Compliance of Systematic Reviews in Plastic Surgery With the PRISMA Statement

Seon-Young Lee, BMedSc; Harkiran Sagoo, BSc(Hons); Katharine Whitehurst, BSc(Hons); Georgina Wellstead, BSc(Hons); Alexander J. Fowler, BSc(Hons), MBBS; Riaz A. Agha, BSc(Hons), MBBS, MSc(Oxf), MRCS_eng, FHEA, FRSPH; Dennis Orgill, MD, PhD

IMPORTANCE Systematic reviews attempt to answer research questions by synthesizing the data in primary articles. They are an increasingly important tool within evidence-based medicine, guiding clinical practice, future research, and health care policy.

OBJECTIVE To determine the reporting quality of recent systematic reviews and meta-analyses in plastic surgery with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement.

METHODS MEDLINE and EMBASE were searched for systematic reviews published between January 1, 2013, and December 31, 2014, in 5 major plastic surgery journals. Screening, identification, and data extraction were performed independently by 2 teams. Articles were reviewed for compliance with reporting of 27 items in the PRISMA checklist. Data analysis was conducted from January 1 to July 30, 2015.

MAIN OUTCOMES AND MEASURES The sum of PRISMA checklist items (1-27) per systematic review.

RESULTS From an initial set of 163 articles, 79 met the inclusion criteria. The median PRISMA score was 16 of 27 items (59%) (range, 6%-26%; 95% CI, 14%-17%). Compliance varied between individual PRISMA items. It was poorest for items related to the use of review protocol (item 5; 4 articles [5%]) and presentation of data on the risk of bias of each study (item 19; 14 articles [18%]). Compliance was the highest for description of rationale (item 3; 78 articles [99%]), sources of funding and other support (item 27; 75 articles [95%]), and inclusion of a structured summary in the abstract (item 2; 75 articles [95%]).

CONCLUSIONS AND RELEVANCE The reporting quality of systematic reviews in plastic surgery requires improvement. Enforcement of compliance through journal submission systems, as well as improved education, awareness, and a cohesive strategy among all stakeholders, is called for.

LEVEL OF EVIDENCE NA.
A systematic review should be an attempt to answer a clearly formulated research question by assessing and synthesizing the relevant and available research evidence. Systematic reviews have become increasingly important in health care by replacing traditional narrative reviews and expert commentaries in summarizing research evidence. They are dependent primarily on the quality of the studies they synthesize but can be more useful than primary studies alone because they may account for bias within individual studies by examining broader populations. Similarly, systematic reviews may quantify the heterogeneity in available data and allow for results from multiple studies to be pooled into a meta-analysis to increase power, and hopefully identify the correct treatment for a particular clinical paradigm.  

Physicians, researchers, and policy makers use systematic reviews and meta-analyses to keep up to date with their field, to inform decision making in clinical practice, and to prepare to conduct future studies. As with any study, the conduct, methodological design, and quality of reporting of a systematic review is paramount. To accurately appraise a study, readers need complete, clear, and transparent information. Several reporting guidelines have been established to improve the reporting quality of clinical research. Previous studies have shown the poor compliance of randomized clinical trials and observational studies in plastic surgery with the CONSORT (Consolidated Standards of Reporting Trials) and STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) statements, respectively.  

In 2009, the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement was published. It consists of a 27-item checklist (Table) covering each section of an article and a 4-phase flow diagram aiming to help authors improve the reporting of systematic reviews and meta-analyses. The aim of our study was to determine the reporting quality of systematic reviews and meta-analyses in plastic surgery with the PRISMA statement.

Methods

This study was conducted according to the recommendations outlined in the Cochrane Handbook for Systematic Reviews of Interventions, version 5.1.0, and reported in line with the PRISMA statement. This study is also registered with Research Registry (unique identifying number: reviewregistry18; http://www.researchregistry.com). Institutional review board approval was not required as the study did not involve human participants.

Search Methods

We searched MEDLINE and EMBASE for systematic reviews published between January 1, 2013, and December 31, 2014, in 5 major plastic surgery journals: Plastic and Reconstructive Surgery; Aesthetic Plastic Surgery; Archives of Facial Plastic Surgery; Journal of Plastic, Reconstructive & Aesthetic Surgery; and Annals of Plastic Surgery. These publications had the highest impact factor among plastic surgery journals according to the Thomson Reuter Impact Factor 2013. Although they are not representative of all plastic surgery publications, they do represent significant output for the field. The search strategies are shown in eAppendix 1 and eAppendix 2 in the Supplement. Data analysis was conducted from January 1 to July 30, 2015.

