GUIDELINES AND GUIDANCE

Individual Participant Data (IPD) Meta-analyses of Randomised Controlled Trials: Guidance on Their Use

Jayne F. Tierney[1]*, **Claire Vale[1]**, **Richard Riley[2]**, **Catrin Tudur Smith[3]**, **Lesley Stewart[4]**, **Mike Clarke[5]**, **Maroeska Rovers[6]**

1 MRC Clinical Trials Unit at University College London, London, United Kingdom, 2 Research Institute of Primary Care and Health Sciences, Keele University, Keele, United Kingdom, 3 MRC North West Hub for Trials Methodology Research, Department of Biostatistics, University of Liverpool, Liverpool, United Kingdom, 4 Centre for Reviews and Dissemination, University of York, York, United Kingdom, 5 Centre for Public Health, Queens University Belfast, Belfast, United Kingdom, 6 Radboud Institute for Health Sciences, Radboud University Medical Center, Nijmegen, the Netherlands

* jayne.tierney@ucl.ac.uk

Summary Points

- Systematic reviews are most commonly based on aggregate data extracted from publications or obtained from trial investigators.

- Systematic reviews involving the central collection and analysis of individual participant data (IPD) usually are larger-scale, international, collaborative projects that can bring about substantial improvements to the quantity and quality of data, give greater scope in the analyses, and provide more detailed and robust results.

- The process of collecting, checking, and analysing IPD is more complex than for aggregate data, and not all IPD meta-analyses are done to the same standard, making it difficult for researchers, clinicians, patients, policy makers, funders, and publishers to judge their quality.

- Following our step-by-step guide will help reviewers and users of IPD meta-analyses to understand them better and recognise those that are well designed and conducted and so help ensure that policy, practice, and research are informed by robust evidence about the effects of interventions.

Background

Systematic reviews provide an objective and reliable way of summarising research evidence. They collate all the studies relevant to a particular research question, using explicit, transparent, and systematic methods in order to minimise bias, and may or may not include a meta-analysis, to combine the results of these studies. Systematic reviews are most commonly based on aggregate data extracted from publications or obtained from investigators [1]. These aggregate data represent a summary of the individual participant or patient data (IPD) for each study and may include, for example, intervention effect estimates (e.g., odds ratios or hazard ratios) for different outcomes and average patient characteristics (e.g., the mean age of

participants or the proportion of women). This limits the analyses that are possible and may also reduce power. Moreover, the availability and quality of such data may vary across studies, and this can affect the reliability of meta-analysis results [2].

Systematic reviews and meta-analyses of IPD have most commonly been large-scale, inter-national, collaborative projects involving the central collection and reanalysis of the original data on each participant, from all of the relevant trials [3–5]. Most have focussed on assessing the efficacy or effectiveness of treatments or other interventions and therefore have been based on randomised controlled trials (RCTs). The IPD approach not only can bring about substantial improvements to the quantity and quality of data—for example, by including more trials, participants, and outcomes—but also enables standardisation of outcomes across trials and detailed data checking (Table 1, [3–5]). IPD also give greater scope and flexibility in the

Table 1. Advantages of using an IPD rather than aggregate data approach to systematic review and meta-analysis of RCTs.

Aspect of Systematic Review/ Meta-analysis	Advantages of the IPD Approach
Trial inclusion	Asking the IPD collaborative group (of trialists and other experts in the clinical field) to supplement the list of identified trials
	Clarify trial eligibility with trial investigators
Data quality	Include trials that are unpublished or not reported in full
	Include unreported data, e.g., more outcomes per trial and more complete information on those outcomes, and data on participants excluded from trials analyses
	Check the integrity of trial IPD and resolve any queries with trial investigators
	Derive standardised outcome definitions across trials or translate different definitions to a common scale
	Derive standardised classifications of participant characteristics or their disease/condition or translate different definitions to a common scale
	Update follow up of time-to-event outcomes beyond that reported
Risk of bias	Clarify trial design, conduct, and analysis methods with trial investigators
	Check risk of bias of trial IPD and obtain extra data when necessary
Analysis	Analyse all important outcomes
	Determine validity of analysis assumptions with IPD, e.g., proportionality of hazards for a Cox model
	Derive measures of effect directly from the IPD
	Use a consistent unit of analysis for each trial
	Apply a consistent method of analysis for each trial
	Conduct more detailed analysis of time-to-event outcomes, e.g., generating Kaplan Meier curves
	Achieve greater power for assessing interactions between effects of interventions and participant or disease/condition characteristics
	Conduct more complex analyses not (usually) possible with aggregate data, e.g., simultaneous assessment of the relationship between multiple trial and/or participant characteristics and effects of interventions
	Use nonstandard models or measures of effect
	Use IPD for secondary clinical questions, e.g., to explore the natural history of disease, prognostic factors, or surrogate outcomes
Interpretation	Discuss the implications for clinical practice and research with a multidisciplinary group of collaborators including trial investigators who supplied the data

doi:10.1371/journal.pmed.1001855.t001

analyses, including, importantly, the ability to investigate whether an intervention is more or less effective for different types of participant (Table 1, [3–5]). These aspects help provide more in-depth exploration and more detailed and robust meta-analysis results, which can differ from those based on aggregate data (e.g., [6–9]). Additionally, collaboration with investigators supplying trial data can lead to more complete identification of relevant trials and a broader interpretation and endorsement of the results (Table 1, [3–5]).

Given their considerable advantages, meta-analyses that are based on IPD have been called the "gold standard" of systematic review [10], and their use to assess the effects of interventions from RCTs has increased [11], across a range of health care areas [5] and in both higher- and lower-resource settings. However, the process of collecting, checking, and analysing IPD is more complex than for aggregate data, and recent evidence suggests that not all IPD meta-analyses are done or reported to the same standard [11,12]. Also, previous guidance, mainly aimed at systematic reviewers, has considered only some of the biases that can arise in aggregate data meta-analyses and how the collection of IPD can help resolve such biases and facilitate understanding (e.g., [3–5]). These issues can make it difficult for researchers, clinicians, patients, policy makers, funders, and publishers to judge the quality of IPD meta-analyses. Therefore, it may hinder their conduct, dissemination, uptake in clinical guidelines [13] policy and practice and their influence on trials [14]. This guidance will help a variety of stakeholders understand, appraise, and make best use of IPD meta-analyses that summarise the efficacy of interventions.

