Reported estimates of diagnostic accuracy in ophthalmology conference abstracts were not associated with full-text publication

Daniël A. Korevaar[a,*], **Jérémie F. Cohen**[a,b], **René Spijker**[c,d], **Ian J. Saldanha**[e], **Kay Dickersin**[e], **Gianni Virgili**[f], **Lotty Hooft**[c], and **Patrick M.M. Bossuyt**[a]

[a]Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Academic Medical Center, University of Amsterdam, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands [b]INSERM U1153, Obstetrical, Perinatal and Pediatric Epidemiology Research Team (EPOPé), Center for Epidemiology and Statistics Sorbonne Paris Cité, Paris Descartes University, 53, avenue de l'Observatoire, 75014 Paris, France [c]Cochrane Netherlands, Julius Center for Health Sciences and Primary Care, University Medical Centre Utrecht, University Utrecht, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands [d]Medical Library, Academic Medical Center, University of Amsterdam, Amsterdam, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands [e]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, 615 North Wolfe Street, Baltimore, MD 21205, USA [f]Department of Translational Surgery and Medicine, Eye Clinic, University of Florence, largo Brambilla, 3, 50134 Florence, Italy

Abstract

Objective—To assess whether conference abstracts that report higher estimates of diagnostic accuracy are more likely to reach full-text publication in a peer-reviewed journal.

Study Design and Setting—We identified abstracts describing diagnostic accuracy studies, presented between 2007 and 2010 at the Association for Research in Vision and Ophthalmology (ARVO) Annual Meeting. We extracted reported estimates of sensitivity, specificity, area under the receiver operating characteristic curve (AUC), and diagnostic odds ratio (DOR). Between May and July 2015, we searched MEDLINE and EMBASE to identify corresponding full-text publications; if needed, we contacted abstract authors. Cox regression was performed to estimate associations with full-text publication, where sensitivity, specificity, and AUC were logit transformed, and DOR was log transformed.

Results—A full-text publication was found for 226/399 (57%) included abstracts. There was no association between reported estimates of sensitivity and full-text publication (hazard ratio [HR] 1.09 [95% confidence interval {CI} 0.98, 1.22]). The same applied to specificity (HR 1.00 [95% CI 0.88, 1.14]), AUC (HR 0.91 [95% CI 0.75, 1.09]), and DOR (HR 1.01 [95% CI 0.94, 1.09]).

---

*Corresponding author. Tel.: 0031-20566-1099. d.a.korevaar@amc.uva.nl (D.A. Korevaar).

Author Manuscript

Conclusion—Almost half of the ARVO conference abstracts describing diagnostic accuracy studies did not reach full-text publication. Studies in abstracts that mentioned higher accuracy estimates were not more likely to be reported in a full-text publication.

1. Introduction

There is abundant evidence that many biomedical studies never reach full-text publication in a peer-reviewed journal [1–3]. Studies with statistically significant results are published more often than those with nonsignificant results [1,2,4–6]. The resulting overrepresentation of "positive" findings in the biomedical literature may introduce publication bias when researchers try to synthesize the available evidence, such as in systematic reviews and clinical practice guidelines [7]. Failure to publish studies jeopardizes adequate patient care and stifles scientific progress, while violating the ethical responsibility to disseminate study findings and use health care and research funds appropriately [8].

Diagnostic accuracy studies assess how well a medical test differentiates between patients with and without a specific target condition. Accurate tests are important in clinical practice because false positive results could expose patients to unnecessary medical interventions, while false negative results could lead to withholding needed treatments. Although there are many investigations of the extent and drivers of publication bias among studies of therapeutic interventions, similar investigations are rare for diagnostic accuracy studies [9,10].

Failure to reach full-text publication has recently been shown to be problematic among diagnostic accuracy studies, but the mechanisms of such failures are largely unknown [11–13]. Statistical significance is unlikely to be a major determinant because most of these studies present their results in terms of accuracy estimates such as sensitivity, specificity, and area under the receiver operating characteristic curve (AUC), without clear hypothesis tests and accompanying *P*-values [14–17]. Nevertheless, when promising findings in conference abstracts, reflecting strong performance of diagnostic tests, more easily reach full-text publication, overoptimistic impressions of a test's accuracy can result. This could invite premature adoption or inappropriate clinical use of tests.

The objective of this study was to evaluate the extent to which diagnostic accuracy studies presented as conference abstracts at an international ophthalmology meeting reached full-text publication in a peer-reviewed journal and to assess associations between reported accuracy estimates and full-text publication. We hypothesized that abstracts reporting higher estimates of diagnostic accuracy would more often lead to full-text publication.

2. Materials and methods

2.1. Selection of conference abstracts

Conference abstracts were considered for inclusion in our study if they were presented between 2007 and 2010 at the annual meeting of the Association for Research in Vision and Ophthalmology (ARVO), the world's largest gathering of eyes and vision researchers. This time frame was selected to ensure that we would include a large number of around 400 abstracts and that abstract authors would have a sufficient amount of time for full-text publication.

