ELSEVIER

Journal of Clinical Epidemiology 60 (2007) 431–439

Journal of Clinical Epidemiology

REVIEW ARTICLE

Evidence synthesis combining individual patient data and aggregate data: a systematic review identified current practice and possible methods

Richard D. Riley[a,*], Mark C. Simmonds[b], Maxime P. Look[c]

[a]*Centre for Medical Statistics and Health Evaluation, School of Health Sciences, University of Liverpool, UK*
[b]*Medical and Pharmaceutical Statistics Research Unit, The University of Reading, UK*
[c]*Department of Medical Oncology, Erasmus MC Rotterdam, Rotterdam, The Netherlands*

Accepted 12 September 2006

Abstract

Objective: Meta-analysis of individual patient data (IPD) is the ''gold-standard'' for synthesizing evidence across several studies. Some studies, however, may only provide aggregate data (AD). In this situation researchers might need to combine IPD with AD to utilize all the evidence available. Here, we review applied IPD meta-analysis articles to assess if and how AD is combined with IPD in practice.

Study Design and Setting: A systematic review of articles identified from bibliographic databases and searches.

Results: We identified 33 applied IPD articles that combined IPD and AD and 166 that did not. For each article, we recorded the proportion of total studies providing IPD, and found that articles combining IPD and AD had, on average, IPD available in only 64% of studies (compared to 90% in articles not combining IPD and AD). Two different methods were used to combine IPD and AD, the two-stage method and analysis of partially reconstructed IPD, but a review of methodological articles identified two further methods, multilevel modeling and Bayesian hierarchical related regression. We summarize each method to aid practitioners.

Conclusion: Combining IPD and AD is a relevant issue for evidence synthesis, and the further development and validation of suitable meta-analysis methods is needed.

Keywords: Aggregate data; Evidence synthesis; Individual patient data; Meta-analysis; Summary data; Systematic review

1. Introduction

Methods for evidence synthesis are common in medical research. In particular, meta-analysis is increasingly used to combine the quantitative evidence across clinical studies and develop results based on a whole body of research [1,2]. A traditional meta-analysis involves the synthesis of aggregate data (AD) available from the individual study publications or directly from the study authors themselves. Meta-analysis of AD is also sometimes referred to as meta-analysis of summary data or of literature-based results. Typical AD include a mean difference for continuous outcomes, a log-hazard ratio for time-to-event outcomes, and the number of events and participants for binary outcomes. Meta-analysis then produces a weighted average of the AD across studies to give an overall measure of (treatment) effect, such as a pooled odds ratio for binary outcomes.

Meta-analysis of individual patient data (IPD), where the raw data from each study is obtained and synthesized directly, is an alternative to the AD approach and is termed the ''gold-standard'' as it has numerous advantages [3]. For example, it avoids the biases of published AD, it allows one to obtain information not available from the published reports, and one can use consistent inclusion/exclusion criteria across studies. Other advantages are more specific to the type of data and health care area under question. For example, in prognostic studies IPD allows a longer follow-up time to be assessed and enables sophisticated modeling techniques [4,5]. In areas like breast cancer, where treatment needs assessment alongside patient characteristics (such as nodal status and menopausal status), IPD allows *patient-level* covariates to be directly modeled. This method has greater statistical power to detect true patient–treatment relationships when compared to a meta-regression of AD, which assesses treatment in relation to *group-level* summaries (e.g., proportion postmenopausal) [6].

Meta-analyses of IPD are increasingly common; Simmonds et al. [7] identified 44 IPD meta-analyses published

* Corresponding author. Centre for Medical Statistics and Health Evaluation, School of Health Sciences, University of Liverpool, Shelley's Cottage, Brownlow Street, Liverpool, England L69 3GS, Tel.: +44-51-7945121; fax: +44-151-7945130.

E-mail address: richard.riley@liv.ac.uk (R.D. Riley).

0895-4356/07/$ – see front matter
doi: 10.1016/j.jclinepi.2006.09.009

during the years 1999–2001. One potential problem for IPD analysis is that IPD may not be available from all the studies, perhaps due to a loss or destruction of data or an unwillingness to collaborate [8]. The Simmonds et al. [7] review found that about one in every four IPD meta-analyses obtain IPD from fewer than 80% of all the randomized patients across studies. In such situations, Stewart and Tierney [8] advise that results from an IPD-only meta-analysis may be biased if unavailability of IPD is related to the study results. It may therefore be important to supplement the available IPD with AD for those studies where IPD are not available. For example, a review [9] of combination chemotherapy versus melphalan plus prednisone as treatment for multiple myeloma involved 20 studies (4,930 patients) with IPD and seven other studies (1,703 patients) with AD from their publications. The incorporation of such AD here allows a larger number of patients and a greater part of the evidence-base to be included.

The issue of combining IPD and AD in evidence synthesis would thus appear important but it has received little attention. To address this, and to aid practitioners, we conducted a systematic review of applied IPD meta-analyses to assess if and how IPD and AD are combined in practice (Section 2). Relevant methodological articles were also identified and assessed. The results of the review are given in Section 3 and the extent to which combining IPD and AD is important to applied evidence synthesis is examined. The relevant meta-analysis methods identified by the review are described and appraised in Section 4, with a discussion of how they were applied in practice. A critique of the review and discussion of the issues raised in relation to future IPD reviews and methodological research are given in Section 5.

2. Methods

2.1. Search strategy

We performed a systematic review to identify relevant articles from the literature that considered the joint synthesis of IPD and AD, either as part of a clinical overview (e.g., in a Cochrane systematic review) or as part of a methodological article (e.g., those in statistical journals).

We searched Medline (from 1966 to week 24 of 2005), Embase (from 1980 to week 24 of 2005), Cochrane Central and the Cochrane Methodology Register (up to Cochrane Library 2005, issue 2), and also MathSciNet (up to the 27th June, 2005), using a search strategy based on a similar review elsewhere [10] (for full details see the journal's web site at www.elsevier.com). The articles identified from each database were merged and any duplicates removed. Articles mostly comprised full journal papers but a small number of conference abstracts were also identified.

2.2. Inclusion and exclusion criteria

The first author read the abstracts of all the identified articles and categorized those potentially relevant into one of two groups, ''applied'' or ''methodological'' (Fig. 1).

For an article to be included in the ''applied'' group it needed to be an applied clinical article disseminating findings from an actual IPD pooled analysis (e.g., an IPD meta-analysis within a Cochrane review). Pooled analyses involving all study types were included (e.g., randomized trials, observational studies). Where multiple articles from the same authors/collaborative group relating to the same study and the same objectives were identified (e.g., due to publication across multiple journals or due to the study being updated), only the most recent article was included. All conference abstracts and articles not written in English were excluded from the ''applied'' group. For an article to be included in the ''methodological'' group it needed to be a methodological paper (or methodological conference abstract) describing one or more methods for evidence synthesis of IPD.

To assess accuracy, the second author also independently classified about 5% of the identified articles, and similarly the third author independently classified a further 5%. Any resulting discrepancies between the first author's classifications and those of the second/third authors were resolved by making a general agreement after the article was obtained and read in full.

