

## DEVELOPMENT OF A QUANTITATIVE SCORING METHOD FOR STROBE CHECKLIST

DNYANESH LIMAYE<sup>1,2,3\*</sup>, RAVI SHANKAR PITANI<sup>4</sup>, VAIDEHI LIMAYE<sup>1</sup>,  
ARLAN SYDYMANOV<sup>1</sup>, CHRISTIAN OTZIPKA<sup>1</sup>, PATRICK ZIESENIS<sup>1</sup>  
and GERHARD FORTWENGEL<sup>1</sup>

<sup>1</sup>UNIRED working group, Faculty III, Hochschule Hannover, Expo Plaza 12, Hannover 30539, Germany

<sup>2</sup>Hannover Medical School, Hannover, Germany

<sup>3</sup>Department of Epidemiology, Helmholtz Centre for Infection Research, Braunschweig, Germany

<sup>4</sup>Department of Community Medicine, Sri Ramachandra Medical College and Research Institute, Sri Ramachandra University, Porur, Chennai 600116, India

**Abstract:** Evidence-based medicine (EBM) has brought a substantial change in medical practice and clinical care approach. Poor or insufficient quality reporting of observational studies limits the data interpretation, generalizability and clinical utility at large. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines do not have a method to assign a quality score to observational study publication which might lead to differences in assessing the quality of the publication. A scoring method can provide quantitative estimates that could improve quality of reporting, eventual conduct of observational studies and can lead to more transparent judgements about the quality of evidence. Authors decided to develop a quantitative scoring method for STROBE checklist and to design a rating scale in order to categorize cohort (CO), case-control (CC) and cross-sectional (CS) study publications into excellent, good, fair, poor quality based on the overall scores. STROBE-M (STROBE Modified) checklist was developed by dividing composite STROBE items into multiple items for ease of comprehension and scoring. For each item, the authors decided to have 3 quality scores option i.e. 0 (not fulfilled), 1 (fulfilled) and NA (not applicable). By using STROBE-M checklist, two reviewers independently graded 45 observational study publications, 15 each from CO, CC and CS. Inter-rater agreement and reliability were determined for grading done by 2 raters. STROBE-M adherence scoring method was developed to assign quality scores to study publications. High level agreement (inter rater reliability) was seen between two reviewers for rating all 3 types of study publications by STROBE-M checklist. CS studies had 6 (40%) good and 9 (60%) fair; CC studies had 1 (7%) good and 14 (93%) excellent; CO studies had all 15 (100%) study publications with excellent grade as per STROBE-M scoring. Cross-sectional studies had overall fair STROBE-M adherence grade while case-control studies and cohort studies had excellent (85%) grade. This study highlighted deficiencies in the reporting of observational studies. Based on our experience in scoring 45 publications, the STROBE-M checklist with the proposed scoring system seems to be a useful tool for assessing the reporting quality of the observational studies.

**Keywords:** STROBE, STROBE-M, observational studies, cross-sectional, cohort, case-control, publication quality, scoring method

Evidence-based medicine (EBM) has brought a substantial change in medical practice and clinical care approach (1-3). It has helped to focus on valid and reliable information to make sound decisions in the area of disease management. Although randomized controlled trials (RCTs) are considered as the reference standard and provide valuable evidence about the efficacy of new treatments, yet they do not answer all important questions about the specific treatment. Most of the knowledge in clinical practice comes from observational studies from which

we can gain further experience to plan interventional studies in regards to efficacy, safety and tolerability of approved medicines (4).

An observational study is defined as a study that provides estimates and examines associations of events in their natural settings without recourse to experimental intervention (5). It includes cohort studies (CO) – prospective and retrospective, cross-sectional studies (CS), and case-control studies (CC) (6). Current evidence suggests that observational studies, such as those discussed earlier, both com-

\* Corresponding author: e-mail: dnyanesh.limaye@hs-hannover.de

plement and build on the evidence base established by RCTs. Poor or insufficient quality reporting of observational studies limits the data interpretation, generalizability and clinical utility (7). It is seen often that healthcare professionals come across publications of inadequate quality which cannot be used as evidence in clinical decision making (4, 8). Poor quality reporting also hinders the process of evidence generation. Incomplete or inadequate reporting also leads to omission of otherwise eligible studies in systematic reviews or meta-analysis (4).

In the area of RCTs, guidelines such as the Consolidated Standards of Reporting Trials (CONSORT) have improved the reporting quality (9-11). Similarly, the 'Strengthening the Reporting of Observational Studies in Epidemiology' (STROBE) Statement was developed in a collaborative effort of epidemiologists, methodologists, statisticians, researchers, and journal editors to improve the quality of reporting of observational studies (4, 12). It consists of a checklist of 22 items, which relate to the title, abstract, introduction, methods, results and discussion sections of articles. The STROBE guideline explains and elaborates the rationale for each item by citing examples from published articles. STROBE guideline has been accepted and used by more than 100 journals as well as by the International Committee of Medical Journal Editors (4). STROBE guideline ensures a clear presentation of what was planned, done, and found in an observational study and the authors of STROBE explicitly state that 'the recommendations are not prescriptions for setting up or conducting studies, nor do they dictate methodology or mandate a uniform presentation (4). The positive effect of reporting guidelines on the quality of reporting has been documented in several reviews (10, 13-20).

So far, there is no method to assign an overall quality score, as well as an individual score to each of the 22 items in the STROBE guideline. This might lead to differences in quality rating when it is done by multiple reviewers. Scoring method for items can provide quantitative estimates of quality that could improve the quality of reporting and lead to more transparent judgements about the quality of evidence (21). It may also improve the data interpretation, generalizability and clinical utility of the reported study (22).

## OBJECTIVES

The first objective of the present study was to develop a valid and reliable quantitative scoring method for STROBE for determining the overall

quality score for a given publication. The second objective was to design a rating scale in order to categorize publications into excellent, good, fair, or poor quality based on the overall scores.

