Using individual participant data to improve network meta-analysis projects

Richard D Riley ,[1] Sofia Dias,[2] Sarah Donegan,[3] Jayne F Tierney,[4] Lesley A Stewart,[2] Orestis Efthimiou,[5] David M Phillippo[6]

[1]School of Medicine, Keele University, Keele, UK
[2]Centre for Reviews and Dissemination, University of York, York, UK
[3]Department of Health Data Science, University of Liverpool, Liverpool, UK
[4]MRC Clinical Trials Unit at UCL, UCL, London, UK
[5]Institute of Social and Preventive Medicine (ISPMU), University of Bern, Bern, Switzerland
[6]Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK

Correspondence to:
Professor Richard D Riley, School of Medicine, Keele University, Keele ST5 5BG, UK; r.riley@keele.ac.uk

Check for updates

Abstract

A network meta-analysis combines the evidence from existing randomised trials about the comparative efficacy of multiple treatments. It allows direct and indirect evidence about each comparison to be included in the same analysis, and provides a coherent framework to compare and rank treatments. A traditional network meta-analysis uses aggregate data (eg, treatment effect estimates and standard errors) obtained from publications or trial investigators. An alternative approach is to obtain, check, harmonise and meta-analyse the individual participant data (IPD) from each trial. In this article, we describe potential advantages of IPD for network meta-analysis projects, emphasising five key benefits: (1) improving the quality and scope of information available for inclusion in the meta-analysis, (2) examining and plotting distributions of covariates across trials (eg, for potential effect modifiers), (3) standardising and improving the analysis of each trial, (4) adjusting for prognostic factors to allow a network meta-analysis of conditional treatment effects and (5) including treatment–covariate interactions (effect modifiers) to allow relative treatment effects to vary by participant-level covariate values (eg, age, baseline depression score). A running theme of all these benefits is that they help examine and reduce heterogeneity (differences in the true treatment effect between trials) and inconsistency (differences in the true treatment effect between direct and indirect evidence) in the network. As a consequence, an IPD network meta-analysis has the potential for more precise, reliable and informative results for clinical practice and even allows treatment comparisons to be made for individual patients and targeted populations conditional on their particular characteristics.

Introduction

A network meta-analysis (NMA) project compares multiple treatments that have been evaluated in existing randomised trials.[1] A conventional NMA uses aggregate data (AD) on treatment effects extracted from study publications, such as estimates of odds ratios (ORs), hazard ratios (HRs) or mean differences and corresponding standard errors, or the total participants and events per treatment group. An alternative approach is an NMA of individual participant data (IPD), in which the participant-level data are obtained from multiple studies then checked, harmonised and synthesised.[2] IPD refers to the raw information recorded for each participant in a study, such as baseline characteristics, prognostic factors, treatments received, outcomes and follow-up details. In this article, we describe the potential benefits that IPD offers for NMA projects compared with using AD, emphasising how IPD enables more precise, reliable and informative results for patients and clinical decision makers. This work adapts and extends our recent book chapter on this topic.[3]

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ A network meta-analysis provides a framework to compare and rank all available treatments for a specific disease, using both direct evidence and indirect evidence. Most network meta-analysis projects use aggregate data obtained from study publications or study authors.

WHAT THIS STUDY ADDS

⇒ The use of individual participant data (IPD), rather than aggregate data, can enhance network meta-analysis projects by increasing the quality and scope of information available; by allowing participant characteristics to be compared across studies; by standardising and improving statistical analyses; and by inclusion of prognostic factors and treatment effect modifiers to reduce concerns of inconsistency and heterogeneity.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ The use of IPD in network meta-analysis projects can produce more precise, reliable and informative results for patients and clinical decision makers. We recommend network meta-analysis projects consider using IPD where possible.

What is an NMA?

A pairwise meta-analysis compares two treatments (eg, *A* and *B*) by combining the evidence only from trials that directly compared those treatments. However, different trials often evaluate different sets of treatments (eg, some compare *A* and *B*, others compare *A* and *C* or *B* and *C*), and then the pairwise meta-analysis approach is unable to provide a coherent comparison of all treatments. Moreover, for some pairs of treatments there may be no studies comparing them directly at all.

To address this, NMA provides a comprehensive framework to coherently compare and rank all available treatments for a specific disease. It does so by synthesising data from all trials in a single analysis, using both *direct evidence* (eg, about *A vs B* from trials that compared *A* to *B* head-to-head) and *indirect evidence* (eg, about *A vs B* from trials that compared *A* to *C* and trials that compared *B* to *C*).[4–6] The premise is as follows. For a randomised trial in a particular population of patients, the following relationship holds exactly

Treatment contrast of *A vs B*

= (treatment contrast of *A vs C*) – (treatment contrast of *B vs C*)

where 'treatment contrast' is the true relative effect between two treatments, measured on a scale such as a log risk ratio, log OR, log HR or mean difference. In an NMA, we assume that this relationship also holds (on average) across the populations of patients in different trials. This notion is often referred to as the consistency assumption (or coherence),[7] and it allows us to combine direct and indirect evidence from trials comparing different sets of treatments.