2.3. Identification of current practice and relevant methods for combining IPD and AD

Each of the articles classified as ''applied'' were obtained and read to establish if they combined IPD with AD and, if so, what methods they used. Where possible, the proportion of the total studies providing IPD and the proportion of the total patients represented by the IPD were recorded. The data extracted for each of the articles that combined IPD and AD were double-checked.

Each of the articles classified as ''methodological'' were also obtained, read in full, and deemed relevant if they proposed a method for combining IPD and AD. If the article was a conference abstract, the first author was e-mailed to ask for any related papers or detailed presentation slides. To identify any other relevant methodological articles not previously identified we then (i) checked the references of ''methodological'' articles; (ii) used Scopus to check which papers cited the ''methodological'' articles; (iii) read three meta-analysis textbooks [1,2,11], and (iv) performed two Internet searches in Google using ''individual data meta'' and ''individual data combining''. The different methods for combining IPD and AD that were suggested by the final set of relevant methodological articles were then documented in detail.

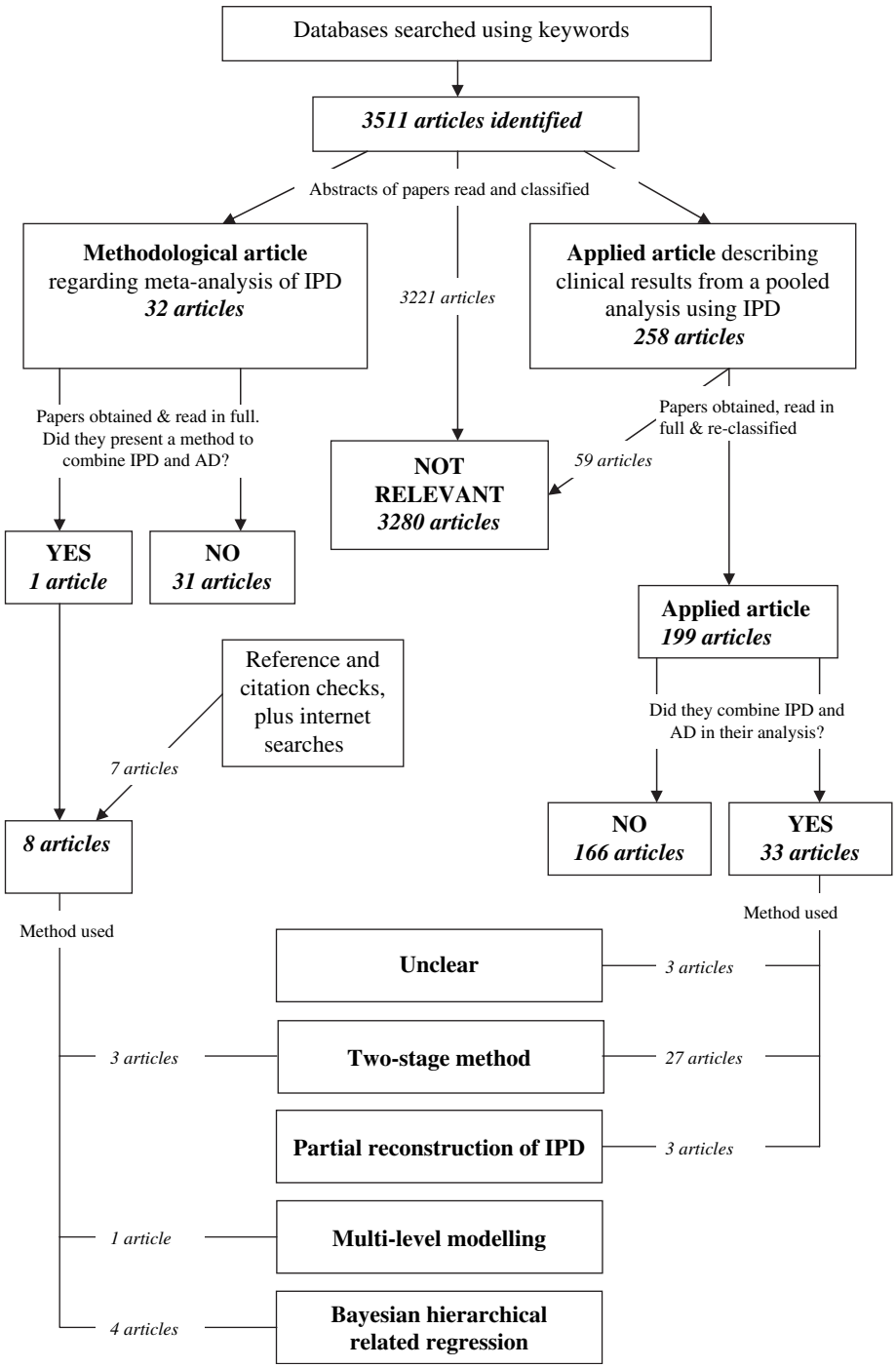

Fig. 1. Description and results of the search and classification process.

3. Results

The first author assessed 3,551 articles identified from the bibliographic databases, and classified 258 articles as ''applied'' and 32 articles as ''methodological'' (Fig. 1). The second and third authors also independently classified two distinct sets of 177 articles. Of these 354 articles, there were six rejected by the first author but were considered potentially relevant by the other authors. However, after further investigation, none of these six articles were found to apply or propose a method for combining IPD and AD.

3.1. Summary of applied IPD meta-analysis articles

After assessing in full the 258 ''applied'' articles, it was found that 59 of these were not relevant IPD meta-analyses (Fig. 1). Of the 199 relevant articles (all references available on request), 166 did not combine IPD and AD and

in 142 of these we could ascertain the proportion of the total studies providing IPD. This proportion ranged from 25% to 100%, with a mean and median of 90% and 100%, respectively; encouragingly 112 (79%) of the 142 articles obtained IPD for 80% or more of the total studies for which IPD were sought (Fig. 2a).

Thirty-three of the 199 articles did combine IPD and AD (references and full details available on the journal's web site at www.elsevier.com). Two different methods were used to do this—the two-stage method and analysis of partially reconstructed IPD (see Section 4). In three articles, the method used was unclear. We note that the observed commonality of methods across studies will have been partially augmented by two of the 33 articles being written by the Advanced Bladder Collaboration [12,13], four by members of the EU Hernia Trialists Collaboration [14–17], and three by Gorman et al. [18–20]; these articles, though similar, were all included as they had different main objectives.

Only 12 (36%) of these 33 articles obtained IPD for 80% or more of the studies (Fig. 2b), and the proportion of studies providing IPD ranged from 10% to 92%, with a mean and median value of 64% and 71%, respectively. In 27 of the 33 articles, we could also ascertain the percentage of patients for whom IPD were obtained. The patients represented by the IPD ranged from 11% to 98% of the total patients, with a mean and median of 67% and 68%, respectively; only 10 (37%) of these 27 articles obtained IPD for over 80% of the patients in the meta-analysis (Fig. 3).