Abstracts were eligible for inclusion if the authors reported on a study that assessed the accuracy of one or more diagnostic tests to establish a clinical diagnosis in humans, and calculated—or announced the calculation of—at least one of the following measures of diagnostic accuracy: sensitivity, specificity, predictive values, likelihood ratios, AUC, diagnostic odds ratio (DOR), or total accuracy. We excluded abstracts reporting on the prognostic or predictive accuracy of tests, evaluated against a future event or the outcome of treatment. We also excluded abstracts for which updated results were reported in another included abstract.

Potentially eligible abstracts were identified by searching ARVO's online abstract proceedings. The full search strategy was developed by two investigators (D.A.K., in consultation with R.S., a medical information specialist). It consists of 34 different terms that are commonly used in reports of diagnostic accuracy studies (Appendix A at www.jclinepi.com). All retrieved abstracts were screened for inclusion by one investigator (D.A.K.). Whenever there was any doubt whether an abstract fulfilled the inclusion criteria, the case was discussed with a second investigator (J.F.C. and/or P.M.M.B.).

2.2. Data extraction

One investigator (D.A.K.) extracted data from the included ARVO abstracts, and a second investigator (J.F.C.) verified all extracted data. Disagreements were resolved through discussion.

We extracted the first author, year of presentation at ARVO, number of authors, continent of first author, international affiliations (authors from multiple countries/all authors from one country). We also extracted declared conflicts of interest (at least one author/none of the authors), acknowledged funding for the study (industry/nonindustry only/none), whether a trial registration number was provided, study design (cohort/case-control/unclear), data collection (prospective/retrospective/unclear), research field (glaucoma/ocular surface and corneal diseases (keratoconus and dry eye)/common chorioretinal diseases (diabetic retinopathy and age-related macular degeneration)/other), and number of participants and eyes.

For each abstract, we also extracted the highest reported estimate of sensitivity, specificity, AUC, and DOR [18]. The DOR is a single statistic summarizing the results of a 2×2 table; higher values represent better performance [19]. Because not all diagnostic accuracy studies report a DOR, we recalculated this from reported pairs of estimates of sensitivity and

specificity, positive and negative predictive value, or positive and negative likelihood ratio, or from the AUC, using standard formulas [19,20]. In this recalculation, a correction needed to be applied to accuracy estimates of 0 or 1; these were considered to be 0.01 and 0.99, respectively [19].

2.3. Identification of full-text publications

Between May and July 2015, at least 5 years after each included abstract was presented at ARVO, we undertook the following steps to identify corresponding full-text publications in peer-reviewed journals. Similar search strategies were used in previous related projects [11,12,21]:

- For each abstract, one investigator (D.A.K.) searched MEDLINE (through PubMed) and EMBASE (through Ovid) by separately using the abstract's first, second, and last authors' name, combined with (synonyms of) the test(s) under investigation and/or (synonyms of) the target condition. First, the titles and abstracts of retrieved articles were screened; if potentially corresponding to the ARVO abstract, the full-text was assessed.

- If unsuccessful, a second investigator (R.S.) repeated the search and additionally searched Google Scholar, using the same strategy.

- If still no full-text publication could be identified, we tried contacting abstract authors via email. One investigator (D.A.K.) searched for e-mail addresses of two abstract authors through their previous publications and institutional web sites. These two authors were successively contacted, each with two reminders, if necessary. If no response was received or if no working e-mail address of any authors could be identified, the abstract was considered to not have reached full-text publication.

We matched ARVO abstracts and full-text publications by comparing authors' names, dates of participant recruitment, participant characteristics, and technical details about the diagnostic tests applied. Abstracts were considered to have reached full-text publication if at least some of the presented diagnostic accuracy data were reported in the corresponding publication. This means that if an abstract reported on the accuracy of two tests and the publication only reported on the accuracy of one of these, the abstract was considered to have reached full-text publication. However, if the abstract and publication corresponded to the same study, but the publication did not report on test accuracy or only reported on the accuracy of a test that was not presented in the abstract, the abstract was considered to not have reached full-text publication.

If an abstract corresponded to a publication, but reported results were discrepant, the abstract was still considered to have reached full-text publication. This was also the case if an abstract corresponded to a publication that described a lower or higher number of participants. If multiple abstracts corresponded to a single publication, all were considered have reached full-text publication.

If there was any doubt whether an abstract and full-text publication matched, the case was discussed within the research team (D.A.K., with G.V. and/or P.M.M.B.). If doubt persisted,

study investigators were contacted by email for confirmation. For each full-text publication, we considered the date the article was added to the PubMed database as the publication date. If multiple full-text publications corresponded to one abstract, the date of the first publication was selected.

2.4. Statistical analysis

We calculated the overall proportion of abstracts reaching full-text publication and the median time from presentation at ARVO to full-text publication.

Univariable Cox proportional hazards regression analyses were performed, and hazard ratios [HR] were calculated, to analyze whether the accuracy estimates reported in ARVO abstracts were associated with full-text publication. For these analyses, sensitivity, specificity, and AUC were logit transformed, and DOR was log transformed. These transformed accuracy estimates were added as continuous variables to the regression model. Abstracts without accuracy estimates were included in the regression model by adding an indicator of missingness. Full-text publications that were published before or at the date of presentation of the abstract were arbitrarily considered published 1 month after presentation. Publication times for abstracts that did not reach full-text publication were considered censored at May 2015, the month in which we started our searches for corresponding publications. A *P*-value of 0.05 was considered statistically significant.

