IPDMA: Individual Patient/Participant Data Meta-Analysis

Marcel Dijkers, Ph.D., FACRM
Icahn School of Medicine at Mount Sinai
Department of Rehabilitation Medicine

KT Update *presents another in a series of brief articles by Dr. Marcel Dijkers. This article describes the steps of individual patient/participant (data) meta-analysis, or IPDMA, and compares it with traditional aggregate data meta-analysis (ADMA).*

In a previous article in *KT Update* (Dijkers, 2014), the development of meta-analysis (MA) was sketched from (1) meta-analysis of individual study findings (e.g., means and standard deviations, correlations) reported in the literature, to (2) meta-analysis of individual subject data harvested from the published literature, to (3) merging of datasets from multiple studies that share subject population (but not the same individuals) and have similar variables, to (4) creation of common data elements (CDEs) to standardize the variables in related studies, to (5) trialists' collaborations, which offer a mechanism to prospectively combine related studies into a single dataset. Of the above, 2, 3, 4, and 5 all require or offer opportunities for individual patient/participant (data) MA, or IPDMA. The major differences between IPDMA and a secondary analysis of a single-site or multi-site dataset is the need to acquire data or datasets, carefully develop a common coding scheme for all variables, and in the analysis address any issues resulting from the origin of the data (e.g., clustering). How much of this painstaking work needs to be done depends largely on the source of one's data – differences between 2, 3, 4 or 5 above, essentially.

IPDMA is most closely associated with 3 above: synthesizing information resulting from a re-analysis of the data produced by two or more studies. Alternative names that have been proposed for this process are "mega-analysis" and "integrative data analysis." These have not caught on; even though one might ask whether for a one-stage IPDMA (explained below), the term "meta" may be appropriate.

The steps of IPDMA

The steps in this process generally involve the following (Abo-Zaid et al., 2013; Broeze et al., 2010; Burdett & Stewart, 2002; Cooper & Patall, 2009):

1. Develop a question or set of questions that can be answered with existing data.
2. Develop a protocol, with as much detail as is possible, on the hypotheses to guide the analyses that will be performed to answer the questions the nature of the relevant datasets, and the statistical methods to be employed.
3. Systematically search in multiple bibliographic databases and by other means for all published and unpublished research that did result in an applicable dataset.
4. Contact the principal investigator (PI) of each study, and request a copy of the dataset, de-identified or anonymized if needed to protect subjects' confidential information.
5. Check each dataset carefully for the quality of the data (e.g., duplicate cases, empty fields) and clean it as much as possible, in collaboration with the original PI and her/his staff.
6. Recode all variables in all datasets to make them compatible (e.g., the variable is consistently called "sex" and coded 1=female, 2=male, and 9=unknown).

The next steps differ for a one-stage vs. a two-stage statistical analysis. In a one-stage IPDMA approach, the merged data are analyzed together, using a statistical procedure that

takes into account the clustered nature of the cases. In a two-stage approach, the individual datasets (or their subsets—for example, male vs. female subjects) are analyzed in step 1, and the resulting means, correlations, or other descriptive parameters are used in step 2, which combines them using traditional aggregate data (AD) MA methods.

Ergo, for a one-stage IPDMA, the steps are:
7. Merge all of the datasets.
8. Analyze all combined cases, and/or subgroups of interest, to answer the research questions.
9. Report the findings, making sure that complete information on the original studies and the new analyses is provided.

In a two-stage IPDMA, the steps are:
7. Analyze all datasets in parallel, producing means and other values of interest for all cases in each, or for subgroups of interest.
8. Consolidate these parameters as appropriate (with or without bringing in parallel information from studies for which individual patient/participant data are not available) and analyze them as one would in any aggregate data meta-analysis (ADMA).
9. Report the findings, providing comprehensive information on the original studies and the MA analyses performed.

Combinations of these two methods are possible, and not unusual (Abo-Zaid, Sauerbrei, & Riley, 2012). For instance, it is uncommon to be able to acquire all datasets in existence, and performing a two-stage analysis that includes newly calculated information for the studies that are present, and literature-reported parallel information for the missing studies allows one to assess the degree to which "missingness" is likely to bias one's findings.