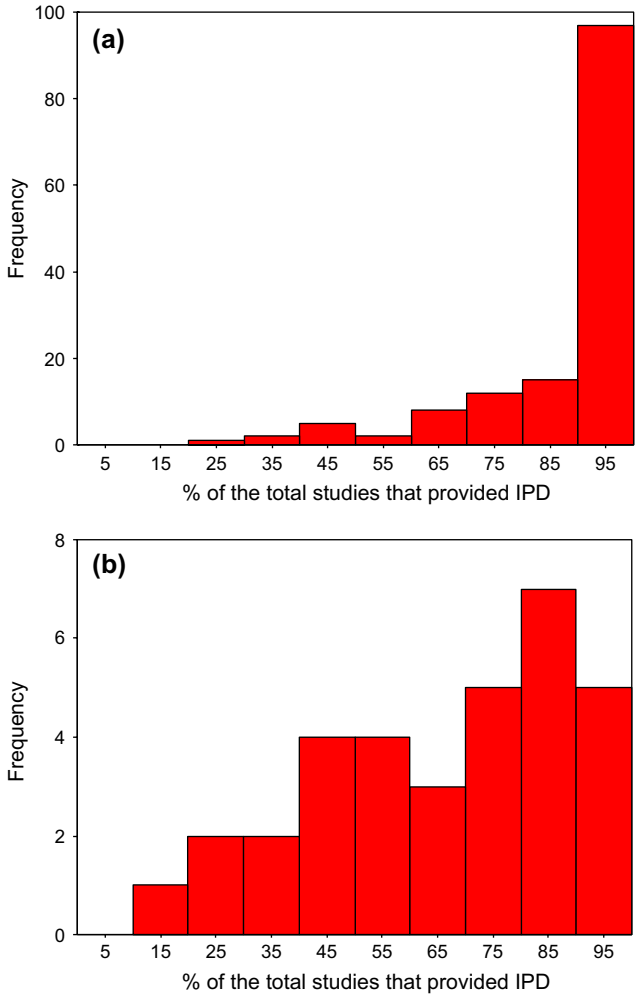

Fig. 2. Histogram showing the % of total studies that provided IPD for (a) 142 applied articles that only used IPD in their meta-analysis and (b) the 33 applied articles that combined IPD and AD in their meta-analysis.

3.2. Summary of methodological articles

Surprisingly, only one of the 32 articles initially classed as "methodological" provided relevant methodology for combining IPD and AD when assessed further; the one relevant article is a currently unpublished paper [21] generously made available to us after we identified the authors' related conference abstract. However, our additional citation, reference, and Internet searches allowed us to identify seven further relevant articles, comprising two journal papers [22,23], one "in press" paper [24], one thesis [25], two unpublished papers [26,27], and one set of conference presentation slides [28].

Hence, in total, a final set of eight relevant methodological articles was established [21–28], and they suggested three different methods for combining IPD and AD in evidence synthesis: the two-stage method, multilevel modeling, and Bayesian hierarchical related regression (see Section 4).

4. Summary of proposed meta-analysis methods to combine IPD and AD

4.1. The two-stage method

In the two-stage method the available IPD are first reduced to AD in each study, and then these AD (from the

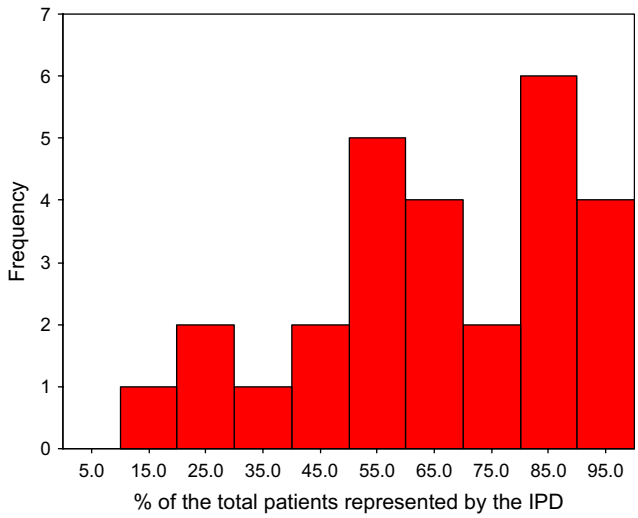

Fig. 3. Histogram showing the % of total patients represented by the IPD for 27 applied articles that combined IPD and AD in their meta-analysis.

IPD studies) are combined with the existing AD (from the AD studies) using standard meta-analysis of AD techniques, such as the Mantel–Haenszel method [29] and the inverse-variance fixed or random-effects approach [30]. The two-stage method is thus relatively simple to apply in practice; however, as it ignores the patient-level information in the IPD, it is best suited to where only the overall pooled (treatment) effect is of interest in relation to study-level covariates (i.e., covariates that are identical for each patient in a study, e.g., method of biomarker measurement in a tumor marker study).

4.1.1. Applied articles using this method

The two-stage method was the most commonly used method for combining IPD and AD in practice, with 27 of the 33 relevant applied articles using this approach. For example, in a meta-analysis of five advanced bladder cancer studies [12], the log-rank expected number of deaths and variance were calculated for each study, using the IPD in four IPD studies and the published hazard ratio and standard error in one AD study. Peto's fixed effect method [31] was then used to combine these AD and produce a pooled hazard ratio. Similar application of standard AD meta-analysis methods occurred in the other 26 articles using the two-stage method, with the exception of Mauer et al. [32] who did not appear to appropriately stratify by trial.

4.1.2. Methodological articles suggesting this method

Three of the eight relevant methodological articles suggest the two-stage method [21,22,25]. The name originates from Simmonds [25] (and concurrently [7]) who considers a variety of settings where the IPD studies are firstly reduced to AD, including where the AD relate to a treatment–covariate interaction. The two other articles [21,22] both show that the two-stage method for combining IPD and AD can have advantages over an IPD-only meta-analysis for time-to-event studies. For example, it enabled Tudur et al. [22] to maximize the number of trials, patients, and events included in the analysis; identify between-study heterogeneity; and gain an increased confidence in some of the results obtained. Collette et al. [21] note though that the two-stage method is limited when detailed time-to-event analyses like Kaplan–Meier curves and subgroup analyses are required, as direct modeling of IPD will usually be necessary.

4.2. Analysis of partially reconstructed IPD

For binary outcome data one can partially reconstruct IPD from the aggregated information in published 2×2 tables. For example, if the number of patients who were dead and alive for each of a treatment group and a control group are known, one can create IPD in a binary data format, where control/treatment group and alive/dead are represented by a series of zeros and ones. This approach allows the reconstructed IPD to then be combined with the IPD already available using standard methods for meta-analysis of IPD with binary outcomes such as logistic regression [33]. For ordinal outcome data, recreating IPD may also be possible given information on the number of responses falling into each category for each treatment group, and methods for meta-analysis of IPD with ordinal outcomes could then be used [34].

