unified coding system before merging [97]. In cases where IPD is unavailable for certain studies/comparators (e.g., binary endpoint), include their aggregate counts via a Binomial likelihood alongside the Bernoulli IPD likelihood in a shared-parameter Bayesian one-stage NMA to enable coherent joint estimation, and if single-arm trials are included, specify assumptions about baseline response and examine their impact on estimates via sensitivity analyses. [95]. Conduct primary syntheses using a prespecified hierarchical NMA framework (one-stage or two-stage, Bayesian or frequentist) with at least a random study intercept to enforce the standard consistency constraints and quantify between-study heterogeneity [94, 97]. One-stage shared-parameter models are often used because they readily accommodate treatment-covariate interactions and multi-arm trials. Model treatment-covariate interactions using IPD to separate within-trial from across-trial information and, where appropriate, allow random interaction slopes [94, 97, 98].

Evaluate assumptions by describing network geometry, checking transitivity, and assessing inconsistency with global/local diagnostics; compare candidate models using appropriate information criteria (e.g., DIC), and predictive checks [44, 94, 97]. If inconsistency is detected, modelling guidance includes decomposition of variability into heterogeneity and inconsistency, with attention to whether inconsistency stems from placement and proportion of IPD (i.e., IPD availability imbalance) [94, 99]. Perform sensitivity analyses on IPD placement and coverage in the network [97, 99], covariate adjustment choices [94, 97, 98], the impact of single arm assumptions (including a two-stage variant as a check), and where relevant, compare alternative parametrizations (contrast versus arm-based) [95] especially when only a few trials supply IPD [99].

Report each study's percentage weight for each parameter (overall effect and interaction) [95], IPD and aggregate data contributions across the network using geometry plots, node/edge counts, and present the planned sensitivity analyses and their results [17, 44, 95, 99].

Table 3: Summary recommendation mapped to the four themes

Primary focus	Subcategories	References
Foundations and study design		
Design and implementation guidance	Rationale and scope, feasibility, statistical framework choices (including IPD-NMA), project management tool kits, ethical and collaborative considerations	13, 52, 56, 94, 97, 99, 109, 134
Standards and best practice	Ethical, methodological and practical standards for IPD conduct and workflow (protocols, collaboration, harmonization, reproducibility), analytical modelling standards, reporting and appraising IPD-Meta-analysis	4, 12, 46, 53, 94, 97, 133, 106
Methodological quality assessment	Quality assessments of the systematic review appraisal of design/conduct/reporting elements unique to IPD-MA, overall methodological quality, RCTs and/or non-randomized.	16, 18, 30, 33

Bias assessment and mitigation	Appropriate ROB tools for included studies, Assessment of the impact of excluding participants, incomplete studies or IPD on estimates. (Publication, selection, small study effect, availability bias)	15, 16, 59, 99, 117, 119, 138,
Reporting	Identification of reporting gap, recommendations on what to report and special considerations for observational studies	17, 46, 23, 94, 95, 97, 101
Theme 2: Study Selection, Data Access & Preparation		
Data access and governance	Data sharing policies, governance, ethics, and legal issues, as well as data access pathways (including IPD requests and transfers), agreements and compliance, and retrieval barriers and feasibility.	47, 58, 143, 158
Data processing, quality and integrity checks.	IPD retrieval & validation; provenance or versioning; quality and integrity assessment; harmonization & derivation; and the impact of checking on meta-analytic estimates	12, 51, 56, 57, 59, 60, 61, 62, 97, 102
Theme 3: Statistical Methods and Analysis Strategy		
Methods for estimation, modelling, and analytic approaches for different endpoints	Binary	22, 63, 66, 71, 113, 123, 154
	Continuous	41, 63, 72, 73
	Time-to-event	42, 64, 75, 76, 77, 78, 116, 120,
	Ordinal	14, 79
	Count	14, 80
Statistical recommendations for special data presentations	Single case experimental design	81
	Network meta-analysis	94-99, 141, 158, 160, 161, 163
	Comparing and combining IPD and AD	14, 43, 65, 95, 103, 110, 115, 118, 137, 146, 148, 163
	Heterogeneity: Testing and handling in all outcome data types	38, 49, 111, 125, 129, 153, 156
	Subgroup Identification Methods	39, 84, 85, 107, 144
	Treatment-covariate interaction: Estimation of participant-level, study-level interaction, non-linear treatment and covariate interaction and interaction evaluation with single case experimental design (SCED)	20, 21, 45, 48, 54, 83, 85, 105, 114, 127, 128, 131, 135, 139, 142, 145, 146, 147, 149, 150, 155, 162
	Missing data (Sporadic, systematic and mixed)	47, 60, 74, 86, 87, 108, 112, 121, 140, 149, 152
Theme 4: Other Statistical and methodological considerations		

Baseline imbalance	Estimation model, analytical approach and estimation method when Baseline imbalance exists in a Continuous outcome.	89, 126
Structural equation modelling	(SEM): For latent variables (e.g., factor scores, mediation paths, latent means)	90
Power calculation	For binary, continuous and survival endpoints to check the feasibility of interaction	54, 91, 92, 93, 135
Repeated measures	Repeated measures/Longitudinal studies	88, 104
Special consideration for the observational study:	Causal inference in complex survey design, clinical interpretation checklist, causal-inference framework and reporting standards for observational IPD meta-analyses	23, 55, 82, 130

Abbreviations: AD, Aggregate Data; SCED, Single-Case Experimental Design; IPD-NMA, Individual Participant Data Network Meta-analysis; ROB; SEM, Structural Equation Modelling.

4. Discussion

4.1 Summary of key findings

In this systematic methodology review, we found a wealth of recommendations in the literature addressing both the general conduct and specific technical aspects of IPD-MA, spanning their lifecycle, from study identification and data acquisition to analysis, reporting, and dissemination. Collectively, the literature provides rich, albeit fragmented, guidance on the planning and conduct of IPD-MAs. We extracted and consolidated these recommendations into a set of quality items across all stages of IPD-MA conduct, which will inform a subsequent Delphi survey aimed at developing a consensus-based critical appraisal tool for reviews with IPD-MA. Given the increasing prevalence of IPD-MAs in health research [16, 44], this review offers a timely synthesis of methodological principles that may help to standardize and improve the conduct of future IPD syntheses. We found relatively limited attention to IPD-MAs involving observational (non-randomized) studies. The existing literature is heavily weighted toward those involving randomized trials, often addressing the added complexities of non-randomized data (e.g., confounding, selection bias) only in passing.