Three of us (S.-Y.L., H.S., and K.W.) then independently sorted the articles by abstract screening. The search was restricted from January 1, 2013, through December 31, 2014, and to systematic reviews and/or meta-analyses. Only nonsystematic reviews and non–meta-analyses were excluded. Any discrepancies in article selection were resolved by consensus first, and if still unsolved, referred to another member of the team (R.A.) for the final decision.

Scoring

Articles were scored independently by 2 of us (H.S. and K.W. or G.W.; articles were divided between K.W. and G.W., as was the workload) against the 27 checklist items in the PRISMA 2009 statement, with each item weighted equally. The resulting score was termed the PRISMA score. Items were scored only if the article had reported all the information detailed for the item, reflecting the latest PRISMA 2009 statement. However, at the same time, any efforts that fulfilled the requirement of one of the checklist items was accepted. For example, even if the article did not specifically mention the word bias, the article gained points if a description relating to the bias could be found. Any disagreements between the scorers were forwarded to another team member (S.-Y.L.), who made the final decision.

Statistical Analysis

Descriptive statistics were calculated, including the median and range for categorical variables; 95% CIs for the median were also calculated. The compliance of individual items with the PRISMA statement was also calculated. Interobserver correlation was determined using Cohen κ, a statistical measure of magnitude of agreement beyond that of chance alone. All statistical analyses were performed using Microsoft Excel 2011 (Microsoft Corporation).

Results

The electronic search identified an initial set of 163 articles, of which 79 met the inclusion criteria (Figure 1). A total of 1231 points were scored for all 79 articles together. Arbitration occurred regarding the inclusion of 22 articles. Disagreement occurred regarding 431 points across the 79 articles, and the Cohen κ statistic was 0.60, indicating substantial agreement between scorers.

The median PRISMA score was 16 of 27 items (59%) (range, 6%-26%; 95% CI, 14%-17%). Compliance with each item in the checklist is shown in the Table, while Figure 2 shows the number of articles according to their overall compliance with the PRISMA criteria.

Compliance varied between PRISMA items. It was poorest for items related to the use of review protocol (item 5; 4 articles [5%]) and presentation of data on the risk of bias of each study (item 19; 14 articles [18%]). Compliance was highest
Discussion

Our study examined 79 articles from 5 plastic surgery journals and found that none of the articles met all 27 criteria in the PRISMA 2009 statement. Overall, there was a relatively low median score of 16 of 27 items (59%), which is lower than in previous studies that examined compliance with the PRISMA checklist in other fields.3,10,11

Compliance varied highly between individual items on the checklist. Five items showed compliance of more than 90%, whereas 11 items had compliance rates below 40%. Items relating to the description of review protocol, description of additional analyses, and risk of bias showed the lowest compliance (Table).

Poor compliance in reporting of bias also has been noted in previous work examining the reporting quality of randomized clinical trials and observational studies.4-6 Moreover, a review of 300 systematic reviews identified that less than

Table. Compliance With PRISMA Checklist Items in 79 Published Systematic Reviews*