Is It Part of a Systematic Review?

A key component of appraising an IPD review or meta-analysis of efficacy is determining whether it is part of a systematic review [3–5]. Some hallmarks are the following:

- clear research question, qualified by unambiguous eligibility criteria relating to participants, interventions, comparisons, and outcomes

- systematic and comprehensive search strategy, to ensure that all relevant trials are identified

- consistent approach to data collection, including an explanation of which items were requested

- assessment of the quality or risk of bias of included trials

To limit bias, ensure transparency, and allow scrutiny, all such methods should have been prespecified in a protocol, and ideally this would have been registered [15], published, or otherwise made available for inspection.

Early systematic reviews and meta-analyses based on IPD were commonly described as "overviews" or "pooled analyses" (e.g., [16,17]), and subsequently, "IPD meta-analysis" became the preferred term (and is still widely used). However, with the emergence of more ad hoc IPD meta-analyses [12], researchers have begun to utilise "systematic review and meta-analysis of IPD" to flag those in the context of a systematic review (e.g., [18]).

Were All Eligible Trials Identified?

Clear eligibility criteria are needed to determine why particular trials were included in an IPD meta-analysis and, therefore, whether it gives an unbiased or otherwise representative view of the evidence (Table 2). For example, a review of 31 IPD meta-analyses found that 29% had an unclear or unduly selective approach to including studies [12]. Of course, there may be situations in which it is reasonable to exclude relevant trials. For example, for cardiovascular disease, in which trials are often very large, it may make sense to exclude small trials that would

Table 2. Biases that can affect systematic reviews and meta-analyses of IPD and aggregate data (AD), and steps taken to investigate and/or minimise these.

Type of Bias	Definition	Steps That Are Taken to Investigate and Minimise Bias			
			Usual with both AD and IPD approaches	Usual with IPD approach but may be possible with AD approach *	Only with IPD approach
Study selection bias	Systematic differences between results of trials that are and are not selected for inclusion	Prospectively define eligibility criteria	✓		
		Clarify eligibility with trial protocol or trialist		✓	
Publication bias	Systematic differences between results of trials that are and are not published	Include all eligible trials irrespective of publication status		✓	
Data availability bias	Systematic difference between the results of trials for which data were and were not available	Include data for all eligible trials		✓	
		Investigate/discuss the impact of trials for which data were not available		✓	
Participant selection bias	Systematic differences between comparison groups in participant characteristics that can lead to differences in prognosis and/or responsiveness to treatment (Prevented by random allocation and allocation concealment)	Clarify the randomisation methods, i.e., sequence generation and allocation concealment with trial protocol or trialist		✓	
		Exclude "nonrandomised" trials		✓	
		Check for unusual allocation patterns or distributions of participant characteristics			✓
		Exclude trials with inappropriate allocation			✓
		Exclude nonrandomised participants from trial IPD			✓
Performance and detection bias	Systematic differences between comparison groups in the care received or provided or in how outcomes are ascertained (Prevented by blinding study participants, care givers, and outcome assessors to the allocated treatment. Note this is not possible for all interventions, e.g., surgery, and is less important for objective outcomes, e.g., mortality)	Obtain more complete information on blinding and outcome assessment from trialist and/or protocol		✓	
Attrition bias	Systematic differences between comparison groups in the dropout or exclusion of participants (Prevented by the maintenance of all participants in the trial and trial analysis)	Include data on all randomised participants, irrespective of whether they were included in trial analyses		✓	
		Analyse all trials according to the allocated intervention ("intention to treat")		✓	

(Continued)

Table 2. (*Continued*)

Type of Bias	Definition	Steps That Are Taken to Investigate and Minimise Bias		
			Usual with both AD and IPD approaches	Usual with IPD approach but may be possible with AD approach *	Only with IPD approach
		Check for "missing" participants and unusual patterns of dropout or exclusion			✓
		Prespecify any reasonable participant exclusions and apply consistently across trials			✓
Outcome reporting or availability bias	Systematic differences between results of reported/available and unreported/unavailable outcomes (Prevented by making results for all study outcomes available)	Check which outcomes were collected in a trial with protocol and/or trialist		✓	
		Include data for all relevant outcomes		✓	

* Only possible with AD reviews if detailed, high-quality data and/or other information can be obtained from trial protocols, trial reports or directly from trial investigators.

doi:10.1371/journal.pmed.1001855.t002

contribute very little evidence and take considerable effort to obtain. Importantly, the intention to exclude trials on this sort of basis should form part of the prospective eligibility criteria.

Given the substantial evidence that publication of research studies, including RCTs, is influenced by the nature of study results [19], all relevant published and unpublished trials should be sought to avoid bias (Table 2). In 11 IPD meta-analyses in cancer, 45 of the 120 included RCTs were either unpublished or published in non-English language journals, book chapters, or conference abstracts [20]; the so-called "grey literature." Not including these trials would have produced meta-analysis results that were more in favour of the research treatments than results based on all available trials [20]. Most differences were modest, but in one meta-analysis, there was stronger evidence that postoperative radiotherapy was detrimental to the survival of patients with non-small cell lung cancer when all trials (hazard ratio = 1.21, 95% CI = 1.05–1.39, $p = 0.001$) rather than only fully published trials were included (hazard ratio = 1.13, 95% CI = 0.95–1.34, $p = 0.066$). Yet, IPD meta-analyses do not necessarily include trials published in the grey literature [12], and in the absence of an explicit search strategy (e.g., [21]), it can be difficult to judge whether all trials were identified. Thus, IPD meta-analyses with inappropriately restrictive or absent eligibility criteria and/or search strategies that do not seek all relevant trials need to be viewed with caution.

Direct contact with trial investigators, as part of the IPD approach, can help identify trials that may not be easily identifiable by other forms of searching [3,4] and clarify the eligibility of individual trials (Table 1, [3–5]). It should be noted, however, that even with researchers' best efforts, all trials may not come to light, and the potential for bias remains. For example, in a systematic review and meta-analysis in which no unpublished trials were identified [22], funnel plot asymmetry suggested that some small studies might be missing [12].

