Scientometrics (2020) 122:989–1001
https://doi.org/10.1007/s11192-019-03317-3

Using STROBE checklist to assess the reporting quality of observational studies affiliated with Shiraz University of Medical Sciences, and its correlates: a scientometric study from Iran

Negin Rahmani[1,2] · Alireza Salehi[3] · Hossein Molavi Vardanjani[1] · Maryam Marzban[4] · Arezoo Behbood[5]

Received: 10 May 2019 / Published online: 7 December 2019

Abstract
The reporting quality of Observational Studies (OSs) is an important measure of their overall quality. We aim to assess the reporting quality of OSs of Shiraz University of Medical Sciences (SUMS) in Iran in the years 2012–2015, using Strengthening the Reporting of Observational Studies checklist. Systematic online search was performed. A random sample of SUMS affiliated published articles was selected. Articles were appraised and scored by two reviewers. Variables such as the study design, publication year, journals' impact factor etc. were retrieved and their correlation with the articles' scores was assessed. Out of 4297 published articles during 2012–2015, 1742 (40.5%) were OSs of which we assessed 171 (~10%) studies. Among these, 87 (50.9%), 74 (43.3%) and 10 (5.8%) articles had a cross-sectional, case–control and cohort design, respectively. Overall score of the reporting quality was 79% ± 0.01. It was at 81% ± 0.1, 77% ± 0.01 and 83% ± 0.02 for cross-sectional, case–control and cohort studies, respectively. A significant correlation was observed between the study design and the score for the reporting quality ($P = 0.015$). Reporting of "flow-diagram" (5%), "sources of bias" (28%) and "study size calculation" (30%) were the most missed items. Although the overall reporting quality of OSs was found to be at an acceptable rate, there are points of concern regarding some of the most important items that deserve the attention of authors as well as reviewers and editors.

Keywords Observational studies · STROBE statement · Quality of reporting · Shiraz University of Medical Sciences · Iran

Introduction

Observational Studies (OSs) are the most reliable source of information on the epidemiology, etiology and prognosis of diseases (Hoppe et al. 2009). They lead to systematic reviews that better interpret their results and have their own assessment tools (Tsakiridis

✉ Alireza Salehi
salehialireza45@yahoo.com

Extended author information available on the last page of the article

et al. 2019). However, OSs are prone to a variety of biases, such as the recall, selection, and information bias, as well as confounding, etc. (Hammer et al. 2009; Bero et al. 2018). Adequate reporting helps readers to critically appraise a study and low quality studies might lead to inappropriate decisions leading to catastrophic results. It has been observed that the reporting of OSs is usually not clear and detailed enough (von Elm et al. 2007), especially in parts like the reporting of confounding variables (Glasziou et al. 2014; Pouwels et al. 2016) and potential biases (Vandenbroucke et al. 2007). Accordingly, a group of 10 renowned researchers from 6 different countries developed a checklist called the Strengthening the Reporting of Observational Studies (STROBE) checklist in 2007. The STROBE statement helps authors to "improve the reporting of observational studies" and "facilitates the critical appraisal and the interpretation of studies by reviewers, journal editors and readers" (Vandenbroucke et al. 2007).

In Iran, universities of medical sciences are responsible for training biomedical researchers as well as conducting biomedical research and monitoring healthcare. Therefore, the assessment of studies affiliated to medical universities provides us with an understanding regarding the quality of their work. This can promote the quality of research and healthcare policies in the long term, if proper actions are taken. Shiraz University of Medical Sciences (SUMS) is a high-ranking university in Iran and to date, no study has evaluated the reporting quality of OSs in this university. Therefore, the results of this study have strong implications for policy makers in this university. In the present study, we aimed to assess the reporting quality of OSs published by the researchers affiliated to SUMS using STROBE checklist as our assessment tool.

Methods

In this cross-sectional study, we examined the reporting quality of 171 OSs affiliated to SUMS in 2012–2015 time period. In order to do so, we initially assessed all the published articles in the mentioned timeframe and categorized them based on their study design. Subsequently, we selected 171 OSs from the total number of OSs and evaluated them in terms of their reporting quality. The following steps were taken.

Search strategy

Systematic online search was performed in the international databases including Scopus, Web of Science, Embase, PubMed/MEDLINE, Cochrane Library, Google Scholar, and national databases such as Science Information Database (SID), Iran Medex and Magiran, using the keyword "Shiraz University of Medical Sciences" both in the affiliation section and the body of documents. Online search results were merged with the list of articles collected by SUMS's scientometry committee. The search included all the studies from March 20th of 2012 until March 20th of 2015.

Selection of articles

After excluding the duplicate articles, we found 4297 articles. In the process of deduplication, we included the more recent duplicated publications. Reports without full text were purchased via the university library. As the designs of most studies were not obvious in the title, two reviewers categorized the articles based on their study design

and type of document. The categories included randomized clinical trials (RCTs), non-RCT interventional studies, case–controls, cross-sectionals, cohorts, qualitative studies, reviews, case reports, case series, short communications, and editorials. We used convenient sampling and included a random sample of 171 studies (~ 10%) out of the total number of observational studies (1742) in the specified timeframe. Out of 171 articles, case–control, cross-sectional and cohort studies were selected, using a stratified random sampling method (Fig. 1).