The validity of combining studies in a pairwise meta-analysis relies on the assumption that relative treatment effects are exchangeable between trials.[8] This is appropriate when trials are sufficiently similar with respect to all study-level and patient-level characteristics that might impact on the relative effect of the treatments being compared (eg, quality, length of follow-up, casemix). Exchangeability also underpins the validity of combining direct and indirect evidence in a network meta-analysis (and thus the statistical consistency, ie, the agreement between direct and indirect evidence), and it implies that the relative effects between any pair of treatments observed directly in some trials would be the same in other trials where they are unobserved. This concept is also known as transitivity,[9 10] and when transitivity does not hold, there is likely to be inconsistency (incoherence) in the NMA such that the direct and indirect evidence disagree.

Between-study heterogeneity in treatment effects is another important concept in NMA, which refers to genuine differences in the true treatment effects (eg, for *A vs B*) across trials, and is caused by differences in treatment effect modifiers across trials making the same comparison(s) (eg, trials of *A vs B*). Effect modifiers are methodological or clinical characteristics of trials that influence the magnitude of relative treatment effects on a given scale, and examples include duration of follow-up, outcome definitions, trial quality (risk of bias) and participant-level characteristics.[11–14]
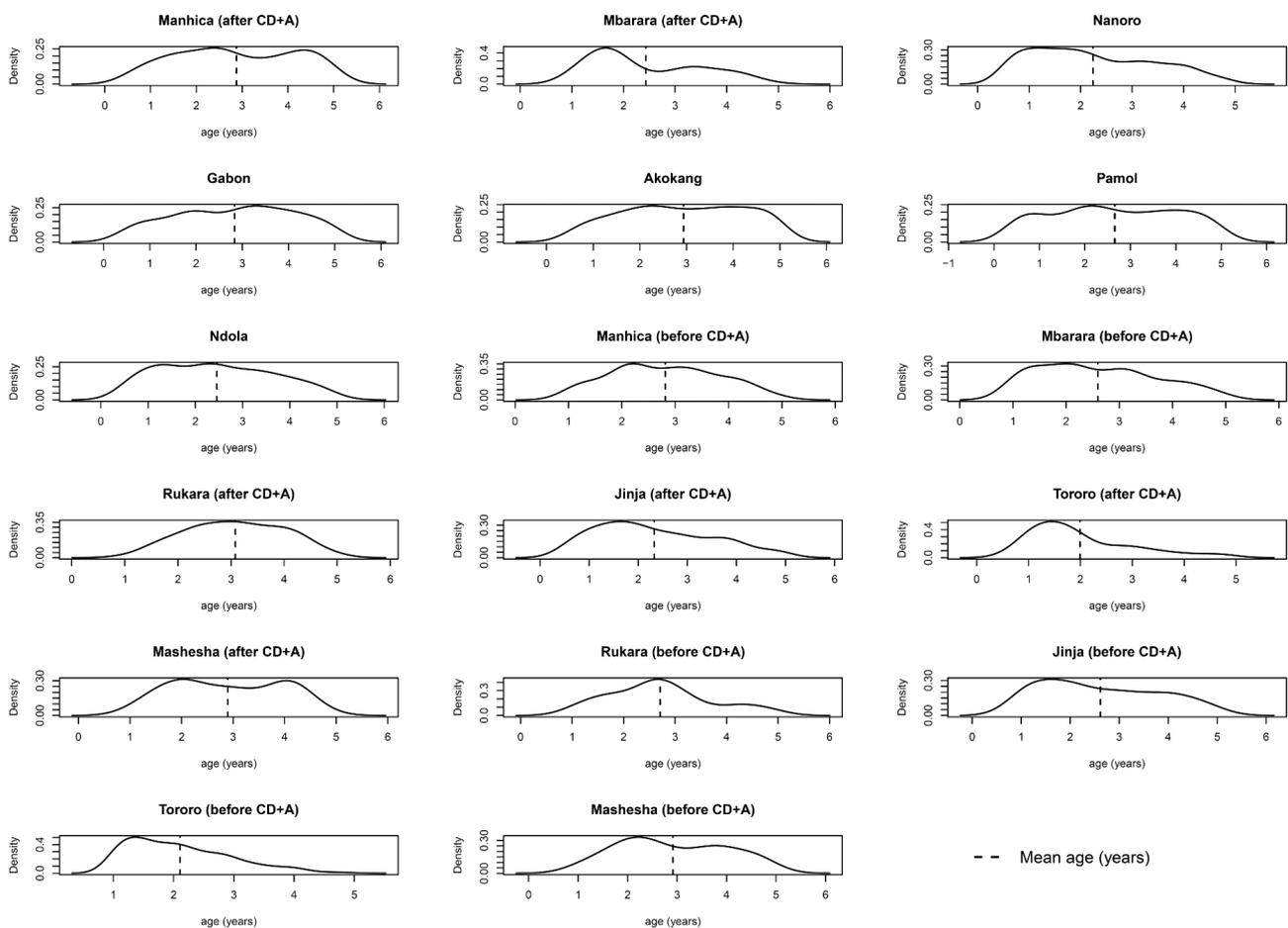
Figure 1 Distribution of age as derived directly from the IPD for each of 17 sites ('studies') included in the IPD NMA of Donegan *et al*.[51] The distribution is similar in most studies, though slightly more skewed in some studies (eg, Mbarara, Tororo), which may lead to heterogeneity or inconsistency in the network due to age. Figure originally presented in online supplemental material of Donegan *et al*.[51] CD+A, chlorproguanil-dapsone plus artesunate; IPD, individual participant data; NMA, network meta-analysis.

