ELSEVIER

Journal of Clinical Epidemiology 63 (2010) 235–237

Journal of Clinical Epidemiology

COMMENTARY

Individual patient meta-analysis—rewards and challenges

Carl van Walraven[a,b,c,*]

[a]Department of Medicine, University of Ottawa, Ottawa, Ontario, Canada
[b]Clinical Epidemiology Program, Ottawa Health Research Institute, Ottawa, Ontario, Canada
[c]Institute for Clinical Evaluative Sciences, Ottawa, Ontario, Canada

Accepted 6 April 2009

Meta-analysis is an essential tool for summarizing medical research and determining the efficacy of therapies and procedures. Meta-analysis has been used in all areas of medical practice and aims to produce an overall estimate of the average treatment effect. Traditionally, this has been accomplished using information and results from published studies. Although such aggregate data meta-analyses (ADMA) are occasionally supplemented with additional data from study investigators, patient information in ADMA is always aggregated to the study level.

The 1990s saw the introduction of individual-patient meta-analyses (IPMA). In IPMA, primary data for each patient in all pertinent randomized trials are acquired and collated. Analyses are then conducted on these data. Compared with ADMA, IPMA hold many rewards, with the most important including the following.

Reward 1: Outcome harmonization

Randomized trials may collect data on a particular outcome but not report it in their publication. In addition, some studies may differ in their definitions of particular outcomes. Both these problems may be addressed by having access to individual patient data.

Reward 2: Analytical harmonization

Even when studies use the same outcome, their analyses may differ significantly by either process or methodology. Process differences include varying exclusion criteria and whether or not patients are analyzed using an intention-to-treat strategy. Depending on the data available, IPMA can standardize study inclusion criteria and ensure that all trials are analyzed using an intention to treat model.

Studies with the same outcome can also differ by their analytical methodology [1]. For example, patient survival can be compared between study groups using a binary outcome (e.g., proportion dead at a particular time point), a count outcome (e.g., number of deaths during observation period), or a time-to-event analysis. When studies use different methodologies to analyze the same outcome, collating results from each trial can be difficult using ADMAs. Having access to the primary data from each lets one harmonize the analysis to better attain an overall estimate of effect.

IPMA significantly simplifies the synthesis of some analytical methods. For time-to-event analyses, ADMA usually compares groups by determining the number of events that occur by a particular time point and calculating an odds ratio (OR). In contrast, IPMA can use survival analytical techniques to describe the intervention effect as a hazard ratio (HR). These two methods can produce startlingly different results. Stewart and Parmar used individual patient data to calculate an OR of survival at each year of follow-up in ovarian cancer patients with or without platinum-containing chemotherapy [2]. They found that the OR value and its statistical significance varied extensively depending on which ‘‘time point’’ was used during the observation period to calculate the OR. They also found important differences between the OR and HR, highlighting the dangers of analyzing survival data in this way.

IPMA is especially helpful for longitudinal or repeat measures of data. With such data, ADMA is difficult because studies frequently vary by the time when outcomes are measured. Even if all time points are measured in each study, meta-analyzing individual time points is inappropriate if serial measures are correlated [3]. Jones et al. [4] found that meta-analyses that ignored correlation between repeated measures had errors in their estimates of both treatment differences and standard errors. Appropriate longitudinal analysis is possible when individual patient data is available and avoids such biases.

* Corresponding author. Clinical Epidemiology Program, Ottawa Health Research Institute, ASB1-003, Ottawa Hospital, Civic Campus, 1053 Carling Avenue, Ottawa, Ontario K1Y 4E9, Canada. Tel.: +613-761-4903; fax: +613-761-5355.

E-mail address: carlv@ohri.ca

0895-4356/10/$ – see front matter
doi: 10.1016/j.jclinepi.2009.04.001

Reward 3: Exploration of efficacy variability

Although some exceptions exist [5], most direct comparisons between ADMA and IPMA show a similar overall treatment effect [2,6,7]. However, IPMA clearly excels in examining the variation of treatment efficacy in distinct patient populations. This capability has several important applications.

