Can we use levels of evidence to make a decision?

The views expressed in this editorial are those of the author and do not necessarily reflect the position of the Canadian Medical Association or its subsidiaries.

Everyone would like a truly objective score ranking the importance of articles in the medical literature. Academia has come to love rating systems as much as movie critics love to judge cinematic glory. Metrics such as the h-index, impact factor, and other rating scores have become increasingly important in academic promotion; however, the ability to score papers like we do authors is a little behind.

Theoretically, the levels of evidence used to grade papers should permit readers to ascertain quickly how relevant the paper is in the overall scope of evidence. From level 1 (meta-analysis) to level 5 (opinion) papers, there seems to be a clear gradient of overall importance to the body of literature. The level of evidence pyramid ranks meta-analyses, systemic reviews and randomized controlled trials (RCTs) at the top, whereas case-control studies, case reports and opinion articles are rated lower. Ranking methods are designed to assign greater weight to the top levels.

The problem with the system is that meta-analyses and systematic reviews depend heavily on large, well-executed RCTs being available in the literature. We know there is a paucity of such trials, and with that comes the adage “garbage in, garbage out.” Meta-analyses that combine a group of poorly designed but published RCTs result in strange findings. It is the reason that there are so few positive Cochrane review subjects. It is also the reason that some positive reviews are not relevant owing to the overweighting of one or two poorly constructed studies. This problem leaves the impression that prospective cohort studies may be the most valuable tools in the literature, with which we can make better-informed medical decisions while we wait for appropriate RCTs based on prospective data to be conducted and published. Furthermore, prospective cohort studies are more feasible to conduct than RCTs, and patients are more likely to participate because prospective studies are observational, with surgical choice left up to the treating surgeon rather than a random number generator. Large cohorts, now possibly available with prospectively performed inclusive studies, would permit evaluation of patient-centred outcomes.

But there is another stumbling block in Canada: the lack of pertinent data in some fields. For example, the Canadian Institute for Health Information has stopped tracking trauma despite it being the number one cause of death in people younger than 45 years. This throws a wet blanket over the prospect of new and interesting research that could be gleaned from looking at a cohort of trauma patients — an exercise that would benefit all Canada. Perhaps if a database existed in which all patients and pertinent data were included, we would be better able to determine important areas for future research.

It is hard to build a level of evidence pyramid to rectify the lack of pertinent data and methodologically sound RCTs. The pyramid might be easier to scale in other specialities, but there is still the lack of proper RCTs to contend with for relevance. We need to interpret the literature carefully — especially when we assign high trust levels to meta-analyses and systematic reviews. The pyramid is not really a pyramid, at least in 2020.

Edward Harvey, MD
Coeditor, Canadian Journal of Surgery

Competing interests: E.J. Harvey is the chief medical officer of Greybox Solutions, the co-founder and head of medical innovation of NXTSens Inc., the co-founder and chief medical officer of MY01 Inc., and the co-founder and director of Strathera Inc. He receives institutional support from J & J DePuy Synthes, Stryker and Zimmer, and he is a board member of the Orthopedic Trauma Association and the Canadian Orthopaedic Association.

DOI: 10.1503/cjs.001920