In addition, we used chi-square tests to explore the association between other abstract characteristics and full-text publication. In this explorative analysis, a Bonferroni correction was applied to adjust for multiple testing, where a *P*-value of 0.004 was considered statistically significant.

All statistical analyses were performed in SPSS version 22 (IBM, Armonk, NY, USA).

3. Results

3.1. Included conference abstracts

In total, 24,497 abstracts were presented at ARVO between 2007 and 2010, of which 958 were identified in our search (Fig. 1). After screening the abstracts, 399 could be included. References of included abstracts are provided in Appendix B at www.jclinepi.com.

Characteristics of the included ARVO abstracts are reported in Table 1. Disagreements between the two reviewers occurred in 1% (75/6,783) of extracted data elements. The median number of authors was 5 (interquartile range [IQR] 4–7), and most first authors were affiliated with organizations in the United States ($n = 151$; 38%), followed by Germany ($n = 34$; 9%) and the UK ($n = 29$; 7%). Some abstracts ($n = 75$; 19%) contained authors from multiple countries. In 133 (33%) abstracts, at least one author declared a conflict of interest, but industry funding for the study was acknowledged in only 37 (9%) abstracts.

A trial registration number was provided in 26 (7%) abstracts, all referring to ClinicalTrials.gov. Most abstracts described a case-control study ($n = 219$; 55%), and almost

half referred to glaucoma research ($n = 186$; 47%). The median number of participants was 107 (IQR 55–223), with a median number of 140 eyes (IQR 75–267).

3.2. Full-text publication

For 226 of 399 (57%) ARVO abstracts, we found a corresponding full-text publication in a peer-reviewed journal. Of these, 220 (97%) were identified through our literature searches, and six (3%) by contacting study authors (Fig. 1). For 15 of 226 (7%) abstracts that reached full-text publication, the number of participants in the abstract was more than 10% greater than the number of participants reported in the corresponding full-text publication. Among abstracts that reached full-text publication, the median time from presentation to publication was 17 months (IQR 8–29) (Fig. 2). Thirteen full-text publications were published before the date of presentation of the corresponding abstract at ARVO.

We confirmed nonpublication by e-mail contact with the authors of 119 of 173 (69%) abstracts for which we were unable to identify a matching full-text publication (Fig. 1). The number of participants was reported for 138 of 173 abstracts that did not reach full-text publication and totaled 50,500.

An overview of proportions of abstracts reaching full-text publication across subgroups defined by abstract characteristics is provided in Table 1.

3.3. Accuracy estimates and full-text publication

We grouped abstracts by quartiles of accuracy estimates (Table 2). Across the ARVO abstracts, 63% of those reporting a sensitivity in the highest quartile reached full-text publication, compared to 48% of those reporting a sensitivity in the lowest quartile. These proportions were 61% and 62% for specificity, 58% and 71% for AUC, and 55% and 56% for DOR, respectively.

In Cox proportional hazards regression analyses, there was no statistically significant association between reported estimates of sensitivity and full-text publication (HR 1.09 [95% confidence interval [CI] 0.98, 1.22]) (Table 3). The same applied to specificity (HR 1.00 [95% CI 0.88, 1.14]), AUC (HR 0.91 [95% CI 0.75, 1.09]), and DOR (HR 1.01 [95% CI 0.94, 1.09]).

4. Discussion

4.1. Comparison with published literature

Almost half of the conference abstracts describing diagnostic accuracy studies presented at the annual ARVO meeting between 2007 and 2010 did not reach full-text publication, 5 years or more after presentation. This represents diagnostic accuracy data collected in at least 50,500 study participants for which findings were not fully reported.

This massive failure to reach full-text publication is in line with previous evaluations of conference abstracts in different fields of biomedical research. A Cochrane systematic review summarized 79 such evaluations and found that, on average, only 45% of abstracts reached full-text publication [2]. One of these evaluations was performed among 327

abstracts that were randomly selected from all studies presented in 1985 at ARVO; a full-text publication could be identified for 63% [22]. Another evaluation was performed among 93 abstracts describing randomized trials that were presented between 1988 and 1989 at ARVO or the American Academy of Ophthalmology; a full-text publication could be identified for 66% [23]. More recently, it was found that among 513 abstracts describing randomized trials that were presented between 2001 and 2004 at ARVO, 45% reached full-text publication [24].

Unfortunately, failure to publish is not a random phenomenon. There is overwhelming evidence of publication bias, caused by an overrepresentation of positive and favorable results in full-text publications, and leading to an overoptimistic literature base [4,6]. Such bias can also be observed for conference abstracts. The Cochrane systematic review cited previously found that conference abstracts reporting at least one statistically significant result were 30% (95% CI 14%, 47%) more likely to reach full-text publication than those that did not [2]. In our analysis, no associations between the accuracy estimates reported in the ARVO abstracts and full-text publication were observed.