IPMA can efficiently measure treatment effects in important patient subgroups. Such analyses are very limited in ADMA, because published individual studies may not report treatment effect in subgroups or the patient subgroups may vary between studies. Koopman et al. [8] found large differences between subgroup analyses by IPMA vs. those by ADMA in antibiotic treatment for acute otitis media. They concluded that these differences were primarily because of a lack of subgroup reporting in several of the studies included in the ADMA. When subgroups are actually reported in an article, such reporting may be a function of an interesting or significant result, and may produce biased results [9]. IPMA makes the examination of outcomes in important subgroups both possible and consistent between studies.

When comparing treatment efficacy between subgroups, the most important categorization is by baseline outcome risk. Individual patient data are uniquely placed to generate and use baseline-risk indexes in patient groups. Trikalinos and Ioannidis [10] illustrated how individual patient data allows one to fine tune such analysis by developing baseline-risk models using individual patient data and then analyzing treatment effect by risk strata. Such a mechanism is likely more influential and clinically useful than presenting treatment effect by levels of several discrete factors that may individually have varying influence on the outcome.

An interaction, or effect modifier, indicates a significant variation of the treatment–outcome association by levels of a categorical variable or over the range of a continuous variable. In the absence of individual patient data, meta-regression is used to identify such interactions. In a meta-regression, trial-level outcomes are regressed against trial-level summaries of patient-level variables (e.g., mean patient age). Meta-regression techniques are very insensitive for identifying interactions between treatment effects and patient characteristics. A simulation study by Lambert et al. [11] showed that meta-regression methods identified a significant interaction only 30% of the times. In addition to being underpowered, meta-regression can produce biased results. Teramukai et al. [12] found a complete reversal of the parameter estimate for the interaction between adjuvant chemotherapy and lung cancer stage when two studies were excluded from the analysis. Such a reversal of results was not seen when the analysis was repeated using individual patient data.

Individual patient data facilitate the examination of heterogeneity in meta-analyses. Heterogeneity between study results may occur if studies differ by patient characteristics that are related to treatment efficacy. Individual patient data facilitate the exploration of such heterogeneity in meta-analysis by the search for important interactions.

Reward 4: Expertise

Getting access to primary data from multiple randomized trials requires buy-in from, and cooperation of, the principal investigators of key randomized trials of a particular treatment. Having the collaboration of all of these primary trialists usually results in a scientific group with unparalleled expertise and content knowledge. Such a group is aware of the research themes that should and can be addressed by the individual patient data. In addition, such a group usually has the gravitas that will ensure recognition of any research produced by the collaboration.

These rewards of IPMA explain why they are widely considered the ‘‘gold standard’’ of systematic review [13]. However, the extensive rewards of IPMA have not translated into its broad use in the medical literature. Koopman et al. [14] estimated that only 2% of the 8,600 systematic reviews published since 1996 were IPMAs. In addition, the incidence of IPMA does not appear to be increasing. Lyman and Kuderer [15] found that the annual proportion of all oncological meta-analyses that were IPMA did not significantly change between 1990 and 2004. This can, at least in part, be explained by some of the challenges that IPMA pose for the meta-analyst. The most important of these challenges include the following.

Challenge 1: Getting the data

This is, by far, the most challenging aspect of conducting IPMA, and can be, obviously, a deal breaker. There are many reasons why a trialist may not want to share his or her data with an IPMA collaboration. Trialists spend an extensive amount of time and energy collecting data for a study. The fact that they become somewhat protective of those data and suspicious of someone who wants access to those data is somewhat understandable. Occasionally, data from studies funded by industry are unavailable when a company believes the potential benefit of the collaboration will not exceed any perceived risk therein.

Achieving buy-in from principal investigators requires trust, explicit rules regarding data use, and potential benefits. Obviously, scientists need to trust those who are spearheading the IPMA collaboration before they will even consider participating. Gaining this trust is easier when one is already a solid member of the research community; it also gets easier when a critical mass of trial data sets has already been amassed. Explicit rules in the form of a data-sharing agreement is absolutely necessary to protect the participating trialists. Such data-sharing agreements usually prohibit any public use of the data, or any summary thereof,

without explicit permission by the trialist. Finally, trialists must perceive some benefit from participating in such a collaborative. This comes both from the ability to conduct groundbreaking research that would otherwise be impossible and the chance to work with other experts in the field.

