Swedish Science Pioneers
Developing World Journal Series
Journal of Clinical Research & Governance
www.jcrg.sciencepub.se

Research article

A Survey on the Amount of Adherence to STARD and CONSORT Standards in the Abstracts of Diagnostic Accuracy Studies and Randomized Controlled Trials Published in International Continence Society Abstract Book 2011

Kasra Kolahdouzan[a], Neda Madani [a], Neda Parniyanfard [a], Hadi Mostafaie[a], Kimia Madanlou[a], Arian Pourmalek[a], Morteza Ghojazadeh[b], Sakineh Hajebrahimi [c]

a: Students Research Committee, Iranian Center For Evidence Based Medicine, Tabriz University of Medical Sciences
b: Medical Education Research Center, Tabriz University of Medical Sciences
c: Iranian Center for EBM, Tabriz University of Medical Sciences

Correspondence

Professor Sakineh Hajebrahimi
Iranian Center for Evidence Based Medicine, Faculty of Medicine
Tabriz University of Medical Sciences
Tabriz, Iran
Tel: +984113342219
E-mail: hajebrahimis@gmail.com

Keywords:

International continence society
CONSORT
STARD

Abstract

Purpose: To assess reporting quality of Randomized Controlled Trials (RCTs) and Diagnostic Accuracy Studies (DASs) published in International Continence Society (ICS) abstract book 2011.

Methods: Reports of RCTs and DASs published in ICS abstract book 2011 were identified. The proportion of adherence to each item in the Consolidated Standards of Reporting Trials (CONSORT) and Standards for the Reporting of Diagnostic accuracy studies (STARD) checklists was evaluated for each article abstract by two independent reviewers masked to each other's results giving one point to each item of the used checklists only if it was clearly defined in the abstract.

Results: Out of 287 articles, 67 articles matched our inclusion criteria (37 RCTs and 30 DASs). The mean scores were 10.70 and 15.03 out of 17 and 25 respectively. The highest scored items were items 5 and 15 (INTERVENTION and CONCLUSION) in the CONSORT group (both being reported in 97.3% of articles). In the STARD group, items 2, 3, and 25 (INTRODUCTION, PARTICIPANTS/POPULATION and DISCUSSION) were fully reported. The lowest scored items were item 2 (AUTHORS) in the consort group and item 11 (TEST METHODS/BLINDING) in the diagnostic accuracy group reported in 0% and 13.3% of articles respectively.

Conclusion: Our study revealed that the adherence level is moderate in both groups. In a high ranking society like ICS, we do believe that adopting these two tools would improve the reporting quality.

Introduction

Current public attention to medical tests has made breakthroughs in the development of new interventions and also has improved the existing technology. Therefore Transparency of these studies is of great concern, since biased or exaggerated results can lead clinicians into making incorrect treatment decisions causing great waste of time and money and also threatening patients' health; hence before incorporating evidence into practice, the validity of the study must be assessed.

A randomized controlled trial (RCT) is a specific type of scientific experiment, and the preferred design for a clinical trial. RCTs are often used to test the efficacy of various types of intervention within a patient population. RCTs may also provide an opportunity to gather useful information about adverse effects, such as drug reactions.

"Diagnostic accuracy" refers to the potentiality of a test to identify a condition of interest. In studies of diagnostic accuracy, the results of one or more tests are compared to a reference ("gold") standard in a group of patients suspected of having the condition of interest. The term "accuracy" in this context thus refers to the amount of agreement between the studied test(s) and the reference Standard.[1]

The importance of RCTs and DASs as the key ways of assessing new methods' efficacy and accuracy is well recognized especially after the movement of Evidence Based Medicine. Since Poor conduct in the quality of reporting of these studies may bring about treatment methods and instruments that are less effective than the studies themselves,[2] so the proper reporting of them is vital. Previous researches have proved some shortcomings and also overestimation in some fields [3-8]. To guard against all of these, the researchers have developed the Consolidated Standards of Reporting Trials (CONSORT) and the Standards for the Reporting of Diagnostic Accuracy studies (STARD) containing some essential items to help to improve the quality of RCT and DAS reports, and ensure the reliability or the relevance of the findings. Over years both STARD and CONSORT have gained wide acceptance by journals[9]. Since articles presented in International Continence Society (ICS) conferences are gathered or presented by most of the urologists around the world, we aimed to assess their abstracts using CONSORT and STARD statements.

