Level V Evidence

Research Pearls: The Significance of Statistics and Perils of Pooling. Part 3: Pearls and Pitfalls of Meta-analyses and Systematic Reviews

Joshua D. Harris, M.D., Jefferson C. Brand, M.D., Mark P. Cote, P.T., D.P.T., M.S.C.T.R., and Aman Dhawan, M.D.

Abstract: Within the health care environment, there has been a recent and appropriate trend towards emphasizing the value of care provision. Reduced cost and higher quality improve the value of care. Quality is a challenging, heterogeneous, variably defined concept. At the core of quality is the patient's outcome, quantified by a vast assortment of subjective and objective outcome measures. There has been a recent evolution towards evidence-based medicine in health care, clearly elucidating the role of high-quality evidence across groups of patients and studies. Synthetic studies, such as systematic reviews and meta-analyses, are at the top of the evidence-based medicine hierarchy. Thus, these investigations may be the best potential source of guiding diagnostic, therapeutic, prognostic, and economic medical decision making. Systematic reviews critically appraise and synthesize the best available evidence to provide a conclusion statement (a "take-home point") in response to a specific answerable clinical question. A meta-analysis uses statistical methods to quantitatively combine data from single studies. Meta-analyses should be performed with high methodological quality homogenous studies (Level I or II) or evidence randomized studies, to minimize confounding variable bias. When it is known that the literature is inadequate or a recent systematic review has already been performed with a demonstration of insufficient data, then a new systematic review does not add anything meaningful to the literature. PROSPERO registration and PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines assist authors in the design and conduct of systematic reviews and should always be used. Complete transparency of the conduct of the review permits reproducibility and improves fidelity of the conclusions. Pooling of data from overly dissimilar investigations should be avoided. This particularly applies to Level IV evidence, that is, noncomparative investigations. With proper technique, systematic reviews and meta-analyses have the potential to be powerful investigations that efficiently assist clinicians in decision making.

From Houston Methodist Orthopedics and Sports Medicine (J.D.H.), Houston, Texas; Heartland Orthopedic Specialists (J.C.B.), Alexandria, Minnesota; UConn Musculoskeletal Institute, Human Soft Tissue Research Laboratory, UConn Health (M.P.C.), Farmington, Connecticut; and Penn State Hershey Bone and Joint Institute (A.D.), Hershey, Pennsylvania, U.S.A.

The authors report the following potential conflicts of interest or sources of funding: J.D.H. received support from Arthroscopy: The Journal of Arthroscopic and Related Surgery, *Smith & Nephew, Magellan, Depuy Synthes, and SLACK Inc. J.C.B. received support from* Arthroscopy: The Journal of Arthroscopic and Related Surgery. *A.D. received support from* Arthroscopy: The Journal of Arthroscopic and Related Surgery, *the Arthroscopy Association of North America, the* Orthopaedic Journal of Sports Medicine, *Biomet, and Smith & Nephew.*

Received December 6, 2016; accepted January 31, 2017.

Address correspondence to Joshua D. Harris, M.D., Houston Methodist Orthopedics and Sports Medicine, Orthopedic Surgery, 6445 Main Street, Outpatient Center, Suite 2500, Houston, TX 77030, U.S.A. E-mail: joshuaharrismd@gmail.com

0749-8063/161227/$36.00

http://dx.doi.org/10.1016/j.arthro.2017.01.055

The delivery of health care is a highly intricate process involving patients, providers, payers, and policy makers. Within the health care environment, there has been a recent and appropriate trend towards emphasizing the value of care provision. Simply defined, value is the quality of care per unit of cost.[1] While cost can be easily defined and quantified, the quality of care is a challenging, heterogeneous, variably defined concept. At the core of quality is the patient's outcome, which may be quantified by a vast assortment of subjective and objective outcome measures. These measures generate scores that can be used to compare a patient's pre- and postintervention health status. Properly developed, valid, reliable, and responsive outcome measurement tools are necessary to quantify this status. These tools are vital components within evidence-based medicine. Within traditional science,

the focus has been inherently at the individual patient level and at the individual study level. However, the recent evolution towards evidence-based medicine has clearly elucidated the role of high-quality evidence across groups of patients and synthesis of studies. Synthetic studies, such as systematic reviews and meta-analyses, are at the top of the evidence-based medicine hierarchy (Fig 1). Thus, if performed rigorously, these investigations may be the best potential source for guiding diagnostic, therapeutic, prognostic, and economic medical decision making.