Inconsistency is also a consequence of treatment effect modifiers, but specifically when there are systematic imbalances in effect modifiers across trials making different comparisons. Thus, an effect modifier might differ across the trials in a network (eg, some trials were performed in younger patients, some in older patients, and age impacts on relative treatment effects); this might cause heterogeneity. However, if there are systematic differences in age across comparisons (eg, all *A vs B* studies are in younger patients and all *A vs C* studies are in older patients), this might cause inconsistency.

Potential benefits of using IPD for NMA

Compared with using AD from existing trials, the availability of IPD brings important advantages for NMA projects, and five key benefits are now described.[3]

Benefit 1: improving the quality and scope of information available

A key benefit of IPD over AD is the potential to improve the quantity, completeness, and validity of data available for each trial, because there is no need to be limited by the study-level information (ie, AD) that has been published for each trial. This enhances data quality,[15–18] provides independent scrutiny of the trial data and enables more information to be available for the NMA. This may lead to more reliable and informative conclusions, with potentially reduced heterogeneity and inconsistency.

For instance, there is greater ability to standardise outcome and covariate definitions across trials, and to include outcomes that were not reported in the original publications, or participants who were inappropriately excluded from the original trial analyses.[15–17] This can help reduce potential trial reporting biases[19] and increase the quantity of information available for the NMA, boosting the statistical power to compare treatments.[20] For example, studies of depression often report the outcome 'response to treatment', with the definition varying across different studies.[21] One study might define response as 20% reduction in baseline severity and another study might use a 50% reduction, while a third study might define response as a different endpoint severity being below another arbitrary threshold (on some scale). When using AD we are limited to using these highly variable definitions of response, which may lead to large heterogeneity and greater scope for inconsistency. If IPD are available, we can better harmonise the definition of response and the severity score can be analysed on its continuous scale, thus avoiding dichotomisation and allowing a more powerful investigation of the outcome as originally measured. For example, in their IPD NMA, Karyotaki *et al* used IPD to convert depression scores to the same Patient Health Questionnaire-9 (PHQ-9) scale,[22] and subsequently chose to analyse PHQ-9 on its continuous scale rather than dichotomising into high-score and low-score groups.

Most IPD NMA projects are collaborative endeavours, involving direct contact with trial investigators, which can help to identify trials not easily identifiable via other forms of searching,[15–17] and to clarify the eligibility of potentially relevant trials. Trial investigators can also provide extra information leading to more reliable risk of bias assessments than what is achievable from trial reports,[23] and if potential biases or errors are identified, they may be able to supply additional data to resolve or minimise these.[15–17] IPD also allow more flexible and detailed modelling of the survival function,[24–26] potentially including longer follow-up times (or even larger numbers of participants), for example, for those trials that continued monitoring (or recruiting) participants beyond the timepoint when the original analyses were conducted.

> ### Box 1 Example of how heterogeneity and inconsistency can arise in an NMA when the distribution of prognostic factors differs across trials
>
> Consider a hypothetical example where, regardless of the trial, the underlying (true) probability (*P*) of having an adverse outcome depends on whether treatment *B* or *A* is used and, in particular, whether a participant smokes, as defined by the equation,
>
> $$\text{logit}(p) = -0.5 + (4 \times smoker) - (0.5 \times B)$$
>
> where *B* = 1 if in treatment group *B* and *B* = 0 if treatment group *A*, and *smoker* = 1 if a participant currently smokes and 0 otherwise. Hence, after adjusting for the strong prognostic effect of smoking, the true treatment effect for *B vs A* is an OR of exp(−0.5)=0.61. Note that there are no treatment effect modifiers (eg, no treatment-smoker interaction).
>
> Now consider that two randomised trials of *B vs A* were conducted: trial 1 used a population where 20% of participants smoked, and trial 2 had a population where 80% of participants smoked. Regardless of the proportion of smokers, the equation tells us that both trials have a true OR of 0.61 for the treatment effect adjusted for the prognostic effect of smoking.
>
> However, when ignoring smoking, the unadjusted treatment effect is considerably different in the two trials. The true unadjusted OR for *B vs A* is 0.69 for trial 1 and 0.77 for trial 2. Hence, there is genuine between-trial heterogeneity in the unadjusted OR for trials 1 and 2 (as available in a typical AD NMA), but no heterogeneity in the OR adjusted for smoking (as available from an IPD NMA).
>
> Similarly, if the distribution of smokers is different in each subset of trials that provide direct and indirect evidence for a particular comparison, then the unadjusted ORs will differ for each subset, leading to inconsistency in the NMA for that comparison. However, with IPD such inconsistency can be removed by synthesising ORs adjusted for smoking for each trial.
>
> AD, aggregate data; IPD, individual participant data; NMA, network meta-analysis.