4.2 Interpretation in the context of other evidence

Our review differs from previous syntheses of the IPD-MA methodology in both scope and depth. While earlier reviews have contributed valuable insights, many have focused on discrete aspects of IPD-MA methodology in isolation. For example, Debray et al. (2015) [14] concentrated primarily on the analytical methods, emphasizing modelling strategies and software for comparative effectiveness questions, with relatively little discussion of upstream elements such as data identification, acquisition, or study quality assessment. Our review extends beyond statistical modelling, incorporating operational and procedural guidance, including collaboration initiation, data verification, harmonization, missing data handling, and reporting. We also included guidance relevant to both IPD-MA of the effects of interventions that use randomized and non-randomized studies. While Debray et al. [14] and others acknowledged the inclusion of non-randomized studies of intervention, their focus remained on clinical trials and efficacy estimation. Other

methodological papers provided in-depth guidance on specific analytical techniques, such as combining aggregate data and IPD [65], exploring treatment-covariate interactions [20, 21], handling non-linear interactions [45], and methods for time-to-event data, including approaches suitable under non-proportional hazards [75-76]. Outcome-specific methods for binary and time-to-event endpoints have also been well-documented [22, 75], including comparative evaluations of one-stage versus two-stage models and their statistical assumptions [64, 66, 68, 69, 72, 80]. Despite the richness of these sources, we did not identify any prior systematic review that integrates methodological guidance across all phases of IPD-MA conduct. This review consolidates scattered insights into a unified framework, including guidance on power calculations for different endpoints and analytical strategies [91-93], IPD-NMA [94-98], and IPD-MA of observational studies evaluating intervention effects [23].

4.3 Strengths and limitations

This review's strength is its comprehensive, systematic approach. We searched multiple databases, including peer-reviewed articles, guidelines, handbook chapters, grey literature, and expert consensus statements, reducing the risk of missing guidance. We identified extensive guidance across all phases of IPD-MA, from planning to reporting. When suitable, we assessed the quality of papers using tools like ADEMP for simulation studies. Including guidance for both randomized and observational studies broadens applicability, aiding diverse research settings. Our review was purpose-driven; the synthesis aimed to inform a Delphi process and develop a practical critical appraisal tool, ensuring guidance was actionable, not just descriptive.

Nonetheless, we acknowledge several limitations. First, the strength of evidence supporting individual recommendations varied considerably. While some guidance is based on simulation studies or formal consensus, others rely primarily on expert opinion or theoretical justification. We chose to be inclusive of all proposed best practices, which means the resulting list contains items for which consensus or empirical validation may still be lacking. This variability limits the overall certainty of the synthesized guidance. It will be addressed during the Consensus process for developing the critical appraisal tool, where experts will assess the importance and clarity of each item. Second, we were unable to assess the quality of 35 formally included studies because these were not traditional studies with quality metrics, but rather guidance papers; consequently, all guidance was weighted equally in our synthesis, even though, for instance, a simulation study or formal consensus guideline might carry more authority than a recommendation on a single method from a single study. Third, our method of synthesizing guidance entailed qualitative interpretation and consolidation. Different sources sometimes present similar concepts using varied terminology or framing, requiring our team to make judgment calls about whether to merge or distinguish guidance points. While this synthesis approach was necessary for practicality, it introduces a degree of subjectivity. Fourth, generalizability: our review aims to inform a tool that applies broadly; however, some guidance may be context-specific (e.g., applicable only to IPD-MAs of certain observational study designs). We hope to filter out or reframe overly narrow items during the consensus. In the interim, our inclusive approach may result in a lengthy or overlapping set of candidate items. However, we believe this breadth is necessary to ensure comprehensiveness at the item-generation phase. Finally, our actual synthesis deviated from the a priori thematic plan outlined in the registered protocol. This is because, in keeping with the nature of qualitative evidence synthesis, where data complexity often necessitates iterative coding and theme

refinement, our approach evolved substantively during execution. This means that what was ultimately conducted differs from what was initially registered.

4.5 Implications for practice and future research

The results of this review carry important implications for both the conduct of IPD-MA and assessments of its quality [16, 100], as well as ongoing efforts to improve it. For practice, the synthesized recommendations provide a consolidated methodological reference that can be used as a best-practice checklist across all phases of an IPD-MA. From the perspective of research quality appraisal, our review offers a foundational evidence base for generating and wording candidate items for the IPD meta-analysis critical appraisal tool. Importantly, this synthesis also highlights key gaps in existing guidance on the methods for IPD-MA. There is a notable scarcity of detailed guidance tailored to IPD-MAs involving observational studies of intervention. Further work is needed on harmonizing and adjusting for differing confounders across studies, as well as design-aware analysis strategies (e.g., handling complex survey designs). Additionally, principled approaches are needed for handling heterogeneous data structures with intrinsically lower data consistency and bias assessment when pooling data from non-randomized sources. Additionally, the impact of adhering to, or deviating from, specific methodological recommendations in IPD-MAs remains underexplored. Evaluative studies could compare how alternative processing and modelling choices influence effect estimates, uncertainties, and conclusions.

5. Conclusion

While the literature offers a wealth of recommendations on IPD-MA methods, these have not been previously collated into a cohesive framework. Our systematic methodology review provides an integrated synthesis of guidance spanning the full cycle of activities in an IPD-MA. We have also created a resource that can inform the development of a consensus critical appraisal and quality assessment tool. Our work serves as a foundation upon which the IPD-MA methodology can continue to be refined, with future research targeting the unresolved questions and new challenges as this research method becomes increasingly prevalent.

Acknowledgement

We want to express our gratitude to the individuals who contributed to this research. Specifically, we thank Roxanne Isard and Alanna Marson for refining the literature search strategy, Sarah Abdunnabi, Ph.D. Candidate and Nana-Ama Tiwaa-Boateng, MSc, for screening titles, abstracts, and for the full-text review.

Credit authorship contribution statement

Edith G. Otalike: Conceptualization; Methodology, Data curation; Formal analysis, Writing - original draft, Writing - review & editing, Project administration. **Mike Clarke**: Conceptualization, Methodology, Supervision, Writing - review & editing. **Farjana Akhter**: Methodology, Validation, Data curation, Writing - review & editing. **Areti Angeliki Veroniki**: Methodology, Writing - review & editing. **Ngianga-Bakwin Kandala**: Conceptualization, Methodology, Writing - review & editing. **Joel J. Gagnier**: Conceptualization, Methodology, Supervision, Writing - review & editing.

Supplementary files

Supplementary Files 1: Full search strategy
Supplementary Files 2: The complete reference list
Supplementary Table 1: Characteristics of included studies
Supplementary Table 2: Result of synthesis (Recommendation catalogue)
Supplementary Table 3: The PRISMA Checklist.