As for the two-stage method, this approach is best suited to where the pooled (treatment) effect is of interest in relation to study-level covariates, because it is difficult to reconstruct patient-level covariates from AD and so modeling patient–treatment relationships is very hard. There has so far been little validation of using reconstructed IPD in practice; reconstructed IPD may not be equivalent to the original IPD from a study and could be biased, perhaps due to patients being excluded, inappropriate analyses being performed, or selective reporting of results.

4.2.1. Applied articles using this method

Three of the 33 relevant applied articles used this method. In one of these, Hukkelhoven et al. [35] use logistic regression to model IPD from 2,664 patients and AD from 2,948 patients to assess the strength of the association between age and outcome in severe traumatic brain injury. In the AD studies, age was typically reported in categories with a range of values, for example, 10–20 years, and so the mean age in each category was used in the modeling. The authors indicate that this may have led to the AD studies providing less reliable information than the IPD studies, and so accordingly they also apply logistic regression models to just the IPD studies. This example highlights the difficulty in analyzing partially reconstructed IPD to assess patient-level covariates.

4.2.2. Methodological articles suggesting this method

None of the eight relevant methodological articles used this method. However, the idea of reconstructing IPD from tabulated binary or ordinal data has been suggested in the literature [33,34,36], although not specifically in the context of combining AD studies and IPD studies. These papers show how to directly model individual binary (or ordinal) patient responses, and Thompson et al. [36] suggest that this is particularly important when the observed event probabilities in a study are close to 0 or 1.

4.3. Multilevel modeling

A meta-analysis of IPD studies can be viewed as a multilevel model, in which the highest level is that of the study and the lowest level that of an observation from an individual patient. This framework therefore allows one to estimate the effect of interest in relation to both study-level covariates and patient-level covariates. For continuous outcome data, an extended multilevel framework has been proposed [23] that also allows AD studies to be incorporated alongside IPD studies by including a dummy variable in

the model that distinguishes AD responses from IPD responses. This enables both the AD studies and the IPD studies to contribute toward estimation of the overall effect and study-level covariates, but appropriately it only lets the IPD studies estimate patient-level covariate effects.

4.3.1. Applied articles using this method

None of the 33 relevant applied articles used this method.

4.3.2. Methodological articles suggesting this method

In one of the eight relevant methodological articles, Goldstein et al. [23] suggest a multilevel model for combining IPD and AD where continuous data responses are of interest. Their application is to studies of class sizes and achievements in schools, where in eight studies data are available at the aggregate level in terms of overall mean values for each class size and in one study data are available at the individual level relating to test scores for students within classes within schools. Their multilevel meta-analysis model has four levels relating to study, school, class size, and students, and the IPD study contributes to all these levels, whereas the AD studies only contribute toward estimating the mean test score across studies for each class. Goldstein et al. indicate their multilevel framework also facilitates modeling of multiple outcomes (e.g., test scores for both mathematics and reading), even if not all studies provide all the outcomes of interest. Multilevel modelling of both IPD and AD is thus potentially advantageous, but as yet it has received little empirical assessment or validation. Indeed, Goldstein et al. note that the inclusion of AD studies may be difficult when each study presents results adjusted for different factors. Consideration of the approach in more clinical settings is also needed.

4.4. Hierarchical related regression using a Bayesian framework

Ecological studies synthesize findings from across different groups, such as geographical areas, and they assume the group-level relationship between mean exposure and mean outcome is the same as the individual-level relationship between individual exposure and individual outcome. Jackson et al. [24,26] show that when the between-group variation in mean exposure is large, relative to the within-study variation in individual exposure, then this assumption is valid. However, in other settings group-level relationships will often not reflect the individual-level relationships [24], a phenomenon known as ecological bias [37].

As in clinical evidence synthesis, for ecological research some individual-level data (which we term IPD for simplicity) may be available alongside AD. However, such IPD usually involves a much smaller number of subjects than those contributing to the AD, and so, an IPD-only analysis may itself have low power to detect small variations in health. It has therefore been suggested [24,26,27] to combine the AD with the IPD using a Bayesian approach termed Hierarchical Related Regression (HRR). This uses Markov Chain Monte Carlo methods to simultaneously estimate two related regression models, the IPD-only model and the AD-only model, which are related by common parameters that assess the exposure-outcome relationship of interest. These common parameters are thus estimated from both the IPD and the AD, reducing the issues associated with analyzing IPD (low power) or AD (ecological bias) separately.

4.4.1. Applied articles using this method

None of the 33 relevant applied articles used this method.

4.4.2. Methodological articles suggesting this method

Four of the eight relevant methodological articles were written by the BIAS Project team (http://www.bias-project.org.uk) and they suggested HRR in the context of ecological research [24,26–28], although there are clear parallels with the issues of combining IPD and AD from clinical studies. For example, Jackson et al. [24] apply HRR to a data set from 255 electoral wards in London regarding the long-term illness among men aged between 45 and 59 years in relation to age, sex, ethnicity, and income. They also perform simulations and show that estimates from HRR can have a smaller bias and mean-square error then estimates from an AD-only analysis [24,26]. Further, even where there is a relatively large amount of IPD they show that HRR can obtain estimates with a smaller mean-square error than those from an IPD-only analysis. Best et al. [28] also show how HRR can be implemented through graphical models using WinBUGS. However, two potential difficulties of applying HRR in practice are [24] (i) forming appropriate prior distributions for the unknown parameters and (ii) ensuring that the IPD-only model is itself specified correctly. Further assessment and validation of HRR is thus needed, especially within more clinical settings where the available IPD are likely to represent a far greater proportion of the total evidence than within ecological settings.

5. Discussion

Meta-analysis of IPD is clearly the ‘‘gold-standard’’ [38] and we encourage researchers to primarily seek IPD rather than AD when conducting evidence synthesis. Our review of 199 applied IPD meta-analyses encouragingly shows that researchers often obtain a large proportion of the IPD required; IPD from 90% or more of the total number of studies were obtained in 102 (58%) of the 175 IPD articles providing this information, which concurs with Simmonds et al. [7] who estimated this to be 62% in a smaller sample. Yet our findings also show that often one may still require AD from a substantial number of studies, as 51 (29%) of

175 articles obtained IPD for less than 80% of the studies. Our review therefore demonstrates the potential importance of appropriate methods for combining IPD with AD in evidence synthesis. Indeed, protocols for future IPD reviews should specify if and how they would include AD alongside IPD in meta-analysis, if the need arose. To this end, the four methods we identified and summarized should provide a helpful starting point for practitioners and also for subsequent methodological research.

5.1. Critique of our review

To identify relevant applied articles in our review we used a thorough search strategy and assessed 3,511 articles, but it is possible some relevant articles have been missed or excluded unintentionally. For example, applied meta-analyses using IPD but not mentioning IPD (or a related term) in their abstract are likely to have been missed; such articles may well include those where the meta-analysis using IPD was not a main objective. Due to time-restraints, the second and third authors double-checked only 10% of the 3,511 classifications and similarly we only double-checked information extracted from 33 of the 199 applied IPD articles, so it is possible a few classification errors have occurred unintentionally. However, these potential limitations are unlikely to change the main conclusions, namely that obtaining AD alongside IPD is a relevant research issue and that the two-stage method is predominately used to combine IPD and AD in practice.