To ensure the reliability and validity of quality assessments, two reviewers were trained similarly. Initially, a random sample of 10 studies was assessed simultaneously by the two reviewers to obviate any ambiguities regarding the STROBE items and how to use them for evaluating the articles. Any uncertainty was clarified via discussing them with a third person (methodologist). Then, pilot assessment of 30 randomly selected articles was performed separately by each reviewer and their concordance was measured. The inter-rater reliability was calculated for these random samples with interclass correlation coefficient. After reaching acceptable concordance (90%, 95% CI 67–97, $P < 0.001$), the evaluation of 171 studies was commenced. While appraising the articles, the two reviewers were blinded to the authors' names and the names of the journals in order to avoid information bias.

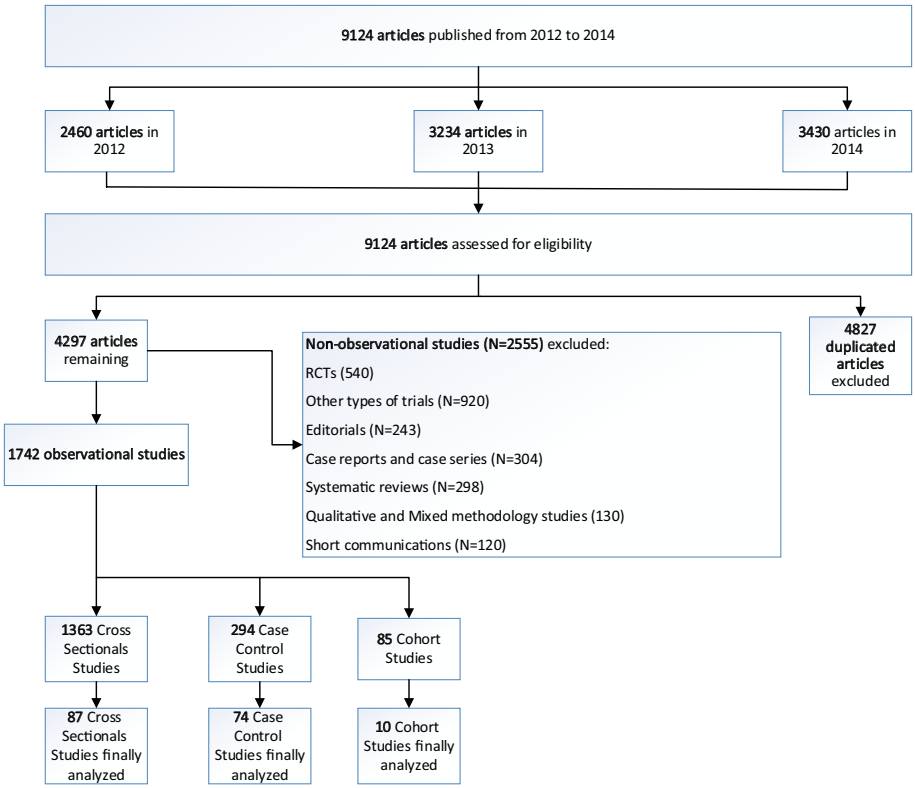

Fig. 1 Flow diagram illustrating the process of selection of 171 observational studies

Instrument

We used STROBE as the assessment tool. The STROBE checklist is comprised of 22 items related to the title and abstract (item 1), introduction (items 2 and 3), methods (items 4–12), results (items 13–17), discussion (items 18–21) and information on funding (item 22). Eighteen items are common to all three categories of OSs while 4 items have subcategories that are related to a specific type of OS (von Elm et al. 2007). Similar to many previous studies and in order to facilitate the scoring, we further subcategorized some items (e.g. item No.5 becomes 5a, 5b, 5c, 5d) and developed 53 smaller items with equal weights (Table 1). If the item was mentioned in the appropriate section of the article, it received one point and if not, zero point was given. Some items might not have been applicable to a specific article; hence, the items were not taken into account. We arbitrarily labeled items with the Average Reporting Percentage (ARP) > 70% as "sufficiently reported", items with ARP 40–70% as "insufficiently reported" and items < 40% as "poorly reported". During the appraisal process, our methodologist was consulted whenever the two reviewers were not certain how to score an item.

Scoring and correlations

The Average Reporting Percentage (ARP) of items—based on STROBE checklist—in each article was calculated. In addition, the ARP of each item was calculated among all the articles. We also calculated the ARP in different sections of articles including "title and abstract", "introduction", "methods", "results", "discussion and conclusion" in different study designs. Subsequently, we measured the correlation between the ARP and the following variables: the article's publication year, the type of OS, the field of study-clinical medicine, dentistry, basic sciences and others-, the number of authors, whether the first author was a faculty member or not, the last impact factor of the journal, whether the article was written in English or Persian, whether it was the result of an international or inter-university collaboration or not, and whether it was the result of a dissertation or not. The data regarding the authors and journals was obtained from the Resource Finder website of the Health Ministry.