<table>
<thead>
<tr>
<th>Section and Topic</th>
<th>No.</th>
<th>Brief Description of the Item</th>
<th>Compliance, No. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td>1</td>
<td>Identification of the report</td>
<td>56 (71)</td>
</tr>
<tr>
<td>Abstract</td>
<td>2</td>
<td>Provide a structured summary</td>
<td>76 (96)</td>
</tr>
<tr>
<td>Introduction</td>
<td>3</td>
<td>Rationale objectives</td>
<td>78 (99)</td>
</tr>
<tr>
<td>Objectives</td>
<td>4</td>
<td>Description of PICOS (participants, interventions, comparisons, outcomes, and study design)</td>
<td>26 (33)</td>
</tr>
<tr>
<td>Methods</td>
<td>5</td>
<td>Indication of review protocol and registration information</td>
<td>4 (5)</td>
</tr>
<tr>
<td>Eligibility criteria</td>
<td>6</td>
<td>Specification of study and review characteristics as eligibility criteria</td>
<td>72 (91)</td>
</tr>
<tr>
<td>Information sources</td>
<td>7</td>
<td>Describe all information sources and date last searched</td>
<td>63 (80)</td>
</tr>
<tr>
<td>Search</td>
<td>8</td>
<td>Present repeatable full electronic search strategy for at least 1 database</td>
<td>40 (51)</td>
</tr>
<tr>
<td>Study selection</td>
<td>9</td>
<td>State the process for selecting studies</td>
<td>62 (78)</td>
</tr>
<tr>
<td>Data collection process</td>
<td>10</td>
<td>Describe method of data extraction</td>
<td>49 (62)</td>
</tr>
<tr>
<td>Data items</td>
<td>11</td>
<td>Report all variables and any assumptions and simplifications made</td>
<td>64 (81)</td>
</tr>
<tr>
<td>Risk of bias in individual studies</td>
<td>12</td>
<td>Describe methods used for assessing risk of bias of individual studies</td>
<td>29 (37)</td>
</tr>
<tr>
<td>Summary measures</td>
<td>13</td>
<td>State the principal summary measures</td>
<td>41 (52)</td>
</tr>
<tr>
<td>Synthesis of results</td>
<td>14</td>
<td>Describe the methods of handling and analyzing data</td>
<td>30 (38)</td>
</tr>
<tr>
<td>Risk of bias across studies</td>
<td>15</td>
<td>Specify any assessment of risk of bias that may affect the cumulative evidence</td>
<td>20 (25)</td>
</tr>
<tr>
<td>Additional analyses</td>
<td>16</td>
<td>Describe methods of additional analyses</td>
<td>20 (25)</td>
</tr>
<tr>
<td>Results</td>
<td>17</td>
<td>Give numbers of studies at each stage of the study</td>
<td>52 (66)</td>
</tr>
<tr>
<td>Study characteristics</td>
<td>18</td>
<td>For each study, present characteristics for which data were extracted</td>
<td>65 (82)</td>
</tr>
<tr>
<td>Risk of bias within studies</td>
<td>19</td>
<td>Present data on risk of bias of each study</td>
<td>14 (18)</td>
</tr>
<tr>
<td>Results of individual studies</td>
<td>20</td>
<td>Report the summary of each data intervention group and estimates of confidence intervals</td>
<td>29 (37)</td>
</tr>
<tr>
<td>Synthesis of results</td>
<td>21</td>
<td>Present results of each meta-analysis</td>
<td>26 (33)</td>
</tr>
<tr>
<td>Risk of bias across studies</td>
<td>22</td>
<td>Present results of any assessment of risk of bias across studies</td>
<td>19 (24)</td>
</tr>
<tr>
<td>Additional analysis</td>
<td>23</td>
<td>Give results of additional analysis</td>
<td>23 (29)</td>
</tr>
<tr>
<td>Discussion</td>
<td>24</td>
<td>Summarize the main findings, including the strength of evidence</td>
<td>70 (89)</td>
</tr>
<tr>
<td>Limitations</td>
<td>25</td>
<td>Discuss limitations at study, outcome and review level</td>
<td>53 (67)</td>
</tr>
<tr>
<td>Conclusions</td>
<td>26</td>
<td>Provide a general interpretation of the results</td>
<td>75 (95)</td>
</tr>
<tr>
<td>Funding</td>
<td>27</td>
<td>Describe sources of funding and other support</td>
<td>75 (95)</td>
</tr>
</tbody>
</table>

Abbreviation: PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.
* Adapted with permission from Moher et al.3

for description of rationale (item 3; 78 articles [99%]), sources of funding and other support (item 27; 75 articles [95%]), and inclusion of a structured summary in the abstract (item 2; 75 articles [95%]).
one-fourth (23%) of authors reported assessing possible publication bias.\textsuperscript{11} This finding does not suggest poor design or conduct of the study itself; the PRISMA statement was not intentionally developed to improve the conduct of systematic reviews and meta-analyses.\textsuperscript{12} The scores were given based solely on the items reported in the article, so even if the study was thoroughly designed and conducted to avoid bias, points were not awarded if these efforts were not described.

In addition, most of the articles involved in our study lost points by lacking the additional details in the description relevant to each item. For example, many articles were not awarded points for item 4 because they failed to refer to PICOS (participants, interventions, comparisons, outcomes, and study design), but most articles stated the aim of the study. This finding is very important because the purpose of publication is to provide clear, transparent information in which all aspects of the work are available for interpretation, and our study shows that many articles have been failing to do so.

Another concerning finding is the relatively low compliance with study selection (item 17). Fifty-two articles (66%) addressed the numbers of studies involved at screening, assessing for eligibility, and included in the review. Although this value is relatively high compared with those of other items, it is low considering the ubiquity of PRISMA criteria.\textsuperscript{3}

Poor reporting in systematic reviews and meta-analyses prevents robust critical appraisal in these fields; clinical judgments resulting from such articles could be misleading and potentially dangerous. There have been increasing calls for an evidence-based approach across all of medicine, and specifically within plastic surgery. It is hoped that the universal implementation of guidelines such as PRISMA will address suboptimal reporting by researchers.\textsuperscript{3,6} Although PRISMA focuses on the effective reporting of meta-analysis of randomized trials, it can be used for all other types of systematic reviews. The use of PRISMA as a guideline can benefit researchers not only by ensuring they will report the basic research components but also by helping them develop research questions as well as the structured summary and rationale of the study.\textsuperscript{3}