Benefit 2: examining and plotting distributions of covariates across trials

Before undertaking an NMA, it is important to select only those trials relevant to the population of clinical interest and to then identify any systematic differences in variables that might affect measures of relative treatment effect. Of particular interest are suspected effect modifiers or, specifically when analysing ORs or HRs, prognostic factors that modify baseline risk, as these may lead to heterogeneity or inconsistency in the NMA if their distribution is different across trials (see also benefits 4 and 5).

When using only AD, statistical summaries of the distributions of participant-level characteristics need to be extracted from the trial publications (or obtained from the trial investigators). Such AD may be the mean and SD for a continuous covariate, or the proportion of participants in each category for a categorical covariate. In contrast, when IPD are available for each trial, researchers can summarise and plot covariate distributions

themselves,[27] [28] and may have access to a broader set of recorded covariates than summarised in the original trial publication. An example is provided in figure 1. This allows better comparison of how covariate distributions differ across comparisons and across trials to help assess their potential to cause heterogeneity and inconsistency in the network.[29] For example, to assess the plausibility of the consistency assumption in their IPD NMA, Karyotaki et al 'checked the distribution of possible effect size modifiers in the studies grouped by comparison' and did not identify any systematic differences across comparisons.

Benefit 3: standardising and improving the analysis of each trial

With IPD, effect measures can be standardised across trials, and this allows NMA researchers to define their own estimands of interest, rather than being limited to those presented in the original trial publications. For example, many published trials report ORs, but with IPD the NMA researcher may rather be able to estimate HRs in order to account for censoring and time-to-event information. Similarly, restricted mean survival time differences could be used rather than HRs if non-proportional hazards are a concern.[30] Such decisions can be informed by using the IPD to check model assumptions (eg, proportional hazards assumption) and model fit,[31] after addressing other issues such as outliers, none of which are likely to be possible without IPD.

IPD also enable more appropriate or advanced analytical approaches, in particular for the modelling of time-to-event data[32] [33]; the analysis of continuous variables on their continuous scale (rather than dichotomised); the examination of non-linear relationships between covariates and the outcome[34]; the inclusion of prognostic factors and treatment–covariate interactions (see benefits 4 and 5); and the handling of missing data, which is especially important for trials that originally used suboptimal or inappropriate methods to deal with missing values.

IPD also facilitate a multivariate NMA approach, to compare treatments across multiple outcomes simultaneously (eg, benefits and harms; systolic and diastolic blood pressure; surrogate outcomes like disease-free and overall survival; pain scores at 6 and 12 months), while accounting for the (participant-level and study-level) correlation among outcomes.[35] This can lead to more precise inferences and even change the ranking of treatments,[36] [37] compared with an NMA of each outcome separately. For example, in an NMA of 68 trials comparing 13 active antimanic drugs and placebo for acute mania,[38] two negatively correlated outcomes of interest were efficacy (defined as the proportion of participants with at least a 50% reduction in manic symptoms from baseline to week 3) and acceptability (defined as the proportion of participants with treatment discontinuation before 3 weeks). When performing a separate NMA for each outcome, carbamazepine ranked as the most effective treatment in terms of response; however, when analysing outcomes jointly in a multivariate NMA, carbamazepine fell to fourth place.

Benefit 4: adjusting for prognostic factors in the analysis of each trial

IPD allow the NMA to compare treatment effects conditional on (ie, adjusting for) prognostic factors, which is important.[2] First, adjustment for prognostic factors in a single trial (eg, using regression) is often preferred to increase power to detect treatment effects (and, for continuous outcomes, to increase the precision of the estimates), as prognostic factors may explain variation in outcomes across participants,[39–42] and is necessary to obtain correct estimates of uncertainty when stratified randomisation has been employed[43]; thus, subsequent NMA results may also be

more powerful and appropriately quantify uncertainty.[44] Second, conditional treatment effects also align more closely with the drive toward personalised medicine tailoring treatment decisions to each patient given their particular characteristics and outcome risks. Third, conditioning on observed prognostic factors improves the plausibility of the missing at random assumption for participants with missing outcomes, and so improves on a complete-case analysis of unadjusted treatment effect estimates.