The advantages of IPDMA compared to ADMA
Compared with "traditional" ADMA (i.e., meta-analysis that relies on published data summaries only), IPDMA has a number of advantages (Abo-Zaid et al., 2012; Broeze et al., 2010; Cooper & Patall, 2009):

1. The data in the various datasets can be checked extensively. For instance, investigators performing IPDMA of randomized control trials (RCTs) can determine whether randomization was performed properly (Burdett & Stewart, 2002), and if so, whether the treatment and comparator groups are actually balanced in terms of the most important outcome predictor variables. (If not, adjustments can be made in the analysis).
2. By duplicating the original analyses, one can determine whether the published literature contains valid evidence, and if not, publish a correction.
3. If the original authors reported selectively on those subgroups or outcome variables that happened to be statistically significant, or if they made a unit-of-analysis error, or failed to use intent-to-treat analysis, the public record can similarly be corrected.
4. Subgroup analyses that the original studies could not execute (because there were too few cases in the groups to offer sufficient statistical power) can be performed.
5. More complex analytical methods that did not exist at the time the original studies were reported or that the original authors failed to use (e.g., latent growth curve analysis) can be used to provide more precise or more certain answers to old questions, or answers to completely new ones.
6. Standardized analyses, whether traditional or more advanced (e.g., time-to-event analysis), can be conducted across all datasets.

7. One can analyze the effects of within-study and between-study moderators of effect sizes even when subgroups in the joined samples are still considered "small."

Some methodologists add the following advantage, although the IPDMA in this case starts to meld into a multi-site longitudinal follow-up dataset:

8. By adding to the dataset outcome information not yet available at the time the original investigators published their findings, the long-term outcomes of treatments, diagnostic decisions, or natural recovery can be analyzed.

The disadvantages of IPDMA compared to ADMA

Everything has a price, and this great power of IPDMA comes with some disadvantages compared to ADMA (Abo-Zaid et al., 2012; Cooper & Patall, 2009):

1. IPDMAs take much more effort and time (and thus money) to complete.
2. Incomparable outcome measures used in studies constitute a major problem, at least for one-stage IPDMA.
3. Traditional ADMA can be performed even if individual patient/participant data are not available, by "just" summarizing the published parameters. Even if IPDMA has a much higher power than ADMA, this advantage may be lost if the ADMA can use many more published and unpublished studies.
4. The proper statistical analysis to be used in IPDMAs, especially one-stage IPDMAs, is still a topic of intense discussion, more so than that of group-data level MA.
5. Compared to ADMA, IPDMA brings tricky authorship issues. Are the investigators of the primary studies to be offered the opportunity to become co-investigators on the IPDMA? Offered the chance to "veto" interpretations of the combined data they disagree with?

Effort. A number of reports have indicated how time-intensive IPDMAs are – possibly five times as high as ADMAs (Steinberg et al., 1997). Finding contact information for PIs, and getting them to create and supply a useable deidentified dataset (with documentation sufficient documentation to allow reanalysis) takes much time. However, the increasingly common requirement by funding sources that data be shared, often by uploading them to a proper data-sharing site, will change the difficulty of obtaining data in the next decade. Sharing is now required by the National Institutes of Health and will be required in the near future by the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR). It also is required by many journals (e.g., *Nature, BMJ*); and professional organizations (e.g., the American Psychological Association, APA).

Checking datasets and making the data compatible can take even more time; the experience of the IMPACT TBI researchers (International Mission for Prognosis and Analysis of Clinical Trials in Traumatic Brain Injury) offers some guidance here (Marmarou et al., 2007). The creation of CDEs, and funding agency expectations that these will be used for all key constructs in a particular scientific domain of research, in the future may also reduce the effort needed to merge datasets.

Different measures for the same construct. Discrepant outcome measures used in various studies (e.g., Beck Depression Inventory-II, or BDI-II, vs. Patient Health Questionnaire 9, PHQ-9) to quantify depression symptomatology is not a serious issue for ADMA, but it creates major problems for a one-stage IPDMA: the apples cannot be mixed with the oranges without making some possibly untenable assumptions in turning both into the same fruit. For instance, for both BDI-II and PHQ-9, one could calculate what percentage of the maximum score a subject had, but a sophisticated analysis of a sample for which both measures are available might find a low correlation between the raw scores of the two or the percentage-of-maximum scores. A Rasch analysis might even show a poor correlation of person trait

estimates resulting from separate scaling efforts, and the impossibility of creating a "crosswalk." If the IPDMA uses dichotomous or dichotomized outcomes, such as death, hospital readmission, or a score above the clinical cut-off for depression, these problems do not exist, or can be resolved more easily.

