Special Review

How to Perform a Systematic Review and Meta-analysis of Diagnostic Imaging Studies

Paul Cronin, MD, MS, Aine Marie Kelly, MD, MS, MA, Duaa Altaee, BA, Bradley Foerster, MD, PhD, Myria Petrou, MD, MS, Ben A. Dwamena, MD

A systematic review is a comprehensive search, critical evaluation, and synthesis of all the relevant studies on a specific (clinical) topic that can be applied to the evaluation of diagnostic and screening imaging studies. It can be a qualitative or a quantitative (meta-analysis) review of available literature. A meta-analysis uses statistical methods to combine and summarize the results of several studies. In this review, a 12-step approach to performing a systematic review (and meta-analysis) is outlined under the four domains: (1) Problem Formulation and Data Acquisition, (2) Quality Appraisal of Eligible Studies, (3) Statistical Analysis of Quantitative Data, and (4) Clinical Interpretation of the Evidence. This review is specifically geared toward the performance of a systematic review and meta-analysis of diagnostic test accuracy (imaging) studies.

Key Words: Diagnostic accuracy; evidence-based medicine; evidence-based radiology; heterogeneity; literature search; meta-analysis; meta-regression; publication bias; receiver operating characteristic analysis; ROC analysis; sensitivity analyses; systematic review; subgroup analysis; threshold effect.

INTRODUCTION

Systematic reviews and meta-analyses have become popular in medicine and are very commonly applied to treatment trials. However, they are still less common for diagnostic imaging studies. Systematic reviews and meta-analyses aim to provide summaries of the average result. In the case of imaging tests, this is diagnostic performance such as sensitivity or specificity, and the uncertainty of this average. In radiology, the smaller patient size and limited methodological quality of the primary studies can limit the quality of the review and meta-analysis. However, systematic reviews and meta-analyses may be the best assessment of the published literature available at any point in time, especially in the absence of large, definitive trials. They may provide important information to guide patient care and direct future clinical research. Performing and interpreting systematic reviews in radiology can be challenging given the paucity of available clinical studies. However, if investigators adhere to proper methodology, systematic reviews may provide useful information from a comprehensive study of the literature with limited bias.

In this review, a 12-step framework for performing systematic reviews (and meta-analyses) is outlined under the four domains: (1) Problem Formulation and Data Acquisition, (2) Quality Appraisal of Eligible Studies, (3) Statistical Analysis of Quantitative Data, and (4) Clinical Interpretation of the Evidence (Table 1). We will subsequently use "systematic review" and "meta-analysis" to represent the whole process of evidence synthesis. The steps in "problem formulation and data acquisition" are "define the question and objective of the review," "establish criteria for including studies in the review," and "conduct a literature search to retrieve the relevant literature." The steps in "quality appraisal of eligible studies" are "extract data on variables of interest," "assess study quality and applicability to the clinical problem at hand," and "summarize the evidence qualitatively and, if appropriate, quantitatively (meta-analysis)." The steps in "statistical analysis of quantitative data" are "estimate summary diagnostic test performance metrics and display the data," "assess heterogeneity," "investigate data for publication bias," "assess the robustness of estimates of diagnostic accuracy using sensitivity analyses," and "explore and explain heterogeneity in test accuracy using subgroup analysis (if applicable)." The steps in "clinical interpretation of the evidence" are "graphically display how the evidence alters the posttest probability using a Fagan plot (Bayes nomogram), likelihood ratio scatter graph, or probability-modifying plot." This review is tailored for radiologists who are new to the process of performing a systematic review and meta-analysis. However, we hope that those with

Acad Radiol 2017; ■:■■–■■

From the Department of Radiology, University of Michigan, B1 132G Taubman Center/5302, 1500 East Medical Center Drive, Ann Arbor, MI 48109 (P.C., A.M.K., B.F., M.P., B.A.D.); Department of Neurology, University of Michigan (D.A.); Nuclear Medicine Service, VA Ann Arbor Health Care System, Ann Arbor, Mississippi (B.A.D.). Received August 10, 2017; revised November 21, 2017; accepted December 6, 2017. **Address correspondence to:** P.C. e-mail: pcronin@med.umich.edu

https://doi.org/10.1016/j.acra.2017.12.007

TABLE 1. An Outline of the Main Steps in Doing a Meta-analysis of Diagnostic Test Accuracy

