Publication bias in abstracts presented to the annual meeting of the American Academy of Orthopaedic Surgeons

IA Harris, M Mourad, A Kadir
Liverpool Hospital, Liverpool, New South Wales, Australia

MJ Solomon, JM Young
Surgical Outcome Research Centre, Camperdown, New South Wales, Australia

ABSTRACT

Purpose. To examine possible causes of publication bias in the orthopaedic literature so as to avoid inappropriate clinical decisions based on reviews of the literature.

Methods. Two reviewers independently reviewed abstracts presented to the 1999 American Academy of Orthopaedic Surgeons annual meeting. Data pertaining to sample size, statistical significance, study setting, country of origin, outcome, study type, and sponsorship were extracted from each abstract. The publication rate was measured after 5 years, by electronic searching and author contact. Predictors of publication were identified using logistic regression analysis.

Results. Of the 318 abstracts listed in the proceedings, 175 (55%) were published within 5 years. Publication was associated with positive rather than neutral outcomes (odds ratio, 1.62; 95% confidence interval, 1.01–2.59; p<0.05) and with the reporting of statistical significance (odds ratio, 2.05; 95% confidence interval 1.24–3.39; p=0.005). Sponsorship, country of origin, sample size, study setting, and study type did not significantly influence the publication rate.

Conclusion. Evidence of publication bias exists in abstracts presented at the 1999 American Academy of Orthopaedic Surgeons annual meeting. Clinical decisions based on the literature may be biased due to an over-representation of studies with positive outcomes.

Key words: orthopedics; publication bias

INTRODUCTION

Publication bias refers to the greater likelihood that studies with positive results are published. Dickersin defines publication bias as “the tendency on the parts of investigators, reviewers, and editors to submit or accept manuscripts for publication based on the direction or strength of the study findings”. Another broader definition describes it as the failure of some studies to be published, for reasons related to the findings. This definition allows for other factors (e.g., sponsorship, sample size, study type, quality, or country of origin) that may influence the publication rate, apart from the clinical or statistical significance of the findings. This paper uses the broader definition...
and considers publication bias to be any form of selective publication related to study factors that may result in an imbalance in the representation of papers in the literature. Therefore, multiple explanatory variables are examined.

Publication bias that relates to the direction of the findings is known as ‘positive outcome bias’. Definitions of positive outcome, however, vary in the literature. A positive outcome has been defined as (1) improved outcome with the study intervention, (2) similar outcome with new treatment compared to established treatments, (3) one in which the objective is supported, (4) one in which statistical significance is present, or (5) combinations of these. Publication bias based on the strength of the findings is easier to measure, as the definition usually relies on the presence of statistical significance. The consequence of publication bias due to an over-representation of positive studies may affect clinical decisions based on reviews of the (biased) literature. The possibility of this bias influencing summary effect estimates in meta-analyses is often stated. This bias may influence conclusions in clinical guidelines or individual physician decisions. There is evidence of publication bias due to reasons other than the strength or direction of the outcome. Industry sponsorship, for example, has been shown to influence publication rates.

Many studies have looked at publication rates of abstracts to examine publication bias. The abstracts are usually selected from major scientific meetings but may also be selected from ethics committee applications and electronic sources. While many of these studies have shown evidence of publication bias based on positive outcome or statistical significance, others have not. Other factors associated with higher publication rates are sample size, oral versus poster presentation, study quality, and prospective versus retrospective studies.

This study aimed to measure the publication rate for papers presented at the 66th annual scientific meeting of the American Academy of Orthopaedic Surgeons (AAOS) in 1999 to examine, through multiple logistic regression analysis, possible causes of publication bias. The null hypothesis states that the publication rate is not significantly predicted by any of the explanatory variables.

MATERIALS AND METHODS

Abstracts of papers presented at the 66th annual scientific meeting of the AAOS held in Anaheim, California in February 1999 were reviewed. Data on positive outcome, statistical significance, presence of a p value, sponsorship, sample size, study setting (laboratory versus clinical), country of origin, and study type were extracted by 2 of the authors independently and a consensus was reached by discussion.

Positive outcome was defined as present when the authors stated that the intervention was of some benefit, regardless of statistical significance. Negative outcome was defined as present when the authors advised against the intervention (e.g., the intervention was described as not useful or not recommended). A neutral study was one in which no opinion regarding the effectiveness was provided, or the outcomes of the study groups were considered the same.

A study was considered significant if the authors stated that the results were statistically significant or if the p value of the appropriate statistical test was <0.05. A p value (regardless of the value) was considered present if a p value was provided in the abstract. Whether the p value was <0.05 or ≥0.05 was also recorded.

Studies were classified according to the (1) type of study (randomised controlled trial, prospective cohort, case-control, case series, or other) and (2) study setting (laboratory or clinical). Laboratory studies included animal and cadaver studies. Country of origin was defined as the country in which the study took place. If multiple countries were involved the country of origin of the senior author was used. Sponsorship was considered present if there was any type of declared sponsorship for any author.

