PRIME-IPD SERIES Part 3. The PRIME-IPD tool fills a gap in guidance for preparing IPD for analysis.

Authors: Omar Dewidar[a,b], Alison Riddle[a], Elizabeth Ghogomu[a], Alomgir Hossain [b,c], Paul Arora[d], Zulfiqar A Bhutta[e,f], Robert E Black[g], Simon Cousens[h], Christine Mathew[a], Jessica Trawin[i], Peter Tugwell[j,k,l], Vivian Welch[a,b*], George A Wells[b,m*]

[a] Bruyère Research Institute, University of Ottawa, 85 Primrose Ave, Ottawa, Ontario, K1R 6M1, Canada

[b] School of Epidemiology and Public Health, University of Ottawa, 600 Peter Morand Crescent, Ottawa, Ontario, K1G 5Z3, Canada

[c] Department of Medicine (Cardiology), The University of Ottawa Heart Institute and University of Ottawa, 40 Ruskin Street, Ottawa, Ontario, K1Y 4W7, Canada

[d] Dalla Lana School of Public Health, University of Toronto, 155 College St Room 500, Toronto, Ontario M5T 3M7, Canada

[e] Centre for Global Child Health, Hospital for Sick Children, 555 University Ave, Toronto, Ontario, M5G 1X8, Canada

[f] Institute for Global Health & Development, Aga Khan University, South-Central Asia, East Africa & United Kingdom, Karachi, Pakistan

[g] Department of International Health, Johns Hopkins School of Hygiene and Public Health, 615 N Wolfe St Suite E8527, Baltimore, Maryland, 21205, United States

[h] Department of Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine (LSHTM), Keppel Street, London, WC1E 7HT, UK

[i] Women's Health Research Institute, 4500 Oak St, Vancouver, British Columbia, V6H 2N9, Canada.

[j] Clinical Epidemiology Program, Ottawa Hospital Research Institute, 501 Smyth Rd, Ottawa, Ontario K1H 8L6, Canada

[k] Department of Medicine, University of Ottawa Faculty of Medicine, Roger Guindon Hall, 451 Smyth Rd #2044, Ottawa, Ontario, K1H 8M5, Canada

[l] WHO Collaborating Centre for Knowledge Translation and Health Technology Assessment in Health Equity, Bruyère Research Institute, 85 Primrose Ave, Ottawa, Ontario, K1R 6M1, Canada

[m] Cardiovascular Research Methods, University of Ottawa Heart Institute, 40 Ruskin St, Ottawa, Ontario, K1Y 4W7, Canada

* Joint senior authors

We thank Levis et al. for their interest in our study and for sharing their perspective on preparing IPD for analysis. We wish to emphasize that our paper's purpose was to provide guidance on the different steps that should be undertaken when preparing IPD for analysis, but not to be prescriptive about how each step should be performed. The examples provided in the manuscript are examples of how we applied PRIME (Preparation, Replication, Imputation, Merging, Evaluation) to our Network Meta-Analysis and are in no way comprehensive (1, 2). The suggestions provided by Levis et al. in their commentary emphasize the importance of having a structured process for managing IPD. We agree that the finer details of that structured process may vary since they should be determined by the nature of the datasets and share Levis et al.'s desire to see our framework refined in order to develop a robust scientific process to answer the research questions with validity.

Processing:

The first action in the processing step is to have the dataset in their computer software of choice to access the dataset. This will be necessary to manipulating the dataset and data verification. Subsequently, we suggested verifying the dataset before proceeding with the IPD analysis by comparing the sample size in the acquired datasets with what was published. The replication step encompasses recalculation of reported descriptive statistics and study findings in which we suggested to proceed the processing step. It is necessary to consider all of the steps that we enlisted in the processing step to replicate the study findings accurately.

As mentioned in PRIME's processing step, datasets may need to have values converted to the preferred standard using appropriate conversion formulae. Discussions between statisticians and

clinicians can help identify appropriate harmonization methods. Levis et al. suggested the use of a master codebook to ensure harmonization of different variables across datasets. We agree that this would be a beneficial tool for guiding the data harmonization process. From our experience, this master codebook can contribute to efficient communication between statisticians and clinicians, minimizing the risk of miscommunication and errors that could have costly repercussions.

Following harmonization of variables, we suggested verifying that the blank cells are missing values and not due to a conversion error. In addition to Levis et al.'s proposition to using distribution plots to identify reverse-coded variables, we suggest referring to the data dictionary, if available.

Replication:

Levis et al. support our call to recalculate the reported baseline and endline study results using the acquired datasets. We agree that it may not always be possible to exactly replicate results due to constraints on time, cost, or lack of methodological clarity; however, we believe some form of replication is essential to ensure complete, accurate and unbiased reporting of results. We suggest in our paper using a 10% difference as an indicator of an important difference between reported and replicated results while acknowledging that there is no universally accepted threshold to indicate a statistically significant imbalance for the standardized difference (3-6). As Levis et al. suggested, IPD project teams may choose a different threshold depending on their investigation.