1. Problem formulation and data acquisition
Step 1. Define the question and objective of the review
Step 2. Establish criteria for including studies in the review
Step 3. Conduct a literature search to retrieve the relevant literature
2. Quality appraisal of eligible studies
Step 4. Extract data on variables of interest
Step 5. Assess study quality and applicability to the clinical problem at hand
Step 6. Summarizing the evidence qualitatively and if appropriate, quantitatively (meta-analysis)
3. Statistical analysis of quantitative data
Step 7. Estimate diagnostic accuracy and display the data
Step 8. Assess heterogeneity
Step 9. Assess for publication bias
Step 10. Assess the robustness of estimates of diagnostic accuracy using sensitivity analyses (if applicable)
Step 11. Explore and explain heterogeneity in test accuracy using subgroup analysis (if applicable)
4. Clinical interpretation of the evidence
Step 12. Graphically display how the evidence alters the posttest probability

experience with systematic review and meta-analysis will also find new information in this article.

PROBLEM FORMULATION AND DATA ACQUISITION

Step 1. Define the Question and Objective of the Review

A good review question addresses a clinical problem for which there is uncertainty. Therefore, the first step is to identify the relevant clinical problem. This includes specifying the patient, the index test(s) and reference test being studied, and the outcome measurements (diagnostic test accuracy) (1). In evidence-based practice, these components can be abbreviated to PICO (Patient, Intervention, Comparator, and Outcome) or in the Cochrane guidelines for diagnostic accuracy tests as PICTS (Patient, Index test, Comparator test, Target disorder and Study design) (2–4). Patients can refer to patients presenting signs and symptoms of the disease (diagnostic studies), patients with the disease (prognostic studies), or population at risk of the disease (screening studies). The index test is the test to be evaluated. A meta-analysis may consider and compare several index tests. The comparator test is standard practice or the reference standard or the "gold standard" that the index tests are compared to. It is the test or procedure used to classify patients as having the target condition or disease or not. The target disorder is the disease that one is trying to diagnose. Examples of PICO questions or statements are shown in Table A1. These include "In patients with symptomatic carotid stenosis, how does computed tomographic angiography (CTA) compare with magnetic resonance angiography (MRA) for the detection and quantification of carotid stenosis?" or "In patients with known or suspected coronary artery disease, how does CT coronary angiography compare with invasive catheter coronary angiography for identifying one (or more) potentially or probably hemodynamically significant (≥50% coronary artery luminal diameter) stenosis in terms of sensitivity, specificity and diagnostic accuracy?" or "In patients with a solitary pulmonary nodule, how well does dynamic contrast material–enhanced CT, dynamic contrast material–enhanced MR imaging, FDG PET, and ^{99m}Tc-depreotide SPECT compare for the diagnosis of malignancy (diagnostic accuracy)?" or "In patients with known or suspected rotator cuff tears, how does ultrasound compare to MRI for diagnosis?" or "Is low-dose CT colonography equivalent to optical colonoscopy in identifying clinically meaningful colonic polyps?" It should be remembered that evidence synthesis can be derailed by not asking a focused question. It is also important to have a focused research question as this is used to direct the search.

Step 2. Establish Criteria for Including Studies in the Review

In the perfect diagnostic imaging study, all patients receive one or more index tests and the "gold" standard test. However, in reality, there can be important deviations from the study described previously. Examples include using different sets of inclusion or exclusion criteria for those with and those without the target disease. Another example is, verification of the index test results based on information that will only be available after inclusion in the study. These important issues should be considered when drawing up inclusion and exclusion criteria.

Bias or variation may be introduced in five aspects of diagnostic imaging study design. First, the criteria used for study population selection; second, comparator test selection; third, index test and comparator test execution; forth, index test and comparator test interpretation; and fifth, result analysis. The inclusion criteria should incorporate all relevant clinical characteristics of the target condition with which such patients would present. It may be important to include the imaging setting, as test accuracy can vary between primary, secondary, and tertiary care, and also between screening and diagnostic uses. The inclusion criteria for the index test may include details of the tests being evaluated, such as, but not limited to, the manufacturer, type of image processing, and generation of technology. This also applies for the comparator test.

Step 3. Conduct a Literature Search to Retrieve the Relevant Literature

Secondary research too is prone to biases, especially selection bias and publication bias. Selection bias, which the researcher has control over, is bias in the published studies included in the review. Publication bias, which the researcher does not have control over, is bias in the primary