Fourth, adjustment for prognostic factors can improve homogeneity and consistency of treatment effects in an NMA. Even in situations where there are no effect modifiers, differences in the distribution of prognostic factors between trials can lead to heterogeneity and inconsistency to when the treatment effect is measured on the OR or HR scale.[3] IPD help to address this by adjusting for prognostic factors in the analysis of each trial, and an example is provided in box 1. Occasionally a prognostic factor may also be an effect modifier,[45] which can also be better modelled using IPD, as considered in benefit 5.

### Benefit 5: including treatment–covariate interactions (effect modifiers)

One of the most important advantages of IPD is that it allows the examination and inclusion of participant-level effect modifiers (treatment–covariate interactions) that would otherwise cause heterogeneity or inconsistency in the NMA. Single trials are rarely powered to detect treatment–covariate interactions, and typically an AD meta-analysis can only examine across-trial relationships, which are prone to aggregation bias and trial-level confounding.[46] [47] In contrast, IPD meta-analysis allows participant-level relationships to be modelled directly and more precisely, for example in a regression that models a covariate's interaction with treatment effect while adjusting for the covariate's prognostic effect (and potentially also other prognostic factors). This allows NMA results to compare treatments for specific patient populations, subgroups or at covariate values defined by participant-level characteristics. Indeed, since the magnitude of treatment effects may change in the presence of effect modification, the best ranking treatment(s) may differ across populations, subgroups, and covariate values defined by effect modifiers.[3] [13] [48–50] In this situation, treatment recommendations will need to be tailored to a chosen target population, or for subgroups or individual patients, according to their corresponding (distribution of) covariate values. This type of analysis (adjustment for effect modifiers, followed by the production of population-targeted treatment effect estimates) is known as 'population adjustment'.[49]

Riley et al describe how to include participant-level treatment–covariate interactions in two-stage or one-stage IPD meta-analysis models with a single pairwise comparison.[2] [47] These can be extended to NMA situations to accommodate treatment–covariate interactions corresponding to the multiple treatment effects. Ideally, these interactions are assumed to be independent (ie, different for each treatment), but to aid model convergence, it may be necessary to assume interactions are exchangeable (eg, by including random effects) or even common for each treatment.[51]

Consider an NMA presented by Donegan et al,[51] who use IPD to examine four artemisinin-based combination therapies for uncomplicated malaria: amodiaquine-artesunate, dihydroartemisinin-piperaquine (DHAPQ), artemether-lumefantrine (AL), and chlorproguanil-dapsone plus artesunate. The binary outcome of interest was treatment success at 28 days. IPD were available from 17 sites, which for simplicity can be considered as 17 'trials' here. Age was prespecified as a potential treatment effect modifier, since in areas with endemic malaria older patients are more likely
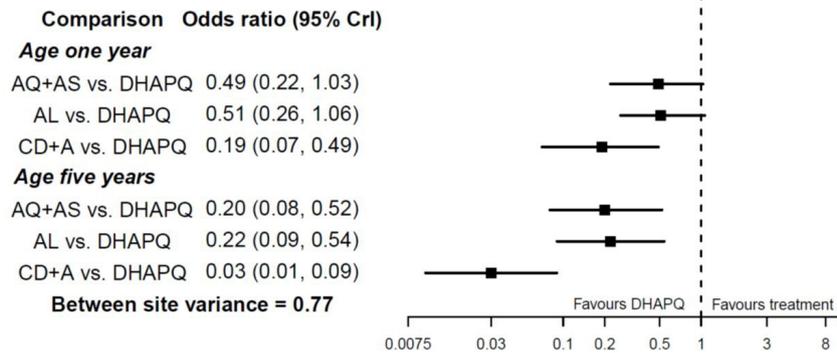
Figure 2 Results reported by Donegan et al[51] after estimation of an NMA allowing for treatment–age interactions, as applied to compare four treatments for malaria using IPD from 11 sites. Figure adapted from Donegan et al,[51] and results assume treatment–age interactions are independent for each treatment comparison. AQ+AS, amodiaquine-artesunat; AL, artemether-lumefantrine; DHAPQ, dihydroartemisinin-piperaquine; IPD, individual participant data; NMA, network meta-analysis.
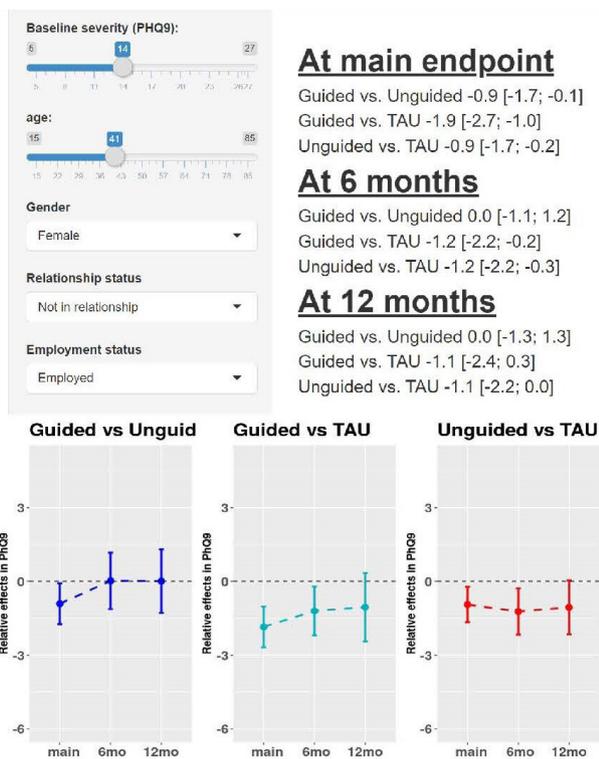
Figure 3 Example output from an online calculator of treatment effects for depression (measured on Patient Health Questionaire-9 (PHQ-9) scale) conditional on patient-level characteristics, as derived from the IPD NMA results of Karyotaki et al.[22] The tool is available here: https://esm.ispm. unibe.ch/shinies/iCBT/). iCBT, internet-based cognitive behavioural therapy; IPD, individual participant data; NMA, network meta-analysis; TAU, treatment as usual.

