# Individual participant data meta-analysis. When? Why? How? A scoping review

*Michail Belias*

*May 13, 2019*

## Abstract

### Background

Individual participant data(IPD) meta-analysis(MA) is considered the gold standard for evidence based inference. It is well established that IPD-MA offers great advantages compared to aggregate MA and single studies. Therefore, several systematic reviews have been conducted in order to investigate current practice. Besides the characteristics of the meta-analyses(size, design, type of outcome) most reviews reported the primary goal (subgroup analysis, risk factor assessment) and subsequently the statistical approaches applied. Nevertheless, last five years no review has been published, while new guidance is present. Furthermore, all reviews were narrowed down to specific the IPD-MA advantages. Particularly, either of effect modification, or risk factor investiagtion, while none reported the modelling approaches perfomed and/or comibnation of studies with different designs.

**Objective:** Therefore, our objective is to conduct a scoping review of existing IPD-MA, and summarise the aforementioned characteristics. Consequently, we aim to inform how IPD-MA are performed, what is their goal, which statistical approaches were used and whether they were clearly described to the level that they can be reproduced.

### Methods

We performed a scoping review to identify IPD-MA performed the last five years. We searched MEDLINE and the Cochrane Library for IPD-MA, with at least their abstract written in English. We included both randomised clinical trials, observational studies and their combinations, but excluded diagnostic IPD-MA. We screened the abstracts and extracted the size of MA, their primary goal, outcome(s), study designs included, statistical analysis and modeling approach. When unclear we consider the full-text.
**(we can include more databases i.e. EMBASE etc)**

### Results

Our search resulted in 1538 articles. We included only IPD-MAs with at least one treatment comparison. We showed an increase per year in IPD-MAs performed. The two most predominant medical fields were Cancer *(16%)*, Cardiovascular diseases *(16%)* and Mental health *(10%)*. An increasing trend in both IPD-MA in general and specifically in one-stage methods per year has been showed. Nevertheless, more information should be provided in both the abstract and the article over the statistical approaches followed.

*Temporary*

**Most of the IPD-MAs had as a goal to investigate for subgroups effects. Reporting type of**

### Conclusions

*Temporary* Goal and statistical approach description is still unclear. One-stage methods are increasing per year. Subgroups analysis is the primary goal of IPD-MAs.

# Introduction

Meta-analysis (MA) is a statistical method that involves combining information from multiple sources. While initially, meta-analyses were limited in aggregated data (AD) in the early 1990s individual participant data meta-analysis (IPD-MA or IPDMA) was introduced (CHALMERS 1993). In IPD-MA the participant level information is available and therefore evidence from multiple studies can be analysed centrally. Collecting the IPD may be a difficult and time consuming task, but nevertheless IPD-MA is considered the gold standard in evidence synthesis (Stewart and Parmar 1993 ; and 1995 ; Stewart and Tierney 2002) and offers great opportunities (Walraven 2010) that in AD-MA are considered impossible. Besides when investigating overall treatment effects where AD-MA and IPD-MA are mathematically equivalent, IPD-MA offers (1) the possibility to standardize subgroup definitions and outcomes across studies, (2) higher validity and credibility of subgroup findings, (3) increased flexibility to search for subgroups based on combinations of patient and/or disease characteristics (4) the possibility to avoid ecological BIAS (5) investigate non-linear functional forms (6) training better prediction models and (7) efficiently synthesizing evidence from different designs. Given these advantages systematic reviews were typically applied to inform of how are IPD from mulitple sources analysed and what for. For instance, Simmonds et al (Simmonds et al. 2005) identified 44 IPD-MAs performed during 2000-2005 time period and 1) summarized whether IPD-MAs obtained all the data they sought 2) reported the types of approaches that were used in the analysis 3) and whether the effects of covariates have been investigated and 4) report which medical field was their topic. On a subsequent paper, 10 years later Simmonds et al. (Simmonds, Stewart, and Stewart 2015) identified 1371 potential IPD-MAs performed during 2010-2015 time period, sampled 184 of them and after obtaining full texts included 100 IPD-MAs. Then they investigated along with the topics investigated in the initial paper they investigated also the quality of IPD-MA reporting. Riley et al. (Riley, Lambert, and Abo-Zaid 2010) identified 383 IPD-MAs performed form instance until 2009 and summarised only: 1) their medical field topic and 2) whether they assessed risk or prognostic factors. Finally, Schuit and Ioannidis (Schuit, Li, and Ioannidis 2018) identified 327 IPD-MAs performed from inception until 2014. Nevertheless, they restricted their interest in subgroup effects investigation. Our objective is to conduct a systematic review of IPD-MA from 2015 onwards and and summarise their properties. Furthermore, we aim to inform when and how IPD-MA are performed, whether state-of the art methods are used and whether they are clearly described.

