Predictive value of Altmetric score on citation rates and bibliometric impact

D. B. T. Robinson 1, A. G. M. T. Powell 1,2*, J. Waterman 1, L. Hopkins 1, O. P. James 1, R. J. Egan 3,4 and W. G. Lewis 1

1 Health Education and Improvement Wales’ School of Surgery, Ty Dgyu, Nantgarw, UK
2 Division of Cancer and Genetics, Cardiff University, Heath Park, Cardiff, UK
3 Morriston Hospital, Cwmymdeirn, Swansea, UK
4 Swansea University, Singleton Park, Sketty, Swansea, UK

*Correspondence to: Division of Cancer and Genetics, Cardiff University, Heath Park, Cardiff CF14 4XW, UK (e-mail: powella16@cardiff.ac.uk)

Abstract

Background: Bibliometric and Altmetric analyses provide different perspectives regarding research impact. This study aimed to determine whether Altmetric score was associated with citation rate independent of established bibliometrics.

Methods: Citations related to a previous cohort of 100 most cited articles in surgery were collected and a 3-year interval citation gain calculated. Citation count, citation rate index, Altmetric score, 5-year impact factor, and Oxford Centre for Evidence-Based Medicine levels were used to estimate citation rate prospect.

Results: The median interval citation gain was 161 (i.q.r. 83–281); 74 and 62 articles had an increase in citation rate index (median increase 2.8 (i.q.r. –0.1 to 7.7)) and Altmetric score (median increase 3 (0–4)) respectively. Receiver operating characteristic (ROC) curve analysis revealed that citation rate index (area under the curve (AUC) 0.86, 95 per cent c.i. 0.79 to 0.93; P < 0.001) and Altmetric score (AUC 0.65, 0.55 to 0.76; P = 0.008) were associated with higher interval citation gain. An Altmetric score critical threshold of 2 or more was associated with a better interval citation gain when dichotomized at the interval citation gain median (odds ratio (OR) 4.94, 95 per cent c.i. 1.99 to 12.26; P = 0.001) or upper quartile (OR 4.13, 1.60 to 10.66; P = 0.003). Multivariable analysis revealed only citation rate index to be independently associated with interval citation gain when dichotomized at the median (OR 18.22, 6.70 to 49.55; P < 0.001) or upper quartile (OR 19.30, 4.23 to 88.15; P < 0.001).

Conclusion: Citation rate index and Altmetric score appear to be important predictors of interval citation gain, and better at predicting future citations than the historical and established impact factor and Oxford Centre for Evidence-Based Medicine quality descriptors.

Introduction

Academic reach and impact are sought after by author and journal editorial boards. Currently, the number of citations achieved by any given article is the most commonly used measure of academic reach, and the foundation of author Hirsch index (h-index) and journal impact factor (IF). In many countries, higher education institutions allocate funding according to academic outputs associated with journals held in high esteem on the basis of these traditional measures 1,2.

Citations, however, take time to accumulate and other faster means of assessment have risen: alternative metrics, or Altmetrics. Altmetrics extend the concept of citation beyond mention in other scholarly articles by recording how often an article is downloaded, and applies to people, journals, books, data sets, presentations, videos, source code repositories, and web pages. The Altmetric score (AS) calculates impact based on diverse online research outputs, including social media, blogs, online news media, and online reference managers 3. Consequently, the concept of impact remains as important and relevant a philosophical question as ever 3,4.

Articles with a higher AS have been reported to be associated with higher citation counts 5, but whether this symbolizes preliminary increased curiosity or sustained citation growth is unknown. The aim of this study was to identify the factors associated with citation accumulation, using a historical control cohort of the 100 most cited articles in surgery (2018) 3. The hypothesis was that AS is directly related to prospective citations and an indicator of citation trajectory.

Methods

Citation counts related to a previously reported cohort of 100 all-time most cited articles from a search performed in 2017 associated with the keyword ‘surgery’ were reanalysed 3. Thomas Reuters’ Web of Science citation index was accessed again, providing a 3-year interval citation gain (ICG). Previously collected bibliometric data, including AS, citation count, citation rate index (CRI; total number of citations divided by number of years since article was published) 6, 5-year IF, and Oxford Centre for Evidence-Based Medicine quality descriptors.
Evidence-Based Medicine (OCEBM)\(^7\) quality of evidence levels were used to classify citation rate prediction.

**Altmetric scores**

An AS is a weighted count of all of the mentions that Altmetric has tracked for an individual research output, and is designed to be an indicator of the degree of notice or attention an item has received. Scores are accumulated in an adjusted fashion to reflect the relative influence of different sources. For example, getting newspaper attention is more difficult than Twitter attention from a single individual; consequently, newspaper articles are more likely to receive attention than tweets and are thus allocated a higher contribution to a research output’s score. AS values are always whole numbers and, importantly, only the first mention from a source contributes to its accumulation. Some sources of attention will contribute less than 1 whole point to an output’s AS; in such instances, the AS is rounded up to the next whole number. The weighted scoring system can be found on the Altmetric\(^8\) website.