Were IPD Obtained from Most Trials?

It may be difficult to obtain IPD for all trials identified, because trial investigators do not respond, do not have access to the IPD, or refuse to participate. If nonparticipation is born of a

desire to bury unfavourable results, not including such trials could bias a meta-analysis, whereas if it is to avoid providing trials of poor quality, then not including these trials might make a meta-analysis more robust. Aiming to obtain a large proportion of the eligible trials and participants will both counter bias and enable exploration of any quality issues and so will likely provide the most reliable and precise assessment of the effects of an intervention. Upwards of 90% of eligible participants has been suggested as a suitable target to achieve [3], but 33% of a sample of IPD meta-analyses included less than 80% of the data requested [12]. Of course, if the available data can be shown to provide sufficient power to detect an effect reliably, then a lower threshold may be acceptable. With substantial data missing, there is no guarantee that the meta-analysis results will be representative of those based on all data, making data availability bias a possibility. This should be investigated (Table 2), preferably by comparing or combining results based on IPD with those based on aggregate data, to check for coherence and consistency [23]. For example, an IPD meta-analysis in non-Hodgkin lymphoma suggested no clear improvement in survival with high-dose compared to conventional-dose chemotherapy (hazard ratio = 1.14, 95% CI = 0.98–1.34, p = 0.401) [24]. Results based on aggregate data were obtained for four of the five trials with missing IPD, and adding these gave greater confidence in the conclusion that high-dose chemotherapy does not affect survival (hazard ratio = 1.05, 95% CI = 0.92–1.19, p = 0.157) [24]. Funnel plots may also help indicate the potential impact of both known and unknown missing trials on the results [25] but would not be appropriate in a meta-analysis including less than ten trials [25] or few trials with statistically significant results or one in which heterogeneity has been observed [26]. As a bare minimum, the proportion of trials and participants for which IPD are available and the reasons for unavailability of data should be a standard part of IPD meta-analysis reporting [27].

Was the Integrity of the IPD Checked?

While there is little opportunity to check the accuracy of aggregate data, IPD can be checked for missing, invalid, out-of-range, or inconsistent items and for discrepancies with any trial publication (Table 1, [3–5]). This would highlight, for example, a participant having an unusually old age or a date of randomisation prior to the date the trial started or instances in which sequentially numbered participants were absent from a dataset. Although similar checks will have been applied during trial conduct, occasionally anomalies still arise. Ideally, these should be queried and, where possible, resolved with the trial's personnel, to improve data quality and ensure that trials are represented accurately.

Were the Analyses Prespecified in Detail?

As IPD meta-analyses offer the potential for substantially greater numbers of analyses, there is a greater risk that data might be interrogated repeatedly until the desired results are obtained, these being the results that are published. Therefore, it is important that the detailed methods of analysis are prespecified in a protocol or analysis plan and are clearly reported upon publication, including the following:

- primary and secondary outcomes and their definitions

- methods for analysis of efficacy/effectiveness, including those for exploring the impact of trial or participant characteristics

- methods for quantifying and accounting for heterogeneity

- methods for checking IPD and assessing the risk of bias of trials

This is not to say that unplanned analyses are unjustified or invalid; rather, they can play an important role in explaining or adding to per-protocol results. However, such exploratory analyses do need to be justified and clearly labelled as such in any report of the results.

Was the Risk of Bias of Included Trials Assessed?

Checking the quality of included trials is now a key component of systematic reviews, so that any issues found can be either rectified or taken into account when interpreting the results, and well-conducted IPD meta-analyses are no exception. For aggregate data reviews, this is frequently achieved using the risk of bias tool developed by Cochrane [28,29], based on information provided in trial reports. Seeking additional information from trialists can improve risk of bias assessments associated with, for example, the randomisation procedures or the completeness of outcome data [30,31]. Requesting such clarification tends to be more a feature of IPD rather than aggregate data reviews. Furthermore, checking the IPD directly can provide more reliable investigations of key potential biases, some of which might be reduced or alleviated in the process (Table 2, [32]). As not all IPD meta-analyses include such assessments or may not describe them explicitly, these are outlined below.

Were Randomisation, Allocation Concealment, and Blinding Assessed?

Checking that the methods of randomisation and allocation concealment are appropriate is vital in ensuring that a trial provides a fair comparison of interventions (Table 2 [28]). The IPD can be used, for example, to check whether participant characteristics are balanced by arm, as would be expected, or to examine how participants were allocated to intervention arms (Table 2). For example, in a trial in multiple myeloma, similar numbers of patients were allocated to each arm until 1985, when, for a period of about six months, all patients were allocated to the chemotherapy arm [3]. The investigators explained that this was because the radiotherapy equipment was unavailable. Having IPD, it was possible to exclude this small group of non-randomised patients, ensuring an unbiased analysis [3]. IPD meta-analyses employing such checks safeguard against the inclusion of both nonrandomised trials and nonrandomised participants.

The contact with trial staff that is intrinsic to collaborative IPD meta-analyses can provide useful clarification of, for example, how participants and other study personnel were blinded, allowing the risk of differential care or outcome assessments to be judged with more certainty (Table 2). However, IPD cannot reduce bias or otherwise alter this aspect of trial conduct.

Were the IPD Checked to Ensure All Randomised Participants Were Included?