Challenge 2: Harmonizing the data

Although getting buy-in from individual trialists may be difficult, one should never underestimate the challenges involved in harmonizing the data from each trial into a common, standard data set. Variations between studies with respect to data formats, data structure, and variable definitions must all be harmonized to create a common data set for the IPMA. Throughout this process, strong and efficient programming skills along with regular contact with the primary investigator, or his or her team, are necessary to make the project possible. Troubleshooting is necessary and can be achieved by using the new data set to recreate analyses, tables, and figures that are compared with those in the original study publication. In general, harmonizing the data is very time consuming and frequently tedious with few immediate payoffs. However, this step is absolutely necessary to ensure high-quality analyses with the data.

Challenge 3: Keeping communication lines open

After data have been collated and standardized, communication with all investigators in the collaboration is necessary for ongoing productivity and success. Each participant needs to be given the opportunity to influence the collaborative's research direction and priorities. In addition, everyone must have the chance to provide input into the interpretation and presentation of research findings coming from the collaborative. Although these steps will prolong the time required to conduct and disseminate analyses, they invariably result in more thoughtful and balanced analysis. These steps are necessary to optimize output from the collaborative, but they do not excessively increase costs. E-mail and teleconferences usually suffice to get input from each collaborative member and obviate the need for more expensive and time-consuming meetings.

These are the most important aspects regarding the rewards and challenges of IPMA. This information will hopefully clarify when meta-analysts should seek individual patient data for their systematic review along with the major hurdles they will need to clear to achieve success.

References

[1] 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.

[2] Stewart LA, Parmar MK. Meta-analysis of the literature or of individual patient data: is there a difference? Lancet 1993;341:418−22.

[3] Ishak KJ, Platt RW, Joseph L, Hanley JA, Caro JJ. Meta-analysis of longitudinal studies. Clin Trials 2007;4:525−39.

[4] Jones AP, Riley RD, Williamson PR, Whitehead A. Meta-analysis of individual patient data versus aggregate data from longitudinal clinical trials. Clin Trials 2009;6:16−27.

[5] McCormack K, Grant A, Scott N. Value of updating a systematic review in surgery using individual patient data. Br J Surg 2004;91: 495−9.

[6] Jeng GT, Scott JR, Burmeister LF. A comparison of meta-analytic results using literature vs individual patient data. Paternal cell immunization for recurrent miscarriage. JAMA 1995;274:830−6.

[7] Steinberg KK, Smith SJ, Stroup DF, Olkin I, Lee NC, Williamson GD, et al. Comparison of effect estimates from a meta-analysis of summary data from published studies and from a meta-analysis using individual patient data for ovarian cancer studies. Am J Epidemiol 1997;145:917−25.

[8] Koopman L, van der Heijden GJ, Hoes AW, Grobbee DE, Rovers MM. Empirical comparison of subgroup effects in conventional and individual patient data meta-analyses. Int J Technol Assess Health Care 2008;24:358−61.

[9] Hahn S, Williamson PR, Hutton JL, Garner P, Flynn EV. Assessing the potential for bias in meta-analysis due to selective reporting of subgroup analyses within studies. Stat Med 2000;19:3325−36.

[10] 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.

[11] 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.

[12] Teramukai S, Matsuyama Y, Mizuno S, Sakamoto J. Individual patient-level and study-level meta-analysis for investigating modifiers of treatment effect. Jpn J Clin Oncol 2004;34:717−21.

[13] 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.

[14] Koopman L, van der Heijden GJ, Glasziou PP, Grobbee DE, Rovers MM. A systematic review of analytical methods used to study subgroups in (individual patient data) meta-analyses. J Clin Epidemiol 2007;60:1002−9.

[15] Lyman GH, Kuderer NM. The strengths and limitations of meta-analyses based on aggregate data. BMC Med Res Methodol 2005;5:14.