A vital consideration pointed out by Levis et al. is to decide on which set of participants from the IPD to use for their analysis. We did not discuss this issue in our paper, but it is an important

decision that must be made before embarking on IPD and may be decided by clinicians and statisticians. We emphasise however, that all PRIME-IPD steps should be conducted using the same sample as the original paper.

Imputation and Merging:

Almost all datasets have missing data, and various methods are available for dealing with this missingness. As we detailed in the PRIME-IPD article, depending on the amount and type of missingness in the processed datasets, imputation may or may not be appropriate (7). If imputation is carried out, we suggested doing so before merging the datasets for the following reasons.

1. Merging the datasets prior to the imputation procedure would increase the likelihood of imputing values different from what the original authors might have imputed.
2. The larger the input file for which imputations are to be performed, the greater the processing power needed to conduct the process. Merging the datasets before carrying out imputations will increase the time needed for imputation.

Imputation within or across studies depends on the goal of the study and the variability of data in the datasets. Differences in databases would skew the results in the imputed datasets. Therefore, we would generally prefer to perform imputation prior to merging to avoid possibly costly errors.

Authors can acquire datasets from different sources. However, access to the data is imperative to conduct a "one step" IPD analysis (i.e. merge the obtained datasets for pooled analysis) (8). As Levis and colleagues have pointed out, limited access to some data may prevent one step IPD analysis. A possible approach would be to conduct a "two-step" IPD analysis, where an individual analysis of each study is conducted, and the effect estimates are meta-

analyze to obtain a pooled estimate. This analysis method may be perfectly adequate for some purposes and therefore, the merging step would not be feasible.

Evaluation:

In the last step of PRIME, we encourage authors to examine data distributions visually or by statistical tests. Levis et al. have suggested visual checking in the first stage of PRIME (processing stage) to identify outliers or wrong values. However, the intended purpose of examining data distribution here is to assess if there is a sufficient distribution to run subgroup analysis. Nonetheless, we agree that it would be a valuable tool to identify outliers or wrong values, indicating that some instructions within the PRIME steps may be adjusted to the users' needs. Following IPD preparation with PRIME, we fully agree with Levis and colleagues on the need to check model fit for IPD analysis.

Conclusions:

IPD analyses are resource extensive. Access to a small proportion of trials makes the interpretation of the results challenging. Project teams and funders need to weigh potential benefits with the additional resource requirements, such as incorporating aggregate results and revisiting published reports for additional information. This highlights the importance of a well-defined and thought-out IPD analysis plan before initiating the project. It is also essential to consider the extra challenges of conducting an IPD, such as unexpected delays due to data checking issues and the inefficient data harmonization process. This further supports PRIME's need, a structured process that minimizes the chances of costly and resource-intensive mistakes.

We agree with Levis et al. that a consensus approach would be required for developing reporting guidance, for instance, a Delphi survey method. Due to the resource restrictions, we were unable to undertake such an exercise, so we shared this method with the scientific community due to the literature's lack of available guidance.

References:

1. Welch VA, Hossain A, Ghogomu E, Riddle A, Cousens S, Gaffey M, et al. Deworming children for soil-transmitted helminths in low and middle-income countries: systematic review and individual participant data network meta-analysis. Journal of Development Effectiveness. 2019;11(3):288-306.
2. Welch VA, Ghogomu E, Hossain A, Riddle A, Gaffey M, Arora P, et al. Mass deworming for improving health and cognition of children in endemic helminth areas: A systematic review and individual participant data network meta-analysis. Campbell Systematic Reviews. 2019;15(4):e1058.
3. Austin PC. Using the Standardized Difference to Compare the Prevalence of a Binary Variable Between Two Groups in Observational Research. Communications in Statistics - Simulation and Computation. 2009;38(6):1228-34.
4. Austin PC. Propensity-score matching in the cardiovascular surgery literature from 2004 to 2006: a systematic review and suggestions for improvement. J Thorac Cardiovasc Surg. 2007;134(5):1128-35.
5. Austin PC. A critical appraisal of propensity-score matching in the medical literature between 1996 and 2003. Stat Med. 2008;27(12):2037-49.
6. Normand ST, Landrum MB, Guadagnoli E, Ayanian JZ, Ryan TJ, Cleary PD, et al. Validating recommendations for coronary angiography following acute myocardial infarction in the elderly: a matched analysis using propensity scores. J Clin Epidemiol. 2001;54(4):387-98.
7. Rubin DB, Schenker N. Multiple imputation in health-care databases: an overview and some applications. Stat Med. 1991;10(4):585-98.
8. Burke DL, Ensor J, Riley RD. Meta-analysis using individual participant data: one-stage and two-stage approaches, and why they may differ. Stat Med. 2017;36(5):855-75.