ELSEVIER

Journal of Clinical Epidemiology 67 (2014) 612–621

Journal of Clinical Epidemiology

REVIEW ARTICLES

Diagnostic tests often fail to lead to changes in patient outcomes

Konstantinos C. Siontis[a], George C.M. Siontis[b], Despina G. Contopoulos-Ioannidis[c,d], John P.A. Ioannidis[e,f,g,*]

[a]*Department of Medicine and Mayo School of Graduate Medical Education, College of Medicine, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, USA*

[b]*Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, University Campus, P.O. Box 1186, 45110 Ioannina, Greece*

[c]*Division of Infectious Diseases, Department of Pediatrics, Stanford University School of Medicine, Stanford, CA 94305, USA*

[d]*Palo Alto Medical Foundation Research Institute, Palo Alto, CA 94301, USA*

[e]*Department of Medicine, Stanford Prevention Research Center, Stanford University School of Medicine, 1265 Welch Rd, MSOB X306, Stanford, CA 94305, USA*

[f]*Department of Health Research and Policy, Stanford University School of Medicine, Stanford, CA 94305, USA*

[g]*Department of Statistics, Stanford University School of Humanities and Sciences, Stanford, CA 94305, USA*

Accepted 9 December 2013; Published online 27 March 2014

Abstract

Objectives: To evaluate the effects of diagnostic testing on patient outcomes in a large sample of diagnostic randomized controlled trials (D-RCTs) and to examine whether the effects for patient outcomes correlate with the effects on management and with diagnostic accuracy.

Study Design and Setting: We considered D-RCTs that evaluated diagnostic interventions for any condition and reported effectiveness data on one or more patient outcomes. We calculated odds ratios for patient outcomes and outcomes pertaining to the use of further diagnostic and therapeutic interventions and the diagnostic odds ratio (DOR) for the accuracy of experimental tests.

Results: One hundred forty trials (153 comparisons) were eligible. Patient outcomes were significantly improved in 28 comparisons (18%). There was no concordance in significance and direction of effects between the patient outcome and outcomes for use of further diagnostic or therapeutic interventions (weighted κ 0.02 and 0.09, respectively). The effect size for the patient outcome did not correlate with the effect sizes for use of further diagnostic ($r = 0.05$; $P = 0.78$) or therapeutic interventions ($r = 0.18$; $P = 0.08$) or the experimental intervention DOR in the same trial ($r = -0.24$; $P = 0.51$).

Conclusion: Few tests have well-documented benefits on patient outcomes. Diagnostic performance or the effects on management decisions are not necessarily indicative of patient benefits. © 2014 Published by Elsevier Inc.

Keywords: Diagnostic tests; Randomized controlled trials; Diagnostic odds ratio; Patient outcomes; Diagnostic accuracy; Empirical assessment

1. Introduction

Far more studies on diagnostic tests focus on diagnostic accuracy rather than the assessment on the clinical impact of testing. Physicians largely rely on diagnostic accuracy information to decide on the usefulness of a test [1] because direct evaluation of patient benefits by testing is difficult to measure [2,3]. However, diagnostic accuracy may not necessarily translate into patient benefits [4]. One would anticipate that a good test would effectively guide further testing and selection of treatments, but it is not always clear whether the increase or decrease of downstream diagnostic or therapeutic interventions translates into improved patient outcomes. There is currently no empirical evidence on the inference of effectiveness of diagnostic tests for improving patient outcomes from information on other outcomes or even from diagnostic accuracy alone.

The most conclusive evidence regarding patient outcomes can be derived from diagnostic randomized controlled trials (D-RCTs), in which participants are randomized to have a new diagnostic test vs. a control or no test [5–7]. Performing such trials is challenging [8,9], but randomized controlled trials (RCTs) represent a rigorous approach to diagnostic test evaluation [10].

D-RCTs evaluating patient outcomes are relatively uncommon [11], but their examination can offer useful insights. Here, we performed an empirical assessment of a large sample of D-RCTs addressing diverse patient outcomes. We investigated how often diagnostic testing is found to have significant effects (and, if so, of what magnitude) on patient outcomes and whether the effects on patient outcomes are

Conflict of interest: None.

Funding: None.

* Corresponding author. Tel.: 650-725-5465.

E-mail address: jioannid@stanford.edu (J.P.A. Ioannidis).