**Web of Science category and quartile rank status**

A journal’s IF category rank was recorded from Web of Science. A percentile rank (\(z\)) was calculated by dividing the rank (\(y\)) by the number of journals in a subject field (\(y\)), where \(z = y/x\)\(^9\). The percentile rank is categorized into four quartiles (Q1–4), with the most prestigious journals in the top quartile. Q1 journals are those belonging to the top 25 per cent, and Q4 are those within the bottom 25 per cent in a given subject field.

**Statistical analysis**

All data were expressed as median (i.q.r.) and non-parametric statistical methods used throughout. The OCEBM scoring system for quality of evidence was reported in a Likert-scale format (level 1, highest level of evidence—systematic review or meta-analysis of RCTs; level 5, lowest level of evidence—expert opinion without critical appraisal). The Spearman correlation coefficient test was used to test the relationship between quality of evidence and iCG. The iCG was clustered into deciles to meet the test assumptions. Receiver operating characteristic (ROC) curve analysis was used to assessed the predictive value of continuous variables, with the primary outcome measure iCG dichotomized at the median and top quartile.\(^10\) Univariable and multivariable logistic regression models were developed to identify independent associations with the primary outcome measures. Variables with \(P < 0.100\) in univariable analysis were included in the multivariable analysis, using a backward conditional model. All statistical analysis was performed using SPSS\(^{15}\) version 25.0 (IBM, Armonk, New York, USA).

**Results**

The total cumulative citations and iCG for all articles to date were 100 127 and 28 990 respectively. Ninety-eight articles were published in Q1 journals, one in a Q2 and one in a Q3 journal; there were no articles published in Q4 journals. Every article received more citations, with a median iCG of 161 (i.q.r. 83–281). Seventy-four articles had an increase in CRI (median gain 2.8 (–0.1 to 7.7)), it remained constant for one, and decreased for 25. AS increased for 62 articles (median gain 3 (0–4)), did not change for 36, and decreased for two. The top 10 articles related to the highest iCG can be found in Table 1\(^{11–20}\), and the full article list is available in Table S1. Of the original 10 most cited articles in 2017, eight maintained their top ten status 3 years later in terms of total citations.

**Bibliometric factors associated with interval citation gain**

Median article iCG when first or last authors featured more than once was 240 (i.q.r. 128–701.5), compared with 153.5 (75.5–257; \(P = 0.011\)) when authors featured only once in the 100 most cited articles. Multicentre studies were not associated with higher iCGs than single-centre studies (median 183 (98–336) versus 143 (80–257); \(P = 0.209\)). Median iCG in relation to the country of the leading institute, OCEBM levels, research subject, disease category, and specialty are shown in Figs S1–S5. The highest and lowest iCG related to these variables were: country—median iCG 1084 (251–5002) for Switzerland versus 57 (42–134) for Japan; OCEBM evidence—291 (98–456) for level 5 versus 110 (83–135) for level 3; subject—456 (291–719) for epidemiology versus 101 (101–101) for nutrition; disease category—222 (166–321) for musculoskeletal versus 54 (54–54) for respiratory, and surgical specialty—235 (74–977) for vascular surgery versus 48 (48–48) for plastic surgery.

The relationship between original AS, number of citations, CRI, 5-year journal IF, and iCG is shown in Fig. 1. Linear regression analysis showed that original citation count (\(R^2 = 0.797\), 95 per cent c.i. 0.90 to 1.10, \(P < 0.001\)) and CRI (\(R^2 = 0.908\), 95 per cent c.i. 12.47 to 14.17, \(P < 0.001\)) were associated with iCG. Original AS (\(R^2 = 0.024\), 95 per cent c.i. –3.40 to 28.66, \(P = 0.121\)), 5-year journal IF (\(R^2 = 0.011\), 95 per cent c.i. –25.82 to 80.24, \(P = 0.311\)), and OCEBM level (\(R^2 = 0.002\), 95 per cent c.i. –124.50 to 78.45, \(P = 0.653\)) were not associated with iCG. When the two outliers were removed from the original AS analysis, AS (\(R^2 = 0.199\), 95 per cent c.i. 0.01 to 0.01; \(P < 0.001\)) was also associated with iCG (Fig. 1b).