References

1. Kuiper, R., Buskens, V., Raub, W., Hoijtink, H., 2012. Combining statistical evidence from several studies: a method using Bayesian updating and an example from research on trust problems in social and economic exchange. Sociol. Methods Res. 42, 60–81. doi: 10.1177/0049124112464867.
2. Rosner AL. Evidence-based medicine: revisiting the pyramid of priorities. J Bodyw Mov Ther. 2012 Jan;16(1):42-9. doi: 10.1016/j.jbmt.2011.05.003.
3. Chavalarias D, Wallach JD, Li AHT, Ioannidis JPA. Evolution of Reporting *P* Values in the Biomedical Literature, 1990-2015. *JAMA.* 2016;315(11):1141–1148. doi:10.1001/jama.2016.1952
4. Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. BMJ. 2010 Feb 5;340:c221. doi: 10.1136/bmj.c221. PMID: 20139215.
5. Showell MG, Cole S, Clarke MJ, DeVito NJ, Farquhar C, Jordan V. Time to publication for results of clinical trials. Cochrane Database of Systematic Reviews 2024;(11):MR000011. doi: 10.1002/14651858.MR000011.pub3.
6. Thompson SG, Higgins JPT. How should meta-regression analyses be undertaken and interpreted? Stat Med. 2002;21(11):1559-1573. doi:10.1002/sim.1187
7. Oxman AD, Clarke MJ, Stewart LA. From science to practice. Meta-analyses using individual patient data are needed. JAMA. 1995 Sep 13;274(10):845-6. doi: 10.1001/jama.274.10.845. PMID: 7650811.
8. Tudur Smith C, Marcucci M, Nolan SJ, Iorio A, Sudell M, Riley R, Rovers MM, Williamson PR. Individual participant data meta-analyses compared with meta-analyses based on aggregate data. Cochrane Database Syst Rev. 2016;9(9): MR000007. doi: 10.1002/14651858.MR000007.pub3.
9. Olkin I, Sampson A. Comparison of meta-analysis versus analysis of variance of individual patient data. Biometrics. 1998 Mar;54(1):317-22. PMID: 9544524.
10. Tierney JF, Fisher DJ, Burdett S, Stewart LA, Parmar MKB. Comparison of aggregate and individual participant data approaches to meta-analysis of randomised trials: An observational study. PLoS Med. 2020 Jan 31;17(1):e1003019. doi: 10.1371/journal.pmed.1003019.
11. Tierney JF, Pignon JP, Gueffyier F, Clarke M, Askie L, Vale CL, Burdett S; Cochrane IPD Meta-analysis Methods Group. How individual participant data meta-analyses have influenced trial design, conduct, and analysis. J Clin Epidemiol. 2015 Nov;68(11):1325-35. doi: 10.1016/j.jclinepi.2015.05.024.
12. Tierney JF, Vale C, Riley R, Smith CT, Stewart L, Clarke M, Rovers M. Individual Participant Data (IPD) Meta-analyses of Randomised Controlled Trials: Guidance on Their Use. PLoS Med. 2015 Jul 21;12(7):e1001855. doi: 10.1371/journal.pmed.1001855. PMID: 26196287; PMCID: PMC4510878.
13. Veroniki AA, Seitidis G, Tsivgoulis G, Katsanos AH, Mavridis D. An Introduction to Individual Participant Data Meta-analysis. Neurology. 2023 Jun 6;100(23):1102-1110. doi: 10.1212/WNL.0000000000207078. Epub 2023 Feb 16. PMID: 36797070; PMCID: PMC10256124.
14. Debray TP, Moons KG, van Valkenhoef G, Efthimiou O, Hummel N, Groenwold RH, Reitsma JB; GetReal Methods Review Group. Get real in individual participant data (IPD) meta-analysis: a methodology review. Res Synth Methods. 2015 Dec;6(4):293-309. doi: 10.1002/jrsm.1160.

15. Ahmed I, Sutton AJ, Riley RD. Assessment of publication bias, selection bias, and unavailable data in meta-analyses using individual participant data: a database survey. BMJ 2012; 344: doi: 10.1136/bmj.d7762
16. Wang H, Chen Y, Lin Y, Abesig J, Wu IX, Tam W. The methodological quality of individual participant data meta-analysis on intervention effects: systematic review. BMJ. 2021 Apr 19;373:n736. doi: 10.1136/bmj.n736.
17. Stewart LA, Clarke M, Rovers M, et al. Preferred Reporting Items for a Systematic Review and Meta-analysis of Individual Participant Data: The PRISMA-IPD Statement. *JAMA.* 2015;313(16):1657–1665. doi:10.1001/jama.2015.3656
18. Jayne F. Tierney, Lesley A. Stewart, Claire L. Vale, and Richard D. Riley. A Tool for the Critical Appraisal of IPD Meta-Analysis Projects (CheckMAP): Individual Participant Data Meta-Analysis: A Handbook for Healthcare Research, First Edition. Edited by Richard D. Riley, Jayne F. Tierney, and Lesley A. Stewart. (2021)
19. Tierney JF, Stewart LA, Clarke M. Chapter 26: Individual participant data. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions* version 6.4 (updated August 2023). Cochrane, 2023
20. Godolphin PJ, Marlin N, Cornett C, Fisher DJ, Tierney JF, White IR, Rogozińska E. Use of multiple covariates in assessing treatment-effect modifiers: A methodological review of individual participant data meta-analyses. Res Synth Methods. 2024 Jan;15(1):107-116.
21. Koopman L, van der Heijden GJ, Glasziou PP, Grobbee DE, Rovers MM. A systematic review of analytical methods used to study subgroups in (individual patient data) meta-analyses. J Clin Epidemiol. 2007 Oct;60(10):1002-9.
22. Thomas D, Radji S, Benedetti A. Systematic review of methods for individual patient data meta-analysis with binary outcomes. BMC Med Res Methodol. 2014 Jun 19;14:79. doi: 10.1186/1471-2288-14-79.
23. Hufstedler H, Mauer N, Yeboah E, Carr S, Rahman S, Danzer AM, et al. Application of causal inference methods in individual-participant data meta-analyses in medicine: addressing data handling and reporting gaps with new proposed reporting guidelines. BMC Med Res Methodol. 2024 Apr 19;24(1):91. doi: 10.1186/s12874-024-02210-9.
24. Clarke M. Guide to the contents of a Cochrane Methodology protocol and review. In: Clarke M, editor. Guide to the contents of Cochrane Methodology protocol and review. 2020.
25. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71. doi:10.1136/bmj.n71
26. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev. 2015 Dec 1;4(1):1.
27. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Statement. J Clin Epidemiol. 2016 Jul;75:40-6. doi: 10.1016/j.jclinepi.2016.01.021.
28. Nevitt SJ, Marson AG, Davie B, Reynolds S, Williams L, Smith CT. Exploring changes over time and characteristics associated with data retrieval across individual participant data meta-analyses: systematic review. BMJ. 2017 Apr 5;357:j1390. doi: 10.1136/bmj.j1390.
29. Covidence systematic review software, Veritas Health Innovation, Melbourne, Australia. Available at www.covidence.org.
30. Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: A critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ (Online). 2017;358.
31. Morris TP, White IR, Crowther MJ. Using simulation studies to evaluate statistical methods. Stat Med. 2019 May 20;38(11):2074-2102. doi: 10.1002/sim.8086.

