A Brief Guide to Performing Review and Meta-analysis of Rare Diseases, Procedures, and Other Low Frequency Pathology

Meta-analysis publications have proliferated in the last 20 years. Publicly available, open access, easy to use software has allowed anyone to perform a systematic review and meta-analysis. Decision makers are increasingly demanding systematic review and/or meta-analysis as high quality evidence for the basis of policy and guidelines. As part of this process there is a widespread acceptance that meta-analysis of randomised trials is the apex of evidence. But from a data quality point of view this is not entirely true. Most clinicians understand that meta-analysis of non-randomised data is of inherently lower quality. The problem is that meta-analyses of all data qualities are often reported in the same way, which may lead to the results and conclusions misleading the reader. Even “higher quality” meta-analyses in higher impact journals fall into this trap and do not have consistent reporting quality. As these studies are used by decision makers, it is extremely important to understand and communicate the limitations of a systematic review and meta-analysis within the study.

Meta-analysis of the non-powered endpoints from randomised trials is a good example of lower quality data that may be misinterpreted as the apex of evidence. The powered outcome data from the randomised trial (usually the primary outcome) is the most valid and robust data, with non-powered outcomes becoming less useful as the event rate decreases. An example is the 2018 meta-analysis of mortality data following paclitaxel balloon angioplasty that changed policy internationally. Even though this was a “meta-analysis of randomised trials”, the trials were small in terms of patient numbers and were powered to detect patency loss of the treated artery, not death, which happened far more infrequently. Meta-analysing the non-powered, low event rate mortality outcome is reasonable, but should be clearly communicated with the caveats of low quality data that may be prone to a number of biases. In this study the authors’ assessment of bias using the Cochrane risk of bias 1 tool for randomised trials was favourable; however, that assessment only really applies to the powered outcome of the trial, not the mortality data. This flaw is corrected in the risk of bias 2 tool, which came out after the paclitaxel study. This limitation can also be mitigated by performing a certainty analysis for each meta-analysed outcome using a framework such as GRADE (Grading of Recommendation, Assessment, Development and Evaluation). GRADE allows authors to focus directly on the result extracted for an individual meta-analysis calculation (e.g., here mortality from a randomised trial powered for patency) and then state the quality, or certainty, of that result for context. This would have been very helpful for the paclitaxel meta-analysis. “There is low certainty evidence for an association between the application of paclitaxel coated balloons and a subsequent risk of death” reads very differently from the published conclusion, “There is an increased risk of death following application of paclitaxel coated balloons and stents...”. Using the word “association” to acknowledge the lower quality result also changes the tone. The GRADE result would be different if the same randomised trials were used for a pooled patency result where the certainty would be higher.

Meta-analysis of cohort or case series data comes with the same problems as non-powered data from randomised trials, as well as being more prone to publication and selection bias. The European Journal of Vascular and Endovascular Surgery is about to publish a systematic review and meta-analysis on the treatment of median arcuate ligament syndrome (MALS). As a rare disease, this topic has a large number of low volume case series in the literature. Again, we are back to unpowered data with low event rates. In addition, studies describing rare diseases or procedures are highly likely to have publication bias from authors with either very good or very bad results, and selection bias because of individual clinical practice. In this situation it is important to look for clinical heterogeneity before making the decision to pool the data. Are the study results reasonable to pool, or is the clinical practice so different that you are including apples with oranges? There are ways to assess for common biases before an analysis begins, commonly used examples are ROBINS-1 and QUADAS-2. These are to non-randomised studies what the Cochrane risk of bias tool is for randomised, with QUADAS-2 being specific for primary diagnostic accuracy studies. If the decision is made to pool the data, these assessments feed into the GRADE framework to allow any bias detected to downgrade the certainty recommendation for the meta-analysis. In the current review on MALS, appraisal of the literature identified on systematic review with the QUADAS-2 tool suggested that none of the articles were of sufficient quality to be regarded as “low risk” for bias. This was due to ill-defined and heterogeneous outcome measures in the studies, and the risk of confounding and bias. As a result, meta-analysis of the data was not performed by the authors. As no synthesis of data was done, the authors could not perform a GRADE evaluation of the findings.
Sometimes when there is a large volume of low quality data in the literature a systematic review with meta-analysis is the wrong choice of study design. Scoping reviews may be used in this situation. While there remains no standardised definition, scoping reviews are broadly a means to identify the types, characteristics, concepts, or definitions of the evidence available for a topic. A good example was when the authors of the European Society for Vascular Surgery acute limb ischaemia guidelines wanted to provide an update in light of COVID 19. They recognised there was a huge amount of low quality data published in a very short space of time on limb outcomes during COVID. A systematic review would have produced a vast number of studies too heterogeneous and biased for meaningful meta-analysis and difficult to summarise. The scoping review allowed them to understand the key concepts of the COVID limb outcomes literature and summarise it in a way that facilitated updated recommendations.11

So if a research group is considering summarising the literature for a rare disease or procedure the first question to ask is whether it should be a systematic or scoping review. If systematic, are the studies too heterogeneous for meta-analysis? Then, finally, if meta-analysis is performed, the results must be presented clearly, communicating the limits of the inputted data, including a bias assessment for each relevant study outcome and a quality or certainty assessment for each meta-analysis. Accurate, informed decisions can then be made by policy makers who will look to these important studies for guidance.

CONFLICT OF INTEREST STATEMENT AND FUNDING

None.

REFERENCES

9 Powell JT, Koelemay MJW. Systematic reviews of the literature are not always either useful or the best way to add to science. EJVES Forum 2022;5:4–6.

Christopher P. Twine*
North Bristol NHS Trust and University of Bristol, Bristol UK

Kevin Mani
Section of Vascular Surgery, Department of Surgical Sciences, Uppsala university, Uppsala, Sweden

*Corresponding author. Department of Vascular Surgery, North Bristol NHS Trust, Southmead Hospital, Southmead Road, Bristol BS10 5NB, UK.
E-mail address: Christopher.twine@bristol.ac.uk (Christopher P. Twine)