In terms of our review of methodological articles, it was surprising that only one of the eight relevant articles was identified from our extensive literature search. We could have made our search strategy even more inclusive, although more time-consuming, by also including articles that specified a ''meta-analysis'' related term in the abstract. However, our reference, citation, and Internet checks should have limited any potential deficiencies in our search strategy. Indeed, in retrospect using an extended strategy is unlikely to have greatly helped, as five of the other seven relevant articles were unpublished or ''in press''.

5.2. Advice for practitioners

The need to incorporate AD alongside IPD is driven by the desire to include all the available evidence in meta-analysis [39]. The concern is that simply ignoring AD studies and performing an IPD-only meta-analysis may distort the truth. Some, however, may conversely argue that AD (including ''reconstructed'' IPD as in Section 4.2) is itself far less reliable and more prone to bias than IPD, and so its inclusion alongside IPD may itself distort the truth. Practitioners should weigh both these arguments in their individual setting, perhaps by ascertaining the potential biases in AD studies (e.g., were AD supplied directly from trial investigators or from the publication?) and by exploring the differences between IPD studies and AD studies. For example, an assessment of hernia trials has recently been undertaken [40] and it was found that AD studies are of comparable quality to IPD studies, which perhaps explains why AD has been included in a number of IPD meta-analysis of such trials recently [14–16].

The argument to include AD undoubtedly becomes stronger as the amount of missing IPD increases, and in the 33 applied IPD articles that did use AD, the mean proportion of studies providing IPD was considerably less than in those articles where only IPD were used (64% compared to 90%). Those conducting IPD meta-analyses should thus consider, and importantly also clearly report, the proportion of the total studies which provided IPD, and perhaps even more crucially the proportion of the total patients represented by that IPD. In time-to-event studies one may also consider the proportion of the total events represented by the IPD.

Where inclusion of AD is deemed important, we recommend that practitioners perform both an IPD-only meta-analysis and a meta-analysis that combines the IPD and the AD. Many of the 33 articles that combined IPD and AD did so in this context, as it allows the robustness of the IPD-only conclusions to be assessed and the impact of including AD studies to be better understood. Indeed, at least until the advanced methods to combine IPD and AD are further developed and validated, IPD-only analyses will usually still be required to properly assess complex factors such as patient level-covariates and between-study heterogeneity. This was evident in the IPD articles we reviewed, with the majority assessing patient-level covariates in an IPD-only analysis. Of course in situations where AD is unavailable or poorly reported [41], an analysis of just the IPD studies may actually be the only viable option. Regardless of the approach taken, we encourage practitioners to clearly document their meta-analysis methods to help ensure transparency of their work.

5.3. Statistical methods

The two-stage method was by far the most common approach for combining IPD and AD in practice, which is perhaps not surprising given its simplicity, that it works for any type of patient data (e.g., binary, continuous, survival), and that it allows established meta-analysis of AD techniques to be used. For binary or ordinal data, analysis of partially reconstructed IPD is another simple approach and facilitates standard IPD methods, such as logistic regression modeling. Yet both the two-stage method and partial reconstruction of IPD are likely to be inadequate where results regarding specific patient-level characteristics (such as the stage of disease in cancer patients) are needed. The two-stage method ignores the patient-level information in the IPD altogether as it is all reduced to AD. The partial reconstruction of IPD method treats the AD studies in the same way as the IPD studies, but does not have patient-level information for the AD studies; thus, the logistic model must

either not include patient-level covariates or, as in Hukkelhoven et al. [35], use the mean patient covariate value for each AD study, which is potentially problematic.

In contrast to the two-stage method and analysis of partially reconstructed IPD, multilevel modeling and HRR distinguish between IPD and AD studies in the analysis. The multilevel approach includes a dummy variable to do this [23], whereas HRR simultaneously fits separate models for the AD and IPD studies [26]. By taking this approach, these methods are able to directly model patient-level information from the IPD and still include AD toward estimation of the overall and study-level effects. They therefore allow more complex modeling, and by specifying separate models the Bayesian HRR framework is particularly flexible and extendable. Indeed, we are aware of work in progress that uses Bayesian HRR to combine IPD and AD from clinical studies that varied in design (Sutton et al., personal communication), with separate AD-only and IPD-only models specified for each type of study. Bayesian approaches have also been used to synthesize multiple sources of evidence in other contexts [42], such as qualitative and quantitative findings [43]. Multilevel modeling and Bayesian HRR are clearly emerging and promising concepts for combining IPD and AD in evidence synthesis, but we recommend their further statistical assessment and validation, especially within more clinical settings, before they can be firmly recommended for practice.

5.4. Further methodological issues for IPD and AD studies

A few applied articles in our review, such as Scott et al. [17], used some AD provided directly by trial investigators and some AD extracted directly from the publication. AD from trial investigators is likely to be closer to IPD in terms of quality than AD from publications, which is much more prone to bias. However, none of the four methods described in Section 4 make any distinction between the type of AD included and so further research of this issue would be interesting. An additional challenge may arise when there is partially missing IPD within a study and therefore one wants to also incorporate AD for the remaining patients in such a study. There may also be the concern of known studies that provide neither IPD nor AD and of course the threat of (unknown) studies affected by publication and other dissemination biases, such as selective outcome reporting. Some authors [44,45] have suggested using correlation in the IPD studies to help reduce the problem of unavailable or poorly reported AD in other studies, and this approach is worthy of further consideration.

5.5. Conclusion

There is an increasing drive to make IPD more generally available for meta-analysis. Our review has shown that IPD meta-analyses are achievable and researchers should thus be encouraged to take this approach. Yet we have also shown that combining IPD and AD is also a relevant issue in practice and that this is currently an underresearched area. The further development, validation, and clinical application of suitable meta-analysis methods are therefore needed, building on the four methods identified in this paper.

Acknowledgments

Richard Riley is funded by the UK Department of Health as a Research Scientist in Evidence Synthesis. We thank David Jones for his helpful advice during the drafting of this paper and also the two reviewers whose comments have greatly improved the content of this paper.