Access to data. Even if the majority of trialists (the opinions of health care researchers who use methods and designs other than clinical trials have not been investigated to any significant degree) agree that research data ought to be publicly available, and state that they personally will make their data available upon request (Rathi et al., 2012; Rathi et al., 2014), the reality is different: various investigators planning or reporting IPDMAs have reported unwillingness of PIs to share their data, with detrimental effects for analyses (Abo-Zaid et al., 2012; Ebrahim et al., 2012; Villain, Dechartres, Boyer, & Ravaud, 2015; Wicherts, Borsboom, Kats, & Molenaar, 2006), up to complete abandonment in at least one instance (Jaspers & Degraeuwe, 2014). In some instances the concerns behind a failure to share are justified—for example, if the dataset has been lost, or, as possible with some case studies and other qualitative research, the data cannot be anonymized so as to maintain patient confidentiality. Another possible reason could be the cost inherent in preparing a deidentified version. (All of these reasons can be put forward, but the complete lack of reply often encountered by candidate IPD meta-analysts is against the collegiality declarations of most professional groups.) Traditional ADMA is not hampered by such problems, unless the research report is unpublished, or the published report is lacking in crucial information that the meta-analyst can obtain only from the primary investigator. As suggested above, ADMA can be used to estimate how much bias is introduced to the IPDMA, because datasets do not go missing "at random"; both analytic strategies, of course, are still subject to publication bias. "Availability bias" is the term that has been coined to describe the problem of selective lack of access to datasets that is unique to IPDMA.

Statistical expertise. Even when the first accounts by statisticians on the proper methods of performing a one-stage IPDMA are at least 20 years old, this is still an area of active research, with reports on methods for handling specific issues being published on a regular basis (Abo-Zaid et al., 2013; Debray, Moons, Ahmed, Koffijberg, & Riley, 2013; Fisher, Copas, Tierney, & Parmar, 2011). The fact that some IPDMA teams do not include a specialized statistician has led to quite a few inappropriate analyses, with the most common issue being a lack of appropriate accounting for the clustering of subjects in one-stage IPDMAs. This is an issue that is not uncommon in primary research involving multiple sites, where the analysis often also omits a proper adjustment. However, there is more of an issue in one-stage IPDMAs: subjects from various studies differ in geographic location, nature of the health care system that served them, specific inclusion/exclusion criteria used in finding them, details of the treatment they received (if intervention research participants), and details in the methods and timing of baseline and outcome data collection methods, for example. Consequently, they cannot be considered a random sample of the population of potential participants, and the analysis should take this into account. (In two-stage IPDMA, clustering is handled "automatically," whether a fixed or random model is used in the analysis.)

Prospects for IPDMA

While no one has completed a comprehensive inventory of IPDMAs, several partial counts suggest that the number of these studies has increased significantly in each decade, after the first ones were published in the 1970s. While intervention research data are still the most common target of researchers using this technique (especially in cancer care and cardiology, and in areas such as surgery where small RCTs or case series are common), there are now also IPDMAs of diagnostic test assessment, prognosis studies, causal investigations, and other observational research. In 2015, the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) group (Stewart et al., 2015) published guidance as to

how this type of secondary research is to be reported. IPDMA, because of its many advantages over ADMA, now is considered the "gold standard" for MA (Simmonds, Stewart, & Stewart, 2015). With more examples in the literature, and more datasets potentially available each year, we can expect continuing growth of the number of publications.

In the area of disability and rehabilitation studies, the number of published IPDMAs is still small, and stroke intervention research seems to be overrepresented. However, there are a few interesting studies in other populations (Karyotaki et al., 2015; Kivimaki et al., 2015), and in the last two years a number of IPDMA protocols have been published (e.g., Buffart et al., 2013; Jonkman et al., 2014). Anyone intrigued by the options this "new" IPDMA method proffers would do well to start exploring some of these to see whether the methodology indeed offers the pay-off its enthusiasts claim.