Although investigations of failure to publish and its determinants in diagnostic research are scarce, our findings are in line with what has been found to date. Among 418 diagnostic accuracy studies that were registered between 2006 and 2010 in ClinicalTrials.gov, a full-text publication could be identified for 54% [11]. In an evaluation of 160 conference abstracts describing diagnostic accuracy studies that were presented between 1995 and 2004 at two international stroke meetings, a full-text publication was found for 76%; no association was observed with reported accuracy estimates [12]. In a similar evaluation of 250 abstracts describing diagnostic accuracy studies that were presented in 2009 at three dementia conferences, a full-text publication was identified for only 39%, but potential associations with reported accuracy estimates were not assessed [13].

4.2. Potential limitations

We found no evidence of publication bias in the process of publishing diagnostic accuracy studies in ophthalmology but also examined the possibility that our findings may have been influenced by limitations in the design of our study. It is possible that the selective reporting of studies with favorable results already took place when deciding to submit an abstract to ARVO. If that is the case, bias would only have been detected if publication proportions had been assessed among (a selection of) all initiated diagnostic accuracy studies, not only those presented at ARVO.

We decided to focus our analysis on the highest accuracy estimates reported in each abstract, but many abstracts contained multiple accuracy outcomes, reporting performance for multiple tests, for different target conditions or across subgroups. It is possible that a study's highest accuracy estimates are not the ones that stimulate writing, submitting, or publishing a corresponding full study report. Ideally, we would have focused our assessment on each abstract's most important accuracy estimate. Unfortunately, this almost always is ambiguous in diagnostic accuracy studies because "primary" or "main" outcomes are rarely explicitly defined in abstracts or in full texts [11,16,25].

Although our sample size was relatively large, not all abstracts reported accuracy estimates which limited the power to detect significant associations with full-text publication. Because abstract selection was done by only one investigator (D.A.K.), some relevant abstracts may have been excluded by mistake.

Despite our efforts to contact authors of abstracts for which we did not find a full-text publication, 30% of those authors did not respond to our requests to confirm nonpublication. This proportion is much lower than in previous related projects, and only 6 of the 125 (5%) authors who responded provided a full-text publication that we had missed in our literature searches. When extrapolating this to the 54 abstracts for which we did not receive a response, it is estimated that we have missed three full-text publications, which is less than 1% overall.

4.3. Study implications and conclusions

If publication bias is much less of a problem for diagnostic accuracy studies, a reason could be that these studies are fundamentally different from other types of studies. Most diagnostic accuracy studies lack an explicit, predefined hypothesis, and corresponding statistical testing of these hypotheses is a rarity [15,16]. It has been suggested that nonsignificant results are regarded as disappointing or uninteresting and that investigators are less likely to spend time writing articles describing such findings [26], whereas journal editors are less inclined to publish them [27]. Yet if a distinction between statistically significant and nonsignificant results is rarely made, authors have far greater freedom in interpreting the results and to "spin" them in a positive way, a phenomenon that is highly prevalent in diagnostic accuracy studies [16,28]. This may explain the absence of a strong association between high accuracy estimates and full-text publication; even lower accuracy estimates may be regarded as positive, not hampering writing a longer study report or submitting it to a journal.

To allow the identification of ongoing, terminated, unpublished, or selectively published clinical trials, registries such as ClinicalTrials.gov have been initiated [29]. In 2005, the International Committee of Medical Journal Editors (ICMJE) decided that, for future clinical trials submitted to its member journals, only those that had been registered in a trial registry before initiation of the study would be considered for publication [30]. Although the implementation of this policy by ICMJE journals can be improved [31], the existence of policy has led to a dramatic increase in the number of registered trials [32,33].

Diagnostic accuracy studies are not generally considered to be clinical trials: only 7% of the diagnostic ARVO abstracts included in this analysis provided a registration number, which is in line with a recent evaluation, in which we reported that only 15% of 351 diagnostic accuracy studies published in high-impact journals had been registered [34]. To prevent research waste, the scientific community should strongly consider enforcing registration of all diagnostic accuracy studies, or at least those that are prospective [9,10,35,36]. This would allow researchers and funders to avoid unnecessary duplication of research efforts and improve collaborations, whereas systematic reviewers and guideline developers can uncover all potentially eligible unpublished studies or study materials, and journals and peer reviewers can help minimize selective publication by identifying discrepancies between the registered record and the submitted study report [9].

Inaccessible research is widely considered to be one of the largest sources of research waste [8]. Evidence is now accumulating that many diagnostic accuracy studies never reach full-text publication [11,12]. Although we found no evidence of publication bias in the process of publishing these studies, this failure to publish them cannot be justified for ethical, economic, and scientific reasons [9,35].Changing this will need concerted action from all stakeholders, but it is an absolute must [8,37].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Schmucker C, Schell LK, Portalupi S, Oeller P, Cabrera L, Bassler D, et al. Extent of non-publication in cohorts of studies approved by research ethics committees or included in trial registries. PLoS One. 2014; 9:e114023. [PubMed: 25536072]

2. Scherer RW, Langenberg P, von EE. Full publication of results initially presented in abstracts. Cochrane Database Syst Rev. 2007 MR000005.