Statistical analysis

The collected data were analyzed by SPSS 22 software. Descriptive statistics including frequency and relative frequency were calculated and Mann–Whitney U test, Kruskal–Wallis test and Pearson correlation were used to determine whether a significant correlation existed between the ARPs in different types of OSs and its association with the other studied variables. $P < 0.05$ was considered to be statistically significant.

Results

In this study, 4297 SUMS-affiliated studies published in 2012–2015 timeframe were assessed. As shown in Fig. 2, the most prevalent study design was the cross-sectional (1363, 31.7%), followed by non-RCT interventional studies (920, 21.4%) and RCTs (540,

Table 1 Average Reporting Percentage (ARP) of STROBE items among observational studies

STROBE items	Overall ARP ± SE	CC ARP ± SE	CS ARP ± SE	C ARP ± SE
1a. Abstract: indicate the study's design:2	0.53 ± 0.04	0.47 ± 0.06	0.57 ± 0.05	0.60 ± 0.163
1a. Abstract: balanced summary:1	1.00 ± 0	1.00 ± 0	1.00 ± 0	1.00 ± 0
2. Introduction: background and rationale:1	0.99 ± 0.01	0.99 ± 0.01	1.00 ± 0	1.00 ± 0
3. Introduction: objectives, including any pre specified hypotheses:1	0.99 ± 0.01	1.00 ± 0	0.99 ± 0.01	1.00 ± 0
4. Methods: key elements of study design:1	0.95 ± 0.02	0.99 ± 0.01	0.93 ± 0.03	0.90 ± 0.100
5a. Methods: describe the setting and locations:1	0.90 ± 0.02	0.85 ± 0.04	0.95 ± 0.02	0.80 ± 133
5b. Methods: describe relevant dates including periods of recruitment and data collection:2	0.67 ± 0.04	0.53 ± 0.06	0.74 ± 0.05	1.00 ± 0
5c. Methods: describe relevant dates including periods of exposure:1	0.74 ± 0.04	0.61 ± 0.09	1.00 ± 0	1.00 ± 0
5d. Methods: describe relevant dates including periods of follow-up (if applicable):1	0.93 ± 0.07	0.67 ± 0.33	1.00 ± 0	1.00 ± 0
6a. Methods: give the eligibility criteria:1	0.86 ± 0.03	0.88 ± 0.04	0.86 ± 0.04	0.80 ± 0.13
6b. Methods: sources and methods of selection of participants:1	0.74 ± 0.04	0.83 ± 0.11	0.76 ± 0.05	0.38 ± 0.18
6c. Methods: describe methods of follow-up:1	0.78 ± 0.15	NA	NA	0.78 ± 0.15
6d. Methods: sources and methods of case ascertainment:1	0.74 ± 0.06	0.74 ± 0.06	NA	NA
6e. Methods: sources and methods of control selection:2	0.57 ± 0.06	0.57 ± 0.06	NA	NA
6f. Methods: give the rationale for the choice of cases and controls:3	0.30 ± 0.06	0.30 ± 0.06	NA	NA
6 g. Methods: for matched studies, give matching criteria:1	0.84 ± 0.07	0.84 ± 0.07	NA	NA
6 h. Methods, cohort: give the number of exposed and unexposed:3	0.33 ± 0.33	NA	NA	0.33 ± 0.33
6i. Methods, case control: give the number of controls per case:1	0.96 ± 0.03	0.96 ± 0.03	NA	NA
7a. Methods: clearly define all outcomes:1	0.98 ± 0.01	0.97 ± 0.02	0.99 ± 0.01	1.00 ± 0
7b. Methods: clearly define all exposures and Predictors:1	0.99 ± 0.01	0.97 ± 0.02	1.00 ± 0	1.00 ± 0
7c. Methods: potential confounders and Effect modifiers:1	0.74 ± 0.09	0.69 ± 0.12	1.00 ± 0	0.67 ± 0.33
7d. Methods: give diagnostic criteria, if applicable:1	0.77 ± 0.05	0.67 ± 0.07	0.90 ± 0.06	1.00 ± 0
8a. Methods:Sources of data and details of methods of assessment:1	0.99 ± 0.01	1.00 ± 0	0.99 ± 0.01	1 ± 0
8b. Methods: describe comparability of assessment methods if more than one group:1	1.00 ± 0	1.00 ± 0	NA	NA
9. Methods: address potential sources of bias:3	0.28 ± 0.04	0.29 ± 0.05	0.24 ± 0.05	0.50 ± 0.17

Table 1 (continued)

