Journal of Clinical Epidemiology 77 (2016) 60−67

Meta-epidemiologic study showed frequent time trends in summary estimates from meta-analyses of diagnostic accuracy studies

Jérémie F. Cohen[a,b,c,*], Daniël A. Korevaar[a], Junfeng Wang[a], Mariska M. Leeflang[a], Patrick M. Bossuyt[a]

[a]*Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Academic Medical Center, University of Amsterdam, PO Box 22700, 1100 DE Amsterdam, The Netherlands*

[b]*INSERM U1153, Obstetrical, Perinatal and Pediatric Epidemiology Research Team (EPOPé), Centre de Recherche Épidémiologie et Statistique Sorbonne Paris Cité (CRESS), Paris Descartes University, 53, avenue de l'Observatoire, 75014 Paris, France*

[c]*Department of Pediatrics, Necker-Enfants Malades Hospital, Assistance Publique-Hôpitaux de Paris, Paris Descartes University, 149, rue de Sèvres, 75015 Paris, France*

Accepted 29 April 2016; Published online 20 May 2016

Abstract

Objectives: To evaluate changes over time in summary estimates from meta-analyses of diagnostic accuracy studies.

Study Design and Setting: We included 48 meta-analyses from 35 MEDLINE-indexed systematic reviews published between September 2011 and January 2012 (743 diagnostic accuracy studies; 344,015 participants). Within each meta-analysis, we ranked studies by publication date. We applied random-effects cumulative meta-analysis to follow how summary estimates of sensitivity and specificity evolved over time. Time trends were assessed by fitting a weighted linear regression model of the summary accuracy estimate against rank of publication.

Results: The median of the 48 slopes was −0.02 (−0.08 to 0.03) for sensitivity and −0.01 (−0.03 to 0.03) for specificity. Twelve of 96 (12.5%) time trends in sensitivity or specificity were statistically significant. We found a significant time trend in at least one accuracy measure for 11 of the 48 (23%) meta-analyses.

Conclusion: Time trends in summary estimates are relatively frequent in meta-analyses of diagnostic accuracy studies. Results from early meta-analyses of diagnostic accuracy studies should be considered with caution.

Keywords: Systematic reviews; Meta-analysis; Cumulative meta-analysis; Diagnostic accuracy; Sensitivity and specificity; Time trends

1. Introduction

Clinicians use diagnostic tests in daily practice to evaluate the likelihood of a disease in a given patient. Test accuracy expresses the ability of a diagnostic test to correctly identify patients as having a target condition. In diagnostic accuracy studies, the results of the test under evaluation are compared with the results obtained with the reference standard, which is the best available clinical method for detecting the target condition.

The diagnostic accuracy is not a fixed property of a medical test. Accuracy often varies across studies, not only because of chance, but also because of differences in study design and setting, participant characteristics, and disease prevalence or severity [1−4]. Such study features may change over time; early, exploratory studies often use methods that differ from those in later, more confirmatory studies [5]. Other authors found time trends in estimates of accuracy using meta-regression [6]. In such cases, summary estimates from meta-analyses of diagnostic accuracy studies may also evolve over time, as new data accumulate.

Because of such dynamics in diagnostic research, the implications for clinical practice and health care policy derived from the results of a meta-analysis may be unstable. A diagnostic test that was previously considered to be sufficiently accurate, based on the first few accuracy studies, may later turn out not to be fit for purpose, considering

Conflict of interest: None.

Funding: J.F.C. is supported by a research grant from Assistance Publique-Hôpitaux de Paris (Année Médaille de l'Internat 2013).

* Corresponding author. Tel.: +31(20)566-8975; fax: +31(20)691-2683.

E-mail address: j.f.cohen@amc.uva.nl (J.F. Cohen).

http://dx.doi.org/10.1016/j.jclinepi.2016.04.013

What is new?

Key findings

- This study provides empirical evidence that time trends are frequent in summary estimates from meta-analyses of diagnostic accuracy studies.

What this adds to what was known?

- Cumulative meta-analysis has proven useful as a way to investigate time trends in summary estimates from meta-analyses of diagnostic accuracy studies.

What is the implication and what should change now?

- Because estimates based on more studies often show different performance, conclusions based on early meta-analyses, with few studies, may be unreliable.
- We call for caution when decision making is based on results from early meta-analyses of diagnostic accuracy studies.

the combined evidence from additional studies. This potential for change constitutes part of the rationale for periodically updating systematic reviews in general.

Cumulative meta-analysis has been proposed as a way to investigate time trends in summary estimates of effect [7–10]. In contrast with conventional meta-regression, the aim of cumulative meta-analysis in such applications is not so much to investigate how diagnostic accuracy estimates evolve over time across primary studies, but to follow how summary estimates evolve over time, as new studies are additionally incorporated in the meta-analysis. Using these methods, several investigators found empirical evidence of time trends in summary effects of medical interventions [11,12]. Here, we report on the use of cumulative meta-analysis to assess temporal changes in summary estimates of diagnostic accuracy in a sample of meta-analyses.

2. Materials and methods

2.1. Empirical data

We used a data set collected in a broad meta-epidemiologic project setup to examine the methods used in systematic reviews and meta-analyses of diagnostic accuracy studies [13,14]. For the purpose of this overarching project, we aimed at collecting a sample of at least 100 consecutively published diagnostic test accuracy systematic reviews. Details of the search strategy are described in Appendix A.1 at www.jclinepi.com. The search targeted MEDLINE-indexed diagnostic accuracy systematic reviews published in English between September 2011 and January 2012, in which at least one meta-analysis was performed.

For the present study, we excluded meta-analyses that did not provide data for the contingency tables, and those for which more than half of the primary studies were not indexed in MEDLINE. We also excluded meta-analyses with less than five primary studies [2,15].

In some systematic reviews, several meta-analyses were performed (e.g., different index tests, different reference standards, different inclusion criteria). Because we wanted to avoid overrepresentation of specific primary studies or specific reviews, we decided to include only a limited number of meta-analyses per review. If multiple meta-analyses were reported within one review, we decided to include only the two with the largest number of participants. We also excluded meta-analyses with a time span of less than 5 years between the first and the last primary study included in the meta-analysis.

2.2. Data extraction

Data extraction was performed by two independent authors (J.F.C. and D.A.K.); disagreements were resolved through discussion. For each included meta-analysis, we extracted the following information: first author, target condition, index test under evaluation, type of index test (imaging, laboratory, other), time span in years from first to last primary study included in the meta-analysis, total number of primary studies, and total number of participants. For each primary study, we extracted the name of the first author, publication date (Appendix A.2 at www.jclinepi.com), and contingency tables (number of true positives, false positives, false negatives, and true negatives). In each meta-analysis, we ranked primary studies by publication date.

2.3. Statistical analysis

2.3.1. Cumulative meta-analysis of diagnostic data

For each meta-analysis, we used data from the contingency tables to perform random-effects meta-analysis of diagnostic accuracy estimates. Our accuracy measures were sensitivity and specificity. We performed random-effects univariate meta-analysis of logit-transformed sensitivity and specificity separately. Between-study heterogeneity was estimated using the DerSimonian-Laird method. In case a primary study had a 0 cell, we added 0.5 to all cells from the contingency table.

Following the principles of cumulative meta-analysis, summary estimates of logit sensitivity and logit specificity and their variance were recalculated iteratively, by incorporating primary studies one by one in the meta-analysis, according to their rank of publication date [7]. In such a cumulative meta-analysis, the first summary estimate corresponds to the estimate from the first published primary