Postscript

In a few instances where the findings of ADMA and IPDMA have been compared, the size and even the direction of the effect sizes have been found to be different. According to Cooper and Patall (2009), the explanation lies in the "ecological fallacy" well known to social science students, and in a related phenomenon called Simpson's paradox. An hypothetical example of the ecological fallacy might be: at the census tract level, tracts with a high percentage of households over the national median income are also the tracts with more police reports of breaking and entering per 1,000 people. Does this mean the rich do more stealing? Or would we see, if we had individual-level or household-level data, that the poor do more breaking and entering, but the tract-level data are distorted because:

- The poor go steal where the rich live.
- The poor do not report break-ins of their residence to the police.

Cooper and Patall offer a more complete description of the ecological fallacy and its "inverse," the Simpson paradox, and discuss how they apply to MA. Berlin et al. (2002) discuss how this applies to the difficulty of generalizing from group analysis findings to individual patient/participant benefits of treatment.

References

Abo-Zaid, G., Guo, B., Deeks, J. J., Debray, T. P., Steyerberg, E. W., Moons, K. G., & Riley, R. D. (2013). Individual participant data meta-analyses should not ignore clustering. *Journal of Clinical Epidemiology, 66*(8), 865-873.e4. doi:10.1016/j.jclinepi.2012.12.017

Abo-Zaid, G., Sauerbrei, W., & Riley, R. D. (2012). Individual participant data meta-analysis of prognostic factor studies: state of the art? *BMC Medical Research Methodology, 12*, 56-2288-12-56. doi:10.1186/1471-2288-12-56

Berlin, J. A., Santanna, J., Schmid, C. H., Szczech, L. A., Feldman, H. I., & Anti-Lymphocyte Antibody Induction Therapy Study Group. (2002). Individual patient- versus group-level data meta-regressions for the investigation of treatment effect modifiers: ecological bias rears its ugly head. *Statistics in Medicine, 21*(3), 371-387. doi:10.1002/sim.1023

Broeze, K. A., Opmeer, B. C., van der Veen, F., Bossuyt, P. M., Bhattacharya, S., & Mol, B. W. (2010). Individual patient data meta-analysis: a promising approach for evidence synthesis in reproductive medicine. *Human Reproduction Update, 16*(6), 561-567. doi:10.1093/humupd/dmq043

Buffart, L. M., Kalter, J., Chinapaw, M. J., Heymans, M. W., Aaronson, N. K., Courneya, K. S., . . . Brug, J. (2013). Predicting OptimaL cAncer RehabIlitation and Supportive care (POLARIS): rationale and design for meta-analyses of individual patient data of randomized controlled trials that evaluate the effect of physical activity and psychosocial interventions on health-

related quality of life in cancer survivors. *Systematic Reviews, 2*, 75-4053-2-75. doi:10.1186/2046-4053-2-75

Burdett, S., & Stewart, L. A. (2002). A comparison of the results of checked versus unchecked individual patient data meta-analyses. *International Journal of Technology Assessment in Health Care, 18*(3), 619-624.

Cooper, H., & Patall, E. A. (2009). The relative benefits of meta-analysis conducted with individual participant data versus aggregated data. *Psychological Methods, 14*(2), 165-176. doi:10.1037/a0015565

Debray, T. P., Moons, K. G., Ahmed, I., Koffijberg, H., & Riley, R. D. (2013). A framework for developing, implementing, and evaluating clinical prediction models in an individual participant data meta-analysis. *Statistics in Medicine, 32*(18), 3158-3180. doi:10.1002/sim.5732

Dijkers, M. (2014). Meta-analysis: The end of the beginning, or the beginning of the end? *KT Update, 2*(5).