References

[1] Egger M, Davey Smith G, Altman DG. Systematic reviews in health care: meta-analysis in context. London: BMJ Publishing Group; 2001.
[2] Whitehead A. Meta-analysis of Controlled Clinical Trials. West Sussex: Wiley; 2002.
[3] Stewart LA, Parmar MK. Meta-analysis of the literature or of individual patient data: is there a difference? Lancet 1993;341:418–22.
[4] 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.
[5] Trikalinos TA, Ioannidis JP. Predictive modeling and heterogeneity of baseline risk in meta-analysis of individual patient data. J Clin Epidemiol 2001;54:245–52.
[6] Lambert PC, Sutton AJ, Abrams KR, Jones DR. A comparison of summary patient-level covariates in meta-regression with individual patient data meta-analysis. J Clin Epidemiol 2002;55:86–94.
[7] Simmonds MC, Higgins JPT, Stewart LA, Tierney JF, Clarke MJ, Thompson SG. Meta-analysis of individual patient data from randomized trials: a review of methods used in practice. Clin Trials 2005;2:209–17.
[8] 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.
[9] Myeloma Trialists' Collaborative Group. Combination chemotherapy versus melphalan plus prednisone as treatment for multiple myeloma: an overview of 6,633 patients from 27 randomized trials. J Clin Oncol 1998;16:3832–42.
[10] Clarke M, Stewart L, Tierney J, Williamson P. Individual patient data meta-analyses compared with meta-analyses based on aggregate data. The Cochrane Database Methodol Rev 2005;Issue 4.
[11] Sutton AJ, Abrams KR, Jones DR, Sheldon TA, Song F. Methods for meta-analysis in medical research. London: John Wiley; 2000.
[12] Advanced Bladder Cancer Collaboration. Neoadjuvant cisplatin for advanced bladder cancer. Cochrane Database Syst Rev 2000;Issue 2.
[13] Advanced Bladder Cancer Collaboration. Neoadjuvant chemotherapy in invasive bladder cancer: a systematic review and meta-analysis. Lancet 2003;361:1927–34.
[14] EU Hernia Trialists Collaboration. Repair of groin hernia with synthetic mesh: meta-analysis of randomized controlled trials. Ann Surg 2002;235:322–32.
[15] Grant AM, Collaboration EHT. Laparoscopic versus open groin hernia repair: meta-analysis of randomised trials based on individual patient data. Hernia 2002;6(1):2–10.

[16] McCormack K, Scott N, Grant A. Are trials with individual patient data available different from trials without individual patient data available? 9th Annual Cochrane Colloquium Abstracts, Lyon. 2001.
[17] Scott NW, McCormack K, Graham P, Go PM, Ross SJ, Grant AM. Open mesh versus non-mesh for repair of femoral and inguinal hernia. Cochrane Database Syst Rev 2002;Issue 4.
[18] Gorman JD, David-Vaudey E, Pai M, Lum RF, Criswell LA. Lack of association of the HLA-DRB1 shared epitope with rheumatoid nodules: an individual patient data meta-analysis of 3,272 Caucasian patients with rheumatoid arthritis. Arthritis Rheum 2004;50:753–62.
[19] Gorman JD, David-Vaudey E, Pai M, Lum RF, Criswell LA. Particular HLA-DRB1 shared epitope genotypes are strongly associated with rheumatoid vasculitis. Arthritis Rheum 2004;50:3476–84.
[20] Gorman JD, Lum RF, Chen JJ, Suarez-Almazor ME, Thomson G, Criswell LA. Impact of shared epitope genotype and ethnicity on erosive disease: a meta-analysis of 3,240 rheumatoid arthritis patients. Arthritis Rheum 2004;2004(50):400–12.
[21] Collette L, Suciu S, Bijnens L, Sylvester R. Including literature data in individual patient data meta-analyses for time-to-event endpoints. Personal communication, based on abstract from the first Symposium on Systematic Reviews: beyond the basics, Oxford, 1998.
[22] Tudur C, Williamson PR, Kahan S, Best LY. The value of the aggregated data approach to meta-analysis with time-to-event outcomes. J Royal Stat Soc A 2001;164:357–70.
[23] Goldstein H, Yang M, Omar RZ, Turner RM, Thompson SG. Meta-analysis using multilevel models with an application to the study of class size effects. Appl Stat 2000;49:399–412.
[24] Jackson C, Best N, Richardson S. Improving ecological inference using individual-level data. Stat Med 2006;25:2136–59.
[25] Simmonds MC. Statistical methodology of individual patient data. PhD Thesis, University of Cambridge; 2005.
[26] Jackson C, Best N, Richardson S. Hierarchical related regression for combining aggregate and survey data in studies of socio-economic disease risk factors (Submitted for publication).
[27] Richardson S, Jackson C, Best N. Bayesian hierarchical models in ecological studies. Invited Paper for the International Workshop for Statistical Modelling, Sydney, 2005.
[28] Best N, Richardson S, Jackson C. Graphical models for combining multiple data sources. Presentation at the National Centre for Research Methods, Oxford, 2005.
[29] Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. J Natl Cancer Inst 1959;22: 719–48.
[30] DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7:177–88.
[31] Yusuf S, Peto R, Lewis J, Collins R, Sleight P. Beta blockade during and after myocardial infarction: an overview of the randomized trials. Prog Cardiovasc Dis 1985;17:335–71.
[32] Mauer DK, Nolan J, Plaisance P, Sitter H, Benoit H, Stiell IG, et al. Effect of active compression-decompression resuscitation (ACD-CPR) on survival: a combined analysis using individual patient data. Resuscitation 1999;41:249–56.
[33] 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.
[34] Whitehead A, Omar RZ, Higgins JP, Savaluny E, Turner RM, Thompson SG. Meta-analysis of ordinal outcomes using individual patient data. Stat Med 2001;20:2243–60.
[35] Hukkelhoven CW, Steyerberg EW, Rampen AJ, Farace E, Habbema JD, Marshall LF, et al. Patient age and outcome following severe traumatic brain injury: an analysis of 5600 patients. J Neurosurg 2003;99:666–73.
[36] Thompson SG, Turner RM, Warn DE. Multilevel models for meta-analysis, and their application to absolute risk differences. Stat Methods Med Res 2001;10:375–92.
[37] Greenland S. Quantitative methods in the review of epidemiologic literature. Epidemiol Rev 1987;9:1–30.
[38] Oxman AD, Clarke MJ, Stewart LA. From science to practice. Meta-analyses using individual patient data are needed. JAMA 1995;274: 845–6.
[39] Ioannidis JP, Contopoulos-Ioannidis DG, Lau J. Recursive cumulative meta-analysis: a diagnostic for the evolution of total randomized evidence from group and individual patient data. J Clin Epidemiol 1999;52:281–91.
[40] McCormack K, Scott N, Grant A. Are trials with individual patient data available different from trials without individual patient data available? 9th Annual Cochrane Colloquium Abstracts, Lyon. 2001.
[41] Riley RD, Abrams KR, Sutton AJ, Lambert PC, Jones DR, Heney D, et al. Reporting of prognostic markers: current problems and development of guidelines for evidence-based practice in the future. Br J Cancer 2003;88:1191–8.
[42] Peters JL, Rushton L, Sutton AJ, Jones DR, Abrams KR, Mugglestone MA. Bayesian methods for the cross design synthesis of epidemiological and toxicological evidence. RSS series C 2005;54:159–72.
[43] Roberts KA, Dixon-Woods M, Fitzpatrick R, Abrams KR, Jones DR. Factors affecting uptake of childhood immunisation: a Bayesian synthesis of qualitative and quantitative evidence. Lancet 2002;360: 1596–9.
[44] Friedberg B, Watine J, Miedouge M. Unresected colorectal liver metastases: prognostic value of laboratory variables. Gastroenterol Clin Biol 2001;25:962–6.
[45] Steyerberg EW, Eijkemans MJ, Van Houwelingen JC, Lee KL, Habbema JD. Prognostic models based on literature and individual patient data in logistic regression analysis. Stat Med 2000;19: 141–60.