STROBE items	Overall ARP ± SE	CC ARP ± SE	CS ARP ± SE	C ARP ± SE
10. Methods: explain how the study size was arrived at:3	0.30 ± 0.04	0.16 ± 0.04	0.42 ± 0.06	0.43 ± 0.20
11a. Methods: explain how quantitative variables were handled in the analyses:1	0.96 ± 0.02	0.94 ± 0.03	0.98 ± 0.02	1.00 ± 0
11b. Methods: if applicable, describe which groupings were chosen and why:1	1.00 ± 0	1.00 ± 0	1.00 ± 0	1.00 ± 0
12a. Methods: describe all statistical methods:1	0.91 ± 0.02	0.95 ± 0.03	0.86 ± 0.04	1.00 ± 0
12b. Methods: describe any methods used to examine subgroups and interactions (if applicable):1	0.91 ± 0.03	0.95 ± 0.03	0.85 ± 0.06	1.00 ± 0
12c. Methods: explain how missing data were addressed (if applicable):1	0.82 ± 0.12	0.80 ± 0.20	1.00 ± 0	0.50 ± 0.50
12d. Methods: Loss to follow-up, matching of cases and controls and sampling strategy for cohort, case-control and cross-sectional studies respectively:3	0.23 ± 0.05	0.11 ± 0.06	0.33 ± 0.09	0.33 ± 0.33
12e. Methods: describe any sensitivity analyses (if applicable):1	1.00 ± 0	1.00	1.00	NA
13a. Results: report numbers of individuals at each stage of study:1	0.78 ± 0.03	0.77 ± 0.05	0.75 ± 0.05	1.00 ± 0
13b. Results: give reasons for non-participation at each stage (if applicable):3	0.31 ± 0.12	0.33 ± 0.21	0.33 ± 0.21	0.25 ± 0.25
13c. Results: consider use of a flow diagram:3	0.06 ± 0.02	0.07 ± 0.04	0.02 ± 0.02	0.50 ± 0.50
14a. Results: give characteristics of study participants (e.g. demographic, clinical, social):1	0.87 ± 0.03	0.81 ± 0.05	0.90 ± 0.03	1.00 ± 0
14b. Results: information on exposures and potential confounders:1	0.89 ± 0.02	0.85 ± 0.04	0.90 ± 0.03	1.00 ± 0
14c. Results: indicate number of participants with missing data(if applicable):2	0.57 ± 0.14	0.60 ± 0.24	0.80 ± 0.20	0.25 ± 0.25
14d. Results: Summarize follow-up time:2	0.64 ± 0.15	1.00	NA	0.60 ± 0.16
15. Results: Report numbers of outcome events or summary measures:1	1.00 ± 0	1.00 ± 0	1.00 ± 0	1.00 ± 0
16a. Results: give unadjusted estimates:1	0.99 ± 0.01	1.00 ± 0	0.99 ± 0.01	1.00 ± 0
16b. Results: if applicable, confounder-adjusted estimates:1	0.82 ± 0.04	0.86 ± 0.05	0.80 ± 0.06	0.60 ± 0.24
16c. Results: report category boundaries when continuous variables were categorized:1	1.00 ± 0	1.00 ± 0	1.00 ± 0	1.00 ± 0
16d. Results: if relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	NA	NA	NA	NA
17. Results: report other analyses done:1	1.00 ± 0	1.00 ± 0	1.00 ± 0	1.00 ± 0
18. Discussion: summarize key results:1	0.99 ± 0.01	0.99 ± 0.01	0.99 ± 0.01	1.00 ± 0
19a. Discussion: discuss limitations of the study:2	0.53 ± 0.04	0.55 ± 0.06	0.49 ± 0.05	0.70 ± 0.15
19b. Discussion: discuss both direction and magnitude of any potential bias:2	0.62 ± 0.04	0.57 ± 0.06	0.66 ± 0.05	0.70 ± 0.15

Table 1 (continued)

STROBE items	Overall ARP ± SE	CC ARP ± SE	CS ARP ± SE	C ARP ± SE
20a. Discussion: give a cautious overall interpretation of results:1	0.99 ± 0.01	0.97 ± 0.02	1.00 ± 0	1.00 ± 0
20b. Discussion: results from similar studies, and other relevant evidence:1	0.98 ± 0.01	0.99 ± 0.01	0.97 ± 0.02	1.00 ± 0
21. Discussion: discuss the generalizability of the study results:2	0.43 ± 0.04	0.47 ± 0.06	0.40 ± 0.05	0.40 ± 0.16
22. Discussion: give the source of funding and the role of the funders for the present study:2	0.61 ± 0.04	0.65 ± 0.06	0.06 ± 0.05	0.50 ± 0.17

CC: Case Control; CS: Cross Sectional; C: Cohort; NA: Not Applicable 1: sufficiently reported; 2: insufficiently reported; 3: poorly reported percentage was estimated for applicable items