3. Hua F, Walsh T, Glenny AM, Worthington H. Thirty percent of abstracts presented at dental conferences are published in full: a systematic review. J Clin Epidemiol. 2016; 75:16–28. [PubMed: 26854259]

4. Dwan K, Gamble C, Williamson PR, Kirkham JJ. Systematic review of the empirical evidence of study publication bias and outcome reporting bias - an updated review. PLoS One. 2013; 8:e66844. [PubMed: 23861749]

5. Dickersin K. The existence of publication bias and risk factors for its occurrence. JAMA. 1990; 263:1385–1389. [PubMed: 2406472]

6. Song F, Parekh-Bhurke S, Hooper L, Loke YK, Ryder JJ, Sutton AJ, et al. Extent of publication bias in different categories of research cohorts: a meta-analysis of empirical studies. BMC Med Res Methodol. 2009; 9:79. [PubMed: 19941636]

7. Sterne, JAC.; Egger, M.; Moher, D. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0. The Cochrane Collaboration; 2011. Chapter 10: addressing reporting biases.

8. Chan AW, Song F, Vickers A, Jefferson T, Dickersin K, Gotzsche PC, et al. Increasing value and reducing waste: addressing inaccessible research. Lancet. 2014; 383:257–266. [PubMed: 24411650]

9. Hooft L, Bossuyt PM. Prospective registration of marker evaluation studies: time to act. Clin Chem. 2011; 57:1684–1686. [PubMed: 21984537]

10. Rifai N, Altman DG, Bossuyt PM. Reporting bias in diagnostic and prognostic studies: time for action. Clin Chem. 2008; 54:1101–1103. [PubMed: 18593957]

11. Korevaar DA, Ochodo EA, Bossuyt PM, Hooft L. Publication and reporting of test accuracy studies registered in ClinicalTrials.gov. Clin Chem. 2014; 60:651–659. [PubMed: 24381228]

12. Brazzelli M, Lewis SC, Deeks JJ, Sandercock PA. No evidence of bias in the process of publication of diagnostic accuracy studies in stroke submitted as abstracts. J Clin Epidemiol. 2009; 62:425–430. [PubMed: 19013759]

13. Wilson C, Kerr D, Noel-Storr A, Quinn TJ. Associations with publication and assessing publication bias in dementia diagnostic test accuracy studies. Int J Geriatr Psychiatry. 2015; 30:1250–1256. [PubMed: 25779466]

14. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. J Clin Epidemiol. 2005; 58:882–893. [PubMed: 16085191]

15. Bachmann LM, Puhan MA, ter Riet G, Bossuyt PM. Sample sizes of studies on diagnostic accuracy: literature survey. BMJ. 2006; 332:1127–1129. [PubMed: 16627488]

16. Ochodo EA, de Haan MC, Reitsma JB, Hooft L, Bossuyt PM, Leeflang MM. Overinterpretation and misreporting of diagnostic accuracy studies: evidence of "spin". Radiology. 2013; 267:581–588. [PubMed: 23360738]

17. van Enst WA, Ochodo E, Scholten RJ, Hooft L, Leeflang MM. Investigation of publication bias in meta-analyses of diagnostic test accuracy: a meta-epidemiological study. BMC Med Res Methodol. 2014; 14:70. [PubMed: 24884381]

18. Linnet K, Bossuyt PM, Moons KG, Reitsma JB. Quantifying the accuracy of a diagnostic test or marker. Clin Chem. 2012; 58:1292–1301. [PubMed: 22829313]

19. Glas AS, Lijmer JG, Prins MH, Bonsel GJ, Bossuyt PM. The diagnostic odds ratio: a single indicator of test performance. J Clin Epidemiol. 2003; 56:1129–1135. [PubMed: 14615004]

20. Walter SD, Sinuff T. Studies reporting ROC curves of diagnostic and prediction data can be incorporated into meta-analyses using corresponding odds ratios. J Clin Epidemiol. 2007; 60:530–534. [PubMed: 17419965]

21. Kho ME, Brouwers MC. Conference abstracts of a new oncology drug do not always lead to full publication: proceed with caution. J Clin Epidemiol. 2009; 62:752–758. [PubMed: 19124219]

22. Juzych MS, Shin DH, Coffey J, Juzych L, Shin D. Whatever happened to abstracts from different sections of the association for research in vision and ophthalmology? Invest Ophthalmol Vis Sci. 1993; 34:1879–1882. [PubMed: 8473126]

23. Scherer RW, Dickersin K, Langenberg P. Full publication of results initially presented in abstracts. A meta-analysis. JAMA. 1994; 272:158–162. [PubMed: 8015133]

24. Saldanha IJ, Scherer RW, Rodriguez-Barraquer I, Jampel HD, Dickersin K. Dependability of results in conference abstracts of randomized controlled trials in ophthalmology and author financial conflicts of interest as a factor associated with full publication. Trials. 2016; 17(1):213. [PubMed: 27113767]

25. Korevaar DA, Cohen JF, Hooft L, Bossuyt PM. Literature survey of high-impact journals revealed reporting weaknesses in abstracts of diagnostic accuracy studies. J Clin Epidemiol. 2015; 68:708–715. [PubMed: 25703213]

26. Scherer RW, Ugarte-Gil C, Schmucker C, Meerpohl JJ. Authors report lack of time as main reason for unpublished research presented at biomedical conferences: a systematic review. J Clin Epidemiol. 2015; 68:803–810. [PubMed: 25797837]