32. Baethge C, Goldbeck-Wood S, Mertens S. SANRA-a scale for the quality assessment of narrative review articles. Res Integr Peer Rev. 2019 Mar 26;4:5. doi: 10.1186/s41073-019-0064-8.
33. Whiting P, Savović J, Higgins JP, Caldwell DM, Reeves BC, Shea B, Davies P, Kleijnen J, Churchill R; ROBIS group. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. J Clin Epidemiol. 2016 Jan;69:225-34. doi: 10.1016/j.jclinepi.2015.06.005. Epub 2015 Jun 16.
34. Gattrell WT, Logullo P, van Zuuren EJ, Price A, Hughes EL, Blazey P, Winchester CC, Tovey D, Goldman K, Hungin AP, Harrison N. ACCORD (ACcurate COnsensus Reporting Document): A reporting guideline for consensus methods in biomedicine developed via a modified Delphi. PLoS Med. 2024 Jan 23;21(1):e1004326. doi: 10.1371/journal.pmed.1004326.
35. Thomas, J., Harden, A. Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Med Res Methodol* **8**, 45 (2008). https://doi.org/10.1186/1471-2288-8-45
36. Otalike EG, Clarke M, Veroniki AA, Tricco AC, Moher D, Shea B, Doherty-Kirby A, Kandala NB, Gagnier J. Development and validation of critical appraisal tool for individual participant data meta-analysis: protocol for a modified e-Delphi study. BMJ Open. 2025 Jun 25;15(6):e097297. doi: 10.1136/bmjopen-2024-097297.
37. Katsahian S, Latouche A, Mary JY, Chevret S, Porcher R. Practical methodology of meta-analysis of individual patient data using a survival outcome. Contemp Clin Trials. 2008 Mar;29(2):220-30. doi: 10.1016/j.cct.2007.08.002.
38. Rudrapatna VA, Ravindranath VG, Arneson DV, Mosenia A, Butte AJ, Wang S. Sequential regression and simulation: a method for estimating causal effects from heterogeneous clinical trials without a common control group. BMC Med Res Methodol. 2023 Oct 3;23(1):218. doi: 10.1186/s12874-023-02020-5.
39. Gavaghan DJ, Moore AR, McQuay HJ. An evaluation of homogeneity tests in meta-analyses in pain using simulations of individual patient data. Pain. 2000 Apr;85(3):415-424. doi: 10.1016/S0304-3959(99)00302-4.
40. Simmonds MC, Tierney J, Bowden J, Higgins JP. Meta-analysis of time-to-event data: a comparison of two-stage methods. Res Synth Methods. 2011 Sep;2(3):139-49. doi: 10.1002/jrsm.44.
41. Papadimitropoulou K, Stijnen T, Dekkers OM, le Cessie S. One-stage random effects meta-analysis using linear mixed models for aggregate continuous outcome data. Res Synth Methods. 2019 Sep;10(3):360-375. doi: 10.1002/jrsm.1331.
42. Crowther MJ, Look MP, Riley RD. Multilevel mixed effects parametric survival models using adaptive Gauss-Hermite quadrature with application to recurrent events and individual participant data meta-analysis. Stat Med. 2014 Sep 28;33(22):3844-58. doi: 10.1002/sim.6191.
43. Yamaguchi Y, Sakamoto W, Goto M, Staessen JA, Wang J, Gueyffier F, Riley RD. Meta-analysis of a continuous outcome combining individual patient data and aggregate data: a method based on simulated individual patient data. Res Synth Methods. 2014 Dec;5(4):322-51. doi: 10.1002/jrsm.1119.
44. Leahy J, O'Leary A, Afdhal N, Gray E, Milligan S, Wehmeyer MH, Walsh C. The impact of individual patient data in a network meta-analysis: An investigation into parameter estimation and model selection. Res Synth Methods. 2018 Sep;9(3):441-469. doi: 10.1002/jrsm.1305.
45. Marlin N, Godolphin PJ, Hooper RL, Riley RD, Rogozińska E. Nonlinear effects and effect modification at the participant-level in IPD meta-analysis part 2: methodological guidance is available, Journal of Clinical Epidemiology, Volume 159, 2023, Pages 319-329, ISSN 0895-4356, https://doi.org/10.1016/j.jclinepi.2023.04.014.
46. Simmonds MC, Higgins JP, Stewart LA, Tierney JF, Clarke MJ, Thompson SG. Meta-analysis of individual patient data from randomized trials: a review of methods used in practice. Clin Trials. 2005;2(3):209-17. doi: 10.1191/1740774505cn087oa. PMID: 16279144.

47. Rogozińska E, Marlin N, Thangaratinam S, Khan KS, Zamora J. Meta-analysis using individual participant data from randomised trials: opportunities and limitations created by access to raw data. Evid Based Med. 2017 Oct;22(5):157-162. doi: 10.1136/ebmed-2017-110775.
48. 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 Sep;64(9):949-67. doi: 10.1016/j.jclinepi.2010.11.016.
49. Gagnier JJ, Morgenstern H, Altman DG, Berlin J, Chang S, McCulloch P, et al,. Consensus-based recommendations for investigating clinical heterogeneity in systematic reviews. BMC Med Res Methodol. 2013 Aug 30;13:106. doi: 10.1186/1471-2288-13-106.
50. Pincus T, Miles C, Froud R, Underwood M, Carnes D, Taylor SJ. Methodological criteria for the assessment of moderators in systematic reviews of randomised controlled trials: a consensus study. BMC Med Res Methodol. 2011 Jan 31;11:14. doi: 10.1186/1471-2288-11-14.
51. Hunter KE, Aberoumand M, Libesman S, Sotiropoulos JX, Williams JG, Aagerup J, et al,. The Individual Participant Data Integrity Tool for assessing the integrity of randomised trials. Res Synth Methods. 2024 Nov;15(6):917-939. doi: 10.1002/jrsm.1738.
52. Maxwell L, Shreedhar P, Carabali M, Levis B. How to plan and manage an individual participant data meta-analysis. An illustrative toolkit. Res Synth Methods. 2024 Jan;15(1):166-174. doi: 10.1002/jrsm.1670.
53. Hollis S, Fletcher C, Lynn F, Urban HJ, Branson J, Burger HU, et al. Best practice for analysis of shared clinical trial data. BMC Med Res Methodol. 2016 Jul 8;16 Suppl 1(Suppl 1):76. doi: 10.1186/s12874-016-0170-y.
54. 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 Jul 10;39(15):2115-2137. doi: 10.1002/sim.8516
55. Ballew SH, Mok Y, Matsushita K. A Practical Guide to Interpret Individual Participant Data Meta-analysis of Observational Studies. Am J Kidney Dis. 2021 Sep;78(3):464-467. doi: 10.1053/j.ajkd.2021.02.333.
56. Nevitt SJ, Tudur Smith C. Practical Considerations and Challenges When Conducting an Individual Participant Data (IPD) Meta-Analysis. Methods Mol Biol. 2022;2345:263-278. doi: 10.1007/978-1-0716-1566-9_16.
57. Hunter KE, Webster AC, Clarke M, Page MJ, Libesman S, Godolphin PJ, et al. Development of a checklist of standard items for processing individual participant data from randomised trials for meta-analyses: Protocol for a modified e-Delphi study. PLoS One. 2022 Oct 11;17(10):e0275893.
58. Ventresca M, Schünemann HJ, Macbeth F, Clarke M, Thabane L, Griffiths G, et al. Obtaining and managing data sets for individual participant data meta-analysis: scoping review and practical guide. BMC Med Res Methodol. 2020 May 12;20(1):113. doi: 10.1186/s12874-020-00964-6.
59. Burdett S, Stewart LA. A comparison of the results of checked versus unchecked individual patient data meta-analyses. Int J Technol Assess Health Care. 2002 Summer;18(3):619-24. PMID: 12391954.
60. Levis B, Hattle M, Riley RD. PRIME-IPD SERIES Part 2. Retrieving, checking, and harmonizing data are underappreciated challenges in individual participant data meta-analyses. J Clin Epidemiol. 2021 Aug;136:221-223. doi: 10.1016/j.jclinepi.2021.05.006.
61. Kalter, J., Sweegers, M.G., Verdonck-de Leeuw, I.M. et al. Development and use of a flexible data harmonization platform to facilitate the harmonization of individual patient data for meta-analyses. BMC Res Notes 12, 164 (2019). https://doi.org/10.1186/s13104-019-4210-7
62. Dewidar O, Riddle A, Ghogomu E, Hossain A, Arora P, Bhutta ZA, et al. PRIME-IPD SERIES Part 3. The PRIME-IPD tool fills a gap in guidance for preparing IPD for analysis. J Clin Epidemiol. 2021 Aug;136:224-226.