Table 1
Search strategy used to identify articles in Medline and Embase

1. (individual patient$ adj6 data).ti,ab.
2. (individual patient$ adj6 report$).ti,ab.
3. (individual patient$ adj6 outcome$).ti,ab.
4. (individual patient$ adj6 level$).ti,ab.
5. (individual participant data).ti.ab.
6. ipd.ti,ab.
7. (individual subject$ adj6 data).ti,ab.
8. (individual subject$ adj6 report$).ti,ab.
9. (individual subject$ adj6 outcome$).ti,ab.
10. (individual subject$ adj6 level$).ti,ab.
11. (raw patient$ adj6 data).ti,ab.
12. (raw patient$ adj6 report$).ti,ab.
13. (raw patient$ adj6 outcome$).ti,ab.
14. (raw patient$ adj6 level$).ti,ab.
15. (raw subject$ adj6 data).ti,ab.
16. (raw subject$ adj6 report$).ti,ab.
17. (raw subject$ adj6 outcome$).ti,ab.
18. (raw subject$ adj6 level$).ti,ab.
19. idiopathic.ti,ab.
20. (immediate pigment darkening).ti,ab.
21. (intermittent peritoneal dialysis).ti,ab.
22. (invasive pneumococcal disease).ti,ab.
23. (indirect photometric detection).ti,ab.
24. (interaural phase disparity).ti,ab.
25. or/1-18
26. or/19-24
27. 25 not 26

Articles identified were those from line 27. The term ''ti,ab'' denotes that the word or phrase had to appear in the title or abstract of the paper; $ indicates that extended forms of the word were also accepted (e.g., ''patients'' was accepted as patient$ was specified); adj6 indicates that the following word must be within six words of the previous phrase; for example, for ''(individual patient$ adj6 data).ti,ab'' the word ''data'' had to be within six words of the phrase ''individual patient$'' for the paper to be included. Papers with any of terms 19–24 were excluded, as they were irrelevant but would otherwise have often been picked up by searching for ''IPD''. The same keywords and strategy as above was also used to search the Cochrane databases and MathSciNet, although the necessary notation was different.

Table 2
Description of the 33 applied articles that were found to combine IPD and AD

First author	Year of publication	No. trials with IPD	No. trials with AD	Proportion of trials with IPD	No. of patients with IPD	No. of patients with AD	Proportion of patients with IPD	Method used to combine IPD and AD
Advanced Bladder cancer [46]	2000	4	1	0.80	479	325	0.60	Two-stage method
Advanced bladder cancer [47]	2003	10	1	0.90	2,688	317	0.89	Two-stage method
Advanced colorectal cancer collaboration [48]	1992	9	1	0.90	1,381	429	0.87	Two-stage method
Amiodarone trials [49]	1997	11	2	0.85	6,252	301	0.95	Two-stage method
Birks [50]	2003	7	10	0.41	721	313	0.70	Two-stage method
Caro [51]	2002	8	3	0.73	233	108	0.68	Unclear
Childhood ALL Collaboration [52]	1996	38	4	0.90	10,752	1,199	0.90	Two-stage method
Craig [53]	2002	19	12	0.61	2,129	2,312	0.48	Two-stage method
EU Hernia Trialists [54]	2002	35	23	0.60	6,901	4,273	0.62	Two-stage method
Flossmann [55]	2003	5	43	0.10	4,643	12,196	0.28	Two-stage method
Flynn [56]	1998	7	16	0.30	?	?	?	Analysis of partially reconstructed IPD
Fried [57]	2001	5	7	0.42	102	282	0.27	Two-stage method
Gorman [58]	2004	22	2	0.92	2,925	347	0.89	Two-stage method
Gorman [59]	2004	12	2	0.86	1,378	190	0.88	Two-stage method
Gorman [60]	2004	25	4	0.86	2,804	436	0.87	Two-stage method
Grant [61]	2002	25	16	0.61	4,165	2,996	0.58	Two-stage method
Grines [62]	2003	10	1	0.91	2,635	90	0.97	Unclear
Hukkelhoven [63]	2003	4	11	0.27	2,664	2,948	0.47	Analysis of partially reconstructed IPD
Klerk [64]	2002	40	12	0.77	23,920	?	?	Two-stage method
Langhorne [65]	2005	9	2	0.82	?	?	?	Two-stage method
Leonardi-Bee [66]	2005	5	2	0.71	11,240	219	0.98	Analysis of partially reconstructed IPD
Marshall [67]	2001	4	5	0.44	594	922	0.39	Unclear
Martin [68]	2002	3	8	0.27	?	?	?	Two-stage method
Mauer [69]	1999	7	1	0.87	2,866	860	0.77	Two-stage method
McCormack [70]	2005	15	22	0.41	2,907	2,653	0.52	Two-stage method
Myeloma trialists [71]	1998	20	7	0.74	4,930	1,703	0.74	Two-stage method
Neal [72]	2000	6	11	0.35	8,472	67,452	0.11	Two-stage method
Rojas [73]	2005	1	1	0.50	1,320	1,243	0.52	Two-stage method
Scott [74]	2002	12	10	0.55	2,688	2,158	0.55	Two-stage method
Shekelle [75]	2003	6	6	0.50	?	?	?	Two-stage method
Simmonds [76]	2000	7	6	0.54	866	499	0.63	Two-stage method
Stocken [77]	2005	4	1	0.80	875	43	0.95	Two-stage method
Wade [78]	1997	6	2	0.75	1,395	?	?	Two-stage method

In most of the 33 articles the method used to combine IPD and AD was not explicitly stated but could be assumed from other available information (e.g., from the description of how the IPD-only meta-analysis was conducted or from the forest plots presented); however, in three articles it was especially unclear. Similarly, the proportion of patients represented by the IPD, and the proportion of studies represented by the IPD often had to be estimated from other information.

? indicates that information was not available and could not be estimated.

References for the 33 articles which combined IPD and AD

[46] Advanced Bladder Cancer Collaboration. Neoadjuvant cisplatin for advanced bladder cancer. Cochrane Database Syst Rev 2000:Issue 2.

[47] Advanced Bladder Cancer Collaboration. Neoadjuvant chemotherapy in invasive bladder cancer: a systematic review and meta-analysis. Lancet 2003;361:1927–34.

[48] Anonymous. Modulation of fluorouracil by leucovorin in patients with advanced colorectal cancer: evidence in terms of response rate. Advanced Colorectal Cancer Meta-Analysis Project. J Clin Oncol 1992;10:896–903.

[49] Amiodarone Trials Meta-Analysis Investigators. Effect of prophylactic amiodarone on mortality after acute myocardial infarction and in congestive heart failure: meta-analysis of individual data from 6500 patients in randomised trials. Lancet 1997;350:1417–24.

[50] Birks J, Flicker L. Selegiline for Alzheimer's disease. Cochrane Database Syst Rev 2003:Issue 1.

