

## Systematic reviews and meta-analyses “For Dummies”

Giuseppe G. L. Biondi-Zoccai<sup>1\*</sup>, MD; Antonio Abbate<sup>2</sup>, MD; Imad Sheiban<sup>1</sup>, MD

1. Division of Cardiology, University of Turin, S. Giovanni Battista “Molinette” Hospital, Turin, Italy; 2. VCU Pauley Heart Center, Virginia Commonwealth University, Medical College of Virginia, Richmond, VA, USA

Dr. Biondi-Zoccai has consulted for Biotronik, Boston Scientific, Cordis, Invatec and Medtronic, and has lectured for Medtronic.

*“I like to think of the meta-analytic process as similar to being in a helicopter. On the ground individual trees are visible with high resolution. This resolution diminishes as the helicopter rises, and in its place we begin to see patterns not visible from the ground.”*

Ingram Olkin

As explained in the metaphor by the meta-analysis pioneer Ingram Olkin, systematic reviews are viewpoints on a given topic quoting primary authors or studies (i.e. reviews), which deliberately use and report a systematic approach to study search, selection, abstraction, and appraisal. Meta-analyses are studies (not necessarily reviews) which use specific statistical methods for pooling data from separate datasets. This issue of EuroIntervention includes two systematic reviews with meta-analytic pooling focusing, respectively, on the impact of cilostazol on restenosis and repeat revascularisation rates after coronary stenting,<sup>1</sup> and devices for prevention of distal embolisation in acute myocardial infarction.<sup>2</sup> The diffusion of this research study type even in the journal of European interventionists is a further proof of the scientific validity, cost-effectiveness,<sup>3</sup> clinical impact, and high quotability of meta-analyses,<sup>4</sup> despite several concerns on their recent plethora.<sup>5</sup> Is publication of any given study in a peer-reviewed journal such as EuroIntervention definitive evidence of its internal validity and usefulness for the clinical practitioner or researcher? Unfortunately not. Peer-review is not very accomplished in judging or improving the quality of scientific manuscripts and many examples of bad or unsuccessful peer-reviewing efforts can be easily found. However, in as much as democracy, in the words of Sir Winston Churchill (“democracy is the worst form of government except all those other forms that have been tried from time to time”), peer-review is the

worst form to evaluate and appraise scientific research except all other methods that have been tried so far. This applies to all clinical research products in general and also, of course, to systematic reviews and meta-analyses. Thus, provided that meta-analyses are accurately and thoroughly reported, the burden of quality appraisal lies largely, as usual, in the eye of the beholder (i.e. the reader).

We provide, hereby, some simple suggestions to appraise the quality of systematic reviews and meta-analyses exploiting the two interesting systematic reviews published in this issue<sup>1,2</sup> (similarly to what has been done in the last few years by the “For Dummies” collection of paperback manuals).<sup>6</sup> Those looking for more precise and structured appraisal tools can refer to the Oxman-Guyatt tool<sup>7</sup> and the AMSTAR index,<sup>8</sup> available in, respectively, Tables 1 and 2.

### The two-step approach

Our simple two-step approach is not very original, being a simplification of the evidence-based medicine (EBM) approach for the evaluation of sources of clinical evidence, but nonetheless quite helpful.<sup>9</sup> Indeed, reminding the reader a brief definition of EBM is timely and appropriate: EBM is “the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients”.<sup>9</sup> It must also be stressed that “the practice of evidence-based medicine requires integration of individual clinical expertise and patient references with the best available external clinical evidence from systematic search”.<sup>9</sup> Systematic reviews and meta-analyses, if well conducted and reported, help us in reducing our effort in looking for, appraising, and summarising the evidence. But the burden of deciding what to do with the evidence obtained for the care of our individual patient remains ours.

\* Corresponding author: Division of Cardiology, University of Turin, S. Giovanni Battista “Molinette” Hospital, Corso Bramante 88-90, 10126 Turin, Italy  
E-mail: gbiondizoccai@gmail.com

Thus, the first step in appraising a systematic review and meta-analysis is trying to find an answer to the question: can I trust it? In other words, is this review internally valid (i.e. does it provide a precise and largely unbiased answer to its scientific question)? Providing a definitive appraisal of the internal validity of a systematic review is not simple (see Tables 1 and 2), but it largely depends on the methods employed and reported regarding study search, selection, abstraction, appraisal and, if appropriate, pooling.

Even if we can conclude that a given meta-analysis is internally valid, we still have to face the second step in its appraisal. This second step mainly involves the external validity of the study. In other words, can I apply the review results to the case I am facing or will shortly face? It more basically means answering the question: so what? Decisions on external validity are highly subjective and may change depending on the clinical, historical, logistical, cultural or ethical context. Nonetheless, systematic reviews and meta-analyses can improve our appraisal of the external validity of any given clinical intervention, by suggesting its overall clinical efficacy (or lack of it).