## METHODS

The scoring system was developed based on discussions in the UNIRED (UNESCO German Unit of the International Network at Hochschule Hannover, The UNESCO Chair In Bioethics, Medical Research in Developing countries) working group of the Faculty III, Hochschule Hannover, University of Applied Sciences and Arts, Hannover, Germany. The UNIRED group consists of professors, senior researchers, and masters students working in the area of public health, epidemiology and ethics. The group convenes twice a month to discuss research proposals, progress of ongoing projects, manuscripts in preparation, or methodological papers from the literature.

### Scoring method

The aim of the scoring method was to obtain an overall quality score for a given publication. All questions were derived directly from the STROBE guideline. It was decided to have scores of 0: if the particular checklist item is not fulfilled, a score of 1: if the particular checklist item is fulfilled and Score of NA: if particular checklist item is not applicable for the specific publication. Maximum possible STROBE scores for observational studies were as follows: cohort = 84, case-control = 83 and cross-sectional = 77 (Appendix 1).

### STROBE-M (STROBE Modified)

New items were added under methods and other information section. Items related to ethics were added in the domain 2 under the methods section i.e. ethics committee permission and written informed consent from study participants. We divided composite STROBE checklist items into multiple items for ease of comprehension and scoring. e.g. "Original STROBE item" – Setting – Describe the setting, location, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection. "Separated STROBE items" – "Study setting mentioned" yes = 1, no = 0; "Study location written" yes = 1, no = 0; "Relevant dates mentioned" (recruitment, exposure, follow-up, data collection) if all 4 dates mentioned = 4, few dates mentioned (1 to 3) = 1 to 3, no = 0.

To get a final percent STROBE-M adherence for a particular publication, total STROBE-M adher-

ence score was counted based on fulfillment of STROBE-M items which was then divided by the applicable maximum possible STROBE-M score (cohort = 84, case-control = 83 and cross-sectional = 77).

$$\text{Percent STROBE-M adherence} = \frac{\text{Total STROBE-M adherence score}}{\text{Applicable maximum possible STROBE-M score}} \times 100$$

For example, if a cross-sectional study publication achieves total STROBE-M adherence score of 35, then this score will be divided by 77 (applicable maximum possible STROBE-M score for a cross-sectional study). In this case, % STROBE-M adherence =  $(35/77) \times 100 = 45.5\%$  is achieved.

#### Piloting of the STROBE-M score checklist

We conducted a piloting of the STROBE-M score checklist. The two reviewers (VL, PR) from UNIREG group independently graded 12 observational study papers (3 each from cohort, case-control, and cross-sectional studies) with draft STROBE-M

score checklist. These papers were published in the year 2016 and were available on the PubMed Medline database. These two reviewers then met with UNIREG group to discuss their assessments and any difficulties they experienced. Based on these discussions, some modifications were done in framing of the sentences for individual STROBE-M checklist items.

#### Inter-rater agreement of STROBE-M score checklist

Following the development and piloting of the STROBE-M score checklist, we conducted a more formal evaluation of its inter-rater agreement. One reviewer (VL) selected 45 observational study publications (15 each from cross-sectional (23-37), cohort (38-52), and case-control study (53-67) from "PubMed database" published in the time period of 1<sup>st</sup> March 2016 to 31<sup>st</sup> May 2016 with the filters of English language, human studies, and full text availability. Irrelevant publications (opinion or critique of previous studies, letters to editors, meta-analysis were excluded (Fig. 1). The two reviewers (RP, FF) independently graded these 45 observational study publications (23-67) with STROBE-M score checklist. Inter-rater agreement (IRR) was determined by using kappa statistic and Intra Class Correlation (ICC). Possible values for kappa statistics range from -1 to 1, with 1 indicating perfect agreement, 0 indicating completely random agreement, and -1 indicating "perfect" disagreement (68-70). In case

Table 1. Publication quality grades.

STROBE-M adherence score (%)	Quality grade for a publication
≥ 85	Excellent
70 to < 85	Good
50 to < 70	Fair
< 50	Poor

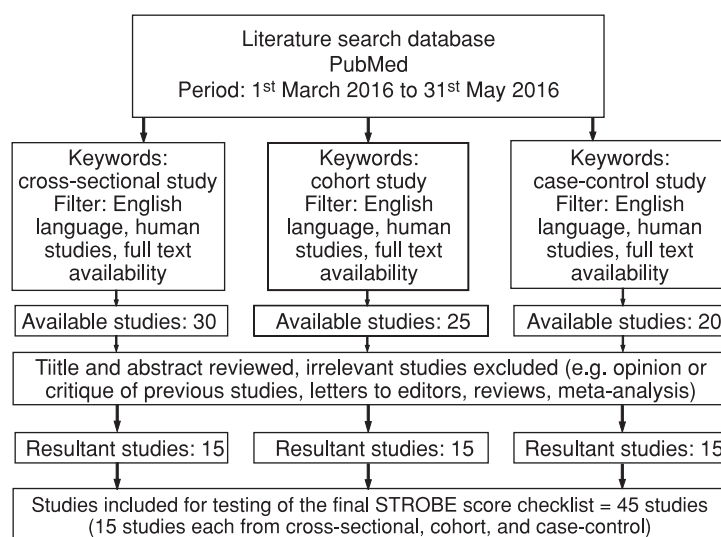


Figure 1. Search strategy for selecting 45 observational study publications

of ICC, poor IRR for ICC values less than 0.40, fair for values between 0.40 and 0.59, good for values between 0.60 and 0.74, and excellent for values between 0.75 and 1.0 was considered (68-70). Each reviewer recorded the time spent in conducting an assessment, and from this data, we estimated the likely resource implications of using this approach.

Two authors (DL, GF) provided formal STROBE-M and STROBE-M scoring training during the course of the study.