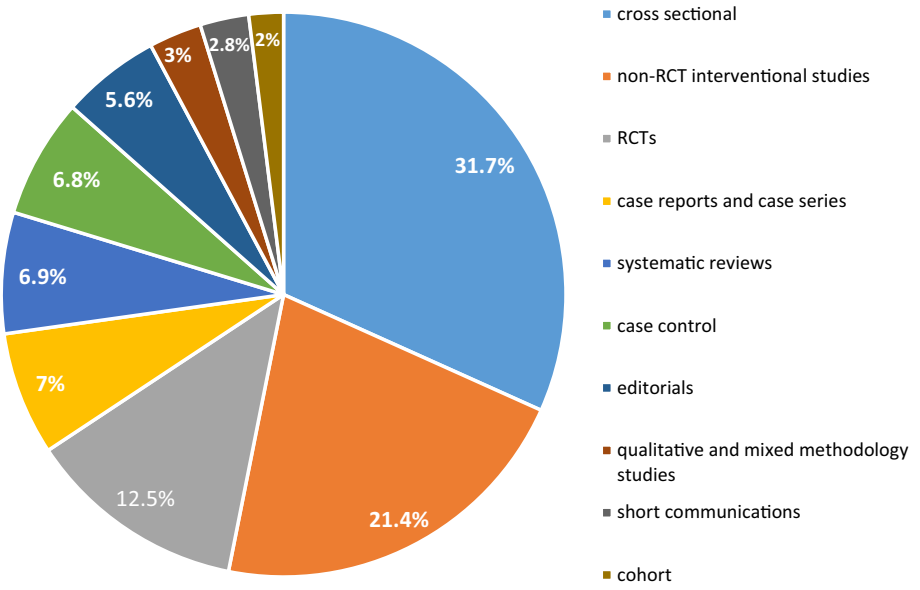

Fig. 2 The distribution of study designs among 4297 articles

12.5%). Among 171 OSs, the number (%) of cross-sectional, case–control and cohort studies was 87 (50.9%), 74 (43.3%) and 10 (5.8%), respectively (Fig. 1). Of all the OSs, 8 (4.6%) were the product of international collaboration. Fifty-two (30.4%) were the result of students ' dissertations and 27 (15.7%) were written in Persian. Regarding the articles' field of study, articles in the field of clinical medicine comprised the majority of articles (133/171, 78%). Dentistry, basic sciences and other categories accounted for 12 (7%), 8 (4.6%) and 18 (10.5%) of all 171 OSs.

Around 69% of the items in our study showed sufficient reporting (ARP > 70%), 17% showed insufficient reporting (ARP = 40–70%) and 13% showed poor reporting (ARP < 40%). The ARP of each STROBE item in OSs and its three subtypes are shown in Table 1. The ARP (± SE) among all OSs was found to be 79% ± 0.01. The ARP in the cross-sectional, case–control and cohort studies was 81% ± 0.01, 77% ± 0.01 and 83% ± 0.02, respectively. Cross-sectional studies exhibited a significantly higher ARP in comparison to case–control studies ($P = 0.015$).

The ARP in the methods section of OSs was 0.78 ± 0.14, which differed significantly ($P = 0.008$) in the three types of OSs. The ARP in the methods section was significantly higher in the cohort and cross-sectional studies compared with case–control studies ($P = 0.039$ and 0.005). As for the publication year, articles published during March 20th of 2013 to March 20th of 2014 were reported better than articles from March 20th of 2012 to March 20th of 2013 ($P = 0.034$). In addition, articles from March 20th of 2014 to March 20th of 2015 were reported better than articles from March 20th of 2012 to March 20th of 2013($P = 0.000$) (Table 2).

The association between the ARP and some variables related to the journal and authors was evaluated as well. Field of study had a significant association with the ARP ($P = 0.040$); articles in the field of clinical medicine were better reported than those related to basic sciences ($P = 0.031$). Furthermore, articles that were the result of an international

Table 2 Average Reporting Percentage (ARP) in different article sections

Article sections	Overall ARP ± SD	ARP ± SD in cross sectional studies	ARP ± SD in case control studies	ARP ± SD in cohort studies
Title and abstract	0.76 ± 0.25	0.78 ± 0.25	0.73 ± 0.25	0.80 ± 0.26
Introduction	0.99 ± 0.05	0.99 ± 0.05	0.99 ± 0.06	1.0 ± 0.0
Methods	0.78 ± 0.14	0.81 ± 0.15	0.74 ± 0.14	0.82 ± 0.07
Results	0.82 ± 0.17	0.84 ± 0.17	0.80 ± 0.17	0.86 ± 0.13
Discussion	0.75 ± 0.16	0.75 ± 0.17	0.75 ± 0.16	0.80 ± 0.17

collaboration (8/171, 4.7%) had significantly higher reporting percentages than those without international collaboration ($P = 0.003$). The following variables did not show a significant correlation with the ARP: The last impact factor of the journal, the language of the article, the number of authors, whether the first author was a faculty member or not, whether the article was the result of an inter-university collaboration, and whether it was the result of a dissertation or not.

Discussion

In the present study, we evaluated the adherence of 171 OSs to STROBE items and also assessed the reporting of items in different article sections. The overall ARP in our study (79%) was close to percentages reported by previous studies (Nagarajan et al. 2018; Fung et al. 2009; Bastuji-Garin et al. 2013). The "introduction" and "results" sections had the highest ARPs among the article sections and the acceptable ARP in our study is a result of the high percentages in these two sections. The lowest ARPs belonged to the "discussion" and "title and abstract" sections, which can be attributed to inadequate reporting in items 1a, 19a and 21, as will be discussed below. We found the reporting quality of the "methods" section to be significantly lower in case–control studies compared with cohort and cross-sectional studies. This is a very important flaw since the methods section is essential in order for readers to "judge whether the methods were adequate to provide reliable and valid answers, and to assess whether any deviations from the original plan were reasonable" (Vandenbroucke et al. 2007). Therefore, attention must be paid to the reporting of the methods section in case–control studies.