Even with adequate randomisation methods, an unbiased comparison of intervention groups can only be guaranteed if all randomised participants are analysed according to the interventions initially assigned, i.e., using an intention-to-treat approach [33–35]. If patients drop out or are excluded from the analysis of a trial in considerable numbers and/or disproportionately by arm, this can cause attrition bias (Table 2, [36]), particularly if the reasons are related to the intervention and outcome. For example, in 14 IPD cancer meta-analyses incorporating 133 trials, between 0.3% and 38% of randomised patients were excluded from the original analyses of 69% of trials, often with more being excluded from the research arms [37]. Around 1,800 excluded patients were reinstated when IPD were collected, and without them the meta-analyses results would have been biased towards the research treatments, albeit to a small degree in most cases [37]. However, for one meta-analysis in soft tissue sarcoma [38], the evidence for a benefit of chemotherapy on survival was stronger when based only on those patients included

in the original trial analyses (hazard ratio = 0.85, 95% CI = 0.72–1.00, p = 0.06), compared to when all possible patients were included (hazard ratio = 0.90, 95% CI = 0.77–1.04, p = 0.16) [37], which could have affected the clinical interpretation. Thus, to help minimise and potentially eliminate attrition bias, IPD should be checked to ensure that data on all or as many randomised participants as possible are included. If there is good reason to exclude certain participants, then this should be prespecified as part of the meta-analysis eligibility criteria and such exclusions applied consistently across trials, to avoid bias [37]. Taking this approach of making consistent (clinically plausible) exclusions across trials in the sarcoma meta-analysis produced very similar results (hazard ratio = 0.91, CI = 0.78–1.07, p = 0.28) to those based on all patients.

Were All Relevant Outcomes Included?

Outcome reporting bias arises if the effect of an intervention varies by outcome, and only particular outcomes are reported, for example, those showing benefit (Table 2). Such bias can be compounded in systematic reviews of published aggregate data [39], because by definition they can only include reported outcomes. Obtaining additional outcome data, as part of the IPD collection process, can resolve this problem, as demonstrated in a systematic review of laparoscopic versus open surgery for the repair of inguinal hernia. Based on published aggregate data from the three trials that reported it, the risk of persistent pain was found to be significantly greater with laparoscopic repair (odds ratio = 2.03, 95% CI = 1.03–4.01) [40]. Yet, when IPD from a further 17 trials (that did not publish results for this outcome) were added, the combined results showed that the risk of persistent pain was actually less with laparoscopic repair (odds ratio = 0.54, 95% CI = 0.46–0.64, p < 0.001) [40]. Moreover, the much narrower confidence intervals show that these results are far more reliable. Therefore, to avoid bias and provide a balanced view of benefits and harms, IPD meta-analyses should include data on all available key outcomes, bearing in mind that there may be some that have been collected but not reported in the original trial publication.

Was the Quality of Time-to-Event Outcome Data Checked?

For time-to-event outcomes, such as survival or time to resolution of symptoms, bias can occur if participants are observed more frequently or for a longer duration on one intervention arm compared to another. This is because it can make it appear that event rates are higher in that arm (Table 2). Even if follow up is similar by arm, trials that stop early or have follow-up that is too short might produce results that are transitory or fail to pick up long-term effects. For example, in an IPD meta-analysis of chemotherapy in addition to local treatment for soft tissue sarcoma [38], median follow up for survival was reported for seven of the included trials and ranged from 16 to 64 months. Investigators supplying IPD were able to extend the median follow up for these trials to between 74 and 204 months [32], which allowed a more reliable examination of the effects of chemotherapy in the longer term [38]. Thus, with IPD, both the pattern and extent of follow up should be checked, and where relevant, additional follow up obtained to rectify any imbalances and provide more up-to-date and reliable outcome data (Table 2). Of course, this may not be practical or necessary if most or all events have already taken place.

Were the Methods of Analysis Appropriate?

The outcomes and effect measures of interest in an IPD meta-analysis are often the same as those used in a meta-analysis of aggregate data, such as relative risks, odds ratios, or risk differences for dichotomous outcomes; mean differences for continuous outcomes; and hazard ratios

for time-to-event outcomes. However, having IPD can allow harmonisation of outcome definitions across trials, analytical assumptions to be checked, and effect measures to be derived directly from the analysis of the collated data (Table 1), with no need to rely on the original trial analyses. Alternative effect measures can be computed and synthesised, if desired. As more complex, nonstandard methods may be used for meta-analyses based on IPD, we outline some key principles to look for, noting that analyses to investigate publication and data availability bias have already been discussed.

Were the Methods of Assessing the Overall Effects of Interventions Appropriate?

As participants in an IPD meta-analysis are recruited according to different trial protocols, combining IPD across trials, as though it comes from a single "mega" trial, could lead to biased comparisons of interventions and overprecise estimates of effect (Table 2). For example, the effect of nicotine gum on smoking cessation was shown to be severely attenuated and too precise when the analysis treated the IPD as though it belonged to one trial (odds ratio = 1.40, 95% CI = 1.02–1.92) rather than multiple trials (odds ratio = 1.80, 95% CI 1.29–2.52) [41]. Thus, a key component of the analysis is to stratify or account for clustering of participants [41]. To date, two-stage meta-analysis of IPD is the most common way of combining data while preserving participants' trial membership [42,43]. In the first stage, estimates of effect are derived from the IPD for each trial, and in the second stage, these are combined using methods analogous to most meta-analyses of aggregate data. Thus, standard fixed-effect (e.g., Peto, Mantel-Haenszel, and inverse-variance [44,45]) and random effects methods (e.g., Dersimonian and Laird [46]) can be utilised to derive estimates of the intervention effect, with the latter helping to account for any differences in effect across trials (heterogeneity), and standard statistics (e.g., study weights and I^2) and forest plots can be produced.

Alternatively, a one-stage model [42] can be used to estimate intervention effects while stratifying or otherwise accounting for differences between trials [47–50]. This is typically a regression model such as linear, logistic, or Cox regression, with either a separate term for each trial or one that varies across trials via a random effect. Based on theoretical considerations [51–53], simulation studies [43,54], and case studies in epilepsy [54], cancer [43], maternal health [55], and childhood infections [56], one-stage and two-stage methods usually produce similar meta-analysis results. For example, the effects of anti-platelets on pre-eclampsia in pregnancy from a two-stage (relative risk = 0.90, 95% CI = 0.83–0.96) and one-stage model (relative risk = 0.90, 95% CI = 0.83–0.97) were almost identical [55]. However, when included studies are small and/or when an effect is large [43,54] or events are rare [57], there can be some bias in a two-stage analysis, because some of the assumptions in the second stage are potentially inappropriate. Thus, for many, the one-stage approach is preferable, although it makes it more difficult to calculate the standard meta-analysis statistics directly, and the complexity can make interpretation of the results less accessible to those who are used to the more familiar two-stage approach. Hence, researchers often will do a one-stage analysis to obtain estimates of effect and use a two-stage analysis to obtain further statistics and forest plots. It is important, therefore, that the choice of one or two-stage analysis is specified in advance or that results for both approaches are reported [55].

