Commentary: Like it and lump it? Meta-analysis using individual participant data

Richard D Riley

Department of Public Health, Epidemiology and Biostatistics, Public Health Building, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK. E-mail: r.d.riley@bham.ac.uk

Accepted 28 June 2010

Meta-analysis—the synthesis of quantitative information from related studies—is one of the most popular research methods of the past 20 years. Though the large majority of meta-analyses synthesise aggregate data (e.g. hazard ratio estimates) obtained from publications or investigators, there have been repeated calls to facilitate meta-analysis of individual participant data (IPD),1 where the raw participant-level data are obtained from each study and synthesized. IPD is the original source material, which brings many opportunities over the aggregate data approach,2 like standardizing statistical analyses in each study; deriving desired summary results directly, independent of study reporting; checking modelling assumptions; and assessing participant-level effects, interactions and non-linear trends. The approach is increasingly applied,2 despite the extra cost, time and complexity required to obtain and manage raw data.3 Commendable examples of IPD meta-analysis are those conducted by the Emerging Risk Factors Collaboration (ERFC),4 who have remarkably gathered IPD from 116 prospective studies and over 1.2 million participants.

While meta-analysis methods using aggregate data are well established and fairly routine, methods for IPD meta-analysis are more complex and not well known. They require statistical models specific to the type of data under investigation and must account for the clustering of patients within studies, while appropriately synthesizing effects of interest across studies. The article in this issue of the IJE by Thompson et al.5 is thus both pertinent and timely. The authors consider IPD meta-analysis to identify risk factors for a time-to-event outcome, as in the ERFC. I will now expand on a few issues in the article and emphasize some methodological challenges that remain.

The authors describe two-step IPD meta-analysis methods, where Cox models are fitted within each study separately and then relevant parameter estimates pooled across studies. This methodology is sound, but I agree with their comment that in principle a one-step approach is preferred, where the IPD are analysed simultaneously in a single model that accounts for clustering of participants within studies. One-step hierarchical Cox models with random-effects are achievable,6 and a single modelling framework is more convenient. The biggest advantage of the one-step approach arises when modelling non-linear trends and dealing with non-proportional hazards, as it more easily allows established procedures like fractional polynomials7 and cubic splines8 to be implemented. However, the one-step approach is computationally intensive,6 especially in situations with large numbers of participants, but this problem will ease as software and computational speed improves. The one-step approach may also be more feasible when using flexible parametric survival models,9 rather than the Cox model. These allow the baseline...
hazard function to be modelled in each study, and so one could also estimate absolute differences in hazard rates, not just relative ones, to inform individual risk prediction. Indeed, the methods proposed by Thompson et al. do not address risk prediction and the development of prediction models (and subsequent clinical decision rules), but IPD from multiple studies can achieve this, as described by Royston et al.

The data set used by the authors involves 31 studies and 154,311 patients from the ERFC; thus they are in a privileged, and rare, position of having a large amount of IPD available. However, should one embark on an IPD meta-analysis when only few studies provide their IPD, such that estimating random effects is difficult? Similarly, is an IPD meta-analysis reliable when only a proportion of existing studies provide IPD? This situation is common, and raises the threat of what I call availability bias—that is, where studies that provide IPD are a biased subset of all existing studies. In particular, IPD may be less obtainable from studies with non-significant findings—such studies are more susceptible to remaining unidentified due to reporting biases; are perhaps more likely to have lost or destroyed their IPD; and, as their findings have less impact, may involve less recognized authors who are less likely to be invited into an IPD meta-analysis collaboration. Thompson et al. refer to this issue as ‘constraints on comprehensiveness in IPD meta-analysis’ but, due to the large number of studies and participants, feel that ‘although publication and reporting biases are potential concerns in all meta-analyses, they may be less so in the ERFC’.

Interestingly, the contour-enhanced funnel plot of the 31 studies included in the meta-analysis (Figure 1) shows evidence of small-study effects (that is, the tendency for smaller studies to provide more positive findings) with smaller studies to the bottom left of the plot (toward non-significance and a hazard ratio of 1) potentially missing. Egger’s test provides evidence of small-study effects at the conventional 10% level for this test ($P = 0.08$). Thus, the summary result is potentially upwardly biased here in favour of fibrinogen. Admittedly, availability bias is only one potential cause of this asymmetry (heterogeneity and the play of chance are two other reasons, for example), but it illustrates that—even in one of the largest and comprehensive IPD meta-analyses in existence—methodological concerns remain.

One way to limit availability bias is to consider combining IPD studies with aggregate data from other ‘non-IPD’ studies. In some of the simpler analyses proposed by Thompson et al., where just one or a few variables are included in the within-study Cox models, non-IPD studies may directly provide suitable hazard ratio estimates to be included in the second-step of the IPD framework. If IPD studies allow adjustment for multiple variables not (fully) available in non-IPD studies, one could implement the novel multivariate framework proposed by Jackson et al. to utilize both IPD and non-IPD studies.

Finally, I must echo the author’s suggestion that assessments of interactions between two or more participant-level variables are best based on within-study information using IPD. Between-study interactions are based on aggregated participant
characteristics (e.g. proportional male) and thus have low power to identify interactions at the participant-level;\textsuperscript{18} they are also susceptible to confounding and ecological bias.\textsuperscript{14–16} If researchers decide to combine within-study and between-study interactions, perhaps to increase statistical precision, then this is best treated as a sensitivity analysis as it makes the (strong) assumption that the between-study information is unbiased.

In summary, I congratulate Thompson et al.\textsuperscript{5} for their excellent and timely guidance on IPD meta-analyses of time-to-event data. I am also a strong advocate of the IPD approach and believe the research community should facilitate it where necessary. But let's not forget that methodological problems and statistical issues remain when conducting IPD meta-analyses; thus—although I like it and lump it—IPD is not the be-all and end-all for meta-analysis just yet.

**Conflict of interest:** None declared.

**References**
\textsuperscript{1} Oxman AD, Clarke MJ, Stewart LA. From science to practice. Meta-analyses using individual patient data are needed. *JAMA* 1995;274:845–46.