Table1: Analysis of Adherence of RCTs to CONSORT Checklist

Items	yes		no		unclear		PV
	number	percent	number	percent	number	percent	
Title	21	56.8	16	43.2	0	0	NS
Authors	0	0	37	100	0	0	NS
Trial Design	13	35.1	24	64.9	0	0	NS
Methods							
Participant	21	56.8	10	27.0	6	16.2	0.008
Intervention	36	97.3	1	2.7	0	0	0.00
Objective	32	86.5	4	10.8	1	2.7	0.00
Outcome	22	59.5	14	37.8	1	2.7	0.00
Randomization	27	73.0	10	27.0	0	0	0.005
Blinding	11	29.7	26	70.3	0	0	0.014
Results							
Numbers	21	56.8	16	43.2	0	0	NS
Recruitment	24	64.9	12	32.4	1	2.7	0.00
Numbers	27	73.0	9	24.3	1	2.7	0.00
Outcome	11	29.8	24	64.9	2	5.4	0.00
Harms	18	48.6	13	35.1	6	16.2	0.05
Conclusion	36	97.3	1	2.7	0	0	0.00
Registration	28	75.7	9	24.3	0	0	0.002
Funding	30	81.1	7	18.9	0	0	0.00

Methods

Evaluation tools

The CONSORT Statement is intended to improve the reporting of a randomized controlled trial (RCT), allowing readers to understand a trial's design, conduct, analysis and interpretation, and to assess the validity of their results. It emphasizes that this can only be achieved through complete transparency from the authors' side. Investigators and editors developed and revised the CONSORT (Consolidated Standards of Reporting Trials) Statement to help authors improve reporting of two-parallel design RCTs by using a checklist and flow diagram. The most up-to-date revision of the CONSORT Statement is CONSORT 2010 that includes 25 items.

CONSORT for abstracts: This extension to the CONSORT Statement provides a minimum list of essential items, that authors should consider when reporting the main results of a randomized trial in any journal or conference abstract. CONSORT for Abstracts aims to improve reporting of abstracts of RCTs published in journal articles and conference proceedings. It will help authors of abstracts of these trials to provide the details and clarity needed by readers wishing to assess a trial's validity and the applicability of its results.

STARD: The objective of the STARD initiative is to improve the accuracy and completeness of reporting of the studies of diagnostic accuracy, to allow readers to assess the potential for bias in the study (internal validity) and to evaluate its generalisability (external validity). The STARD statement consists of a checklist of 25 items and recommends the use of a flow diagram which describes the design of the study and the flow of patients. We have to note here that as the abstracts published in ICS abstract book are extended abstracts, unavailability of an extension of STARD checklist for abstracts in the same way CONSORT checklist's extension was of little importance. Beacuse extended abstracts report about the same important data as in a full article.

Prior to the study the researchers participated in several workshops including: RCT and diagnostic studies workshops and the international workshop of critical appraisal of therapeutic and diagnostic studies. They also took part in Coordination meetings where they were instructed by experts about how to use STARD and CONSORT checklists for critical appraisal of related medical articles. Two groups of researchers conducted this study; One for reviewing RCT abstracts and the other, for reviewing the diagnostic accuracy articles' abstracts.

From January to March 2012, of 287abstracts published in the abstract book of ICS conference of 2011, all of the RCTs (n=37) and diagnostic accuracy studies (n=30) were selected by a urologist expert in research methodology. Each RCT and DAS was inspected by two independent reviewers using the CONSORT for abstracts and STARD checklists. The related abstracts were then assessed by the reviewers separately. If any of the items in the abstracts live up to the criteria defined in the checklists, the reviewers scored 1 point for it, whereas if they were unclear or fail to have the criteria defined by the checklist items, the reviewers gave 0 point to them. Cases of disagreement were discussed in discussion meetings. Further discordes were referred to a third reviewer in each group. Then the overall reporting quality and the differences between high and low quality studies were explored.