Were the Methods of Assessing If Effects of Interventions Varied by Trial or Participant Characteristics Appropriate?

Exploring whether intervention effects vary or are modified by trial characteristics, such as the duration of the intervention or dose of drugs, and the nature of these interactions is another

important aspect of meta-analysis. For both aggregate data and IPD, it is reasonable to investigate these sorts of trial-level interactions using subgroup analyses, whereby intervention effects are compared between groups of trials, or meta-regression in which the change in overall intervention effect in relation to trial characteristics is investigated [58].

A frequent motivation, however, for collecting IPD is the ability to investigate interactions between overall intervention effects and individual participant-level characteristics, such as age or stage of disease or risk of an outcome. This is important for clinical decision-making and policy, in which the aim is to target interventions at those who are most likely to benefit or are least likely to be harmed by them. Single trials are typically underpowered to detect such interactions, and meta-regression of aggregate data is usually inadequate for the purpose, because it also has low power and only permits trial-level interactions between overall intervention effects and summaries of participant characteristics, such as mean age. The latter may not reflect the genuine relationships between intervention effects and individual participant characteristics (Table 2, [59]). Similarly, comparing a two-stage IPD meta-analysis of intervention effects in one participant group with another, for example, effects in men and women, using a test for interaction (sometimes called subgroup analyses) is common in IPD meta-analysis but prone to bias and best avoided [60]. Rather, the focus should be on estimating the individual participant-level interactions, which is usually only possible with IPD. Again, this can be achieved in two stages, by estimating the interactions separately for each trial in the first stage and then pooling these across trials in the second stage, using meta-analysis techniques [60], For example, using this methodology, there was no clear evidence of a relationship between lymph node involvement and the effect of postoperative radiotherapy on survival in non-small cell lung cancer and (interaction hazard ratio = 0.91, 95% CI = 0.74–1.11, p = 0.34) [60]. In contrast, using the flawed subgroup analysis approach, the effect of postoperative radiotherapy on survival was erroneously found to vary with the number of lymph nodes affected (chi-square test for trend, df = 1, p = 0.016) [61].

Alternatively, a one-stage approach that incorporates an intervention by participant covariate interaction term in a regression model while also accounting for clustering of patients within trials (as described above) can be used. This must carefully separate out this individual participant-level interaction from any trial-level interactions [60,62]. Uniquely, the one-stage approach allows the relative influence of multiple trial and patient characteristics on any intervention effect to be considered simultaneously, which can help deal with confounding [63] and may provide greater clinical insight. When there is no particular evidence of a differential effect by trial or participant characteristic, emphasis should be placed on the overall result.

To allow appropriate interpretation, results of all such analyses should be reported in full and should include effect sizes and confidence intervals for each meta-analysis, measures of statistical heterogeneity, and interactions between intervention effects and trial or participant characteristics [27].

Discussion

We have provided guidance to help those using the results of IPD meta-analyses to inform policy, practice, or research, and those reviewing grant applications or manuscripts relating to IPD meta-analyses, to assess critically their quality (summarised in Box 1). Expert opinion may be required to navigate the more complicated statistical methods. Given that collecting IPD can be a lengthy and costly endeavour, additionally funders might want to ask researchers to explain what value IPD will bring over aggregate data and whether applicants have established the feasibility of the IPD approach, for example, by contacting trial investigators to see whether they are willing and able to provide data [64]. Realistically, many IPD meta-analyses will not be

Box 1. Key Questions to Ask When Appraising an IPD Meta-analysis

1. Is the IPD meta-analysis part of a systematic review?

   a Does it have a clear research question qualified by explicit eligibility criteria?

   b Does it have a systematic and comprehensive search strategy for identifying trials?

   c Does it have a consistent approach to data collection?

   d Does it assess the "quality" or risk of bias of included trials?

   e Are all the methods prespecified in a protocol?

   f Has the protocol been registered or otherwise made available?

2. Were all eligible trials identified?

   a Were fully published trials identified?

   b Were trials published in the grey literature identified?

   c Were unpublished trials identified?

3. Were IPD obtained for most trials?

   a Were IPD obtained for a large proportion of the eligible trials?

   b Was an assessment of the potential impact of missing trials undertaken?

   c Were the reasons for not obtaining IPD provided?

4. Was the integrity of the IPD checked?

   a Were the data checked for missing, invalid out-of-range, or inconsistent items?

   b Were there any discrepancies with the trial report (if available)?

   c Were any issues queried and, if possible, resolved?

5. Were the analyses prespecified in detail?

   a Were the detailed analysis methods included in a protocol or analysis plan?

   b Were the outcomes and methods for analysing the effects of interventions, quantifying and accounting for heterogeneity, and assessing risk of bias included?

6. Was the risk of bias of included trials assessed?

   a Were the randomisation, allocation concealment, and blinding assessed?

   b Were the IPD checked to ensure all (or most) randomised participants were included?

   c Were all relevant outcomes included?

   d Was the quality of time-to-event-outcome data checked?

7. Were the methods of analysis appropriate?

   a Were the methods of assessing the overall effects of interventions appropriate?

    i. Did researchers stratify or account for clustering of participants within trials using either a one- or two-stage approach to meta-analysis?

    ii. Was the choice of one- or two-stage analysis specified in advance and/or results for both approaches provided?

b Were the methods of assessing whether effects of interventions varied by trial characteristics appropriate?

    i. Did researchers compare treatment effects between subgroups of trials or use meta-regression to assess whether the overall treatment effect varied in relation to trial characteristics?

c Were the methods of assessing whether effects of interventions vary by participant characteristics appropriate?

    i. Did researchers estimate an interaction separately for each trial and combine these across trials in a two-stage fixed effect or random effects meta-analysis? Or;

    ii. Did researchers incorporate one or more a treatment by participant covariate interaction terms in a regression model, whilst also accounting for clustering of participant, separating out this individual participant-level interaction from any trial-level interactions?

d If there was no evidence of a differential effect by trial or participant characteristic, was emphasis placed on the overall result?

e Were exploratory analyses highlighted as such?