[51] Caro JJ, Huybrechts KF, Green TC. Estimates of the effect on hepatic iron of oral deferiprone compared with subcutaneous desferrioxamine for treatment of iron overload in thalassemia major: a systematic review. BMC Blood Disord 2002; 2:4.

[52] Childhood ALL Collaborative Group. Duration and intensity of maintenance chemotherapy in acute lymphoblastic leukaemia: overview of 42 trials involving 12000 randomised children. Lancet 1996;347: 1783–8.

[53] Craig JV, Lancaster GA, Taylor S, Williamson PR, Smyth RL. Infrared ear thermometry compared with rectal thermometry in children: a systematic review. Lancet 2002; 360:603–9.

[54] Collaboration EHT. Repair of groin hernia with synthetic mesh: meta-analysis of randomized controlled trials. Ann Surg 2002;235:322–32.

[55] Flossmann E, Rothwell PM. Prognosis of vertebrobasilar transient ischaemic attack and minor stroke. Brain 2003;126:1940–54.

[56] Flynn JT, Schiffman J, Feuer W, Corona A. The therapy of amblyopia: an analysis of the results of amblyopia therapy utilizing the pooled data of published studies. Trans Am Ophthalmol Soc 1998;96: 431–50.

[57] Fried LF, Orchard TJ, Kasiske BL. Effect of lipid reduction on the progression of renal disease: a meta-analysis. Kidney Int 2001; 59:260–9.

[58] Gorman JD, David-Vaudey E, Pai M, Lum RF, Criswell LA. Lack of association of the HLA-DRB1 shared epitope with rheumatoid nodules: an individual patient data meta-analysis of 3,272 Caucasian patients with rheumatoid arthritis. Arthritis Rheum 2004;50:753–62.

[59] Gorman JD, David-Vaudey E, Pai M, Lum RF, Criswell LA. Particular HLADRB1 shared epitope genotypes are strongly associated with rheumatoid vasculitis. Arthritis Rheum 2004;50:3476–84.

[60] Gorman JD, Lum RF, Chen JJ, Suarez-Almazor ME, Thomson G, Criswell LA. Impact of shared epitope genotype and ethnicity on erosive disease: a meta-analysis of 3,240 rheumatoid arthritis patients. Arthritis Rheum 2004;50:400–12.

[61] Grant AM, Collaboration EHT. Laparoscopic versus open groin hernia repair: meta-analysis of randomised trials based on individual patient data. Hernia 2002;6:2–10.

[62] Grines C, Patel A, Zijlstra F, Weaver WD, Granger C, Simes RJ, et al. Primary coronary angioplasty compared with intravenous thrombolytic therapy for acute myocardial infarction: six-month follow-up and analysis of individual patient data from randomized trials. Am Heart J 2003;145:47–57.

[63] Hukkelhoven CW, Steyerberg EW, Rampen AJ, Farace E, Habbema JD, Marshall LF, et al. Patient age and outcome following severe traumatic brain injury: an analysis of 5600 patients. J Neurosurg 2003;99:666–73.

[64] Klerk M, Verhoef P, Clarke R, Blom HJ, Kok FJ, Schouten EG, et al. MTHFR 677C–>T polymorphism and risk of coronary heart disease: a meta-analysis. JAMA 2002;288:2023–31.

[65] Langhorne P, Taylor G, Murray G, Dennis M, Anderson C, Bautz-Holter E, et al. Early supported discharge services for stroke patients: a meta-analysis of individual patients' data. Lancet 2005;365: 501–6.

[66] Leonardi-Bee J, Bath PMW, Bousser M-G, Davalos A, Diener H-C, Guiraud-Chaumeil B, et al. Dipyridamole for preventing recurrent ischemic stroke and other vascular events: a meta-analysis of individual patient data from randomized controlled trials. Stroke 2005;36: 162–8.

[67] Marshall M, Crowther R, Almaraz-Serrano A, Creed F, Sledge W, Kluiter H, et al. Systematic reviews of the effectiveness of day care for people with severe mental disorders: (1) acute day hospital versus admission; (2) vocational rehabilitation; (3) day hospital versus outpatient care. Health Technol Assess 2001;5:1–75.

[68] Martin J, Donaldson AN, Villarroel R, Parmar MK, Ernst E, Higginson IJ. Efficacy of acupuncture in asthma: systematic review and meta-analysis of published data from 11 randomised controlled trials. Eur Respir J 2002;20:846–52.

[69] Mauer DK, Nolan J, Plaisance P, Sitter H, Benoit H, Stiell IG, et al. Effect of active compression-decompression resuscitation (ACD-CPR) on survival: a combined analysis using individual patient data. Resuscitation 1999;41:249–56.

[70] McCormack K, Wake B, Perez J, Fraser C, Cook J, McIntosh E, et al. Laparoscopic surgery for inguinal hernia repair: systematic review of effectiveness and economic evaluation. Health Technol Assess 2005;9:iii–80.

[71] Group MTC. Combination chemotherapy versus melphalan plus prednisone as treatment for multiple myeloma: an overview of 6,633 patients from 27 randomized trials. J Clin Oncol 1998;16:3832–42.

[72] Neal B, MacMahon S, Chapman N, Blood Pressure Lowering Treatment Trialists' Collaboration. Effects of ACE inhibitors, calcium antagonists, and other blood-pressure lowering drugs: results of prospectively designed overviews of randomised trials. Blood Pressure Lowering Treatment Trialists' Collaboration. Lancet 2000;356: 1955–64.

[73] Rojas MP, Telaro E, Russo A, Moschetti I, Coe L, Fossati R, et al. Follow-up strategies for women treated for early breast cancer. Cochrane Database Syst Rev 2005; Issue 1.

[74] Scott NW, McCormack K, Graham P, Go PM, Ross SJ, Grant AM. Open mesh versus non-mesh for repair of femoral and inguinal hernia. Cochrane Database Syst Rev 2002; Issue 4.

[75] Shekelle PG, Rich MW, Morton SC, Atkinson CS, Tu W, Maglione M, et al. Efficacy of angiotensin-converting enzyme inhibitors and beta-blockers in the management of left ventricular systolic dysfunction according to race, gender, and diabetic status: a meta-analysis of major clinical trials. J Am Coll Cardiol 2003;41:1529–38.

[76] Simmonds PC. Palliative chemotherapy for advanced colorectal cancer: systematic review and meta-analysis. Colorectal Cancer Collaborative Group. BMJ 2000;321:531–5.

[77] Stocken DD, Buchler MW, Dervenis C, Bassi C, Jeekel H, Klinkenbijl JH, et al. Meta-analysis of randomised adjuvant therapy trials for pancreatic cancer. Br J Cancer 2005;92:1372–81.

[78] Wade CE, Grady JJ, Kramer GC, Younes RN, Gehlsen K, Holcroft JW. Individual patient cohort analysis of the efficacy of hypertonic saline/dextran in patients with traumatic brain injury and hypotension. J Trauma-Injury Infect Crit Care 1997;42:S6–5.