The publication rate was determined by searching MEDLINE and EMBASE, and by contacting the authors. The electronic searching took place during March 2004 using authors’ names and, where necessary, key words from the title. Authors of studies not found on the electronic databases were contacted by e-mail or facsimile during April and May 2004. Repeated e-mails or facsimiles were sent one month later if the authors had not responded. No further attempts were made to contact authors.

The follow-up period was over 5 years from the date of abstract presentation, as over 90% of publications occur within that period. The primary outcome was publication in a journal. Listing in a supplement (as an abstract from a conference proceeding) was regarded as unpublished.

Logistic regression was performed with the SAS statistical software (Cary [NC], US) using backward stepwise regression. Published/unpublished was the dichotomous outcome. The significance level was set at 0.05.
RESULTS

Of the 318 abstracts listed in the proceedings of the 66th annual scientific meeting of the AAOS, 175 (55%) were published within 5 years. Electronic searching by MEDLINE and EMBASE found all 175 published articles. The 143 authors of the unpublished papers were contacted by e-mail or facsimile; 87 (61%) of whom responded; none stated that their paper had been published.

Univariate analysis showed reporting of significance, presence of a p value, and positive outcome to be associated with publication (Table). Although the publication rate was higher in abstracts that reported p values of <0.05 as opposed to ≥0.05 (odds ratio [OR], 1.83; 95% confidence interval [CI], 0.52–6.45), this difference was not statistically significant (p=0.35). However, the numbers were small—only 110 (35%) abstracts included a p value, and of these, only 11 (10%) were reported as ≥0.05.

Papers reporting a positive outcome had a significantly higher publication rate, compared to those describing a neutral result (OR, 2.02; 95% CI, 1.20–3.41; p=0.008). The publication rate was not significantly different for papers reporting a negative as opposed to a neutral (or no) result (OR, 1.48; 95% CI, 0.66–3.31; p=0.34). A post hoc analysis comparing studies with a positive result to those with a neutral or negative result revealed that the publication rate remained significantly higher in studies with a positive result (OR, 1.80; 95% CI, 1.13–2.86; p=0.013).

Country of origin, study type, study setting, and sponsorship were not significantly associated with publication. Sample size was only recorded in 289 studies and its distribution was positively skewed. A logarithmic transformation was performed to provide normally distributed data. The new variable, logN, was tested as a continuous variable. Univariate analysis showed no significant correlation between sample size and publication (t statistic=−1.44, p=0.15). Backward regression model did not show any significant association. Sample size was not included

<table>
<thead>
<tr>
<th>Variables</th>
<th>No.</th>
<th>Published (%)</th>
<th>Odds ratio</th>
<th>95% CI</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country of origin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North America</td>
<td>283</td>
<td>54.1</td>
<td>1.00*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Europe</td>
<td>20</td>
<td>70.0</td>
<td>1.98</td>
<td>0.74–5.31</td>
<td>0.17</td>
</tr>
<tr>
<td>Other</td>
<td>15</td>
<td>60.0</td>
<td>1.28</td>
<td>0.44–3.68</td>
<td>0.65</td>
</tr>
<tr>
<td>Significance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>217</td>
<td>49.3</td>
<td>1.00*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>101</td>
<td>68.3</td>
<td>2.22</td>
<td>1.35–3.64</td>
<td>0.002</td>
</tr>
<tr>
<td>P value</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not recorded</td>
<td>208</td>
<td>49.0</td>
<td>1.00*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recorded</td>
<td>110</td>
<td>67.3</td>
<td>2.14</td>
<td>1.32–3.46</td>
<td>0.002</td>
</tr>
<tr>
<td>P value (when recorded)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥0.05</td>
<td>11</td>
<td>54.5</td>
<td>1.00*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;0.05</td>
<td>99</td>
<td>68.7</td>
<td>1.83</td>
<td>0.52–6.45</td>
<td>0.35</td>
</tr>
<tr>
<td>Study type</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Randomised controlled trial</td>
<td>18</td>
<td>61.1</td>
<td>1.00*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cohort</td>
<td>77</td>
<td>58.4</td>
<td>0.90</td>
<td>0.31–2.56</td>
<td>0.84</td>
</tr>
<tr>
<td>Case series</td>
<td>184</td>
<td>51.6</td>
<td>0.68</td>
<td>0.25–1.83</td>
<td>0.44</td>
</tr>
<tr>
<td>Other</td>
<td>39</td>
<td>64.1</td>
<td>1.14</td>
<td>0.36–3.59</td>
<td>0.83</td>
</tr>
<tr>
<td>Result</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neutral</td>
<td>81</td>
<td>43.2</td>
<td>1.00*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>203</td>
<td>60.6</td>
<td>2.02</td>
<td>1.20–3.41</td>
<td>0.008</td>
</tr>
<tr>
<td>Negative</td>
<td>34</td>
<td>52.9</td>
<td>1.48</td>
<td>0.66–3.31</td>
<td>0.34</td>
</tr>
<tr>
<td>Study setting</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinical</td>
<td>244</td>
<td>54.5</td>
<td>1.00*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laboratory</td>
<td>74</td>
<td>58.1</td>
<td>1.16</td>
<td>0.68–1.96</td>
<td>0.59</td>
</tr>
<tr>
<td>Sponsorship</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>228</td>
<td>55.7</td>
<td>1.00*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Present</td>
<td>90</td>
<td>54.4</td>
<td>0.95</td>
<td>0.58–1.55</td>
<td>0.84</td>
</tr>
</tbody>
</table>

* Reference group
in the final regression model because significance was lacking and the value was missing in 29 papers, limiting the power of the regression equation.