Regarding the type of OS, cross-sectional studies were better reported than case–control studies. This difference can be attributed to the difference in the reporting of items in the methods section as mentioned above. Since case–control studies provide a great proportion of evidence in the field of medicine, more attention should be given to reporting these studies, especially in the methods section. Contrary to our results, in the study by Nagarajan et al. in which the reporting quality of OSs in Indian journals was evaluated, the ARP was 50.5%, 49.11% and 44.39% for cross-sectional, cohort and case–control studies, respectively and the differences were not statistically significant (Nagarajan et al. 2018).

Reporting "the study design in the title or abstract" helps readers to "easily identify the design that was used" and also helps with "the indexing of articles in electronic databases" (Vandenbroucke et al. 2007). The ARP of this item in our study was 53% (insufficiently reported), which was close to percentages from some previous studies (Bastuji-Garin et al. 2013; Nagarajan et al. 2018). There are also studies that show this item to be well reported (86% and 98%) (Fung et al. 2009; Langan et al. 2010).

"Potential sources of bias" were reported in 28% of the evaluated articles. Cohort and cross-sectional studies had the highest (50%) and lowest (24%) ARP regarding this item. Inadequate reporting of this item is also seen in previous studies (Fung et al. 2009; Bastuji-Garin et al. 2013; Jeelani et al. 2014; Irani et al. 2018); hence, making it difficult for readers to trust the study's conclusion (Vandenbroucke et al. 2007).

Adequate sample size augments the statistical power of the study, and authors must "indicate the considerations that determined the study size or sample size calculations if they were done" (Vandenbroucke et al. 2007). Generally, this item is underreported in articles ranging from 4.5–34% in various studies (Bastuji-Garin et al. 2013; Nagarajan et al. 2018; Langan et al. 2010). In our study, 30% of studies reported this item and the least ARP was amongst case controls (16%).

Use of a flow diagram was the least reported STROBE item in OSs (6%) as well as in case–control (7%) and cross-sectional studies (2%). This pattern is observed in most studies with reporting percentages ranging from 2 to 21% (Ramke et al. 2017; Fung et al. 2009; Langan et al. 2010; Irani et al. 2018).

The items related to the reporting of confounding in our study were: "potential confounders and effect modifiers (7c)" and "information on exposures and potential confounders (14b)". The ARP of these items was 80.5% in our study. This was higher than percentages reported by some previous studies (Pouwels et al. 2016; Groenwold et al. 2008; Langan et al. 2010). Langan et al. (2010) had reported a mean percentage of 37% for item 7c. Unfortunately, the different items and sub-items adopted by different studies makes it difficult to accurately compare the percentages among them.

"The limitations of the study" and "the generalizability of the study results" were reported in 53% and 43% of studies. Sufficient reporting of these items is essential for the assessment of a study's internal and external validity. The ARP observed in our study was close to percentages in previous studies (Nagarajan et al. 2018; Fung et al. 2009; Langan et al. 2010).

Two case–control-specific items that were poorly reported were: "the rationale for the choice of cases and controls" and "explaining how matching of cases and controls was addressed", which were reported adequately in 30% and 11% of case–control studies, respectively. The choice of cases and controls is "crucial to interpreting the results" and matching "ensures similarity in the distribution of potential confounding variables between cases and controls". Both of these items are directly relevant to the efficiency of a study and their lack of reporting questions the study's validity (Vandenbroucke et al. 2007).

"Addressing loss to follow-up" was a poorly reported cohort-specific item (33%). The validity of the study will be affected "if loss to follow-up occurs selectively in exposed individuals, or in persons at high risk of developing the disease" (Vandenbroucke et al. 2007). The underreporting of this critical item is seen in previous cohort studies (Rao et al. 2016; Poorolajal et al. 2011).

Cross-sectional studies were the second best reported type of OSs (81%). The ARP for cross-sectionals in our study is slightly higher than percentages from previous studies on the reporting of cross-sectional studies (Ramke et al. 2017; Irani et al. 2018).

Our study indicates that the overall reporting quality of OSs affiliated to SUMS has improved over a 3-year period from March, 2012 to March, 2015. Similarly, Bastuji-Garin et al. reported an improvement in the reporting quality of OSs between 2004 and 2010 (Nagarajan et al. 2018) while no such improvement was observed in the study by Rao et al. (2016). on cohort studies between 2002 and 2013 Because of the limited timeframe of our study, it is difficult to draw an accurate conclusion regarding the observed positive trend.

Regarding the articles' field of study, we found articles in the field of clinical medicine to be of higher reporting quality than articles in the field of basic sciences. However, the high number of clinical medicine articles (133/171) and the small number of basic science articles (8/171) in our sample of OSs decreases the accuracy of our comparison. Since the quality of reporting of basic science articles, especially in methods and results sections, affects the reproducibility of the related research (Han et al. 2017), it is suggested that the reporting quality of OSs in the field of basic sciences be evaluated in a separate study.

Another studied variable that showed a significant correlation with ARP was the article's status regarding international co-authorship. Articles with international collaboration demonstrated a higher reporting quality than those without international collaboration. It has been reported that internationally co-authored articles are published in journals with higher impact factors (Low et al. 2014) and are cited more often than articles without international co-authorship (Leydesdorff and Wagner 2008). It is therefore recommended to take steps to initiate international research collaboration with the ultimate goal of improving the quality of research.