#### Determining the publication quality

Further to our efforts of formulating a quantitative scoring method for STROBE checklist, we

Table 2. STROBE-M Scores, mean, ICC, kappa for cross-sectional study publications.

Publication	STROBE-M Score (out of 77) Rater 1	STROBE-M Score (out of 77) Rater 2	STROBE-M mean score (out of 77)	ICC	Kappa
1 <sup>23</sup>	60.0	57.0	58.5	0.954	0.884
2 <sup>24</sup>	50.0	48.0	49.0	0.967	0.936
3 <sup>25</sup>	55.0	55.0	55.0	0.966	0.930
4 <sup>26</sup>	54.0	52.0	53.0	0.963	0.929
5 <sup>27</sup>	57.0	54.0	55.5	0.938	0.882
6 <sup>28</sup>	58.0	60.0	59.0	0.978	0.924
7 <sup>29</sup>	56.0	54.0	55.0	0.961	0.923
8 <sup>30</sup>	55.0	55.0	55.0	0.960	0.923
9 <sup>31</sup>	48.0	51.0	49.5	0.950	0.903
10 <sup>32</sup>	54.0	53.0	53.5	0.943	0.891
11 <sup>33</sup>	53.0	51.0	52.0	0.927	0.862
12 <sup>34</sup>	55.0	52.0	53.5	0.943	0.891
13 <sup>35</sup>	52.0	53.0	52.5	0.945	0.895
14 <sup>36</sup>	54.0	51.0	52.5	0.945	0.895
15 <sup>37</sup>	52.0	55.0	53.5	0.943	0.891

Table 3. STROBE-M Scores, mean, ICC, kappa for case-control study publications.

Publication	STROBE-M Score (out of 83) Rater 1	STROBE-M Score (out of 83) Rater 2	STROBE-M mean score (out of 83)	ICC	Kappa
1 <sup>53</sup>	76	76	76	1.000	1.000
2 <sup>54</sup>	69	69	69	0.957	0.910
3 <sup>55</sup>	78	79	78.5	0.924	0.852
4 <sup>56</sup>	71	73	72	0.945	0.888
5 <sup>57</sup>	73	73	73	1.000	1.000
6 <sup>58</sup>	75	74	74.5	0.965	0.928
7 <sup>59</sup>	73	74	73.5	0.901	0.806
8 <sup>60</sup>	76	74	75	0.923	0.846
9 <sup>61</sup>	71	74	72.5	0.912	0.825
10 <sup>62</sup>	76	76	76	1.000	1.000
11 <sup>63</sup>	72	71	71.5	0.920	0.839
12 <sup>64</sup>	77	77	77	1.000	1.000
13 <sup>65</sup>	74	74	74	1.000	1.000
14 <sup>66</sup>	72	75	73.5	0.901	0.807
15 <sup>67</sup>	77	75	76	0.910	0.823

Table 4. STROBE-M Scores, mean, ICC, kappa for cohort study publications.

Publication	STROBE-M Score (out of 84) Rater 1	STROBE-M Score (out of 84) Rater 2	STROBE-M mean score (out of 84)	ICC	Kappa
1 <sup>38</sup>	75.5	78	76.75	0.899	0.775
2 <sup>39</sup>	72	72	72	0.950	0.897
3 <sup>40</sup>	78	76	77	0.910	0.823
4 <sup>41</sup>	78	77	77.5	0.843	0.712
5 <sup>42</sup>	76	77	76.5	0.869	0.753
6 <sup>43</sup>	75	76	75.5	0.888	0.784
7 <sup>44</sup>	76	75	75.5	0.888	0.784
8 <sup>45</sup>	77	79	78	0.890	0.790
9 <sup>46</sup>	77	77	77	0.909	0.823
10 <sup>47</sup>	76	74	75	0.932	0.864
11 <sup>48</sup>	72	74	73	0.945	0.888
12 <sup>49</sup>	77	77	77	0.909	0.823
13 <sup>50</sup>	75	74	74.5	0.969	0.936
14 <sup>51</sup>	79	79	79	0.859	0.740
15 <sup>52</sup>	77	77	77	0.909	0.823

decided to grade the quality of publications (71, 72) on the basis of STROBE-M adherence score as shown in Table 1.

## RESULTS

As detailed in the methods section, as a first step towards developing a quantitative scoring method for STROBE checklist, we divided composite STROBE items into multiple items for ease of comprehension and scoring. Our STROBE-M checklist consists of 73 cross-sectional, 78 case-control and 79 cohort STROBE items. These items reflect all 3 domains of observational study publications i.e. domain 1: abstract, domain 2: introduction, methods, results, discussion and domain 3: other information (Appendix 1). The STROBE scoring checklist was applied to a set of 45 observational study publications (15 each for cross-sectional, case-control and cohort studies). The results are presented in Tables 2-4.

### Inter-rater reliability

A high level of agreement (inter-rater reliability) was seen between two raters for all types of study publications as depicted by ICC and kappa values shown in Tables 2-4.

### Time required for conducting assessments

The median time to conduct the assessment was 30 min (range = 15–40 min). Raters' feedback

suggested that it was relatively straightforward to use the STROBE-M checklist after formal training on STROBE-M and STROBE-M scoring method.

### Quality of study publications

Figure 2 shows the distribution of percent STROBE-M adherence scores for all domains of the three types of observational studies. For the cross-sectional study, domain one score is 96.2%, whereas, for the rest of the two domains, the score is below 85%. i.e. below excellent grade (Table 1). The overall grade for case-control and cohort studies is excellent (more than 85%) but the cross-sectional studies' overall scores are less than 85%. At the same time, it is observed that the domain three score is quite variable and also scores less than 85% across all the three study designs.