Presence of a \( p \) value (regardless of its actual value) was also excluded from the final analysis. It was highly associated with the variable significance, i.e. significance and presence of a \( p \) value were measuring the same property, because significance was considered present if the words statistically significant or a \( p \) value of <0.05 was provided. Most articles reporting statistical significance reported a \( p \) value of <0.05. Similarly, 90% of the abstracts stating a \( p \) value reported a value of <0.05. The 2 variables (significance and presence of a \( p \) value) showed 95.9% agreement (kappa=0.91) and could not be included in the same multiple logistic regression analysis.

Multiple logistic regression analysis confirmed reporting of statistical significance and a positive outcome (compared to a neutral [or no] outcome) were the only 2 significant factors predicting publication. Reporting of statistical significance was the strongest predictor of publication (OR, 2.05; 95% CI, 1.24–3.39; \( p = 0.005 \)), followed by reporting of a positive outcome (as opposed to a neutral [or no] outcome) [OR, 1.62; 95% CI, 1.01–2.59; \( p = 0.046 \)]. Substituting presence of a \( p \) value for significance revealed similar results (OR, 2.00; 95% CI, 1.23–3.26, \( p = 0.005 \)). The effect of a positive outcome remained unchanged. We used reporting of significance instead of presence of a \( p \) value in the final model, as it has been used in studies on publication bias and was considered a better measure of the strength of the findings.

**DISCUSSION**

The overall publication rate for the AAOS meeting compares favourably to similar studies of medical specialty meetings which report publication rates from 25 to 60%, \(^{3,8,11,21,22,27-33}\)

Sponsorship, which was previously shown to predict publication, \(^ {14,20,23,34} \) was not associated with an increased publication rate in this study. However, this may reflect the broad definition of sponsorship used in this study.

The increased publication rate associated with positive outcomes is clear evidence of publication bias, but its association with reporting of significance may simply reflect study quality, in which those that had provided a statistical analysis, and those that were appropriately powered to find significance, were more likely to be published. If this were the case, any summary of the literature would be based on better quality studies, not just favourable studies.

Papers reporting a neutral outcome may be more likely to be underpowered, and the lower publication rate may be due in part to this methodological deficiency rather than the lack of a positive finding. Also findings in published papers may differ considerably (including the direction, strength, and significance of the findings) from the findings reported in the abstracts from which they arose.\(^ {30,36} \)

We suggest caution before accepting recommendations based on summaries of the orthopaedic literature. Although trial registers have been proposed as a safeguard against publication bias (through registration of all trials at inception and reporting of all results regardless of the outcome), this may not be sufficient in orthopaedics, as most of the abstracts, and later publications, are from case series and not from controlled trials.

Contact with the 143 authors of unpublished papers (87 responded) found no further publications beyond those found in MEDLINE and EMBASE. This is in contrast to previous studies reporting that author contact was responsible for identifying 9%,\(^3\) 11%,\(^9\) 24%,\(^{15}\) and 26%\(^{31}\) of published articles. The low retrieval rate from author contact may be because publications from this internationally recognised meeting are more likely to appear in journals listed in electronic databases. Without proof of this, the variation in the response rate from author contact (compared to previous studies) precludes a recommendation on the usefulness of author contact.

Our author response rate of 61% is comparable to the reported rates of 54 to 62%,\(^ {11,15,19} \) Timmer et al.\(^ {11} \) contacted all authors regardless of whether their abstracts were subsequently published. They showed a significant respondent bias, as responding authors were more likely to have had a publication. Authors who do not respond are less likely to have had their paper published, thus minimising the possible error due to authors not responding.

Further information regarding publication bias in the orthopaedic literature may be obtained from an analysis of articles submitted to these meetings (rather than only those presented), and from an analysis of papers submitted to major orthopaedic journals. There may be bias in the selection of abstracts for presentation, as most (64%) presented papers reported a positive outcome. Acceptance of abstracts submitted to scientific meetings has been shown to be more likely if the outcome is positive.\(^ {3,22} \)

Editorial bias (biased selection of papers submitted to a journal) has been examined in the *Journal of the American Medical Association*,\(^4 \) but not found to be significant.
REFERENCES

16. Loannidis JP. Effect of the statistical significance of results on the time to completion and publication of randomized efficacy trials. JAMA 1999;279:281–6.