to achieve success on treatment because they have greater immunity. There is strong evidence of an interaction between treatment effects and age. This leads to larger summary treatment effects at higher ages (figure 2). For example, the summary OR for AL versus DHAPQ is 0.51 (95% CI 0.26 to 1.06) at 1 year of age, and 0.22 (95% CI 0.09 to 0.54) at 5 years of age, when assuming interactions are independent for each treatment comparison.

Karyotaki et al[22] provide an online calculator for estimating and comparing treatment effects conditional on participant-level characteristics of baseline PHQ-9 score, age, gender, relationship status and employment status following their IPD NMA. The tool

is illustrated in figure 3. Such predictions often require shrinkage and penalisation techniques,[47 52 53] in order to mitigate against overfitting (extreme predictions), and Karyotaki et al used the least absolute shrinkage and selection operator (LASSO) for this purpose.[22]

Potential challenges

IPD meta-analysis projects also face challenges.[2] They are a considerable undertaking, often taking upwards of 2 years to obtain, check, harmonise and synthesise IPD. Negotiating and maintaining collaboration with trial investigators takes considerable effort, and care is needed to arrange and adhere to data-sharing agreements, including how IPD are transferred and stored. To safeguard against future conflicts, data-sharing agreements should make clear that the central IPD NMA research team are responsible for making final decisions (eg, about design, IPD included, risk of bias judgements, analysis methods), while still valuing advice from the trial investigators. An independent advisory group may facilitate this.

Inevitably, the requested IPD may not be available from all studies, leading to availability bias concerns,[54] and the need for methods to combine IPD and AD.[48 50 55] Multilevel network meta-regression extends the IPD NMA framework to incorporate IPD and AD (with full-IPD NMA as a special case),[48] and is implemented in the multinma R package.[56] The approach avoids aggregation bias by integrating the individual-level regression model over the covariate distributions in each aggregate study population, and can produce estimates in any target population of interest. Other methods including matching-adjusted indirect comparison, simulated treatment comparison and predictive-adjusted indirect comparison have also been proposed for 'population adjustment' with limited IPD, but are limited to a two-study indirect comparison (one IPD and one AD study) and can only produce estimates relevant to the population of the AD study.[49]

Concluding remarks

In summary, the use of IPD adds value to NMA projects by improving quality (eg, through improved homogeneity and consistency in the network) and scope (eg, additional outcomes and longer follow-up), leading to more reliable and tailored NMA results for clinical practice. In the coming years, we anticipate further methodological research to improve and extend IPD NMA projects, and the website www.ipdma.co.uk provides signposts to new methodological developments.

ORCID iD
Richard D Riley http://orcid.org/0000-0001-8699-0735

References

1 **Dias S**, Ades AE, Welton NJ. *Network meta-analysis for decision making.* Chichester: Wiley, 2018.

2 **Riley RD**, Tierney JF, Stewart LA, eds. *Individual Participant Data Meta-Analysis: A Handbook for Healthcare Research.* Chichester, West Sussex: Wiley, 2021.

3 **Riley RD**, Dias S, Phillippo DM. Network Meta-Analysis Using IPD. In: **Riley RD**, Tierney JF, Stewart LA, eds. *Individual participant data meta-analysis: a Handbook for healthcare research.* Chichester, West Sussex: Wiley, 2021.

4 **Salanti G**. Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool. *Res Synth Methods* 2012;3:80–97.

5 **White IR**. Multivariate Random-effects meta-regression: updates to Mvmeta. *Stata J* 2011;11:255–70.

6 **White IR**. Network meta-analysis. *Stata J* 2015;15:951–85.