Ebrahim, S., Montoya, L., Truong, W., Hsu, S., Kamal El Din, M., Carrasco-Labra, A., . . . Guyatt, G. H. (2012). Effectiveness of cognitive behavioral therapy for depression in patients receiving disability benefits: a systematic review and individual patient data meta-analysis. *PloS One, 7*(11), e50202. doi:10.1371/journal.pone.0050202

Fisher, D. J., Copas, A. J., Tierney, J. F., & Parmar, M. K. (2011). A critical review of methods for the assessment of patient-level interactions in individual participant data meta-analysis of randomized trials, and guidance for practitioners. *Journal of Clinical Epidemiology, 64*(9), 949-967. doi:10.1016/j.jclinepi.2010.11.016

Jaspers, G. J., & Degraeuwe, P. L. (2014). A failed attempt to conduct an individual patient data meta-analysis. *Systematic Reviews, 3*, 97-4053-3-97. doi:10.1186/2046-4053-3-97

Jonkman, N. H., Westland, H., Trappenburg, J. C., Groenwold, R. H., Effing-Tijdhof, T. W., Troosters, T., . . . Schuurmans, M. J. (2014). Towards tailoring of self-management for patients with chronic heart failure or chronic obstructive pulmonary disease: a protocol for an individual patient data meta-analysis. *BMJ Open, 4*(5), e005220-2014-005220. doi:10.1136/bmjopen-2014-005220

Karyotaki, E., Kleiboer, A., Smit, F., Turner, D. T., Pastor, A. M., Andersson, G., . . . Cuijpers, P. (2015). Predictors of treatment dropout in self-guided web-based interventions for depression: an 'individual patient data' meta-analysis. *Psychological Medicine, 45*(13), 2717-2726. doi:10.1017/S0033291715000665

Kivimaki, M., Singh-Manoux, A., Virtanen, M., Ferrie, J. E., Batty, G. D., & Rugulies, R. (2015). IPD-Work consortium: pre-defined meta-analyses of individual-participant data strengthen evidence base for a link between psychosocial factors and health. *Scandinavian Journal of Work, Environment & Health, 41*(3), 312-321. doi:10.5271/sjweh.3485

Marmarou, A., Lu, J., Butcher, I., McHugh, G. S., Mushkudiani, N. A., Murray, G. D., . . . Maas, A. I. (2007). IMPACT database of traumatic brain injury: design and description. *Journal of Neurotrauma, 24*(2), 239-250. doi:10.1089/neu.2006.0036

Rathi, V., Dzara, K., Gross, C., Hrynaszkiewicz, I., Joffe, S., Krumholz, H., . . . Ross, J. (2012). Sharing of clinical trial data among trialists: a cross sectional survey. *BMJ (Clinical Research Ed.), 345*, e7570. doi:10.1136/bmj.e7570

Rathi, V., Strait, K., Gross, C., Hrynaszkiewicz, I., Joffe, S., Krumholz, H., . . . Ross, J. (2014). Predictors of clinical trial data sharing: exploratory analysis of a cross-sectional survey. *Trials, 15*, 384-6215-15-384. doi:10.1186/1745-6215-15-384

Simmonds, M., Stewart, G., & Stewart, L. (2015). A decade of individual participant data meta-analyses: A review of current practice. *Contemporary Clinical Trials, 45*(Pt A), 76-83. doi:10.1016/j.cct.2015.06.012

Steinberg, K. K., Smith, S. J., Stroup, D. F., Olkin, I., Lee, N. C., Williamson, G. D., & Thacker, S. B. (1997). 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. *American Journal of Epidemiology, 145*(10), 917-925.

Stewart, L. A., Clarke, M., Rovers, M., Riley, R. D., Simmonds, M., Stewart, G., . . . PRISMA-IPD Development Group. (2015). Preferred Reporting Items for Systematic Review and Meta-Analyses of individual participant data: the PRISMA-IPD Statement. *JAMA, 313*(16), 1657-1665. doi:10.1001/jama.2015.3656

Villain, B., Dechartres, A., Boyer, P., & Ravaud, P. (2015). Feasibility of individual patient data meta-analyses in orthopaedic surgery. *BMC Medicine, 13*, 131-015-0376-6. doi:10.1186/s12916-015-0376-6

Wicherts, J. M., Borsboom, D., Kats, J., & Molenaar, D. (2006). The poor availability of psychological research data for reanalysis. *The American Psychologist, 61*(7), 726-728. doi:2006-12925-016