Tables 5, 6 and 7 show percentage STROBE-M adherence and quality score of study publications. All 3 types of study publications' STROBE-M adherence was of an excellent grade in domain 1 (Title and abstract) of STROBE checklist items: CS: 96.2 (SD = 8.5), CC: 100, CO: 96.2 (SD = 6.5). In the case of domain 2 (Introduction, methods, results, and discussion) of STROBE-M checklist items, CS study publications achieved fair grade 69.2 (2.6) whereas CC 90.7 (1.7) and CO 92.2 (1.6) study publications STROBE-M adherence was of excellent grade. CS study publications achieved poor grade 41.3 (27.7) while CC 54.7 (30.7) and CO 60 (22.7) study publications achieved a fair grade in domain 3 (other infor-

mation) of STROBEs checklist items. CS had 6 (40%) good and 9 (60%) fair; CC had 1 (7%) good and 14 (93%) excellent; CO had all 15 (100%) study publications with excellent grade as shown in Figure 3.

## DISCUSSION

STROBE-M guideline was developed 10 years back (4) to ensure a clear presentation of what was

planned, done, and found in an observational study. The main difficulty in applying the STROBE guideline to observational study publications is the complexity of the approach. The STROBE guideline has 22 items and each item has more than one sub-item / recommendation. For e.g. 1(b) of STROBE guideline recommends providing in the abstract an Informative and balanced summary of what was done and what was found. We separated STROBE

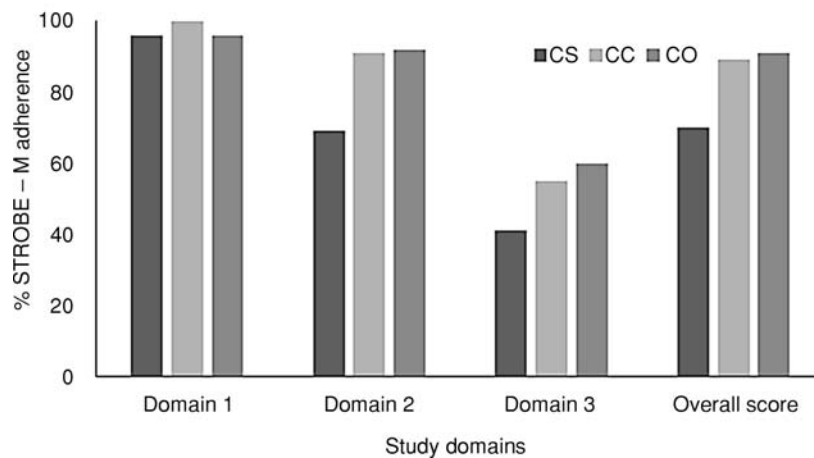


Figure 2. Percentage STROBE-M adherence for 3 types of study publications

Table 5. Percentage STROBE-M adherence and quality grades for cross-sectional study publications.

Publication	Domain 1 (Total = 7)			Domain 2 (Total = 65)			Domain 3 (Total = 5)			Overall (Total = 77)		
	Score	%	Grade	Score	%	Grade	Score	%	Grade	Score	%	Grade
1 <sup>23</sup>	7	100.0	E	47.5	73.1	G	4	80	G	58.5	76.0	G
2 <sup>24</sup>	7	100.0	E	42	64.6	F	0	0	P	49	63.6	F
3 <sup>25</sup>	7	100.0	E	46	70.8	G	2	40	P	55	71.4	G
4 <sup>26</sup>	5	71.4	G	45	69.2	F	3	60	F	53	68.8	F
5 <sup>27</sup>	7	100.0	E	45.5	70.0	G	3	60	F	55.5	72.1	G
6 <sup>28</sup>	7	100.0	E	49	75.4	G	3	60	F	59	76.6	G
7 <sup>29</sup>	6	85.7	E	45	69.2	F	4	80	G	55	71.4	G
8 <sup>30</sup>	6	85.7	E	45	69.2	F	4	80	G	55	71.4	G
9 <sup>31</sup>	7	100.0	E	42.5	65.4	F	0	0	P	49.5	64.3	F
10 <sup>32</sup>	7	100.0	E	45.5	70.0	G	1	20	P	53.5	69.5	F
11 <sup>33</sup>	7	100.0	E	44	67.7	F	1	20	P	52	67.5	F
12 <sup>34</sup>	7	100.0	E	44.5	68.5	F	2	40	P	53.5	69.5	F
13 <sup>35</sup>	7	100.0	E	44.5	68.5	F	1	20	P	52.5	68.2	F
14 <sup>36</sup>	7	100.0	E	44.5	68.5	F	1	20	P	52.5	68.2	F
15 <sup>37</sup>	7	100.0	E	44.5	68.5	F	2	40	P	53.5	69.5	F
		Mean (SD) 96.2 (8.5)	E		Mean (SD) 69.2 (2.6)	F		Mean (SD) 41.3 (27.7)	P		Mean (SD) 69.9 (3.6)	F

Table 6. Percentage STROBE-M adherence and quality grades for case-control study publications.

Publication	Domain 1 (Total = 7)			Domain 2 (Total = 71)			Domain 3 (Total = 5)			Overall (Total = 83)		
	Score	%	Grade	Score	%	Grade	Score	%	Grade	Score	%	Grade
1 <sup>53</sup>	7	100	E	65	91.5	E	4	80.0	G	76	91.6	E
2 <sup>54</sup>	7	100	E	62	87.3	E	0	0.0	P	69	83.1	G
3 <sup>55</sup>	7	100	E	66.5	93.7	E	5	100.0	E	78.5	94.6	E
4 <sup>56</sup>	7	100	E	64	90.1	E	1	20.0	P	72	86.7	E
5 <sup>57</sup>	7	100	E	65	91.5	E	1	20.0	P	73	88.0	E
6 <sup>58</sup>	7	100	E	64.5	90.8	E	3	60.0	F	74.5	89.8	E
7 <sup>59</sup>	7	100	E	64.5	90.8	E	2	40.0	P	73.5	88.6	E
8 <sup>60</sup>	7	100	E	64	90.1	E	4	80.0	G	75	90.4	E
9 <sup>61</sup>	7	100	E	63.5	89.4	E	2	40.0	P	72.5	87.3	E
10 <sup>62</sup>	7	100	E	65	91.5	E	4	80.0	G	76	91.6	E
11 <sup>63</sup>	7	100	E	63.5	89.4	E	1	20.0	P	71.5	86.1	E
12 <sup>64</sup>	7	100	E	67	94.4	E	3	60.0	F	77	92.8	E
13 <sup>65</sup>	7	100	E	64	90.1	E	3	60.0	F	74	89.2	E
14 <sup>66</sup>	7	100	E	63.5	89.4	E	3	60.0	F	73.5	88.6	E
15 <sup>67</sup>	7	100	E	64	90.1	E	5	100.0	E	76	91.6	E
		Mean (SD) 100.0 (0.0)	E		Mean (SD) 90.7 (1.7)	E		Mean (SD) 54.7 (30.7)	F		Mean (SD) 89.3 (2.9)	E