The low level of compliance with PRISMA criteria has been recognized across other specialties as well. Gagnier and Kellam\textsuperscript{10} assessed compliance with the PRISMA criteria of 76 systematic reviews and meta-analyses from 5 orthopedic journals with the highest impact factor. The study showed the average PRISMA compliance was 68%; major deficiencies identified included the description of risk of bias across studies (9%), the analysis of risk of bias in the Results section (37%), and the reporting of a protocol and registration (18%). Fleming et al\textsuperscript{14} assessed compliance with PRISMA criteria of 109 systematic reviews published in 5 orthodontic journals between January 2000 and July 2011, and demonstrated an overall compliance of 64%. Key deficiencies included description of risk of bias across the studies (3%), methods used for assessing the risk of bias (26%), and the information of review protocol and registration number (27%). These items overlap with the key deficiencies identified in our study as well as with the deficiencies identified with other types of study design, such as randomized clinical trials and observational studies.\textsuperscript{4,6}

The reporting of a protocol and registration information (item 5) showed very low compliance in our study (5%), perhaps owing to the fact that registration of systematic reviews is not yet widely recognized; however, it is still important, as providing the protocol creates better communication between the authors and the readers and registration of the systematic review helps to improve transparency of the review process.\textsuperscript{3,15} We suggest that the use of a web database registration program for protocols, such as PROSPERO\textsuperscript{15} or Research Registry,\textsuperscript{16} has good potential to improve this issue.

Enforcing the guidelines at the time of article submission to a journal may lead to an improvement in the quality of reporting; endorsement of the PRISMA checklist by journals has been shown to improve compliance.\textsuperscript{22} It has recently been demonstrated that there has been an increase between 2011 and...
in the quality of reporting in systematic reviews over time. We suggest the use of the PRISMA checklist as a mandatory item during the electronic submission of systematic reviews and meta-analyses. The PRISMA checklist should go forward into the peer review process along with the manuscript and be open to scrutiny by peer reviewers. The checklist can be published as a supplementary item online, allowing for greater transparency and scrutiny by readers.

Further study is required to assess any other potential barriers to compliance with the PRISMA checklist among key stakeholders in plastic surgery: authors, journal reviewers and editors, funders, institutions, and readers. We support previous statements calling for better education of plastic surgeons at all levels in evidence-based medicine.

This study has several limitations. Using the results from 5 journals means that this study does not represent the full breadth of literature related to plastic surgery. However, the use of journals with the highest impact factor is expected to represent a broad cross-section of plastic surgery research. Also, by restricting the search to articles published in 2013 and 2014, our review is cross-sectional in nature, which provides a snapshot of the literature, preventing us from commenting on trends in the quality of reporting in systematic reviews over time. However, the significance of this study is in its evaluation of the current state of the quality of reporting of systematic reviews in plastic surgery. Restricting the search to the MEDLINE and EMBASE databases means that we may have missed some studies if they were not listed in our electronic searches, although this possibility is unlikely. We say this because, as of March 2015, MEDLINE indexes more than 5000 journals and more than 20 million citations, and EMBASE includes more than 29 million indexed records; therefore, we hope that these databases covered a sufficient breadth of the plastic surgery literature. Furthermore, by using 2 independent data extractors and a third to make decisions on scoring judgment, we hope to have reduced the bias and accurately scored all articles meeting the inclusion criteria.

Conclusions

The reporting quality of systematic reviews and meta-analyses in plastic surgery needs improvement. Enforcement of compliance through journal submission systems, as well as improved education, awareness, and a cohesive strategy among all stakeholders, is called for. Further study assessing barriers to compliance with the PRISMA statement is required.

ARTICLE INFORMATION
Accepted for Publication: September 23, 2015.
Published Online: December 17, 2015.

Author Contributions: Mr Lee had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Study concept and design: Lee, Fowler, Agha, Orgill. Acquisition, analysis, or interpretation of data: Lee, Sagoo, Whitehurst, Wellstead, Agha. Drafting of the manuscript: Lee, Whitehurst, Wellstead, Fowler, Agha. Critical revision of the manuscript for important intellectual content: Lee, Sagoo, Whitehurst, Fowler, Agha, Orgill. Statistical analysis: Lee, Whitehurst, Wellstead, Fowler. Administrative, technical, or material support: Lee, Whitehurst, Wellstead, Agha, Orgill. Study supervision: Fowler, Agha, Orgill.

Conflict of Interest Disclosures: Mr Agha is on the Editorial Board of the International Journal of Surgery. No other disclosures were reported.

REFERENCES