

A STUDY ON THE RATE OF ADHERENCE TO STARD STANDARDS IN THE ABSTRACTS OF DIAGNOSTIC STUDIES, PUBLISHED IN INTERNATIONAL CONTINENCE SOCIETY ABSTRACT BOOK 2011.

Hypothesis / aims of study

New diagnostic methods are rapidly being developed keeping up with the pace of the modern world. Since diagnostic tests play an important role in making decisions dealing with patients' health, assessing them before incorporating their results is vital. It's obvious that any deceptive results caused by biased or poorly designed studies will cause great waste of time and money and may also threaten patients' health.(1)Studies done up to now indicate that adherence to standards for reporting of diagnostic accuracy(STARD)statement is highly variable. Since the articles published by ICS have great influence on experts and are followed by them we aimed to assess them using STARD statement.

Study design, materials and methods

Diagnostic accuracy studies (DAS) were identified by searching ICS 2011 journal, which was done by a urologist who was an expert in research methodology. Among 287 abstracts, 30 articles matched with our inclusion criteria (indicating a diagnostic study focusing on human subjects). Animal studies were excluded. Then the selected articles were distributed among 3 independent reviewers who assessed articles using the STARD checklist consisting of 25 items. Each article was reviewed by 2 reviewers masked to each other's results. In this process reviewers had to determine whether each item was reported adequately. Each item could be considered yes (1 score), no or unclear (both 0 score).Disagreements were discussed and resolved and if final agreement was not reached, cases were referred to a third reviewer who had not reviewed the article before and final decision was then made. Scores for each item and also the total score for each article were calculated and compared by using nonparametric tests in SPSS 13.

Results

The mean score of the 30 DAS articles published in ICS 2011 abstract book was 15.03 out of 25. The total score for each article ranged from a minimum of 9 to a maximum of 22(SD: 3.746). The least scored items were blinding and number and training which mainly assess executives and also item 17 which is about time interval between index test and reference standard, mentioned in 13.3%, 16.7% and 16.7% of the abstracts respectively. The highest scored items were items 2(stating research aims), the study population and discussion of clinical applicability of study findings which were mentioned in 100% of articles. 63.3% of articles (n=19) reported more than 50% of STARD items. Table 1 illustrates the scores for each item

Interpretation of results

A diagnostic test study like any clinical research studies should be started with consideration of the research question, study design, study subjects, variables, and outcome(s). Therefore any problems with study design may affect the interpretation or credibility of the results, and a well designed study has to be reported in the best way as well. In the current study, total adherence to STARD standards was less than optimal. Unfortunately most of the articles' titles did not mention MeSH (medical subject heading) headings which are also recommended by STARD statement. We scored item 1 in case we could infer that the study was diagnostic from similar keywords to accuracy or reproducibility, like "determining potential, assessing accuracy, etc". On the other hand to prevent any incorporation bias, the index test must be compared to an independent gold standard, but at least in ten abstracts, it was not counted in study design. Item 11 which assesses the researchers' blinding to the results of the index test and reference standard, is one of the most vital items in the checklist; however it's one of the least included items in the articles and this can lead to an increase in potential bias and it may lead to increasing of sensitivity and reducing of specificity. Recruitment, participants and sampling method were not reported clearly which can cause even a spectrum bias. Finally the authors have to mention what they have found at the end of the study. It is better to include whether results varied in different groups of patients or examiners. For studies of accuracy: sensitivity, specificity, predictive value, LRs, and for reproducibility: kappa, Bland-Altman plots, etc all with confidence intervals are necessary, but most of the studies had not emphasized on this important part (PV=0.00).

Concluding message

The lack of using keywords as mentioned by MeSH makes determining DASs very difficult, so we suggest that authors obey STARD and MeSH rules for writing titles. We do also suggest that ICS label DASs just like RCTs in the end of the published abstracts. Finally we must insist that ICS make inclusion of STARD checklist in the authors' diagnostic accuracy studies obligatory for publication, thus taking an important pace in increasing the quality of these studies.

Disclosures

Funding: Students research committee, Tabriz University of Medical Sciences **Clinical Trial:** No **Subjects:** NONE

Table 1: Results of assessing 30 diagnostic accuracy articles using STARD statement.

Items	yes		no		unclear		PV
	number	percent	number	percent	number	percent	
1.TITLE/ABSTRACT/ KEY WORDS	27	90	3	10	0	0	0.000
2.INTRODUCTION	30	100	0	0	0	0	0.000
METHODS							
3.PARTICIPANTS/POPULATION	30	100	0	0	0	0	0.000
4.PARTICIPANTS/RECRUITMENT	26	86.7	4	13.3	0	0	0.000
5.PARTICIPANTS/SAMPLING	13	43.3	17	56.7	0	0	NS
6.PARTICIPANTS/DATA COLLECTION	23	76.7	1	3.3	6	20	0.000
7.TEST METHODS/REFERENCE STANDARD	20	66.7	10	33.3	0	0	0.001
8. TEST METHODS/TECHNICAL SPECIFICATION...	29	96.7	1	3.3	0	0	0.00
9. TEST METHODS/UNIT DEFINITION...	19	63.3	10	33.3	1	3.3	0.001
10. TEST METHODS/EXECUTIVE NUMBER, TRAINING...	5	16.7	25	83.3	0	0	0.000
11. TEST METHODS/BLINDING	4	13.3	26	86.7	0	0	0.001
12. STATISTICAL METHODS/CALCULATING DIAGNOSTIC ACCURACY...	18	60	11	36.7	1	3.3	NS
13. STATISTICAL METHODS/TEST REPRODUCIBILITY	13	43.3	9	30	8	26.7	NS
RESULTS							
14.PARTICIPANTS/RECRUITMENT BEGINNING AND END DATES	9	30	21	70	0	0	0.000
15. PARTICIPANTS/CLINICAL CHARACTERISTICS...	29	96.7	0	0	1	3.3	0.000
16. PARTICIPANTS/THE NUMBER OF SATISFYING...	14	46.7	14	46.7	2	6.7	NS
17.TEST RESULTS/TIME INTERVAL	5	16.7	23	76.7	2	6.7	0.001
18. TEST RESULTS/DISEASE SEVERITY DISTRIBUTION	23	76.7	4	13.3	3	10	0.000
19. TEST RESULTS/CROSS TAB...	12	40	18	60	0	0	NS
20. TEST RESULTS/ADVERSE EVENTS...	16	53.3	11	36.7	3	10	NS
21.ESTIMATES/DIAGNOSTIC ACCURACY ESTIMATES	19	63.3	11	36.7	0	0	0.001
22. ESTIMATES/HANDLING MISSING RESULTS	11	36.7	13	43.3	6	20	NS
23. ESTIMATES/VARIABILITY ESTIMATES	14	46.7	13	43.3	3	10	NS
24. ESTIMATES/TEST REPRODUCIBILITY	11	36.7	9	30	10	33.3	NS
25.DISCUSSION	30	100	0	0	0	0	0.000