63. Riley RD, Legha A, Jackson D, Morris TP, Ensor J, Snell KIE, et al. One-stage individual participant data meta-analysis models for continuous and binary outcomes: Comparison of treatment coding options and estimation methods. Stat Med. 2020 Aug 30;39(19):2536-2555. doi: 10.1002/sim.8555.
64. Bowden J, Tierney JF, Simmonds M, Copas AJ, Higgins JP. Individual patient data meta-analysis of time-to-event outcomes: one-stage versus two-stage approaches for estimating the hazard ratio under a random effects model. Res Synth Methods. 2011 Sep;2(3):150-62. doi: 10.1002/jrsm.45.
65. Riley RD, Simmonds MC, Look MP. Evidence synthesis combining individual patient data and aggregate data: a systematic review identified current practice and possible methods. J Clin Epidemiol. 2007 May;60(5):431-9. doi: 10.1016/j.jclinepi.2006.09.009.
66. Debray TP, Moons KG, Abo-Zaid GM, Koffijberg H, Riley RD. Individual participant data meta-analysis for a binary outcome: one-stage or two-stage? PLoS One. 2013 Apr 9;8(4):e60650. doi: 10.1371/journal.pone.0060650.
67. Riley RD, Ensor J, Hattle M, Papadimitropoulou K, Morris TP. Two-stage or not two-stage? That is the question for IPD meta-analysis projects. Res Synth Methods. 2023 Nov;14(6):903-910. doi: 10.1002/jrsm.1661.
68. 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 Feb 28;36(5):855-875. doi: 10.1002/sim.7141.
69. Stewart GB, Altman DG, Askie LM, Duley L, Simmonds MC, et al. Statistical Analysis of Individual Participant Data Meta-Analyses: A Comparison of Methods and Recommendations for Practice. PLoS ONE. 2012; 7(10): e46042. doi:10.1371/journal.pone.0046042
70. Scotti L, Rea F, Corrao G. One-stage and two-stage meta-analysis of individual participant data led to consistent summarized evidence: lessons learned from combining multiple databases. J Clin Epidemiol. 2018 Mar;95:19-27. doi: 10.1016/j.jclinepi.2017.11.020.
71. Thomas D, Platt R, Benedetti A. A comparison of analytic approaches for individual patient data meta-analyses with binary outcomes. BMC Med Res Methodol. 2017 Feb 16;17(1):28. doi: 10.1186/s12874-017-0307-7.
72. Legha A, Riley RD, Ensor J, Snell KIE, Morris TP, Burke DL. Individual participant data meta-analysis of continuous outcomes: A comparison of approaches for specifying and estimating one-stage models. Stat Med. 2018 Dec 20;37(29):4404-4420. doi: 10.1002/sim.7930.
73. Morris TP, Fisher DJ, Kenward MG, Carpenter JR. Meta-analysis of Gaussian individual patient data: Two-stage or not two-stage? Stat Med. 2018 Apr 30;37(9):1419-1438. doi: 10.1002/sim.7589.
74. Jolani S. Hierarchical imputation of systematically and sporadically missing data: An approximate Bayesian approach using chained equations. Biom J. 2018 Mar;60(2):333-351. doi: 10.1002/bimj.201600220.
75. 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 Mar;11(2):148-168. doi: 10.1002/jrsm.1384.
76. Crowther MJ, Riley RD, Staessen JA, Wang J, Gueyffier F, Lambert PC. Individual patient data meta-analysis of survival data using Poisson regression models. BMC Med Res Methodol. 2012 Mar 23;12:34. doi: 10.1186/1471-2288-12-34.
77. Tudur Smith C, Williamson PR. A comparison of methods for fixed effects meta-analysis of individual patient data with time to event outcomes. Clin Trials. 2007;4(6):621-30. doi: 10.1177/1740774507085276.
78. Lueza B, Rotolo F, Bonastre J, Pignon JP, Michiels S. Bias and precision of methods for estimating the difference in restricted mean survival time from an individual patient data meta-analysis. BMC Med Res Methodol. 2016 Mar 29;16:37. doi: 10.1186/s12874-016-0137-z.

79. Whitehead A, Omar RZ, Higgins JP, Savaluny E, Turner RM, Thompson SG. Meta-analysis of ordinal outcomes using individual patient data. Stat Med. 2001 Aug 15;20(15):2243-60. doi: 10.1002/sim.919.
80. Huh D, Baldwin SA, Zhou Z, Park J, Mun EY. Which is Better for Individual Participant Data Meta-Analysis of Zero-Inflated Count Outcomes, One-Step or Two-Step Analysis? A Simulation Study. Multivariate Behav Res. 2023 Nov-Dec;58(6):1090-1105. doi: 10.1080/00273171.2023.2173135.
81. Declercq L, Jamshidi L, Fernández CB, Moeyaert M, Beretvas SN, Ferron JM, et al. Multilevel Meta-Analysis of Individual Participant Data of Single-Case Experimental Designs: One-Stage versus Two-Stage Methods. Multivariate Behav Res. 2022 Mar-May;57(2-3):298-317. doi: 10.1080/00273171.2020.1822148.
82. Brunner M, Keller L, Stallasch SE, Kretschmann J, Hasl A, Preckel F, et al. Meta-analyzing individual participant data from studies with complex survey designs: A tutorial on using the two-stage approach for data from educational large-scale assessments. Res Synth Methods. 2023 Jan;14(1):5-35. doi: 10.1002/jrsm.1584.
83. 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 Mar 3;356:j573. doi: 10.1136/bmj.j573.
84. Gil-Herrera E, Tsalatsanis A, Kumar A, Mhaskar R, Miladinovic B, Yalcin A, Djulbegovic B. Identifying homogenous subgroups for individual patient meta-analysis based on Rough Set Theory. Annu Int Conf IEEE Eng Med Biol Soc. 2014;2014:3434-7. doi: 10.1109/EMBC.2014.6944361.
85. Huber, C., Benda, N. & Friede, T. Subgroup identification in individual participant data meta-analysis using model-based recursive partitioning. Adv Data Anal Classif 16, 797–815 (2022). https://doi.org/10.1007/s11634-021-00458-3.
86. Burgess S, White IR, Resche-Rigon M, Wood AM. Combining multiple imputation and meta-analysis with individual participant data. Stat Med. 2013 Nov 20;32(26):4499-514
87. Quartagno M, Carpenter JR. Multiple imputation for IPD meta-analysis: allowing for heterogeneity and studies with missing covariates. Stat Med. 2016 Jul 30;35(17):2938-54. doi: 10.1002/sim.6837.
88. 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 Feb 14;19(1):33. doi: 10.1186/s12874-019-0676-1.
89. Riley RD, Kauser I, Bland M, Thijs L, Staessen JA, Wang J, et al. Meta-analysis of randomised trials with a continuous outcome according to baseline imbalance and availability of individual participant data. Stat Med. 2013 Jul 20;32(16):2747-66. doi: 10.1002/sim.5726.
90. Groot LJ, Kan KJ, Jak S. Checking the inventory: Illustrating different methods for individual participant data meta-analytic structural equation modeling. Res Synth Methods. 2024 Nov;15(6):872-895. doi: 10.1002/jrsm.1735.
91. Kontopantelis, E., Springate, D. A., Parisi, R., & Reeves, D. (2016). Simulation-Based Power Calculations for Mixed Effects Modeling: ipdpower in Stata. *Journal of Statistical Software*, *74*(12), 1–25. https://doi.org/10.18637/jss.v074.i12
92. 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 Oct 30;41(24):4822-4837. doi: 10.1002/sim.9538.
93. 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 May 18;18(1):41. doi: 10.1186/s12874-018-0492-z.
94. Debray TP, Schuit E, Efthimiou O, Reitsma JB, Ioannidis JP, GetReal Workpackage et al. An overview of methods for network meta-analysis using individual participant data: when do benefits arise? Stat Methods Med Res. 2018 May;27(5):1351-1364. doi: 10.1177/0962280216660741.