Table 7. Percentage STROBE-M adherence and quality grades for cohort study publications.

Publication	Domain 1 (Total = 7)			Domain 2 (Total = 72)			Domain 3 (Total = 5)			Overall (Total = 84)		
	Score	%	Grade	Score	%	Grade	Score	%	Grade	Score	%	Grade
1 <sup>38</sup>	7	100.0	E	67.8	94.1	E	2	40	P	76.75	91.4	E
2 <sup>39</sup>	6	85.7	E	66	91.7	E	0	0	P	72	85.7	E
3 <sup>40</sup>	7	100.0	E	66	91.7	E	4	80	G	77	91.7	E
4 <sup>41</sup>	7	100.0	E	66.5	92.4	E	4	80	G	77.5	92.3	E
5 <sup>42</sup>	7	100.0	E	66.5	92.4	E	3	60	F	76.5	91.1	E
6 <sup>43</sup>	7	100.0	E	66.5	92.4	E	2	40	P	75.5	89.9	E
7 <sup>44</sup>	6	85.7	E	66.5	92.4	E	3	60	F	75.5	89.9	E
8 <sup>45</sup>	7	100.0	E	67	93.1	E	4	80	G	78	92.9	E
9 <sup>46</sup>	7	100.0	E	68	94.4	E	2	40	P	77	91.7	E
10 <sup>47</sup>	7	100.0	E	65	90.3	E	3	60	F	75	89.3	E
11 <sup>48</sup>	6	85.7	E	64	88.9	E	3	60	F	73	86.9	E
12 <sup>49</sup>	7	100.0	E	66	91.7	E	4	80	G	77	91.7	E
13 <sup>50</sup>	6	85.7	E	64.5	89.6	E	4	80	G	74.5	88.7	E
14 <sup>51</sup>	7	100.0	E	68	94.4	E	4	80	G	79	94.0	E
15 <sup>52</sup>	7	100.0	E	67	93.1	E	3	60	F	77	91.7	E
		Mean (SD) 96.2 (6.5)	E		Mean (SD) 92.2 (1.6)	E		Mean (SD) 60.0 (22.7)	F		Mean (SD) 90.6 (2.2)	E

items into multiple items and developed a modified STROBE checklist (STROBE-M) applicable to observational studies (CS, CC, and CO).

Ethics is not listed in STROBE guideline as a recommended item. We are of the opinion that failure to consider ethics in guidelines for reporting research is an omission. The consideration of ethical issues by researchers and the review by ethics committees reflect comprehensive modern ethical thinking. The rights of research participants and the safeguarding of the records of individuals being researched are ethical requirements that are well documented (73). No research should proceed without consideration of ethical issues anticipated by researchers and without review by an independent and competent ethics review body. These two actions should be considered as integral parts of the research process, and are thus logical and indispensable elements of the study report checklist (73).

Inappropriate authorship (honorary and ghost authorship) and the resulting lack of transparency and accountability have been substantial concerns for the academic community for decades (74-77). In recent years, there has been much debate about the increasing list of authors in published papers, the order of authors as well as responsibilities. Ghost authors who might be professional writers or staff members might have a conflict of interest in the outcome of the research (or at least the presentation of the research at publication). By omitting their names, the paper affiliations look more neutral (78).

A recent Nature editorial entitled “Ethics and fraud”: “...no one should argue ever again that ... promiscuous authorship on scientific papers ... can be tolerated ... Research ethics matter immensely to the health of the scientific enterprise. Anyone who thinks differently should seek employment in another sphere” (77). To have an unbiased view while reviewing the publication, based on the most prominent and widely utilized authorship guidelines established by the International Committee of Medical Journal Editors (ICMJE), we added an item on authors’ and co-authors’ contribution to the STROBE-M checklist (79).

Public trust in the scientific process and the credibility of published articles depend in part on how transparently competing interests are handled during the planning, implementation, writing, peer review, editing, and publication of scientific work. Competing interest exists when professional judgment concerning a primary interest such as patients’ welfare or the validity of research may be influenced by a secondary interest such as financial gain, personal relationships, academic competition, and intellectual beliefs (79). As a statement about competing interest supports to maintain the credibility of the journal, the authors, and the science itself; we added an item about competing interest in STROBE-M checklist (79).