Nevertheeo far systematic reviews over the IPD-MA practices are limited until 2014.

**Reporting** IPD-MA may be conducted in either one stage or two stages. In one-stage IPD-MA, a statistical model of choice is applied and IPD from all studies are analysed simultaneously, whilst accounting for within-studies clustering of the participants. On the other hand, in two-stage IPD-MA a statistical model of choice is fitted per study. Subsequently the estimates extracted are pooled using inverse-variance meta-analytical methods. Both approaches have a variety of parameters and results that should be reported in the abstract, the methods and the results section. An extended version of PRISMA for IPD (Stewart et al. 2015) offers guidance on how to report results in IPD-MA. For instance, in two-stage IPD-MA 1) heterogeneity measures ( $I^2$ , Cochran's Q $\tau^2$) 2) and their corresponding methods used 3) forest plots (if applicable) and 4) use of fixed or random effects models and any other model assumptions should be described in the Methods section. Furthermore, in't Hout (IntHout et al. 2016) suggested that prediction intervals of estimates are also a valuable information and should be included. On the other hand, in one-stage IPD-MA 1) specification of one-stage models 2) use of fixed-effect, stratified or random-effects in the terms of the model and 3) how clustering of patients within studies was accounted for should be reported in the methods section.

**Effect modification** Simmonds et al. (Simmonds, Stewart, and Stewart 2015) showed that IPD-MA are frequently performed in order to detect treatment effect modification. The approaches that were mostly used were one aggregated data meta-analysis approach 'meta-regression' and three IPD-MA approaches, per-subgroup meta-analysis, meta-analysis of interaction terms and one-stage IPD-MA. Guidance on which method to choose is available. Specifically, Simmonds and Higgins (Simmonds and Higgins 2007) mathematically proved that, given some unrealistic assumptions, one-stage IPD-MA is always more powerful than meta-analysis of interaction terms and meta-regression. Fisher et al. (Fisher et al. 2011) also critically reviewed all four approaches. They concluded that one-stage IPD-MA allows for more complex analysis, but is more difficult to perform than pooling within-trial interaction terms. Furthermore, Hua et al. (Hua

et al. 2016) noted that these one-stage IPD-MA using mixed-effects modelling should also centre the effect modifiers to their mean, in order to separate across and within trial information and therefore accounting for ecological bias.

**Modeling functional forms** IPD-MA may be performed in order to investigate the role of risk-factors in the prevalence of a disease. In that case observational studies are typically meta-analysed. Thereto, IPD-MA may involve modelling also non-linear functional forms. Sauerbrei and Royston (Sauerbrei and Royston 2011) suggested the use of a two stage approach. As a first stage a fractional polynomial is selected and pooling their estimates through a point-wise weighted meta-analytical process. Subsequently they extended these non-linear associations to include interactions (Royston and Sauerbrei 2013). Furthermore, splines may also be applied to detect non-linear associations.

# Methods

This study is a scoping review of current practices in IPD-MAs . We report our study according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. No formal protocol exists for this study.

## IPD-MA search and selection strategy

We searched in MEDLINE, PubMed, Cochrane library **(we can place more)** for IPD-MA conducted between 01/01/2015 and 01/05/2019, see in **Supplementary material** for search strategy. We included IPD-MAs of either RCTs or observational studies, reporting at least one intervention in human subjects. We excluded IPD-MAs not written in English. We also excluded diagnostic and prediction IPD-MAs. We removed duplicate papers and from series we included only the last article. When unclear we considered full-text assessment. Any disagreement was solved by a third reviewer. **if we have to find one it is suggest though**

## Data extraction and statistical analysis

For all eligible papers we extracted information that was classified in three categories: 1) the reporting of the IPD-MA (such as the medical field, number of included studies and participants, the goal described in the abstract); 2) the statistical analysis of the IPD-MA (type of outcome,type of statistical approach,descriptions of methods, and assessment of heterogeneity); and 3) the reporting of the results (forest-plots, how many comparisons were performed, proportion of statistically significant, reporting of subgroup analyses and/or interaction terms) and 4) when