95. Singh, J., Gsteiger, S., Wheaton, L. et al. Bayesian network meta-analysis methods for combining individual participant data and aggregate data from single arm trials and randomised controlled trials. BMC Med Res Methodol 22, 186 (2022). https://doi.org/10.1186/s12874-022-01657-y
96. Riley RD, Ensor J, Jackson D, Burke DL. Deriving percentage study weights in multi-parameter meta-analysis models: with application to meta-regression, network meta-analysis and one-stage individual participant data models. Statistical Methods in Medical Research. 2017;27(10):2885-2905. doi:10.1177/0962280216688033.
97. Riley RD, Dias S, Donegan S, Tierney JF, Stewart LA, Efthimiou O, et al. Using individual participant data to improve network meta-analysis projects. BMJ Evid Based Med. 2023 Jun;28(3):197-203. doi: 10.1136/bmjebm-2022-111931.
98. Freeman SC, Fisher D, Tierney JF, Carpenter JR. A framework for identifying treatment-covariate interactions in individual participant data network meta-analysis. Res Synth Methods. 2018 Sep;9(3):393-407. doi: 10.1002/jrsm.1300.
99. Kanters, S., Karim, M.E., Thorlund, K. et al. When does the use of individual patient data in network meta-analysis make a difference? A simulation study. BMC Med Res Methodol 21, 21 (2021). https://doi.org/10.1186/s12874-020-01198-2
100. Liu M, Gao Y, Yang K, Cai Y, Xu J, Dai D, et al. Reporting quality and risk of bias of Cochrane individual participant data meta-analyses: A cross-sectional study. Journal of Evidence-Based Medicine 2023;16(2):141-151. doi: 10.1111/jebm.12521.

Table 1: Literature databases, coverage years, and update information for the search strategy

S/N	Database	Coverage years	Update description
1	Cochrane Library (CDSR)	1996 to 2025	Updated quarterly; last update on June 11, 2025
2	MEDLINE (via Ovid)	1946 to 2025	Updated quarterly; last update on June 11, 2025
3	EMBASE (via Ovid)	1947 to 2025	Updated quarterly; last update on June 11, 2025
4	CINAHL (via EBSCOhost)	1981 to 2025	Initially searched June 9-10, 2024; updated June 11, 2025
5	Scopus	2004 to 2025	Initially searched June 9-10, 2024; updated June 11, 2025
6	Web of Science (Core Collection)	1900 to 2025	Initially searched June 9-10, 2024; updated June 11, 2025
7	Health Technology Assessment Database (HTA)	1996 to 2025	Initially searched June 9-10, 2024; updated June 11, 2025
8	Research Synthesis Methods (manual search)	2010 to 2025	Initially searched June 9-10, 2024; updated June 11, 2025

Table 2: Descriptive characteristics of included resources that reported recommendations for IPD-MA (N = 141)

Category	Characteristic	Number of studies (n,%)
Year of publication	1995 to 1999	2 (1.4)
	2000 to 2004	8 (5.7)
	2005 to 2009	14 (9.9)
	2010 to 2014	26 (18.4)
	2015 to 2019	43 (30.5)
	2020 to 2025	48 (34.1)
Type of study	Simulation	54 (38.3)
	Empirical	29 (20.6)
	Methodological guide	17 (12.1)
	Statistical guide	10 (7.1)
	Systematic Methodology Review	9 (6.4)
	Consensus	8 (5.6)
	Critical/Narrative Review	6 (4.3)
	Handbook chapters	3 (2.1)
	Tutorial	2 (1.4)
	Empirical meta research	2 (1.4)
	Scoping review	1 (0.7)
Research setting	Europe	105 (74.5)
	North America	27 (19.1)
	Asia	6 (4.3)
	Oceanic	3 (2.1)
Type of funding	Public	71 (50.3)

	Not reported	44 (31.2)
	None	8 (5.7)
	Institutional/Internal	8 (5.7)
	Mixed	8 (5.7)
	Non-Governmental Organization (NGO)	1 (0.7)
	Industry	1 (0.7)
Methodological focus	Treatment-covariate interaction	26 (18.4)
	Analytical approach (e.g., one-stage vs two-stage)	20 (14.2)
	Comparing/combining IPD and AD	13 (9.3)
	Other statistical methods (e.g., Network meta-analysis, baseline imbalance, repeated measures, structural equation modelling)	10 (7.1)
	Missing data	10 (7.1)
	General guidance (e.g., timepoints, clustering, practical challenges, planning and conduct)	10 (7.1)
	Data processing	10 (7.1)
	Effect estimation models	9 (6.4)
	Heterogeneity	6 (4.3)
	Identification of subgroups	5 (3.6)
	Special consideration for observational studies	4 (2.8)
	Methodological quality assessment	4 (2.8)
	Standards and best practices	4 (2.8)
	Power calculation	4 (2.8)
	Bias assessment in IPD-MA	3 (2.1)
	Reporting quality	3 (2.1)