7 **Lu G**, Ades AE. Assessing evidence inconsistency in mixed treatment comparisons. *J Am Stat Assoc* 2006;101:447–59.

8 **Lu G**, Ades A. Modeling between-trial variance structure in mixed treatment comparisons. *Biostatistics* 2009;10:792–805.

9 **Cipriani A**, Higgins JPT, Geddes JR, *et al*. Conceptual and technical challenges in network meta-analysis. *Ann Intern Med* 2013;159:130–7.

10 **Jansen JP**, Trikalinos T, Cappelleri JC, *et al*. Indirect treatment comparison/network meta-analysis study questionnaire to assess relevance and credibility to inform health care decision making: an ISPOR-AMCP-NPC good practice Task force report. *Value Health* 2014;17:157–73.

11 **Song F**, Xiong T, Parekh-Bhurke S, *et al*. Inconsistency between direct and indirect comparisons of competing interventions: meta-epidemiological study. *BMJ* 2011;343:d4909.

12 **Jansen JP**, Naci H. Is network meta-analysis as valid as standard pairwise meta-analysis? it all depends on the distribution of effect modifiers. *BMC Med* 2013;11:159.

13 **Donegan S**, Williamson P, D'Alessandro U, *et al*. Combining individual patient data and aggregate data in mixed treatment comparison meta-analysis: individual patient data may be beneficial if only for a subset of trials. *Stat Med* 2013;32:914–30.

14 **Debray TP**, Schuit E, Efthimiou O, *et al*. An overview of methods for network meta-analysis using individual participant data: when do benefits arise? *Stat Methods Med Res* 2018;27:1351–64.

15 **Stewart LA**, Clarke MJ. Practical methodology of meta-analyses (overviews) using updated individual patient data. Cochrane Working group. *Stat Med* 1995;14:2057–79.

16 **Stewart LA**, Tierney JF. To IPD or not to IPD? advantages and disadvantages of systematic reviews using individual patient data. *Eval Health Prof* 2002;25:76–97.

17 **Tierney JF**, Vale C, Riley R, *et al*. Individual participant data (IPD) meta-analyses of randomised controlled trials: guidance on their use. *PLoS Med* 2015;12:e1001855.

18 **Tierney JF**, Stewart LA, Clarke M. Individual Participant Data. In: **Higgins JPT**, Chandler TJ, Cumpston M, eds. *Cochrane Handbook for systematic reviews of interventions.* London: Cochrane, 2019.

19 **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, eds. *Publication bias in meta-analysis: prevention, assessment and adjustments.* Chichester: John Wiley & Sons, 2005: 261–86.

20 **Tierney JF**, Fisher DJ, Burdett S, *et al*. Comparison of aggregate and individual participant data approaches to meta-analysis of randomised trials: an observational study. *PLoS Med* 2020;17:e1003019.

21 **Courtney DB**, Watson P, Chan BW, *et al*. Forks in the road: definitions of response, remission, recovery, and other dichotomized outcomes in randomized controlled trials for adolescent depression. A scoping review. *Depress Anxiety* 2021;38:1152–68.

22 **Karyotaki E**, Efthimiou O, Miguel C, *et al*. Internet-Based cognitive behavioral therapy for depression: a systematic review and individual patient data network meta-analysis. *JAMA Psychiatry* 2021;78:361–71.

23 **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.

24 **Latimer NR**. Survival analysis for economic evaluations alongside clinical trials-extrapolation with patient-level data: inconsistencies, limitations, and a practical guide. *Med Decis Making* 2013;33:743–54.

25 **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;33:3844–58.

26 **Jansen JP**. Network meta-analysis of survival data with fractional polynomials. *BMC Med Res Methodol* 2011;11:61.

27 **Donegan S**, Dias S, Tudur-Smith C, *et al*. Graphs of study contributions and covariate distributions for network meta-regression. *Res Synth Methods* 2018;9:243–60.

28 **Batson S**, Score R, Sutton AJ. Three-Dimensional evidence network plot system: covariate imbalances and effects in network meta-analysis explored using a new software tool. *J Clin Epidemiol* 2017;86:182–95.

29 **Cope S**, Zhang J, Saletan S, *et al*. A process for assessing the feasibility of a network meta-analysis: a case study of everolimus in combination with hormonal therapy versus chemotherapy for advanced breast cancer. *BMC Med* 2014;**12**:93.

30 **Petit C**, Blanchard P, Pignon JP, *et al*. Individual patient data network meta-analysis using either restricted mean survival time difference or hazard ratios: is there a difference? A case study on locoregionally advanced nasopharyngeal carcinomas. *Syst Rev* 2019;**8**:96.

31 **Leahy J**, O'Leary A, Afdhal N, *et al*. The impact of individual patient data in a network meta-analysis: an investigation into parameter estimation and model selection. *Res Synth Methods* 2018;**9**:441–69.