# Results

We identified 1538 published IPD-MAs. Of these

| General Medical Field | Frequencies | Percentage |
| --- | --- | --- |
| Anaesthesiology | 1 | 0.48% |
| Cancer | 20 | 9.62% |
| Cardiovascular disease | 35 | 16.83% |
| Child health | 18 | 8.65% |
| Ear, nose and throat | 1 | 0.48% |
| Endocrine and metabolic | 10 | 4.81% |
| Gastroenterology | 8 | 3.85% |
| Generic Care | 3 | 1.44% |
| Geriatrics | 3 | 1.44% |
| Gynaecology | 1 | 0.48% |
| Infectious disease | 3 | 1.44% |
| Lungs and airways | 5 | 2.4% |
| Mental health | 24 | 11.54% |
| Neurology | 30 | 14.42% |
| Nutrition | 1 | 0.48% |
| Orthopedics | 6 | 2.88% |
| Other | 1 | 0.48% |
| Pregnancy and childbirth | 18 | 8.65% |
| Psychology | 3 | 1.44% |
| Renal Disease | 1 | 0.48% |
| Review | 2 | 0.96% |
| Statistical | 11 | 5.29% |
| Vaccines | 2 | 0.96% |
| Wound | 1 | 0.48% |

# Number of IPD–MAs performed per year



Figure 1.