27. Wager E, Williams P. "Hardly worth the effort"? Medical journals' policies and their editors' and publishers' views on trial registration and publication bias: quantitative and qualitative study. BMJ. 2013; 347:f5248. [PubMed: 24014339]

28. Lumbreras B, Parker LA, Porta M, Pollan M, Ioannidis JP, Hernandez-Aguado I. Overinterpretation of clinical applicability in molecular diagnostic research. Clin Chem. 2009; 55:786–794. [PubMed: 19233907]

29. Dickersin K, Rennie D. The evolution of trial registries and their use to assess the clinical trial enterprise. JAMA. 2012; 307:1861–1864. [PubMed: 22550202]

30. De Angelis CD, Drazen JM, Frizelle FA, Haug C, Hoey J, Horton R, et al. Clinical trial registration: a statement from the International Committee of Medical Journal Editors. JAMA. 2004; 292:1363–1364. [PubMed: 15355936]

31. Hooft L, Korevaar DA, Molenaar N, Bossuyt PM, Scholten RJ. Endorsement of ICMJE's Clinical Trial Registration Policy: a survey among journal editors. Neth J Med. 2014; 72(7):349–355. [PubMed: 25178768]

32. Zarin DA, Tse T, Ide NC. Trial Registration at ClinicalTrials.gov between May and October 2005. N Engl J Med. 2005; 353:2779–2787. [PubMed: 16382064]

33. Viergever RF, Li K. Trends in global clinical trial registration: an analysis of numbers of registered clinical trials in different parts of the world from 2004 to 2013. BMJ Open. 2015; 5(9):e008932.

34. Korevaar DA, Bossuyt PM, Hooft L. Infrequent and incomplete registration of test accuracy studies: analysis of recent study reports. BMJ Open. 2014; 4(1):e004596.

35. Altman DG. The time has come to register diagnostic and prognostic research. Clin Chem. 2014; 60:580–582. [PubMed: 24520099]

36. Rifai N, Bossuyt PM, Ioannidis JP, Bray KR, McShane LM, Golub RM, et al. Registering diagnostic and prognostic trials of tests: is it the right thing to do? Clin Chem. 2014; 60:1146–1152. [PubMed: 24855278]

37. Moher D, Glasziou P, Chalmers I, Nasser M, Bossuyt PM, Korevaar DA, et al. Increasing value and reducing waste in biomedical research: who's listening? Lancet. 2016; 387:1573–1586. [PubMed: 26423180]

What is new?

Key findings

- Almost half of the conference abstracts describing diagnostic accuracy studies did not reach full-text publication.

- Studies in abstracts that mentioned higher accuracy estimates were not more likely to be reported in a full-text publication.

What this adds to what was known?

- Previous evaluations have shown that failure to reach full-text publication is highly prevalent among studies of therapeutic interventions, and the results of our evaluation indicate that this also applies to diagnostic accuracy studies.

- Studies of therapeutic interventions with statistically significant results are more likely to reach full-text publication, but we found no evidence of a similar selective publication pattern among diagnostic accuracy studies.

What is the implication and what should change now?

- To prevent waste from unpublished diagnostic accuracy studies, the scientific community should strongly consider enforcing prospective registration of these studies in publically accessible registries.
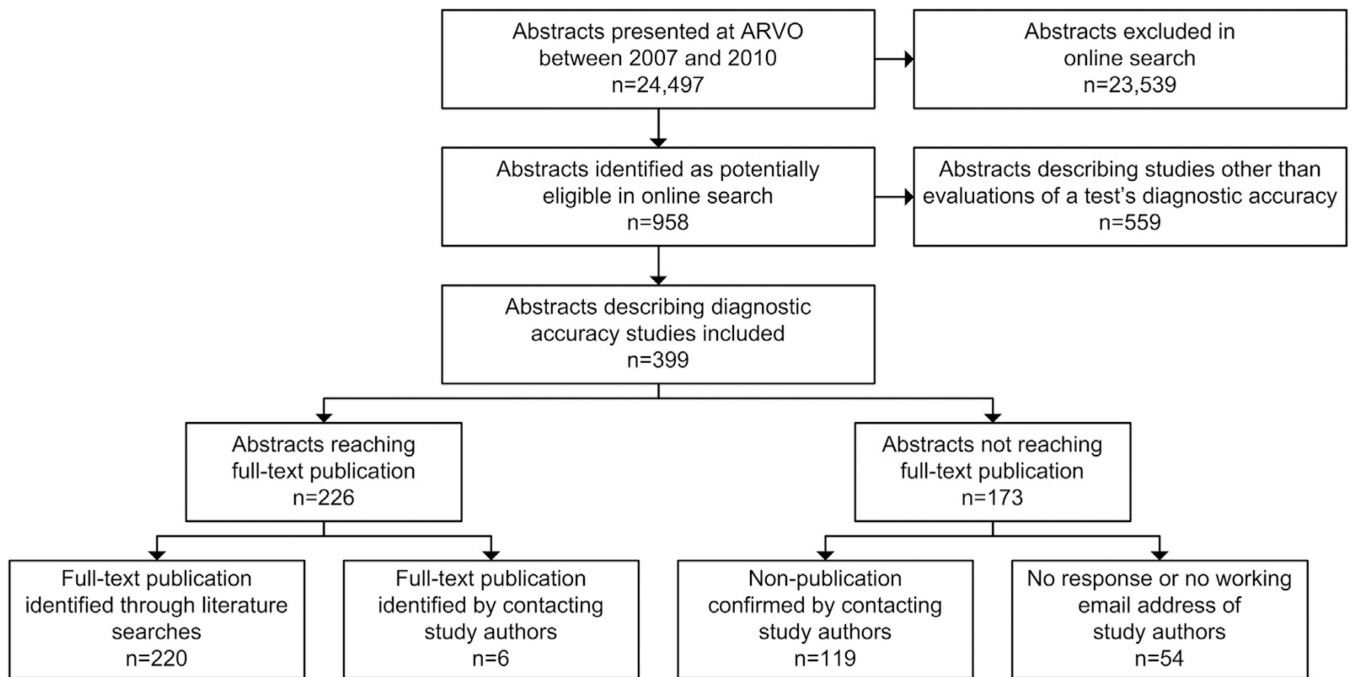
Fig. 1.
Selection of ARVO abstracts and identification of full-text publications. ARVO, Association for Research in Vision and Ophthalmology.