Regarding the relationship between the last impact factor of the journal and the quality of reporting of the associated article, we did not find a significant correlation. In one study, the impact factor of the journal was positively correlated with the reporting of some but not all STROBE items (Hemkens et al. 2016). However, since only few studies have been conducted to investigate this association, we can not draw an accurate conclusion. In our study, since we assessed the association between the last impact factor of the journal and only one of the articles in that journal, we can not say with certainty that this relationship does not exist.

We would like to remind and emphasize that our results reflect the *reporting* quality of studies and not the quality of the research or the study design. This is in accordance with recommendations from STROBE creators who stated that "STROBE statement was not developed as a tool for assessing the quality of published observational research". However, transparent reporting of research leads to better assessment of their methodological quality and the improvement of the methodology as well as the design of the articles over time (von Elm et al. 2007). We highly recommend researchers to use STROBE checklist while writing research manuscripts and recommend university research departments to monitor research papers as well as educate researchers regarding transparent reporting. As SUMS is one of the top ten universities in Iran, the study results might be generalized to other Iranian universities and institutions.

Limitations and strengths

We included only published studies which may increase the risk of publication bias. Another limitation is related to the limited time frame of 3 years, which makes it difficult to assess the reporting quality of articles over time. Furthermore, in line with previous studies, we gave each STROBE item equal weight. Since some items might be more important than others for a specific type of OS, our reporting percentages might not be ideal.

Among the strengths of our study are the blinding of the two reviewers to the articles' authors and the names of journals and the presence of inter-rater agreement in our pilot study. In addition, comparing reporting percentages in different study designs and different article sections is an advantage of this study over the previous studies.

Conclusion

The quality of reporting of OSs assessed in this study is generally acceptable. However, items such as "use of a flow diagram", "sources of bias" and "sample size calculation" among all OSs and "loss to follow-up" among cohort studies are poorly reported items that deserve the attention of authors as well as reviewers and editors.

Acknowledgements The authors wish to thank Mr. H. Argasi at the Research Consultation Center (RCC) of Shiraz University of Medical Sciences for his invaluable assistance in editing this manuscript.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Bastuji-Garin, S., Sbidian, E., Gaudy-Marqueste, C., Ferrat, E., Roujeau, J.-C., Richard, M.-A., et al. (2013). Impact of STROBE statement publication on quality of observational study reporting: Interrupted time series versus before-after analysis. *PLoS ONE, 8*(8), e64733–e64733. https://doi.org/10.1371/journal.pone.0064733.

Bero, L., Chartres, N., Diong, J., Fabbri, A., Ghersi, D., Lam, J., et al. (2018). The risk of bias in observational studies of exposures (ROBINS-E) tool: Concerns arising from application to observational studies of exposures. *Systematic Reviews, 7*(1), 242. https://doi.org/10.1186/s13643-018-0915-2.

Fung, A. E., Palanki, R., Bakri, S. J., Depperschmidt, E., & Gibson, A. (2009). Applying the CONSORT and STROBE statements to evaluate the reporting quality of neovascular age-related macular degeneration studies. *Ophthalmology, 116*(2), 286–296. https://doi.org/10.1016/j.ophtha.2008.09.014.

Glasziou, P., Altman, D. G., Bossuyt, P., Boutron, I., Clarke, M., Julious, S., et al. (2014). Reducing waste from incomplete or unusable reports of biomedical research. *Lancet, 383*(9913), 267–276. https://doi.org/10.1016/s0140-6736(13)62228-x.

Groenwold, R. H., Van Deursen, A. M., Hoes, A. W., & Hak, E. (2008). Poor quality of reporting confounding bias in observational intervention studies: A systematic review. *Annals of Epidemiology, 18*(10), 746–751. https://doi.org/10.1016/j.annepidem.2008.05.007.

Hammer, G. P., du Prel, J.-B., & Blettner, M. (2009). Avoiding bias in observational studies: Part 8 in a series of articles on evaluation of scientific publications. *Deutsches Arzteblatt International, 106*(41), 664–668. https://doi.org/10.3238/arztebl.2009.0664.

Han, S., Olonisakin, T. F., Pribis, J. P., Zupetic, J., Yoon, J. H., Holleran, K. M., et al. (2017). A checklist is associated with increased quality of reporting preclinical biomedical research: A systematic review. *PLoS ONE, 12*(9), e0183591. https://doi.org/10.1371/journal.pone.0183591.

Hemkens, L. G., Benchimol, E. I., Langan, S. M., Briel, M., Kasenda, B., Januel, J.-M., et al. (2016). The reporting of studies using routinely collected health data was often in sufficient. *Journal of Clinical Epidemiology, 79,* 104–111. https://doi.org/10.1016/j.jclinepi.2016.06.005.

Hoppe, D. J., Schemitsch, E. H., Morshed, S., Tornetta, P., 3rd, & Bhandari, M. (2009). Hierarchy of evidence: Where observational studies fit in and why we need them. *Journal of Bone and Joint Surgery, 91*(Suppl 3), 2–9. https://doi.org/10.2106/jbjs.H.01571.