Data sharing (safeguarding confidentiality and anonymity) benefits numerous research-related activities: reproducing analyses, testing secondary

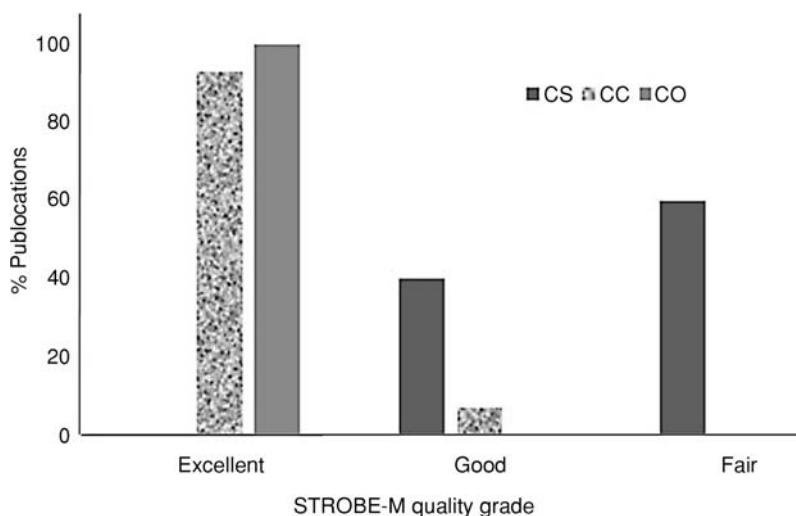


Figure 3. STROBE – M quality grade for 3 types of study publications



hypotheses, developing and evaluating novel statistical methods, aiding the design of studies, meta-analyses; advancing medical knowledge, helping to prevent error, fraud and selective reporting (80).

As an inherent principle of publication that researchers should benefit from, is to be able to replicate and build upon the authors' published claims, we added an item about data sharing in STROBE-M checklist (81).

We presented a scoring system for the STROBE-M that can be used to obtain an overall score for the quality of observational study publications. Three response options for each item of STROBE-M checklist were defined – score of 0: if the particular checklist item is not fulfilled, score of 1: if the particular checklist item is fulfilled and score of NA: if particular checklist item is not applicable for the publication. Subsequently, based on the percentage adherence to STROBE-M score, we decided for 4 quality grades for publications which are, excellent (= 85%), good (70% – < 85%), fair (50% – < 70%) and poor (< 50%). A quality score is highly desirable, as it allows to present conclusions on the quality of the publication based on the various levels of evidence. We also decided to focus not only on the overall STROBE-M adherence score / quality but also on score / quality of each domain of the publication. We identified 3 domains in the publications as domain 1: abstract, domain 2: introduction, methods, results, discussion and domain 3: other information. Individual domain STROBE-M adherence and quality score offered a focused understanding of the positives or lacunas in the publications. The STROBE-M checklist as presented in this manuscript showed a high-level inter-rater reliability when applied to 45 observational study publications (15 each from CS, CC and CO).

Our analysis of 45 observational study publications (15 each from CS, CC and CO), showed that CS study publications achieved an overall fair grade as compared to excellent grade for CC and CO. In the case of domain 2 (introduction, methods, results, and discussion) of STROBE-M checklist items, CS study publications achieved fair grade whereas CC and CO study publications achieved an excellent grade. In the case of domain 3 (other information), neither of the study publications even achieved a good grade. Our results are consistent with other studies assessing deficiencies in reporting of individual STROBE items and also more specifically in the domain 2 such as sample size, bias, selection of participants, use of flow diagram and reporting of missing data (82-86).

One of the main goals of reporting guidelines is to improve reporting clarity and not necessarily

improve the quality of research, but in due course achieve it as an indirect effect. Inadequate reporting not only hinders critical assessment by others of the strengths and weaknesses in study design, conduct, and analysis, but it also affects the judgement of whether and how results can be included in systematic reviews and impacts on the readers' assessment of studies' generalizability (21). This study has shown that it is feasible to develop a scoring method for STROBE checklist that can be used to assess the reporting quality of observational (cross-sectional, case-control and cohort) studies. It has also shown that it is possible to simplify the STROBE checklist for improving transparency and comprehensibility.

## CONCLUSION

This study highlighted deficiencies in the reporting of observational studies. With continued efforts from researchers and with a particular focus on the domains identified as deficient by the STROBE statement, this presents an opportunity to improve the reporting of observational research so as to make it useful in evidence-based medicine. With increased awareness by authors and editors regarding compliance of manuscripts to the STROBE-M checklist and journal endorsement of the same, we hope that not only reporting but also the design and conduct of future studies will be improved (21). Finally, we will like to mention that, based on our experience in scoring 45 publications, we firmly believe that the STROBE-M checklist with the proposed scoring system is a useful tool for assessing the reporting quality of the observational studies.

### Strengths and limitations of this study

- STROBE guidelines do not have a method to assign the quality score to observational study publication which might lead to differences in assessing the quality of the publication.
- A scoring method can provide quantitative estimates that could improve the quality of reporting.
- Our STROBE-M checklist with the proposed scoring system is a useful tool for assessing the reporting quality of the observational studies.
- STROBE – M also will improve quality of reporting, eventual conduct of observational studies and can lead to more transparent judgements about the quality of evidence.

### Competing interests

The authors declare that they have no competing interest.

### Contributorship statement

All authors had full access to all of the data in the study and can take responsibility for the integrity of the data and the accuracy of the data analysis.

### Transparency declaration

The lead author (Dnyanesh Limaye) affirm that our manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

### Ethics approval and consent to participate

Not applicable

### Funding

We did not receive any funding for this project.

### Patient involvement statement

There are no patients involved in this study.

### Data sharing statement

We have shared data as Appendix 1: Limaye Dnyanesh. (2017). STROBE-M quality rating sheet for observational studies [Data set]. Zenodo. <http://doi.org/10.5281/zenodo.824560>.