Fig. 2.
Time from presentation to full-text publication among ARVO abstracts ($n = 399$) describing diagnostic accuracy studies. Full-text publications that were published before or at the date of presentation were considered published 1 month after presentation. ARVO, Association for Research in Vision and Ophthalmology.

Table 1

Characteristics of ARVO abstracts describing diagnostic accuracy studies and association with full-text publication
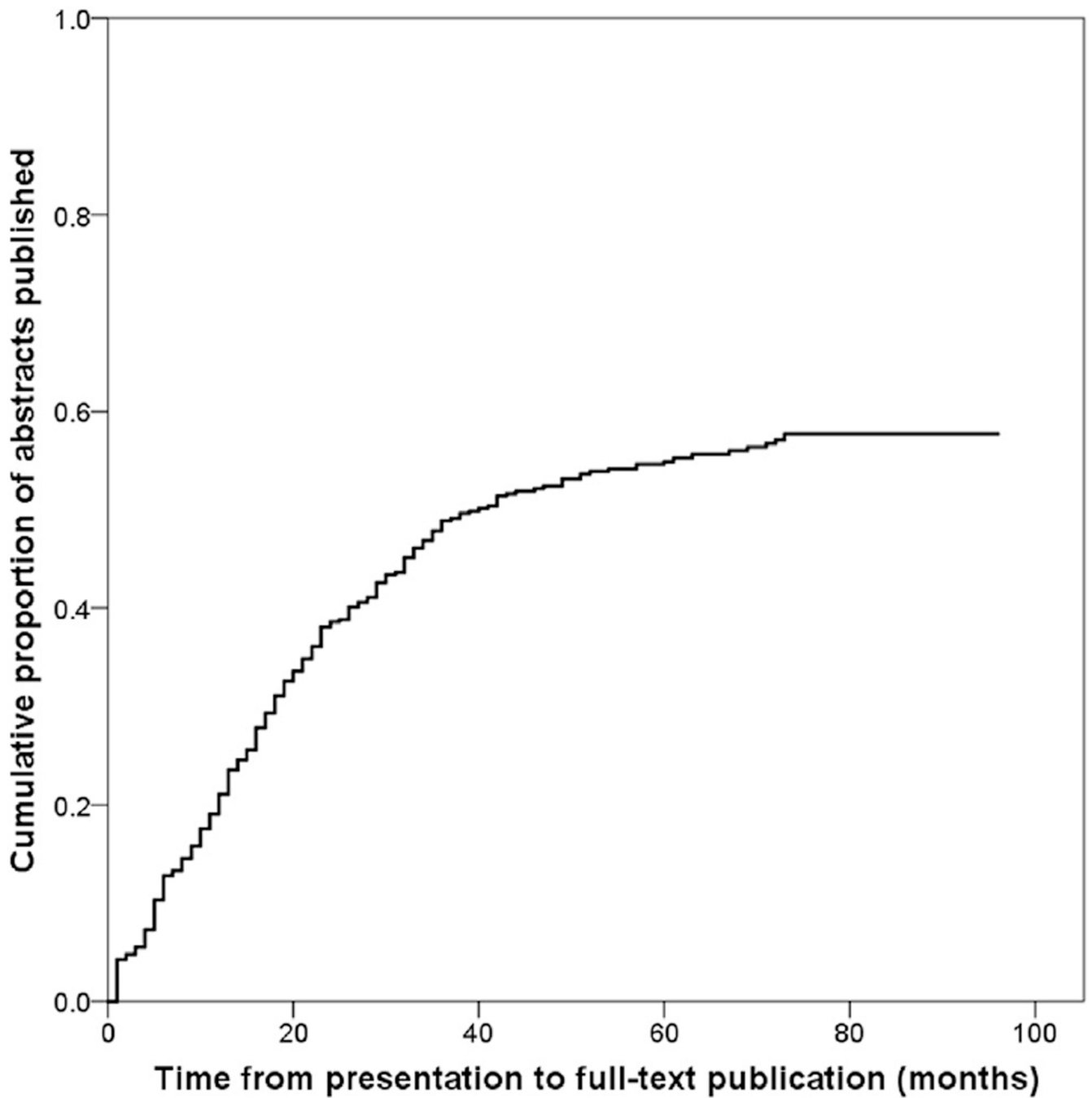
Abstract characteristic	All abstracts, *n* (%)	Abstracts that reached full-text publication, *n*	Abstracts that reached full-text publication,[a] %
Overall	399 (100)	226	57
Year of presentation at ARVO			
2007	75 (19)	41	55
2008	102 (26)	65	64
2009	96 (24)	52	54
2010	126 (32)	68	54
Number of authors			
<5	144 (36)	91	63
5	255 (64)	135	53
Continent of first author			
Asia	64 (16)	44	69
Europe	130 (33)	66	51
North America	165 (41)	86	52
Oceania	18 (5)	15	83
South America	22 (6)	15	68
International affiliations			
Authors from multiple countries	75 (19)	45	60
All authors from one country	324 (81)	181	56
Conflicts of interest			
At least one author	133 (33)	74	56
None of the authors	266 (67)	152	57
Funding for study			
Industry	37 (9)	20	54
Nonindustry only	194 (49)	117	60
None	168 (42)	89	53
Trial registration number provided			
Yes	26 (7)	16	62
No	373 (94)	210	56
Study design			
Cohort	139 (35)	83	60
Case-control	219 (55)	124	57
Unclear	41 (10)	19	46
Data collection			
Prospective	54 (14)	35	65
Retrospective	37 (9)	20	54
Unclear	308 (77)	171	56
Research field			