8. Does any report of the results adhere to the Preferred Reporting Items for a Systematic review and Meta-analysis of IPD (The PRISMA-IPD Statement) [27]?

perfect in all respects, but if they are conducted in the context of a systematic review, include a high-proportion of good-quality data and use appropriate analyses, they will most likely provide reliable assessments of the effects of interventions.

With the expected increase in the use of IPD in systematic reviews, as a consequence of the drive for greater sharing of data held within trials, guidance on appropriate methodology will become ever more valuable. Many of the advantages and pitfalls we describe in relation to standard IPD meta-analyses, are transferrable to network IPD meta-analysis of efficacy, and elements of this guidance relating to trial identification, data quality and risk of bias will frequently apply. However, additional research and guidance on appropriate analytical approaches is needed. IPD is also being collected to improve the quality of systematic reviews of diagnostic and prognostic studies [65–67], and to allow necessarily complex, multi-parameter, meta-analysis methods [68] in these reviews. However, such methodology is less well established, and will differ in additional respects. Hence, more specific guidance is needed to facilitate appraisal of these sorts of IPD meta-analyses.

Increasingly, planned methods for IPD meta-analyses can be found in protocols that are registered in the PROSPERO International Prospective Register of Systematic Reviews [69],

published or made available by other means. This practice needs to be encouraged, as there is empirical evidence that reporting of search strategies, data availability [12,42] and methods of analysis [42] is variable, and anecdotal evidence that the methods for checking data and risk of bias are frequently absent from IPD meta-analyses publications. Although missing information can be sought from researchers, to enable discerning appraisal, clearly reporting should also be improved. Researchers conducting PD meta-analyses often target high impact print journals [70], and restrictions on the number of words, figures and tables, may have led them to condense or omit particular methods or results. They can now utilise online journals, and the increasing opportunities offered by print journals to publish material online, to alleviate these issues. Journals can help by being responsive to the need for full reporting of IPD meta-analyses. Moreover, those conducting IPD meta-analyses should follow the PRISMA reporting guidelines developed specifically for the IPD approach [27].

Taking steps to identify well designed and conducted IPD meta-analyses of efficacy will help ensure that clinical practice and research is informed by robust evidence about the effect of interventions.

Acknowledgments

Author Contributions

Contributed to the writing of the manuscript: JFT CV RR CTS LS MC MR. Wrote the first draft of the manuscript: JFT CV RR CTS LS MC MR. Agree with the manuscript's results and conclusions: JFT CV RR CTS LS MC MR. All authors have read, and confirm that they meet, ICMJE criteria for authorship.

References

1. Higgins JPT, Green SJ, editors. Cochrane Handbook for Systematic Reviews of Interventions. Chichester: John Wiley & Sons Ltd; 2008.

2. Sterne JAC, Egger M, Moher D, on behalf of the Cochrane Bias Methods Group. Addressing reporting biases. In: Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions. Chichester: Wiley & Sons Ltd; 2008. p. 297–333.

3. Stewart LA, Clarke MJ, on behalf of the Cochrane Working Party Group on Meta-analysis using Individual Patient Data. Practical methodology of meta-analyses (overviews) using updated individual patient data. Stat Med. 1995; 14:2057–79. PMID: 8552887

4. Stewart LA, Tierney JF. To IPD or Not to IPD? Advantages and disadvantages of systematic reviews using individual patient data. Eval Health Prof. 2002; 25(1):76–97. PMID: 11868447

5. Stewart LA, Tierney JF, Clarke M, on behalf of the Cochrane Individual Patient Data Meta-analysis Methods Group. Reviews of individual patient data. In: Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions. Chichester: Wiley & Sons Ltd; 2008. p. 547–58.

6. Duchateau L, Pignon J-P, Bijnens L, Bertin S, Bourhis J, Sylvester R. Individual patient-versus literature-based meta-analysis of survival data: Time to event and event rate at a particular time can make a difference, an example based on head and neck cancer. Control Clin Trials. 2001; 22(5):538–47. PMID: 11578787

7. Stewart LA, Parmar MKB. Meta-analysis of the literature or of individual patient data: is there a difference? Lancet. 1993; 341:418–22. PMID: 8094183

8. Jeng GT, Scott JR, Burmeister LF. A comparison of meta-analytic results using literature versus individual patient data: paternal cell immunization for recurrent miscarriage. Journal of the American Medical Association. 1995; 274:840–36.

9. Collaboration EHT. Value of updating a systematic review in surgery using individual patient data. Br J Surg. 2004; 91:495–9. PMID: 15048755

10. Chalmers I. The Cochrane Collaboration: preparing, maintaining and disseminating systematic reviews of the effects of health care. Annals of New York Academy of Sciences. 1993; 703:156–65.

11. Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. BMJ. 2010; 340:c221. doi: 10.1136/bmj.c221 PMID: 20139215

12. 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:d7762. doi: 10.1136/bmj.d7762 PMID: 22214758

13. Vale CL, Rydzewska LHM, Rovers MM, Emberson JR, Gueyffier F, Stewart LA. Uptake of systematic reviews and meta-analyses based on individual participant data in clinical practice guidelines: descriptive study. BMJ. 2015; 350;h1088. doi: 10.1136/bmj.h1088 PMID: 25747860

14. Tierney JF, Pignon J-P, Gueffyier F, Clarke M, Askie L, Vale CL, et al. How individual participant data meta-analyses can influence trial design and conduct J Clin Epidemiol. 2015; In press.