### REFERENCES

- Guyatt G., Drummond R.: Part 1. The Basics: Using the Medical Literature. 1A. Introduction: The philosophy of evidence based medicine. In Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. Second edition. American Medical Association, Chicago. pp 3-12, (2009).
- Sackett D.L.: Evidence-Based Medicine. *Semin Perinatol.* 21, 3 (1997).
- Haynes R.B., Sackett D.L., Gray J.M., Cook D.J., Guyatt G.H.: *ACP J. Club* 125, A14 (1996).
- Vandenbroucke J.P., von Elm E., Altman D.G., Gøtzsche P.C., Mulrow C.D. et al.: *Epidemiology* 18, 805 (2007).
- Papanikolaou P.N., Christidi G.D., Ioannidis J.P.: *CMAJ* 174, 635 (2006).
- Mann C.J.: *Emerg. Med. J.* 20, 54 (2003).
- Ligthelm R.J., Borzž V., Gumprecht J., Kawamori R., Wenying Y. et al.: *Clin. Ther.* 29, 1284 (2007).
- da Costa B.R., Cevallos M., Altman D.G., Rutjes A.W., Egger M.: *BMJ Open* 1:e000048. doi :10.1136/bmjopen-2010-000048 (2011).
- Glasziou P., Meats E., Heneghan C., Shepperd S.: *BMJ* 336, 1472 (2008).
- Devereaux P.J., Manns B.J., Ghali W.A.: *Control. Clin. Trials* 23, 380 (2002).
- Moher D., Jones A., LePage L.: *JAMA* 285, 1992 (2001).
- STROBE statement. ISPM – University of Bern (2009).
- Plint A.C., Moher D., Morrison A., Schulz K., Altman D.G. et al.: *Med. J. Aust.* 185, 263e7 (2006).
- Smidt N., Rutjes A.W., van der Windt D.A., Ostelo R.W., Bossuyt P.M. et al.: *Neurology* 67, 792e7 (2006).
- Prady S.L., Richmond S.J., Morton V.M., MacPherson H.: *PLoS One* 3, e1577 (2008).
- Moher D., Ocampo M., Altman D.J.: 6<sup>th</sup> International Congress on Peer Review and Biomedical Publication (2009).
- Bastuji-Garin S., Sbidian E., Gaudy-Marqueste C., Ferrat E., Roujeau J.C. et al.: *PLoS ONE.* 8, e64733 (2013).
- Cobo E., Cortés J., Ribera I.M.: *BMJ* 22, d6783 (2011).
- Mannocci A., Saulle R., Colamesta V., D'Aguzzo S, Giraldi G.: *J. Public Health (Oxf)* 37, 737 (2015).
- Yoon U., Knobloch K.: *BMC Med. Res. Methodol.* 12, 47(2012).
- Rao A., Brück K., Methven S., Evans R., Stel V.S. et al.: A Systematic Review. *PLoS ONE.* 11, e0155078 (2016).
- Thornton J., Alderson P., Tan T., Turner C., Latchem S. et al.: *J. Clin. Epidemiol.* 66,124 (2013).
- Sakamoto R., Okumiya K., Norboo T., Tsering N., Yamaguchi T. et al. *Psychiatry Res.* 249, 51 (2017).
- Swope J.J., Couey M.A., Wilson J.W., Jundt J.S.: *J. Oral. Maxillofac Surg.* 75, 1010 (2017).
- Oliveira E.F., Nacif S.R., Urbano J.J., Silva A.S., Oliveira C.S. et al.: *Neuromuscul Disord.* 27, 120 (2017).
- Tsabedze N., McCutcheon K., Mkhwanazi L., Garda R., Vachiat A. et al.: *Int. J. Cardiol.* 230, 175 (2017).
- Acosta-Navarro J., Antoniazzi L., Oki A.M., Bonfim M.C., Hong V. et al.: *Int. J. Cardiol.* 230, 562 (2017).