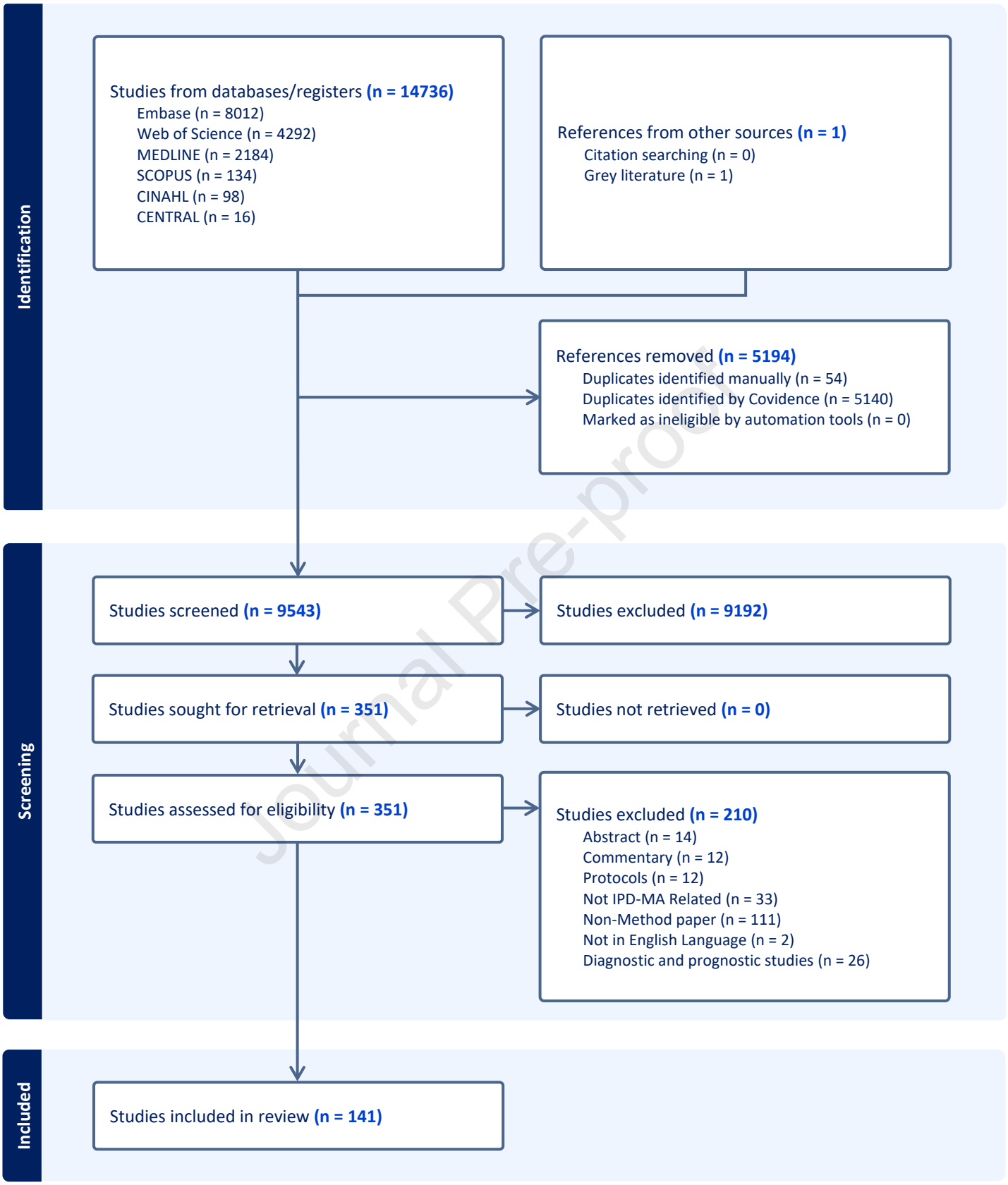
Table 3: Summary recommendation mapped to the four themes

Primary focus	Subcategories	References
Foundations and study design		
Design and implementation guidance	Rationale and scope, feasibility, statistical framework choices (including IPD-NMA), project management tool kits, ethical and collaborative considerations	13, 52, 56, 94, 97, 99, 109, 134
Standards and best practice	Ethical, methodological and practical standards for IPD conduct and workflow (protocols, collaboration, harmonization, reproducibility), analytical modelling standards, reporting and appraising IPD-Meta-analysis	4, 12, 46, 53, 94, 97, 133, 106
Methodological quality assessment	Quality assessments of the systematic review appraisal of design/conduct/reporting elements unique to IPD-MA, overall methodological quality, RCTs and/or non-randomized.	16, 18, 30, 33
Bias assessment and mitigation	Appropriate ROB tools for included studies, Assessment of the impact of excluding participants, incomplete studies or IPD on estimates. (Publication, selection, small study effect, availability bias)	15, 16, 59, 99, 117, 119, 138,
Reporting	Identification of reporting gap, recommendations on what to report and special considerations for observational studies	17, 46, 23, 94, 95, 97, 101
Theme 2: Study Selection, Data Access & Preparation		
Data access and governance	Data sharing policies, governance, ethics, and legal issues, as well as data access pathways (including IPD requests and transfers), agreements and compliance, and retrieval barriers and feasibility.	47, 58, 143, 158
Data processing, quality and integrity checks.	IPD retrieval & validation; provenance or versioning; quality and integrity assessment; harmonization & derivation; and the impact of checking on meta-analytic estimates	12, 51, 56, 57, 59, 60, 61, 62, 97, 102
Theme 3: Statistical Methods and Analysis Strategy		
Methods for estimation, modelling, and analytic	Binary	22, 63, 66, 71, 113, 123, 154
	Continuous	41, 63, 72, 73

approaches for different endpoints		
	Time-to-event	42, 64, 75, 76, 77, 78, 116, 120,
	Ordinal	14, 79
	Count	14, 80
Statistical recommendations for special data presentations	Single case experimental design	81
	Network meta-analysis	94-99, 141, 158, 160, 161, 163
	Comparing and combining IPD and AD	14, 43, 65, 95, 103, 110, 115, 118, 137, 146, 148, 163
	Heterogeneity: Testing and handling in all outcome data types	38, 49, 111, 125, 129, 153, 156
	Subgroup Identification Methods	39, 84, 85, 107, 144
	Treatment-covariate interaction: Estimation of participant-level, study-level interaction, non-linear treatment and covariate interaction and interaction evaluation with single case experimental design (SCED)	20, 21, 45, 48, 54, 83, 85, 105, 114, 127, 128, 131, 135, 139, 142, 145, 146, 147, 149, 150, 155, 162
	Missing data (Sporadic, systematic and mixed)	47, 60, 74, 86, 87, 108, 112, 121, 140, 149, 152
Theme 4: Other Statistical and methodological considerations		
Baseline imbalance	Estimation model, analytical approach and estimation method when Baseline imbalance exists in a Continuous outcome.	89, 126
Structural equation modelling	(SEM): For latent variables (e.g., factor scores, mediation paths, latent means)	90
Power calculation	For binary, continuous and survival endpoints to check the feasibility of interaction	54, 91, 92, 93, 135
Repeated measures	Repeated measures/Longitudinal studies	88, 104
Special consideration for	Causal inference in complex survey design, clinical interpretation checklist, causal-inference	23, 55, 82, 130

the observational study:	framework and reporting standards for observational IPD meta-analyses	

Abbreviations: AD, Aggregate Data; SCED, Single-Case Experimental Design; IPD-NMA, Individual Participant Data Network Meta-analysis; ROB; SEM, Structural Equation Modelling.