Evidence-Based Medicine Hierarchy

The Problem With the Randomized Trial—Is It Really the Gold Standard?

In the evaluation and management of patients, evidence-based medicine should be used to guide treatment options, discussions, and decisions. A well designed, conducted, and reported randomized controlled trial has long been revered as the gold standard of evidence for evaluating the effect of an intervention. This type of study is designed to produce valid results by limiting bias and confounding through the use of techniques including randomization and blinding. However, no study is without flaws or limitations. Recent investigations have demonstrated that a single randomized controlled trial may not produce reliable results. There are several reasons why a single randomized clinical trial may reach a different conclusion than other similar studies. Flexible statistical approaches, selective reporting, industry funding, trials that are stopped early on account of observing large positive effects (overestimation of the effect), financial and nonfinancial conflicts of interest, and differences in patient populations from one to study to another are but a few reasons that drive inconsistency and inflate treatment effects.[2] The challenge of irreproducibility is what makes synthesis of multiple publications very useful.

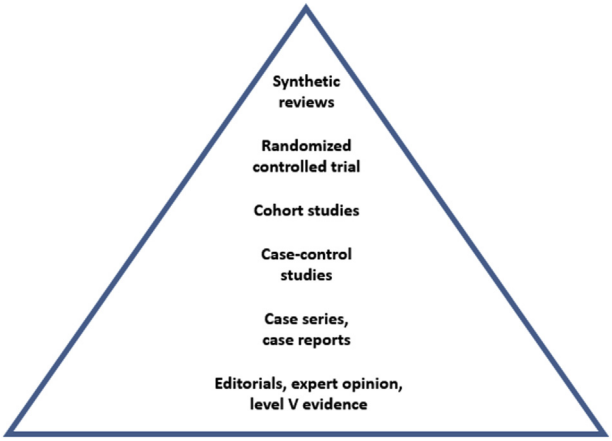

Fig 1. Evidence-based medicine study hierarchy. Synthetic reviews include systematic reviews and meta-analyses and are at the top of the pyramid.

Replication of study results is a fundamental activity in quantitative research.[2] Studies reporting positive findings often contradict one another (e.g., [1] eggs are good for you, then they're not; [2] red wine is unhealthy, then it's not; [3] vitamin C cures the common cold, then it doesn't).[3] In orthopaedic surgery, clavicle fractures should be treated nonsurgically,[4] then they should be treated surgically[5]; humerus fractures should be treated nonsurgically,[6] then they should be treated surgically.[7] These results highlight the need for similarly designed and executed studies to confirm or refute novel findings.

The above reasons provide support for well-executed systematic reviews and meta-analyses. The power of the systematic review lies in its ability to statistically combine patient outcomes from distinct, yet similar, research studies. This process allows the consistency, or lack thereof, of randomized trials to be examined and quantified. When executed correctly, the "true" effect of an intervention can be estimated with more precision than with a single trial. Equally importantly, reasons for inconsistencies (selective reporting, poor study design) and differences in the treatment effect among subgroups can be explored. Thus, the potential exists to provide more information than single high-quality studies and therefore make more powerful evidence-based conclusions. However, the quality and strength of the recommendations from a systematic review are only as strong as the quality of the individual studies included in the analysis. Just as with randomized controlled trials, great care must be exercised by journal reviewers and editors in the peer review process and readers of the publication in interpretation of the bias and extrapolation of the review's findings to translation to clinical practice. Given the recent rapid proliferation of both written and electronic publication outlets, well-done systematic reviews and meta-analyses are highly useful in their analysis and presentation of large bodies of evidence to busy clinicians unable to peruse the entire body of literature because they can answer a clinical question in a short amount of time (Fig 2).[8] In fact, systematic reviews are used by the American Academy of Orthopaedic Surgeons as the evidence to support creation of clinical practice guidelines and appropriate use criteria for common clinical conditions in orthopaedic surgery.[9]

Benefits of Systematic Reviews

Systematic reviews critically appraise and synthesize the best available evidence to provide a conclusion statement (a "take-home point") in response to a specific answerable clinical question. The execution of a systematic review must be transparent, so that any person (not just an author, scientist, researcher, clinician,