Analyzing method

Finally the collected data were descriptively analyzed by SPSS ver.13.

Results

Reporting quality of RCTs

Of 37 RCTs evaluated by CONSORT for abstracts statement, the mean score of the articles was 10.70 out of 17 ranging from a minimum of 6 to a maximum of 15 (Std. Deviation= 2.22) with median of 11. The most scored items were Interventions and Conclusions, which were mentioned in 36 (97.3%) of abstracts. The items Objective and Funding were properly reported in more than 80% of articles. About 81.08% of the abstracts met positively with more than 50% of the items in CONSORT statement. About 65% of the abstracts scored between10 to 13 (59% to 76% of the items of the checklist) with a mode of 11. The least scoring items were items 2 and 9 (Authors and Blinding) which scored 0 and 29.7% respectively. The reporting quality of the items of the CONSORT for abstracts statement is represented in Table 1 in more details.

21 (56.8%) articles were identified as randomized study in title (item1). Contact details of corresponding author which should be included in conference abstracts were not mentioned in the selected abstracts (item2). Trial Design (e.g. Parallel, cluster non-inferiority) was mentioned in 35.1 % of articles. The best scoring items in Methods section (items 4 to 9) were Intervention 97.3% and Objective 86.5%. Allocation of participants to interventions (Randomization) was reported in 73.0% of articles. Eligibility criteria for participants and data collection settings (item4), Outcomes (Clearly defined primary outcome), and Blinding were reported 56.8, 59.5%, and 29.7% respectively.

The items of Results section scored as follows: Numbers Randomized 56.8%, Recruitment (Trial Status) 64.9%, Numbers Analyzed 73.0%, Outcome (results for each group and the estimated effect size and its precision) 29.8% and Harms 48.6%. Risk difference, Relative risk (RR) and NNT were not reported clearly in the result section of most of the articles and estimation of the results was incomplete (PV=0.00). Only few studies reported the application of intention to treat analysis (PV<0.01). Conclusion (General interpretation of the results) was mentioned in 36 (97.3%) articles. Additional details such as Registration and Funding were included in 75.7% and 81.1% of articles.

Reporting quality of DASs

The mean score Of 30 DASs scored by STARD Statement was 15.03 out of 25 ranging from 9 to 22 (Std. Deviation= 3.746). The median was 14.50. The best scoring items which were reported in all of the abstracts (100%) were items 2, 3, and 25 (Introduction\Aim of research, Study Population and Discussion respectively). Items 1, 8, and 15 (Title/Abstract/Keywords, Technical Specification, and Clinical Characteristics respectively) were included in over 90% of articles. The least scoring items were Blinding which was reported in 13.3% of abstracts and Executive Number and Time Interval which were both included in 5 (16.7%) articles. The results are illustrated in Table 2.

Table 2: Analysis of adherence of Diagnostic Accuracy Studies to STARD Checklist

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 Methods	30	100	0	0	0	0	0.000
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 Results	13	43.3	9	30	8	26.7	NS
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

Title/Abstract/Keywords (item 1) which is about identifying the article as a study of diagnostic accuracy scored in 27 (90%) articles. We scored this item only if we could find some similar keywords to accuracy or reproducibility, like "determining potential, assessing accuracy, etc.", however MeSH (Medical Subject Heading) heading was ignored in most of the articles' titles in spite of STARD Statement recommendation. In 100% of articles the aim of research became obvious after reading introduction so the second item (*Introduction*) was fully marked.

The results of the Methods section (items 3 to 13) were as follows: The Participants section was almost well reported except the item Participants Sampling which was reported in 13 (43.3%) articles. The overall percentage of Participants reporting was 76.6%. The best scoring item in Test Method section was item 8 (Technicalspecifications of material and methods involved) which was mentioned in 29 (96.7%) articles. The items "Calculating Diagnostic Accuracy and statistics methods (e.g. 95% CI)" (item12) and "Calculating Test Reproducibility" (item13) scored 60% and 43.3% respectively.

The items in Results section scored as follows: Participants 54.4 %, Test Results 46.675% and Estimates 45.85 %. In Discussion section all of the articles discussed the clinical applicability of the study findings.