Irani, M., Hassanzadeh Bashtian, M., Khadivzadeh, T., Ebrahimipour, H., & Asghari Nekah, S. M. (2018). Weaknesses in the reporting of cross-sectional studies in accordance with the strobe report (the case of congenital anomaly among infants in iran): A review article. *Iranian Journal of Public Health, 47*(12), 1796–1804.

Jeelani, A., Malik, W., Haq, I., Aleem, S., Mujtaba, M., & Syed, N. (2014). Cross-sectional studies published in indian journal of community medicine: Evaluation of adherence to strengthening the reporting of observational studies in epidemiology statement. *Annals of Medical And Health Sciences Research, 4*(6), 875–878. https://doi.org/10.4103/2141-9248.144889.

Langan, S., Schmitt, J., Coenraads, P. J., Svensson, A., von Elm, E., & Williams, H. (2010). The reporting of observational research studies in dermatology journals: A literature-based study. *Archives of Dermatology, 146*(5), 534–541. https://doi.org/10.1001/archdermatol.2010.87.

Leydesdorff, L., & Wagner, C. S. (2008). International collaboration in science and the formation of a core group. *Journal of Informetrics, 2*(4), 317–325. https://doi.org/10.1016/j.joi.2008.07.003.

Low, W. Y., Ng, K. H., Kabir, M. A., Koh, A. P., & Sinnasamy, J. (2014). Trend and impact of international collaboration in clinical medicine papers published in Malaysia. *Scientometrics, 98*(2), 1521–1533. https://doi.org/10.1007/s11192-013-1121-6.

Nagarajan, V. B., Bhide, S., Kanase, H. R., Potey, A., & Tadavi, F. (2018). Adherence of observational studiesindian journals to STROBE statement. *Journal of the Association of Physicians of India, 66,* 39.

Poorolajal, J., Cheraghi, Z., Irani, A. D., & Rezaeian, S. (2011). Quality of Cohort studies reporting post the strengthening the reporting of observational studies in epidemiology (STROBE) statement. *Epidemiology and Health, 33,* e2011005–e2011005. https://doi.org/10.4178/epih/e2011005.

Pouwels, K. B., Widyakusuma, N. N., Groenwold, R. H., & Hak, E. (2016). Quality of reporting of confounding remained suboptimal after the STROBE guideline. *Journal of Clinical Epidemiology, 69,* 217–224. https://doi.org/10.1016/j.jclinepi.2015.08.009.

Ramke, J., Palagyi, A., Jordan, V., Petkovic, J., & Gilbert, C. E. (2017). Using the STROBE statement to assess reporting in blindness prevalence surveys in low and middle income countries. *PLoS ONE, 12*(5), e0176178. https://doi.org/10.1371/journal.pone.0176178.

Rao, A., Bruck, K., Methven, S., Evans, R., Stel, V. S., Jager, K. J., et al. (2016). Quality of reporting and study design of CKD Cohort studies assessing mortality in the elderly before and after STROBE: A systematic review. *PLoS ONE, 11*(5), e0155078. https://doi.org/10.1371/journal.pone.0155078.

Tsakiridis, I., Arvanitaki, A., & Zintzaras, E. (2019). Assessing the reporting quality of systematic reviews of observational studies in preeclampsia. *Archives of Gynecology and Obstetrics, 299*(3), 689–694. https://doi.org/10.1007/s00404-018-5023-y.

Vandenbroucke, J. P., Von Elm, E., Altman, D. G., Gøtzsche, P. C., Mulrow, C. D., Pocock, S. J., et al. (2007). Strengthening the reporting of observational studies in epidemiology (STROBE): Explanation and elaboration. *PLoS Medicine, 4*(10), e297.

von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C., Vandenbroucke, J. P., et al. (2007). The strengthening the reporting of observational studies in epidemiology (STROBE) statement: Guidelines for reporting observational studies. *PLoS Medicine, 4*(10), e296–e296. https://doi.org/10.1371/journal.pmed.0040296.

Affiliations

Negin Rahmani[1,2] · Alireza Salehi[3] · Hossein Molavi Vardanjani[1] · Maryam Marzban[4] · Arezoo Behbood[5]

Negin Rahmani
rahmanine88@yahoo.com

Hossein Molavi Vardanjani
hosseinmolavi@ymail.com

Maryam Marzban
marzbanh@gmail.com

Arezoo Behbood
mahanbehbood@gmail.com

[1] MPH Department, Shiraz Medical School, Shiraz University of Medical Sciences, Shiraz, Iran

[2] Institute for Cancer Research, Shiraz Medical School, Shiraz University of Medical Sciences, Shiraz, Iran

[3] Research Center for Traditional Medicine and History of Medicine, Shiraz Medical School, Shiraz University of Medical Sciences, Shiraz, Iran

[4] Department of Public Health, School of Public Health, Bushehr University of Medical Sciences, Bushehr, Iran

[5] MPH Department, Student Research Committee, Shiraz Medical School, Shiraz University of Medical Sciences, Shiraz, Iran