15. Booth A, Clarke M, Ghersi D, Moher D, Petticrew M, Stewart L. An international registry of systematic-review protocols. Lancet. 2011; 377(9760):108–9. doi: 10.1016/S0140-6736(10)60903-8 PMID: 20630580

16. Advanced Ovarian Cancer Trialists Group. Chemotherapy in advanced ovarian cancer: an overview of randomised clinical trials. BMJ. 1991; 303:884–93. PMID: 1834291

17. Early Breast Cancer Trialists' Collaborative Group. Tamoxifen for early breast cancer: an overview of the randomised trials. Lancet. 1998; 351:1451–67. PMID: 9605801

18. Chemoradiotherapy for Cervical Cancer Meta-analysis Collaboration Reducing uncertainties about the effects of chemoradiotherapy for cervical cancer: a systematic review and meta-analysis of individual patient data from 18 randomized trials. J Clin Oncol. 2008; 26(35):5802–12. doi: 10.1200/JCO.2008.16.4368 PMID: 19001332

19. Dickersin K. Publication bias: recognising the problem, understanding its origins and scope, and preventing harm. In: Rothstein H, Sutton A, Borenstein M, editors. Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments. Chichester: John Wiley & Sons Ltd; 2005. p. 261–86.

20. Burdett S, Stewart LA, Tierney JF. Publication bias and meta-analysis (A practical example). Int J Technol Assess Health Care. 2003; 19(1):129–34. PMID: 12701945

21. Sakamoto J, Hamada C, Yoshida S, Kodaira S, Yasutomi M, Kato T, et al. An individual patient data meta-analysis of adjuvant therapy with uracil-tegafur (UFT) in patients with curatively resected rectal cancer. British Journal of Cancer. 2007; 96(8):1170–7. PMID: 17375049

22. DeLuca G, Gibson CM, Bellandi F, Murphy S, Maioli M, Noc M, et al. Early glycoprotein IIb-IIIa inhibitors in primary angioplasty (EGYPT) cooperation: An individual patient data meta-analysis. Heart. 2008; 94(12):1548–58. doi: 10.1136/hrt.2008.141648 PMID: 18474534

23. 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; 60(5):431–9. PMID: 17419953

24. Greb A, Bohlius J, Schiefer D, Schwarzer G, Schulz H, Engert A. High-dose chemotherapy with autologous stem cell transplantation in the first line treatment of aggressive non-Hodgkin lymphoma (NHL) in adults. Cochrane Database Syst Rev. 2008(1:):CD004024. doi: 10.1002/14651858.CD004024.pub2 PMID: 18254036

25. Sterne JA, Sutton AJ, Ioannidis JP, Terrin N, Jones DR, Lau J, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. BMJ. 2011; 343:d4002. doi: 10.1136/bmj.d4002 PMID: 21784880

26. Ioannidis JP, Trikalinos TA. The appropriateness of asymmetry tests for publication bias in meta-analyses: a large survey. CMAJ. 2007; 176(8):1091–6. PMID: 17420491

27. Stewart LA, Clarke M, Rovers M, Riley RD, Simmonds M, Stewart G, et al. Preferred Reporting Items for a Systematic Review and Meta-analysis of Individual Participant Data: The PRISMA-IPD Statement. JAMA. 2015; 313(16):1657–65. doi: 10.1001/jama.2015.3656 PMID: 25919529

28. Higgins JPT, Altman DG, on behalf of the Cochrane Statistical Methods Group and the Cochrane Bias Methods Group. Assessing risk of bias in included studies. In: Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions. Chichester: John Wiley & Sons Ltd; 2008.

29. Higgins JP, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ. 2011; 343:d5928. doi: 10.1136/bmj.d5928 PMID: 22008217

30. Mhaskar R, Djulbegovic B, Magazin A, Soares HP, Kumar A. Published methodological quality of randomized controlled trials does not reflect the actual quality assessed in protocols. J Clin Epidemiol. 2012; 65(6):602–9. doi: 10.1016/j.jclinepi.2011.10.016 PMID: 22424985

31. Vale CL, Tierney JF, Burdett S. Can trial quality be reliably assessed from published reports of cancer trials: evaluation of risk of bias assessments in systematic reviews. BMJ. 2013; 346:f1798. doi: 10.1136/bmj.f1798 PMID: 23610376

32. Stewart L, Tierney J, Burdett S. Do systematic reviews based on individual patient data offer a means of circumventing biases associated with trial publications? In: Rothstein H, Sutton A, Borenstein M, editors. Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments. Chichester: John Wiley & Sons; 2005. p. 261–86.

33. Altman DG. Randomisation. BMJ. 1991; 302:1481–2. PMID: 1855013

34. Lachin JM. Statistical considerations in the intent-to-treat principle. Control Clin Trials. 2000; 21:167–89. PMID: 10822117

35. Schulz K, F., Grimes DA. Sample size slippages in randomised trials: exclusions and the lost and wayward. Lancet. 2002; 359:781–5. PMID: 11888606

36. Juni P, Altman DG, Egger M. Systematic reviews in health care: Assessing the quality of controlled clinical trials. BMJ. 2001; 323(7303):42–6. PMID: 11440947

37. Tierney JF, Stewart LA. Investigating patient exclusion bias in meta-analysis. Int J Epidemiol. 2005; 34:79–87. PMID: 15561753

38. Sarcoma Meta-analysis Collaboration. Adjuvant chemotherapy for localised resectable soft tissue sarcoma in adults: meta-analysis of individual patient data. Lancet. 1997; 350:1647–54. PMID: 9400508

39. Kirkham JJ, Dwan KM, Altman DG, Gamble C, Dodd S, Smyth R, et al. The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. BMJ. 2010; 340:c365. doi: 10.1136/bmj.c365 PMID: 20156912

40. McCormack K, Grant A, Scott N. Value of updating a systematic review in surgery using individual patient data. Br J Surg. 2004; 91(4):495–9. PMID: 15048755

41. Abo-Zaid G, Guo B, Deeks JJ, Debray TP, Steyerberg EW, Moons KG, et al. Individual participant data meta-analyses should not ignore clustering. J Clin Epidemiol. 2013; 66(8):865–73 e4. doi: 10.1016/j.jclinepi.2012.12.017 PMID: 23651765

42. Simmonds MC, Higgins JPT, Stewart LA, Tierney JF, Clarke MJ, Thompson SG. Meta-analysis of individual patient data from randomised trials—a review of methods used in practice. Clinical Trials. 2005; 2(3):209–17. PMID: 16279144

43. Bowden J, Tierney JF, Simmonds M, Copas AJ. 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. Research synthesis methods. 2011; 2(3):150–62. doi: 10.1002/jrsm.45 PMID: 26061783