Discussion

In the CONSORT group the overall scoring of the abstracts shows that the chosen articles for publication match moderately with CONSORT for abstracts checklist items. However comparing the results of this study with those of others assessing reporting quality of RCTs is a proof of the fact that the general reporting quality is higher in ICS conference abstracts than other published abstracts. One of the main issues is adverse effects or harms that can change the significance of the results clinically. This item was reported appropriately in about half of the abstracts which although not as high as expected of ICS, was anyway higher than other similar reported studies such as the one conducted by Autorino R et al [19] to assess the quality of reporting of randomized controlled trials (RCTs) presented in abstract form at the annual World Congress of Endourology (WCE) which mentioned that adverse events were reported in 38% of cases. The poor reporting of the Authors' part might also be due to ICS abstract rolls. A lack of accuracy in matching with the criteria of CONSORT for abstracts checklist in some other important items such as Blinding, Participants, Trial Design, Harms, and method of the analysis which are essential in extending even more the reporting quality of the gold standard RCTs was observed to some extent. The result has to be presented clearly, objectively, and in sufficient details to enable the readers to draw their own conclusions, which in some of the abstracts was not emphasized. It was not clear how "P values" were calculated and whether or not they were interpreted appropriately. It wasn't clear what level of difference between the groups, outcomes, or interventions constituted a statistically and clinically significant effect. Also confidence intervals were not calculated, and the authors' conclusions did not reflect them. Few studies similar to our paper in the field of urology have been done including a survey by Cavadas v et al [20] of full report RCTs relating to Pelvic Organ Prolapse which mentions that the implementation of randomization, recruitment, blinding, outcomes with effect size and precision, trial registration, and full protocol availability were reported in less than half of the trials and that they found the reporting quality suboptimal.

In the STARD group total adherence to STARD standards was moderate. Except for some items that we would be discussing in the following lines, other items were reported in an acceptable fashion. Unfortunately MeSH headings were ignored in most of the articles. 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 was one of the least included items in the articles and this can lead to an increase in potential bias and it may lead to an increase in sensitivity and a reduction of specificity. Recruitment, Participants and Sampling method were not reported clearly which can cause even a spectrum bias. 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 some of the studies had not emphasized on this important part (PV=0.00). Studies similar to ours in the field of urology have rarely been conducted but we can refer to the one by Selman TJ et al [21] which had aimed to determine the quality of reporting in diagnostic test accuracy studies in obstetrics and gynecology using the STARD checklist and claims that the overall reporting quality of included studies to the STARD criteria was poor but the reporting quality of papers in obstetrics and gynecology is improving.

Limitations: We have to refer to one of our limitations here that might have affected the study results for STARD group. As there is no extension of STARD checklist for conference abstracts like CONSORT for abstracts which has summarized the most important items of the main checklist because of the word count and other limitations in abstracts, and although ICS abstracts are in extended form including more details than other abstracts in general, the lack of sufficient reporting of some of the checklist items might have been due to this fact.

Acknowledgements: This study was conducted by the students research committee, Iranian center for EBM, Tabriz university of medical sciences, Islamic Republic of Iran.

Conflict of interests: The authors declare no conflict of interest.

Abbreviations:

ICS: International Continence Society
DAS: Diagnostic Accuracy Study
RCT: Randomized Controlled Trial
STARD: Standards for Reporting of Diagnostic Accuracy
CONSORT: Consolidated Standards for Reporting of Trials