Abstract characteristic	All abstracts, *n* (%)	Abstracts that reached full-text publication, *n*	Abstracts that reached full-text publication,[a] %
Glaucoma	186 (47)	99	53
Ocular surface and corneal diseases	35 (9)	17	49
Common chorioretinal diseases	44 (11)	24	55
Other	134 (34)	86	64
Number of participants			
<100	150 (38)	88	59
100–1,000	146 (37)	81	56
1,000	31 (8)	20	65
Not reported	72 (18)	37	51
Number of eyes			
<100	71 (18)	35	49
100–1,000	118 (30)	70	59
1,000	18 (5)	12	67
Not reported	192 (48)	109	57

Abbreviation: ARVO, Association for Research in Vision and Ophthalmology.

[a]None of these abstract characteristics were significantly associated with full-text publication, after applying a Bonferroni correction to adjust for multiple testing.

Table 2

Accuracy estimates reported in ARVO abstracts describing diagnostic accuracy studies and association with full-text publication

Accuracy measure	Total number of abstracts, *n* (%)	Abstracts that reached full-text publication, *n*	Abstracts that reached full-text publication, %
Overall	399 (100)	226	57
Sensitivity[a]			
<0.78	62 (16)	30	48
0.78–0.87	61 (15)	33	54
0.87–0.95	64 (16)	35	55
0.95	65 (16)	41	63
No sensitivity reported	147 (37)	87	59
Specificity[a]			
<0.82	55 (14)	34	62
0.82–0.90	45 (11)	24	53
0.90–0.98	81 (20)	39	48
0.98	61 (15)	37	61
No specificity reported	157 (39)	92	59
AUC[a]			
<0.86	38 (10)	27	71
0.86–0.91	31 (8)	17	55
0.91–0.96	42 (11)	28	67
0.96	43 (11)	25	58
No AUC reported	245 (61)	129	53
DOR[a]			
<16.1	78 (20)	44	56
16.1–48.6	86 (22)	51	59
48.6–168.6	80 (20)	46	58
168.6	84 (21)	46	55
No DOR reported	71 (18)	39	55

Abbreviations: ARVO, Association for Research in Vision and Ophthalmology; AUC, area under the receiver operating characteristic curve; DOR, diagnostic odds ratio.

[a] Accuracy estimates were categorized by quartiles.

Table 3

Accuracy estimates reported in ARVO abstracts ($n = 399$) describing diagnostic accuracy studies and hazard ratios of full-text publication

Accuracy measure	Hazard ratio,[a] (95% CI)	P-value
Sensitivity		
Sensitivity (logit transformed)	1.09 (0.98–1.22)	0.126
No sensitivity reported[b]	1.31 (0.91–1.89)	0.151
Specificity		
Specificity (logit transformed)	1.00 (0.88–1.14)	0.951
No specificity reported[b]	1.07 (0.70–1.62)	0.763
AUC		
AUC (logit transformed)	0.91 (0.75–1.09)	0.291
No AUC reported[b]	0.61 (0.36–1.03)	0.065
DOR		
DOR (log transformed)	1.01 (0.94–1.09)	0.753
No DOR reported[b]	0.99 (0.62–1.58)	0.971

Abbreviations: ARVO, Association for Research in Vision and Ophthalmology; CI, confidence interval; AUC, area under the receiver operating characteristic curve; DOR, diagnostic odds ratio.

[a] Estimated using Cox proportional hazards regression analyses.

[b] Abstracts without accuracy estimates were included in the regression model by adding an indicator of missingness.