In the following sections we will focus more practically on the systematic reviews published in this issue of EuroIntervention,<sup>1,2</sup> and provide the results of our two-step approach in their appraisal.

### The first step: can I trust it?

Can I trust the review by Tamhane et al<sup>1</sup> on the beneficial anti-restenotic effect of cilostazol after coronary stenting? I would conclude positively. They have searched several databases and conference proceedings, study selection was performed by two reviewers, with data abstraction followed by a thorough study quality appraisal. Endpoints were clearly spelled out and appropriately pooled with random and fixed-effect models, with the addition of several tests for small study bias. Despite the reliance on angiographic follow-up, which may inflate clinical restenosis rates and lead to more non-ischaemia driven revascularisations, their stance that cilostazol reduces restenosis and repeat revascularisations after coronary stenting is trustworthy.

Can I also trust the meta-analysis by Inaba et al<sup>2</sup> on devices for distal embolisation in acute myocardial infarction, showing that most beneficial findings in favour of these devices stem from single-centre studies? Even a quick look at their paper clearly suggests a positive answer. They performed a comprehensive literature search without language restrictions, with two reviewers independently selecting only high-quality studies, abstracting the data and scoring the internal validity of the included randomised trials. Finally, statistical pooling was performed with established methods complemented by meta-regression analyses exploring for the impact of major covariates. Their main focus was on looking for explanations of the apparent differences between individual studies focusing on this clinical research question. They compared the role of several covariates on the results of devices for distal protection, finding that single-centre studies were more likely to report beneficial results for these devices than multicentre studies. Thus, despite the post hoc feature of most of their comparisons and the risk of ecological fallacy (whereby inferences about the nature of specific individuals are based solely upon aggregate statistics collected for the group to which those individuals belong), their conclusions are trustworthy.

### The second step: so what?

So what? In other words, what is the external validity of the review by Tamhane et al<sup>1</sup> This question is more difficult to answer, and depends on several factors. For instance, cilostazol is not available in Italy, nor in other European countries. Second, the cost and duration of any cilostazol treatment should be factored in and formal cost-effectiveness studies should be completed. Third, I would consider cilostazol currently only in patients at very high risk of restenosis (e.g. diabetics or those with chronic renal failure), whereas a larger usage could be foreseen if the antiplatelet effects of cilostazol are confirmed, so that it could be used to further reduce the risk of stent thrombosis in those more prone to this adverse event.

**Table 1. The Oxman-Guyatt index<sup>7</sup> for the appraisal of systematic reviews and meta-analyses.\***

Question number	Details	Question number	Details
1	Were the search methods used to find evidence stated?	7	Were the methods used to combine the findings of the relevant studies reported?
2	Was the search for evidence reasonably comprehensive?	8	Were the findings of the relevant studies combined appropriately relative to the primary question the overview addresses?
3	Were the criteria for deciding which studies to include in the overview reported?	9	Were the conclusions made by the author(s) supported by the data and/or analyses reported in the overview?
4	Was bias in the selection of studies avoided?	10	This question summarises the previous ones and, specifically, asks to rate the scientific quality of the review from 1 (being extensively flawed) to 3 (carrying major flaws) to 5 (carrying minor flaws) to 7 (minimally flawed)
5	Were the criteria used for assessing the validity of the included studies reported?		
6	Was the validity of all studies referred to in the text assessed using appropriate criteria?		

\*This index evaluates the internal validity of a review on nine separate questions for which three distinct answers are eligible ('yes', 'partially/cannot tell', 'no'). The developers of the index specify that if the 'partially/cannot tell' answer is used one or more times in questions 2, 4, 6 or 8, a review is likely to have minor flaws at best and is difficult to rule out major flaws (i.e. a score≤4). If the 'no' option is used on question 2, 4, 6 or 8, the review is likely to have major flaws (i.e. a score≤3).