0895-4356/$ - see front matter © 2014 Published by Elsevier Inc.

http://dx.doi.org/10.1016/j.jclinepi.2013.12.008

What is new?

Key findings

- In a sample of 140 diagnostic test studies with patient outcomes, in only one in five was there evidence of significant changes in patient outcomes.
- The effects of testing on patient outcomes did not correlate with the effects on further diagnostic and therapeutic interventions or with the diagnostic accuracy of tests.

What this adds to what was known?

- Our findings contrast to the much higher rate of patient outcome improvement seen in the usual randomized controlled trials of drug interventions.

What is the implication and what should change now?

- Outcome reporting in diagnostic testing evaluation should become more patient centered. Reporting of the effects of testing on different types of outcomes should be part of routine evaluation of diagnostic tests.

concordant with the effects on the use of diagnostic and therapeutic interventions and test accuracy.

2. Methods

2.1. Literature searches

We aimed at generating a large reproducible sample of D-RCTs, acknowledging that it is impossible to identify all such trials. We searched MEDLINE using the Clinical Queries tool in PubMed (last update, February 2012). We used the narrow and specific option (''Diagnosis/Narrow [filter]'') and then specifically for the six general medical journals with the highest impact factors (ISI Journal Citation Reports 2010), that is, *New England Journal of Medicine*, *Lancet*, *JAMA*, *Annals of Internal Medicine*, *PLoS Medicine*, and *BMJ*, we used the ''broad, sensitive'' option (''Diagnosis/Broad[filter]''). Searches used the *Randomized Controlled Trial*, *English*, and *Humans* filters. We further searched for eligible trials among reviews of diagnostic interventions included in the Cochrane Database of Systematic Reviews (2010, Issue 1) using the terms *diagnos**, *screen**, and *monitor** (in the title). Finally, we perused the references of all eligible trials for additional trials not captured by the aforementioned literature searches. Searches were performed without year restrictions.

2.2. Study eligibility

Titles and abstracts were scrutinized, and potentially eligible studies were assessed further. Eligible studies were RCTs that randomized participants to a test vs. another test or no testing for diagnosis, monitoring, or screening of any medical condition and those that had assessed efficacy for at least one patient outcome that occurred during follow-up after the performance of a diagnostic test and after any other diagnostic and therapeutic procedures that might have been performed. Eligible outcomes included objectively assessed clinical events (eg, death), change of a relevant laboratory or imaging measure, and quality-of-life assessments. Any type of diagnostic intervention was deemed eligible for consideration, including clinical examination, patient-reported scores, and devices monitoring clinical signs and symptoms. For screening tests, we did not consider conditions that were detected during scheduled screening but focused on subsequent outcomes (eg, we did not consider documentation of cervical intraepithelial neoplasia during Pap smear screening but did consider subsequent cancer deaths). Immediate adverse events of testing (eg, bleeding from liver biopsy) and the impact of testing on further diagnostic or therapeutic procedures did not qualify as eligible patient outcomes.

Studies were eligible regardless of whether they compared single tests or diagnostic strategies/algorithms (eg, algorithms for pulmonary embolism diagnosis). When more than one publications from the same trial reported data for the main outcome, we included the longest follow-up. We excluded crossover trials in which patients underwent the compared tests in random order.

2.3. Data extraction

From each eligible trial, we extracted several demographic, reporting, and methodological characteristics (blinding, allocation concealment, and losses to follow-up) [12]. For trials with three or more arms, all possible pairwise comparisons were considered separately. For each comparison, we recorded the experimental and control tests compared and the number of participants randomized per arm. We adhered to the research hypothesis of each trial to decide which arm is the experimental one. Whenever available, we also recorded metrics of diagnostic performance of the tests that would allow calculation of the diagnostic odds ratio (DOR) [13], that is, the ratio of the products of the diagonals in a 2×2 table or ratio of positive likelihood ratio over negative likelihood ratio.

For each trial, we recorded the type of primary outcome: diagnostic use (ie, effect on the use of further diagnostic tests), therapeutic use (ie, effect on the use of therapies), and patient outcome (as defined previously). In each trial, we selected one patient outcome, and when available, we also selected one diagnostic use and one therapeutic use outcome. The diagnostic and therapeutic use outcomes

Download English Version:

https://daneshyari.com/en/article/10513513

Download Persian Version:

https://daneshyari.com/article/10513513

Daneshyari.com