# Number of trials included per meta–analysis
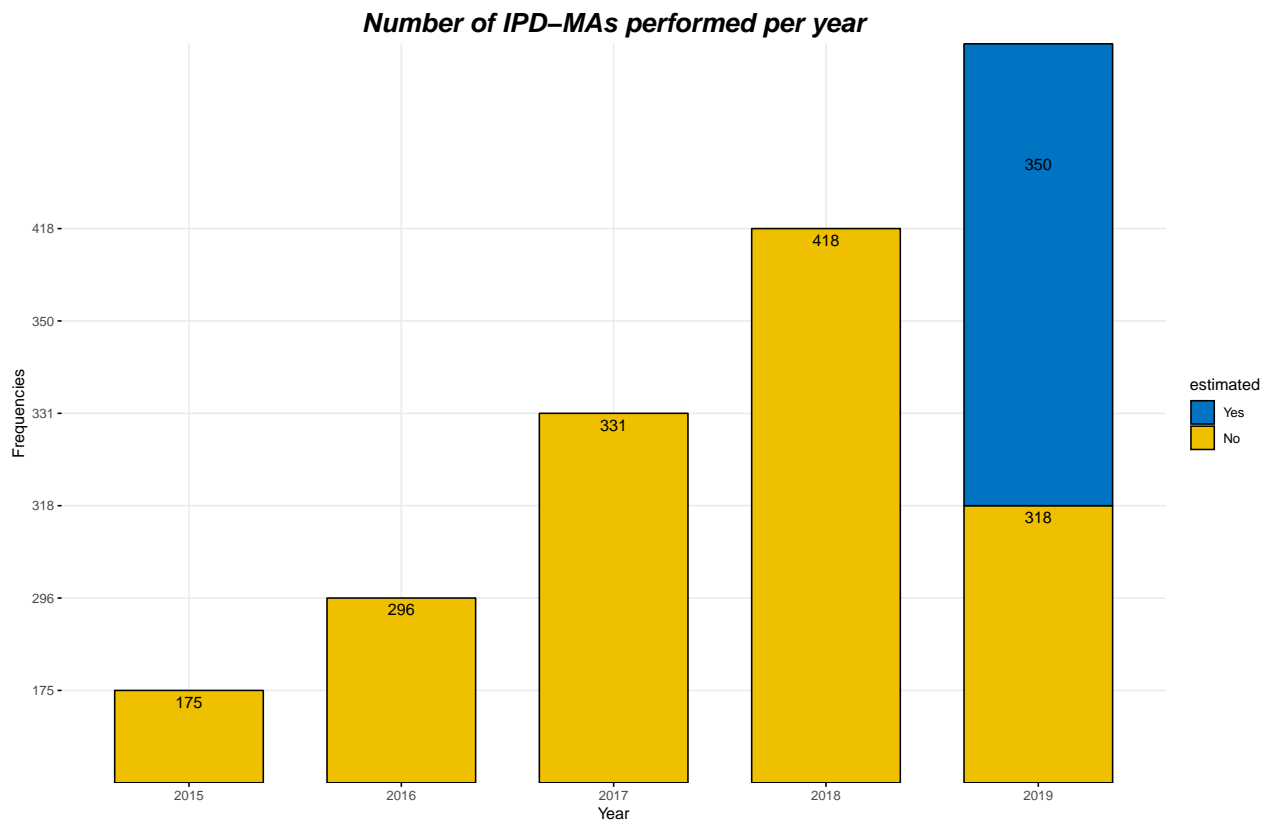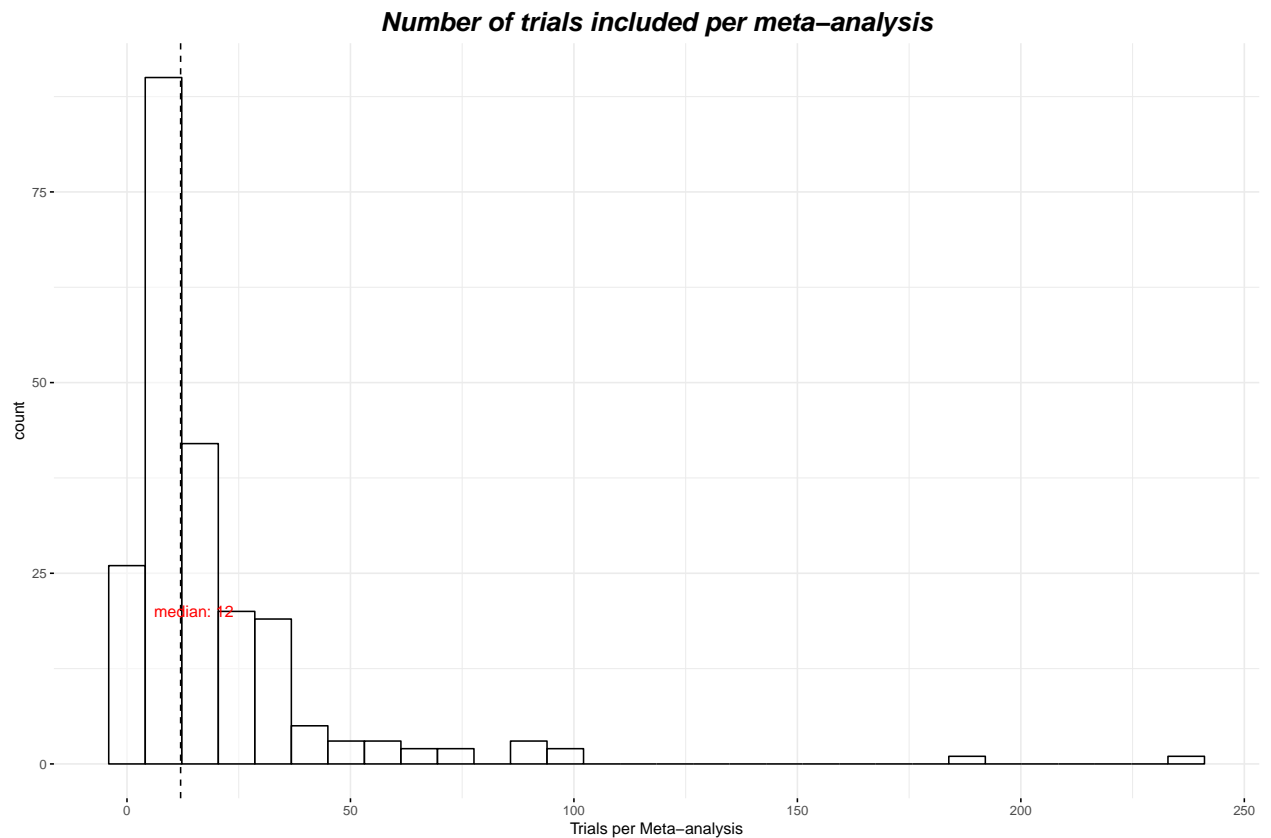
Table 2: Table 2

| Type of primary outcome | Frequencies |
|---|---|
| Binary | 11 |
| Continuous | 5 |
| Time-to-Event | 4 |