44. Yusuf S, Peto R, J.A. L, Collins R, Sleight P. Beta blockade during and after myocardial infarction: an overview of the randomized trials. Prog Cardiovasc Dis. 1985; 27:335–71. PMID: 2858114

45. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. J Natl Cancer Inst. 1959; 22(4):719–48. PMID: 13655060

46. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986; 7:177–88. PMID: 3802833

47. Turner RM, Omar RZ, Yang M, Goldstein H, Thompson SG. A multilevel model framework for meta-analysis of clinical trials with binary outcomes. Stat Med. 2000; 19:3417–32. PMID: 11122505

48. Higgins JPT, Whitehead A, Turner RM, Omar RZ, Thompson SG. Meta-analysis of continuous outcome data from individual patients. Stat Med. 2001; 20:2219–41. PMID: 11468761

49. Whitehead A, Omar RZ, Higgins JPT, Savaluny E, Turner RM, Thompson SG. Meta-analysis of ordinal outcomes using individual patient data. Stat Med. 2001; 20:2243–60. PMID: 11468762

50. Tudur Smith C, Williamson PR, Marson AG. Investigating heterogeneity in an individual patient data meta-analysis of time to event outcomes. Stat Med. 2005; 24:1307–19. PMID: 15685717

51. Olkin I, Sampson A. Comparison of meta-analysis versus analysis of variance of individual patient data. Biometrics. 1998; 54:317–22. PMID: 9544524

52. Mathew T, Nordstrom K. On the equivalence of meta-analysis using literature and using individual patient data. Biometrics. 1999; 55(4):1221–3. PMID: 11315071

53. Mathew T, Nordstrom K. Comparison of one-step and two-step meta-analysis models using individual patient data. Biometrical journal Biometrische Zeitschrift. 2010; 52(2):271–87. doi: 10.1002/bimj.200900143 PMID: 20349448

54. 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. PMID: 18042571

55. Stewart GB, Altman DG, Askie LM, Duley L, Simmonds MC, Stewart LA. Statistical analysis of individual participant data meta-analyses: a comparison of methods and recommendations for practice. PLoS ONE. 2012; 7(10):e46042. doi: 10.1371/journal.pone.0046042 PMID: 23056232

56. Koopman L, van der Heijden GJ, Hoes AW, Grobbee DE, Rovers MM. Empirical comparison of subgroup effects in conventional and individual patient data meta-analyses. Int J Technol Assess Health Care. 2008; 24(3):358–61. doi: 10.1017/S0266462308080471 PMID: 18601805

57. Stijnen T, Hamza TH, Ozdemir P. Random effects meta-analysis of event outcome in the framework of the generalized linear mixed model with applications in sparse data. Stat Med. 2010; 29(29):3046–67. doi: 10.1002/sim.4040 PMID: 20827667

58. Deeks JJ, Higgins JPT, Altman DG, on behalf of the Cochrane Statistical Methods Group. Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions2008. p. 243–96.

59. Berlin JA, Santanna J, Schmid CH, Szczech LA, Feldman HI, Anti-Lymphocyte Antibody Induction Therapy Study G. Individual patient- versus group-level data meta-regressions for the investigation of treatment effect modifiers: ecological bias rears its ugly head. Stat Med. 2002; 21(3):371–87. PMID: 11813224

60. Fisher DJ, Copas AJ, Tierney JF, Parmar MKB. A critical review of methods for the assessment of patient-level interactions in individual patient data (IPD) meta-analysis of randomised trials, and guidance for practitioners. J Clin Epidemiol. 2011; 64:949–67. doi: 10.1016/j.jclinepi.2010.11.016 PMID: 21411280

61. PORT Meta-analysis Trialists Group. Postoperative radiotherapy in non-small-cell lung cancer: systematic review and meta-analysis of individual patient data from nine randomised controlled trials. Lancet. 1998; 352:257–63. PMID: 9690404

62. Riley RD, Lambert PC, Staessen JA, Wang J, Gueyffier F, Thijs L, et al. Meta-analysis of continuous outcomes combining individual patient data and aggregate data. Stat Med. 2008; 27(11):1870–93. PMID: 18069721

63. Groenwold RH, Donders AR, van der Heijden GJ, Hoes AW, Rovers MM. Confounding of subgroup analyses in randomized data. Arch Intern Med. 2009; 169(16):1532–4. doi: 10.1001/archinternmed.2009.250 PMID: 19752412

64. Rydzewska L, Tierney J, Vale C. Obtaining funding for IPD meta-analyses: Top tips for a successful application. 19th Cochrane Colloquium; Madrid, Spain: Wiley-Blackwell; 2011. p. 65.

65. Abo-Zaid G, Sauerbrei W, Riley RD. Individual participant data meta-analysis of prognostic factor studies: state of the art? BMC medical research methodology. 2012; 12:56. doi: 10.1186/1471-2288-12-56 PMID: 22530717

66. Riley RD, Dodd SR, Craig JV, Thompson JR, Williamson PR. Meta-analysis of diagnostic test studies using individual patient data and aggregate data. Stat Med. 2008; 27(29):6111–36. doi: 10.1002/sim.3441 PMID: 18816508

67. Debray TP, Moons KG, Ahmed I, Koffijberg H, Riley RD. A framework for developing, implementing, and evaluating clinical prediction models in an individual participant data meta-analysis. Stat Med. 2013; 32(18):3158–80. doi: 10.1002/sim.5732 PMID: 23307585

68. Riley R, Price M, Jackson D, Wardle M, Gueyffier F, Wang J, et al. Multivariate meta-analysis using individual participant data. Research Synthesis Methods. 2015;In press. doi: 10.1002/jrsm.1129

69. PROSPERO: International prospective register of systematic reviews http://www.crd.york.ac.uk/PROSPERO/prospero.asp.

70. Rydzewska L, Tierney J, Clarke M, Stewart L, Rovers M. Publication of IPD meta-analyses in The Cochrane Database of Systematic Reviews: Potential barriers and opportunities. 19th Cochrahe Colloquium; Madrid, Spain: Wiley-Blackwell; 2011. p. 150.