References

1. Bruns DE: The STARD initiative and the reporting of studies of diagnostic accuracy. *Clin Chem* 2003, 49(1):19-20.
2. Borg Debono V, Zhang S, Ye C, Paul J, Arya A, Hurlburt L, Murthy Y, Thabane L: The quality of reporting of RCTs used within a postoperative pain management meta-analysis, using the CONSORT statement. *BMC Anesthesiol* 2012,4:12:13.
3. Al-Namankany AA, Ashley P, Moles DR, Parekh S: Assessment of the quality of reporting of randomized clinical trials in paediatric dentistry journals. *Int J Paediatr Dent 2009*,19(5):318-24.
4. Sargeant JM, O'Connor AM, Gardner IA, Dickson JS, Torrence ME, Dohoo IR, Lefebvre SL, Morley PS, Ramirez A, Snedeker K: The REFLECT statement: reporting guidelines for randomized controlled trials in livestock and food safety: explanation and elaboration. *J Food Prot* 2010,73(3):579-603.
5. Rios LP, Odueyungbo A, Moitri MO, Rahman MO, Thabane L: Quality of reporting of randomized controlled trials in general endocrinology literature. *J Clin Endocrinol Metab* 2008, 93(10):3810-6.
6. Zhang D, Yin P, Freemantle N, Jordan R, Zhong N, Cheng KK: An assessment of the quality of randomised controlled trials conducted in China. *Trials* 2008, 9:22.
7. Can OS, Yilmaz AA, Hasdogan M, Alkaya F, Turhan SC, Can MF, Alanoglu Z. Has the quality of abstracts for randomised controlled trials improved since the release of Consolidated Standards of Reporting Trial guideline for abstract reporting? A survey of four high-profile anaesthesia journals. *Eur J Anaesthesiol* 2011, 28(7):485-92.
8. Cobos-Carbó A; CONSORT group. [Randomized clinical trials (CONSORT)]. *Med Clin (Barc)* 2005, 125 Suppl 1:21-7.
9. Sargeant JM, Elgie R, Valcour J, Saint-Onge J, Thompson A, Marcynuk P, Snedeker K: Methodological quality and completeness of reporting in clinical trials conducted in livestock species. *Prev Vet Med* 2009, 91(2-4):107-15.
10. Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG, Schulz KF: CONSORT Group. CONSORT for reporting randomized controlled trials in journal and conference abstracts: explanation and elaboration. *PLoS Med* 2008, 22;5(1):e20.
11. Altman DG, Moher D: Developing guidelines for reporting healthcare research: scientific rationale and procedures. *Med Clin (Barc)* 2005, 125 Suppl 1:8-13.
12. Zintzaras E, Papathanasiou A, Ziogas D, Voulgarelis M: The reporting quality of studies investigating the diagnostic accuracy of anti-CCP antibody in rheumatoid arthritis and its impact on diagnostic estimates. *BMC Musculoskelet Disord* 2012, 13(1):113.
13. Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, Pitkin R, Rennie D, Schulz KF, Simel D, Stroup DF: Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA* 1996, 276:637–639.
14. Moher D, Schulz KF, Altman DG: CONSORT Group (Consolidated Standards of Reporting Trials). The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *Ann Intern Med* 2001, 134:657–662.
15. Mills EJ, Wu P, Gagnier J, Devereaux PJ: The quality of randomized trial reporting in leading medical journals since the revised CONSORT statement. *Contemp Clin Trials* 2005, 26:480–487.
16. Lai R, Chu R, Fraumeni M, Thabane L. Quality of randomized controlled trials reporting in the primary treatment of brain tumors. *J Clin Oncol* 2006, 24:1136–1144.
17. Scales CD, Norris RD, Keitz SA, Peterson BL, Preminger GM, Vieweg J, Dahm P: A critical assessment of the quality of reporting of randomized controlled trials in the urology literature. *J Urol* 200, 177:1090–1094.
18. Plint AC, Moher D, Morrison A, Schulz K, Altman DG, Hill C, Gaboury I: Does the CONSORT checklist improve the quality of reports of randomized controlled trials? A systematic review. *Med J Aust* 2006,185:263–267.
19. Autorino R, Borges C, White MA, Altunrende F, Perdoná S, Haber GP, De Sio M, Khanna R, Stein RJ, Kaouk JH: Randomized clinical trials presented at the World Congress of Endourology: how is the quality of reporting? *J Endourol* 2010, 24(12):2067-73.
20. Cavadas V, Branco F, Carvalho FL, Osório L, Gomes MJ, Silva-Ramos M: The quality of reporting of randomized controlled trials in pelvic organ prolapse. *Int Urogynecol J* 2011, 22(9):1117-25.
21. Selman TJ, Morris RK, Zamora J, Khan KS: The quality of reporting of primary test accuracy studies in obstetrics and gynecology: application of the STARD criteria. *BMC Women's Health* 2011,11:8.