Table 3: Table 3

| | One-stage IPD-MA | Meta-analysis of interaction terms | Per subgroup meta-analysis | Me |
|---|---|---|---|---|
| | | Fixed vs | random effects | |
| Fixed effect | x | x | x | x |
| Random effects | x | x | x | x |
| | | Reporting of | heterogeneity | |
| I^2 | x | x | x | x |
| Cochran's Q (without I^2 ) | x | x | x | x |
| tau^2 | x | x | x | x |
| Prediction intervals | x | x | x | x |
| From one-stage model | x | x | x | x |
| Other | x | x | x | x |
| Not reported | x | x | x | x |



**Number of participants included per meta–analysis**
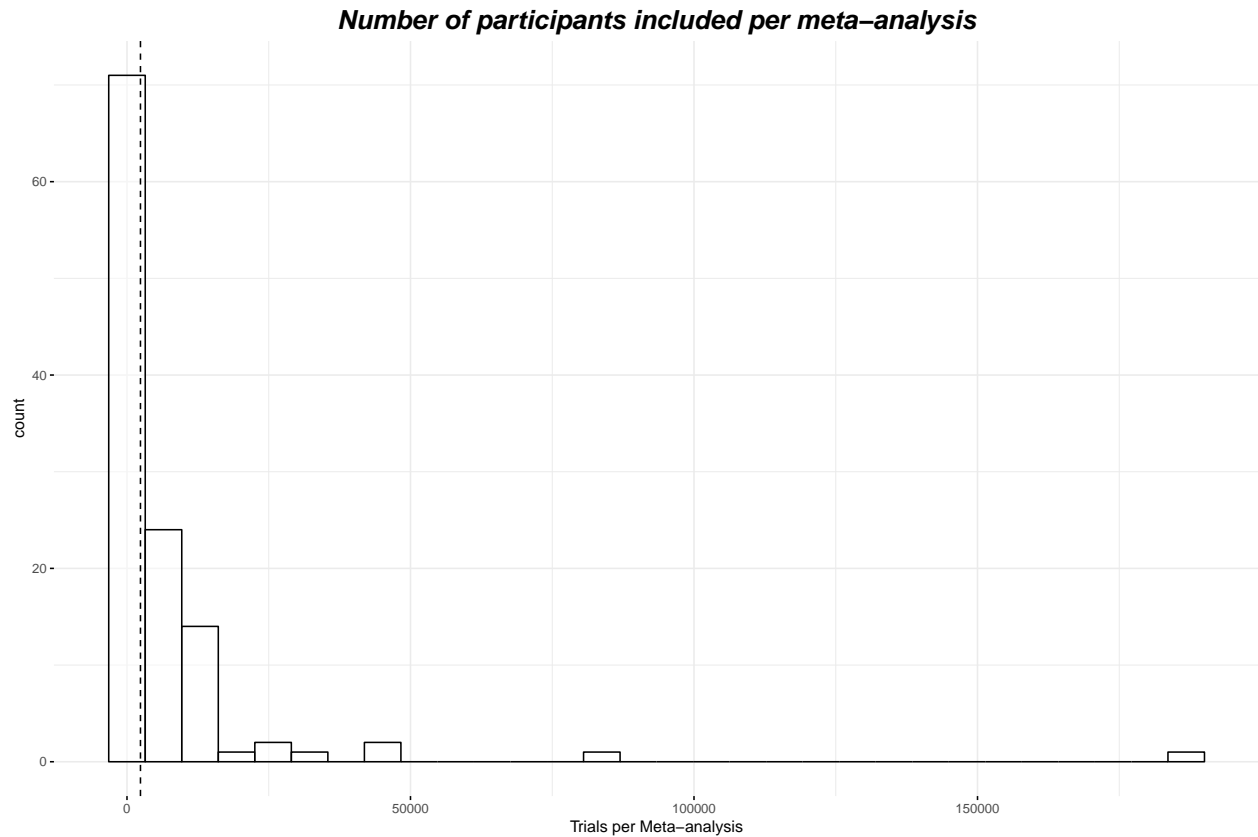
# References

and, Lesley A. Stewart. 1995. "Practical Methodology of Meta-Analyses (Overviews) Using Updated Individual Patient Data." *Statistics in Medicine* 14 (19): 2057–79. https://doi.org/10.1002/sim.4780141902.

CHALMERS, IAIN. 1993. "The Cochrane Collaboration: Preparing, Maintaining, and Disseminating Systematic Reviews of the Effects of Health Care." *Annals of the New York Academy of Sciences* 703 (1 Doing More Go): 156–65. https://doi.org/10.1111/j.1749-6632.1993.tb26345.x.