28. Lee S.Y., Cho N.M., Jung Y.O., Seo Y., Kim H.A.: *J. Korean Neurosurg. Soc.* 60, 67 (2017).
29. da Costa A.G., Vargas S., Clode N., Graça L.M.: *Acta Med. Port.* 29, 514 (2016).
30. Semedo D., Ladeiro F., Ruivo M., d'Oliveira C., de Sousa F. et al.: *Acta Med. Port.* 29, 507 (2016).
31. Tyoalumun K., Abubakar S., Christopher N.: *Int. J. MCH AIDS.* 5, 146 (2016).
32. Alemayehu A., Gedefaw L., Yemane T., Asres Y.: *Anemia* 9817358 (2016).
33. Kizilyildiz B.S., Sönmez B., Karaman K., Beger B., Mercen A.: *Pediatr. Rep.* 8, 6112 (2016).
34. Wolide A.D., Zawdie B., Alemayehu T., Tadesse S.: *Diabetes Metab. Syndr.* 9, 417 (2016).
35. Gyasi R.M., Agyemang-Duah W., Mensah C.M., Arthur F., Torkornoo R. et al.: *J. J. Altern. Complement Med.* 7, 126 (2017).
36. Kumar M.S., Singh A., Jaryal A.K., Ranjan P., Deepak K.K. et al.: *Int. J. Hepatol.* 5160754 (2016).
37. Kubo M.N., Kayima J.K., Were A.J., Ezzi M.S., McLigeyo S.O. et al.: *Int. J. Hypertens.* 8450596 (2016).
38. Käkälä J., Marttila R., Keskinen E., Veijola J., Isohanni M. et al.: *Psychiatry Res.* 249, 16 (2017).
39. Fagin A.P., Gelesko S., Wax M.K., Petrisor D.: *J. Maxillofac Oral Surg.* 75, 1549 (2017).
40. Abbasi A., Kieneker L.M., Corpeleijn E., Gansvoort R.T., Gans R.O. et al.: *Clin. Chem.* 63, 278 (2017).
41. Ou H.T., Chang K.C., Li C.Y., Yang C.Y., Ko N.Y.: *Int. J. Cardiol.* 230, 592 (2016).
42. Nguyen V., Cimadevilla C., Arangalage D., Dehoux M., Codogno I. et al.: *Int. J. Cardiol.* 230, 371 (2016).
43. Wu C., Huo W., Li Y., Zhang B., Wan Y. et al.: *Chemosphere* 172, 29 (2017).
44. Stevelink S. A., Fear N.T.: *JRSM Open.* 7, 2054270416639481 (2016).
45. Goyette M.S., Wilson K.S., Deya R., Masese L.N., Shafi J. et al.: *J. Acquir. Immune Defic. Syndr.* 74, 488 (2017).
46. Greiver M., Wintemute K., Aliarzadeh B., Martin K., Khan S. et al.: *J. Innov. Health Inform.* 23, 580 (2016).
47. Fitzsimons E., Goodman A., Kelly E., Smith J.P.: *Soc. Sci. Med.* 175, 43 (2017).
48. Kamel M., Kadian M., Srinivas T., Taber D., Salas M.A.: *World J. Transplant.* 6, 697 (2016).
49. Ebonyi A.O., Oguche S., Agbaji O.O., Sagay A.S., Okonkwo P.I. et al.: *Germs.* 6, 139 (2016).
50. Kim D.W., Lee B., Lim J.H., Park J.H., Nam W. et al.: *J. Korean Assoc. Oral Maxillofac. Surg.* 42, 358 (2016).
51. Jinnin R., Okamoto Y., Takagaki K., Nishiyama Y., Yamamura T. et al.: *Neuropsychiatr Dis Treat.* 13, 25 (2017).
52. Rogawski E.T., Platts-Mills J.A., Seidman J.C., John S., Mahfuz M.: *Bull World Health Organ.* 95, 61 (2017).
53. Hedström A.K., Mowry E.M., Gianfrancesco M.A., Shao X., Schaefer C.A. et al.: *J. Neurol. Neurosurg. Psychiatry* 87, 454 (2016).
54. Gonçalves S., Machado B.C., Martins C., Hoek H.W., Machado P.P.: *Eur. Eat. Disorders Rev.* 24, 197 (2016).
55. Alemu Y.M., Awoke W., Wilder-Smith A.: *BMJ Open* 6, e009058 (2016).
56. Clark B.A., Osadchuk L., John J., Culver T., Marcus R.: *Transfusion* 56, 933 (2016).
57. Larouche V., Correa J.A., Cassidy P., Beau-regard C., Garfield N. et al.: *Pituitary* 19, 202 (2016).
58. Nocturne G., Virone A., Ng W.F., Le Guern V., Hachulla E. et al.: *Arthritis Rheum.* 68, 977 (2016).
59. Brill J.B., Calvo R.Y., Wallace J.D., Lewis P.R., Bansal V. et al.: *J. Trauma Acute Care Surg.* 80, 625 (2016).
60. Cramer D.W., Vitonis A.F., Terry K.L., Welch W.R., Titus L.J.: *Epidemiology* 27, 334 (2016).
61. Asencio M.A., Huertas M., Carranza R., Tenias J.M., Celis J. et al.: *Int. Ophthalmol.* 36, 185 (2016).
62. Chen C.Y., Tsai J., Hsu T.Y., Lai W.Y., Chen W.K. et al.: *Medicine* 95, e3204 (2016).
63. Hedriana H.L., Wiesner S., Downs B.G., Pelletreau B., Shields L.E.: *Int. J. Gynaecol. Obstet.* 132, 137 (2016).
64. Wang G., Teng F., Chen Y., Liu Y., Li Y.: *J. Stroke Cerebrovasc Dis.* 25, 556 (2016).
65. Antonelli A., Ferri C., Ferrari S., Di Domenicoantonio A., Giuggioli D. et al.: *Rheumatology* 55, (2016).
66. Gupta K., Wang H., Amin S.B.: *J. Parenter. Enteral Nutr.* 40, 335 (2016).
67. Funaki T., Inoue E., Miyairi I.: *BMC Infect. Dis.* 16, 73 (2016).
68. Landis J.R., Koch G.G.: *Biometrics* 33, 159 (1977).
69. Hallgren K.A.: *Tutor. Quant. Methods Psychol.* 8, 23 (2012).
70. Mandrekar J.N.: *J. Thorac. Oncol.* 6, 6 (2011).

71. O'Connor S.R., Tully M.A., Ryan B., Bradley J.M., Baxter G.D. et al.: BMC Res. Notes. 8, 224 (2015).
72. Loney P.L., Stratford P.W.: Phys. Ther. 79, 384 (1999).
73. Edginton M., Enarson D., Zachariah R., Reid T., Satyanarayana S. et al.: Public Health Action 2, 21 (2012).
74. Strange K.: Am. J. Physiol. Cell Physiol. 295, C567 (2008).
75. Wislar J.S., Flanagan A., Fontanarosa P.B., DeAngelis C.D.: BMJ 25, 343 (2011).
76. Welker J.A., Mccue J.D.: J. Am. Med. Inform. Assoc. 14, 16 (2007).
77. Editorial. Ethics and fraud. Nature 439, 117 (2006).
78. Mulligan A., Taylor M., Newsum L.: <https://www.elsevier.com/authors-update/story/publishing-trends> (2014).
79. Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals. ICJME <http://www.icmje.org/icmje-recommendations.pdf> (2016).
80. Why share data? Information for authors. National institute for health research. <https://www.journalslibrary.nihr.ac.uk/information-for-authors/data-sharing/> (2017).
81. Availability of data, material and methods. Authors and referees policies. Nature. <http://www.nature.com/authors/policies/availability.html> (2017).
82. Tapia J.C., Ruiz E.F., Ponce O.J, Malaga G., Miranda J.: Colomb. Med. 46, 168 (2015).
83. Jeelani A., Malik W.R., Haq I., Aleem S., Mujtaba M. et al.: Ann. Med. Health Sci. Res. 4, 875 (2014).
84. Poorolajal J., Cheraghi Z., Irani A.D., Rezaeian S.: Epidemiol. Health 33, e2011005 (2011).
85. Ding G.: 21<sup>st</sup> Cochrane colloquium. Ottawa. P2.047 (2013).
86. Sorensen A.A., Wojahn R.D., Manske M.C., Calfee R.P.: J. Hand Surg. Am. 38, 1584 (2013).

*Received: 28. 11. 2017*