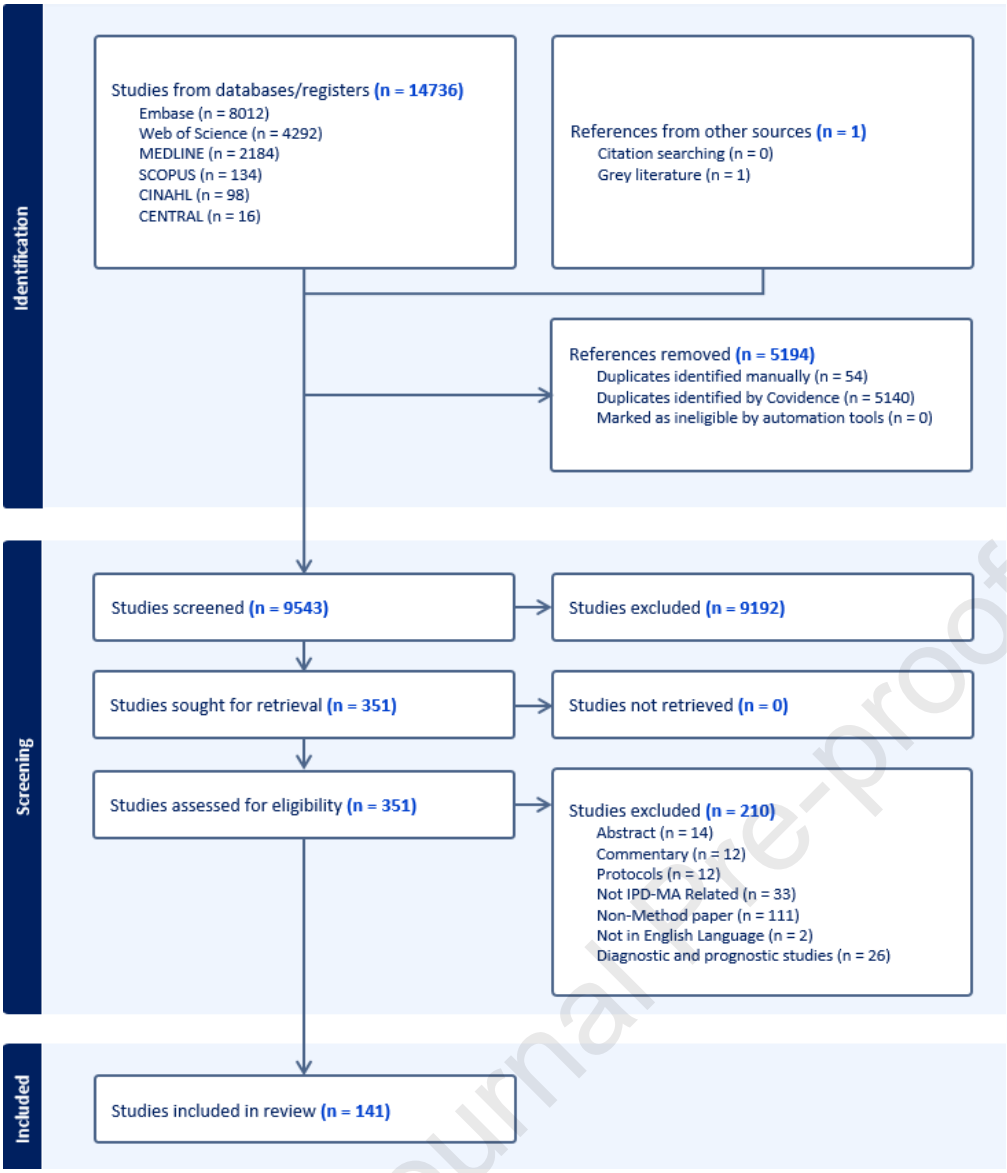
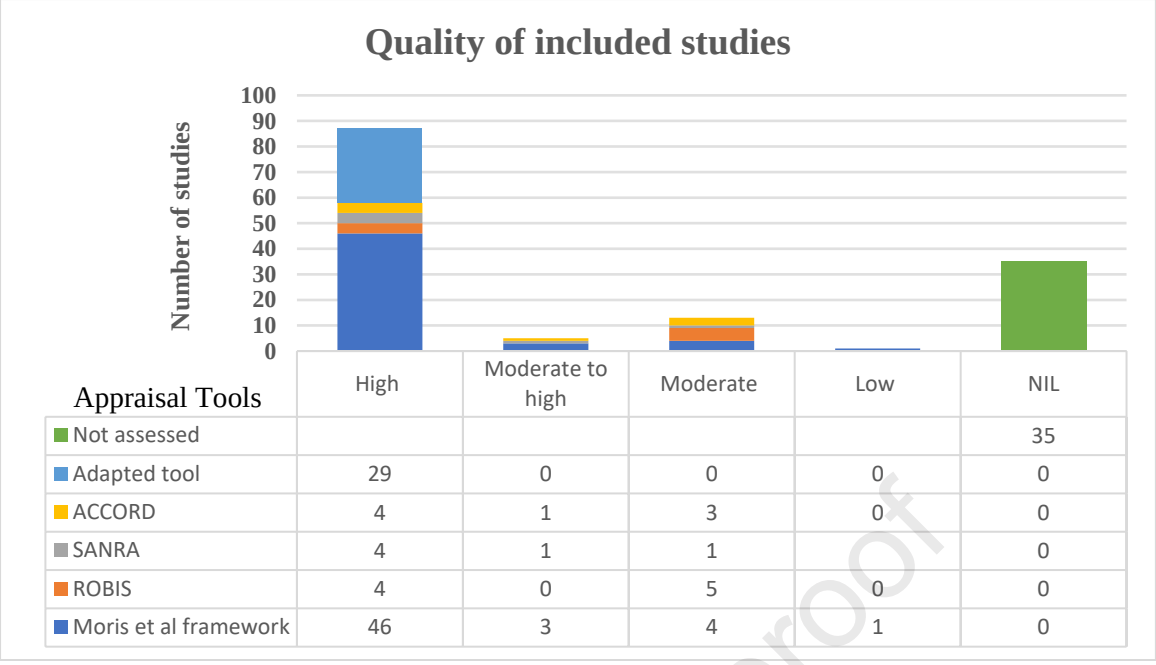


Figure 1: PRISMA flowchart Diagram



Appraisal Tools	High	Moderate to high	Moderate	Low	NIL
Not assessed					35
Adapted tool	29	0	0	0	0
ACCORD	4	1	3	0	0
SANRA	4	1	1		0
ROBIS	4	0	5	0	0
Moris et al framework	46	3	4	1	0

Figure 1: Quality of included studies

Highlights

What is new?

Key findings

- We synthesized methodological guidance for IPD meta-analyses of interventions across the full lifecycle, from planning, data access and preparation, analysis (endpoint-specific), and reporting, by reviewing 141 method papers (1995-2025). We organized guidance into four themes and 25 primary topic areas, yielding a practical set of recommendations and highlighting gaps, especially methods for IPD-MAs based on observational studies.

What this adds to what was known

- Existing guidance was fragmented and often focused on a single technique, such as interaction modelling or endpoint-specific models. Our review provides an end-to-end consolidation that integrates operational procedures with analytic guidance across outcomes and study designs. We also map guidance to conceptual appraisal domains and translate it into candidate items for a consensus-based critical appraisal tool for completed IPD-MAs.

What is the implication, and what should change now?

- The synthesized recommendations provide a consolidated methodological reference that can be used as a best-practice checklist when planning, conducting or reviewing IPD-MAs. Methods work should now target the identified gaps, particularly for IPD-MAs of observational studies to strengthen data management, design-specific analyses, and causal inference.

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: None