Fisher, D. J., A. J. Copas, J. F. Tierney, and M. K.B. Parmar. 2011. "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." *Journal of Clinical Epidemiology* 64 (9): 949–67. https://doi.org/10.1016/j.jclinepi.2010.11.016.

Hua, Hairui, Danielle L. Burke, Michael J. Crowther, Joie Ensor, Catrin Tudur Smith, and Richard D. Riley. 2016. "One-Stage Individual Participant Data Meta-Analysis Models: Estimation of Treatment-Covariate Interactions Must Avoid Ecological Bias by Separating Out Within-Trial and Across-Trial Information." *Statistics in Medicine* 36 (5): 772–89. https://doi.org/10.1002/sim.7171.

IntHout, Joanna, John P A Ioannidis, Maroeska M Rovers, and Jelle J Goeman. 2016. "Plea for Routinely Presenting Prediction Intervals in Meta-Analysis." *BMJ Open* 6 (7): e010247. https://doi.org/10.1136/bmjopen-2015-010247.

Riley, R. D., P. C. Lambert, and G. Abo-Zaid. 2010. "Meta-Analysis of Individual Participant Data: Rationale, Conduct, and Reporting." *BMJ* 340 (feb05 1): c221–c221. https://doi.org/10.1136/bmj.c221.

Royston, Patrick, and Willi Sauerbrei. 2013. "Interaction of Treatment with a Continuous Variable: Simulation Study of Significance Level for Several Methods of Analysis." *Statistics in Medicine* 32 (22): 3788–3803. https://doi.org/10.1002/sim.5813.

Sauerbrei, Willi, and Patrick Royston. 2011. "A New Strategy for Meta-Analysis of Continuous Covariates in Observational Studies." *Statistics in Medicine* 30 (28): 3341–60. https://doi.org/10.1002/sim.4333.

Schuit, Ewoud, Alvin H Li, and John P A Ioannidis. 2018. "How Often Can Meta-Analyses of Individual-Level Data Individualize Treatment? A Meta-Epidemiologic Study." *International Journal of Epidemiology* 48 (2): 596–608. https://doi.org/10.1093/ije/dyy239.

Simmonds, Mark C, Julian P T Higginsa, Lesley A Stewartb, Jayne F Tierneyb, Mike J Clarke, and Simon G Thompson. 2005. "Meta-Analysis of Individual Patient Data from Randomized Trials: A Review of Methods Used in Practice." *Clinical Trials: Journal of the Society for Clinical Trials* 2 (3): 209–17. https://doi.org/10.1191/1740774505cn087oa.

Simmonds, Mark, Gavin Stewart, and Lesley Stewart. 2015. "A Decade of Individual Participant Data Meta-Analyses: A Review of Current Practice." *Contemporary Clinical Trials* 45 (November): 76–83. https://doi.org/10.1016/j.cct.2015.06.012.

Simmonds, M. C., and J. P. T. Higgins. 2007. "Covariate Heterogeneity in Meta-Analysis: Criteria for Deciding Between Meta-Regression and Individual Patient Data." *Statistics in Medicine* 26 (15): 2982–99. https://doi.org/10.1002/sim.2768.

Stewart, L. A, and M. K.B Parmar. 1993. "Meta-Analysis of the Literature or of Individual Patient Data: Is There a Difference?" *The Lancet* 341 (8842): 418–22. https://doi.org/10.1016/0140-6736(93)93004-k.

Stewart, Lesley A., Mike Clarke, Maroeska Rovers, Richard D. Riley, Mark Simmonds, Gavin Stewart, and Jayne F. Tierney. 2015. "Preferred Reporting Items for a Systematic Review and Meta-Analysis of Individual Participant Data." *JAMA* 313 (16): 1657. https://doi.org/10.1001/jama.2015.3656.

Stewart, Lesley A., and Jayne F. Tierney. 2002. "To IPD or Not to IPD?" *Evaluation & the Health Professions* 25 (1): 76–97. https://doi.org/10.1177/0163278702025001006.

Walraven, Carl van. 2010. "Individual Patient Meta-Analysisrewards and Challenges." *Journal of Clinical Epidemiology* 63 (3): 235–37. https://doi.org/10.1016/j.jclinepi.2009.04.001.