32 **Ollier E**, Blanchard P, Le Teuff G, *et al*. Penalized Poisson model for network meta-analysis of individual patient time-to-event data. *Stat Med* 2022;**41**:340–55.

33 **Freeman SC**, Carpenter JR. Bayesian one-step IPD network meta-analysis of time-to-event data using Royston-Parmar models. *Res Synth Methods* 2017;**8**:451–64.

34 **Gasparrini A**, Armstrong B, Kenward MG. Multivariate meta-analysis for non-linear and other multi-parameter associations. *Stat Med* 2012;**31**:3821–39.

35 **Hong H**, Fu H, Price KL, *et al*. Incorporation of individual-patient data in network meta-analysis for multiple continuous endpoints, with application to diabetes treatment. *Stat Med* 2015;**34**:2794–819.

36 **Riley RD**, Jackson D, Salanti G, *et al*. Multivariate and network meta-analysis of multiple outcomes and multiple treatments: rationale, concepts, and examples. *BMJ* 2017;**358**:j3932.

37 **Efthimiou O**, Mavridis D, Riley RD, *et al*. Joint synthesis of multiple correlated outcomes in networks of interventions. *Biostatistics* 2015;**16**:84–97.

38 **Cipriani A**, Barbui C, Salanti G, *et al*. Comparative efficacy and acceptability of antimanic drugs in acute mania: a multiple-treatments meta-analysis. *Lancet* 2011;**378**:1306–15.

39 **Hernández AV**, Eijkemans MJC, Steyerberg EW. Randomized controlled trials with time-to-event outcomes: how much does prespecified covariate adjustment increase power? *Ann Epidemiol* 2006;**16**:41–8.

40 **Maas AIR**, Steyerberg EW, Marmarou A, *et al*. IMPACT recommendations for improving the design and analysis of clinical trials in moderate to severe traumatic brain injury. *Neurotherapeutics* 2010;**7**:127–34.

41 **Roozenbeek B**, Maas AIR, Lingsma HF, *et al*. Baseline characteristics and statistical power in randomized controlled trials: selection, prognostic targeting, or covariate adjustment? *Crit Care Med* 2009;**37**:2683–90.

42 **Hernández AV**, Steyerberg EW, Habbema JDF. Covariate adjustment in randomized controlled trials with dichotomous outcomes increases statistical power and reduces sample size requirements. *J Clin Epidemiol* 2004;**57**:454–60.

43 **Kahan BC**, Morris TP. Improper analysis of trials randomised using stratified blocks or minimisation. *Stat Med* 2012;**31**:328–40.

44 **Kanters S**, Karim ME, 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* 2021;**21**:21.

45 **Riley RD**, van der Windt D, Croft P. *Prognosis research in healthcare: concepts, methods and impact*. Oxford, UK: Oxford University Press, 2019.

46 **Fisher DJ**, Carpenter JR, Morris TP, *et al*. Meta-analytical methods to identify who benefits most from treatments: daft, deluded, or deft approach? *BMJ* 2017;**356**:j573.

47 **Riley RD**, Debray TPA, Fisher D, *et al*. Individual participant data meta-analysis to examine interactions between treatment effect and participant-level covariates: statistical recommendations for conduct and planning. *Stat Med* 2020;**39**:2115–37.

48 **Phillippo DM**, Dias S, Ades AE, *et al*. Multilevel network meta-regression for population-adjusted treatment comparisons. *J R Stat Soc Ser A Stat Soc* 2020;**183**:1189–210.

49 **Phillippo DM**, Ades AE, Dias S. NICE DSU technical support document 18: methods for population-adjusted indirect comparisons in submission to NICE, 2016. Available: wwwnicedsuorguk

50 **Phillippo DM**. Calibration of treatment effects in network meta-analysis using individual patient data University of Bristol; 2019.

51 **Donegan S**, Williamson P, D'Alessandro U, *et al*. Assessing the consistency assumption by exploring treatment by covariate interactions in mixed treatment comparison meta-analysis: individual patient-level covariates versus aggregate trial-level covariates. *Stat Med* 2012;**31**:3840–57.

52 **Hoogland J**, IntHout J, Belias M, *et al*. A tutorial on individualized treatment effect prediction from randomized trials with a binary endpoint. *Stat Med* 2021;**40**:5961–81.

53 **Seo M**, White IR, Furukawa TA, *et al*. Comparing methods for estimating patient-specific treatment effects in individual patient data meta-analysis. *Stat Med* 2021;**40**:1553–73.

54 **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.

55 **Thom HHZ**, Capkun G, Cerulli A, *et al*. Network meta-analysis combining individual patient and aggregate data from a mixture of study designs with an application to pulmonary arterial hypertension. *BMC Med Res Methodol* 2015;**15**:34.

56 multinma: Bayesian network meta-analysis of individual and aggregate data. R package version 0.4.2, 2022. Available: https://dmphillippo.github.io/multinma/