**Table 2. The AMSTAR (a measurement tool to assess systematic reviews) index<sup>8</sup> for the appraisal of systematic reviews and meta-analyses.\***

Question number	Details	Question number	Details
1	Was an 'a priori' design provided? The research question and inclusion criteria should be established before the conduct of the review.		chose to include only randomised, double-blind, placebo controlled studies, or allocation concealment as inclusion criteria); for other types of studies alternative items will be relevant.
2	Was there duplicate study selection and data extraction?		
3	Was a comprehensive literature search performed? At least two electronic sources should be searched. The report must include years and databases used (e.g. Central, EMBASE, and MEDLINE). Key words and/or MESH terms must be stated and where feasible the search strategy should be provided. All searches should be supplemented by consulting current contents, reviews, textbooks, specialised registers, or experts in the particular field of study, and by reviewing the references in the studies found.	8	Was the scientific quality of the included studies used appropriately in formulating conclusions? The results of the methodological rigour and scientific quality should be considered in the analysis and the conclusions of the review, and explicitly stated in formulating recommendations.
4	Was the status of publication (i.e. grey literature) used as an inclusion criterion? The authors should state that they searched for reports regardless of their publication type. The authors should state whether or not they excluded any reports (from the systematic review), based on their publication status, language etc.	9	Were the methods used to combine the findings of studies appropriate? For the pooled results, a test should be done to ensure the studies were combinable, to assess their homogeneity (i.e., chi-squared test for homogeneity, I <sup>2</sup> ). If heterogeneity exists a random effects model should be used and/or the clinical appropriateness of combining should be taken into consideration (i.e., is it sensible to combine?).
5	Was a list of studies (included and excluded) provided?	10	Was the likelihood of publication bias assessed? An assessment of publication bias should include a combination of graphical aids (e.g., funnel plot, other available tests) and/or statistical tests (e.g., Egger regression test).
6	Were the characteristics of the included studies provided? In an aggregated form such as a table, data from the original studies should be provided on the participants, interventions and outcomes.	11	Was the conflict of interest included? Potential sources of support should be clearly acknowledged in both the systematic review and the included studies.
7	Was the scientific quality of the included studies assessed and documented? A priori methods of assessment should be provided (e.g., for effectiveness studies if the authors		

\*This index evaluates the internal validity of a review on 11 separate questions for which four distinct answers are eligible ('yes', 'no', 'can't answer', 'not applicable').

After reading the article by Inaba et al,<sup>1</sup> what should I practically do? In other words, what is the external validity of their work? Should I use devices for distal protection only when I am participating in single-centre trial, because otherwise they are unlikely to be beneficial? Of course this simplistic attitude is almost never correct. More humbly, I could conclude that single-centre studies may provide more significant results because expertise and experience with a given device are much greater, but come at the expense of a higher risk of bias (e.g. lack of allocation concealment or appropriate blinding of assessors), in comparison to multicentre trials.

### Take home messages

It is clear that appraising the internal validity, and even more the external validity, of any research endeavour, including the meta-analyses by Tamhane et al<sup>1</sup> and by Inaba et al,<sup>2</sup> is highly subjective, and thus we leave ample room to the reader to enjoy them and appraise them on his or her own. The only issue that is worth being further stressed is that only collective and constructive, but critical post-publication appraisal of scientific studies can put and maintain them into the appropriate context for their correct and practical exploitation by the clinical researcher and the clinical practitioner.

### Acknowledgements

This work is part of a senior investigator research program of the Meta-analysis and Evidence-based medicine Training in Cardiology (METCARDIO) Center, based in Turin, Italy (<http://www.metcardio.org>).

### References

1. Tamhane U, Meier P, Chetcuti S, Chen KY, Rha SW, Grossman M-P, Gurm H. Efficacy of cilostazol in reducing restenosis in patients undergoing contemporary stent based PCI: a meta-analysis of randomized controlled trials. *EuroIntervention* 2009;5:384-393.
2. Inaba Y, Chen JA, Mehta N, Bergmann SR. Impact of single or multicenter study design on the results of trials examining the efficacy of adjunctive devices to prevent distal embolisation during acute myocardial infarction. *EuroIntervention* 2009;5:375-383.
3. Glasziou P, Djulbegovic B, Burls A. Are systematic reviews more cost-effective than randomised trials? *Lancet* 2006;367:2057-8.
4. Patsopoulos NA, Analatos AA, Ioannidis JP. Relative citation impact of various study designs in the health sciences. *JAMA* 2005;293:2362-6.
5. Agostoni P, Ribichini F, Wijns W. On the current epidemic outburst of meta-analytic rage in interventional cardiology. *EuroIntervention* 2009;5:19-23.
6. How-To Help and Videos - For Dummies. Available online at <http://www.dummies.com/> (last accessed on 16 June, 2009).
7. Oxman AD, Guyatt GH. Validation of an index of the quality of review articles. *J Clin Epidemiol* 1991;44:1271-8.
8. Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, Porter AC, Tugwell P, Moher D, Bouter LM. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol* 2007;7:10.
9. Guyatt G, Rennie D, Meade M, Cook D, eds. Users' guides to the medical literature: a manual for evidence-based clinical practice. 2nd edition. Columbus, OH: McGraw-Hill Professional, 2008.