**Creation of dichotomization thresholds for confounding adjustment**

ROC curve analysis was used to identify possible critical thresholds for variables associated with higher iCG (Fig. 2). Table 2 outlines the critical thresholds associated with higher iCG. No definitive critical thresholds were identified for 5-year IF (Fig. 2d) or OCEBM levels (Fig. 2e). The critical thresholds for original AS, citation count, and CRI were 1.5, 520.5, and 34.4 respectively. Articles with an AS of at least 2 (median iCG 209 (i.q.r. 143–393) versus 133 (71–234.5), \(P < 0.001\)), citation count 521 or more (192 (116–300) versus 120 (70–173), \(P = 0.004\)), and CRI at least 34.4 (250 (163–383) versus 95 (61–157), \(P < 0.001\)) were associated with higher median iCGs than those with values below these thresholds.

**Univariable and multivariable analysis of factors associated with above median and upper quartile of interval citation gain**

Univariable and multivariable analyses of factors associated with iCG are shown in Table 3. In univariable analysis, multicentre studies (\(P = 0.023\)), articles with first or last author featured more than once in the 100 most cited articles (\(P = 0.044\)), an original AS of at least 2 (\(P = 0.001\)), original citation count 521 or more (\(P = 0.008\)), or original CRI at least 34.4 (\(P < 0.001\)) were associated with higher iCG dichotomized around the median. When dichotomized around the upper quartile value, articles where the first or last author featured more than once (\(P = 0.041\)), those with an original AS of at least 2 (\(P = 0.003\)), original citation count 521 or more (\(P = 0.048\)), or original CRI at least 34.4 (\(P < 0.001\)) were associated with higher iCG. In multivariable analysis, only CRI was associated with higher citation accrual when iCG was dichotomized.
around the median (odds ratio (OR) 18.22, 95 per cent c.i. 6.70 to 49.55; P < 0.001) or upper quartile (OR 19.30, 4.23 to 88.15; P < 0.001).

Discussion

The principal finding of this study was that articles that amass citations at faster rates continue to grow their citations in an autonomous fashion that reflects the type of study, quality of evidence, affiliation of author or academic productivity. Articles with a CRI of at least 34.4 were 18 times more likely to accrue more interval citations, irrespective of academic affiliation. Certain authors, or teams of authors, appeared to be associated with a positive multiplication coefficient or constant which, when allied to a novel project, good fit, and timeliness, accrued. Certain authors, or teams of authors, appeared to be associated with a positive multiplication coefficient or constant which, when allied to a novel project, good fit, and timeliness, accrued. Certain authors, or teams of authors, appeared to be associated with a positive multiplication coefficient or constant which, when allied to a novel project, good fit, and timeliness, accrued. Certain authors, or teams of authors, appeared to be associated with a positive multiplication coefficient or constant which, when allied to a novel project, good fit, and timeliness, accrued.

Wang and co-workers\(^\text{33}\) reported an arithmetically derived mechanical model which showed that ultimate article impact, characterized by total lifetime citations, was largely predicted by a single factor described as fitness. Fitness is the perceived novelty and relevance that an article attracts from the research community at large. This may explain why the Clavien–Dindo classification article\(^\text{12}\), which received the greatest number of citations and had the highest iCG in this study, was one of the highest ranking articles in bibliometric analyses of surgery\(^\text{35}\), visceral surgery\(^\text{36}\), and laparoscopic surgery\(^\text{37}\). When Wang et al\(^\text{38}\) compared articles from journals with an impact factor of 3.26, 10.48, and 33.62, articles with similar fitness received initial citations in a journal-dependent manner, but ultimately received similar lifetime citation counts, which may help to clarify why journal IF did not predict citation count in the original 2018
study, or iCG in the present study. The findings described by Wang and colleagues justify how proximity of citation gain, which augments journal IF, is supported by social media platforms. Whether the impact of social media simply shifts mean citation count or boosts total citation count and impact is unclear.

With increasing competition for research funding, emphasis has migrated towards proving impact using bibliometric profiles and, historically, article citation count has always been synonymous with impact. Bibliometric use is supported by the present findings; in particular, CRI derived from citation count was associated independently with iCG. Academic reach is now facilitated by electronic devices, and journals have embraced this paradigm shift with a greater online presence. Zerrweck and colleagues reported that 31.5 per cent (51.8 per cent recreational use) of general surgeons, and 39.7 per cent (68.6 per cent recreational use) of bariatric surgeons used social media on a daily basis for academic purposes, making it possible for researchers to play an active role in dissemination of their academic outputs, likely amplifying citations. Alternative metrics are part of this era in measuring impact. It remains to be seen whether preprint article repositories, which enable early sharing of articles on social media, result in higher AS and citations, providing journal editors with additional and early impact metrics.

This study has a number of inherent limitations. Articles included were published at different times and were therefore placed at different stages of citation accrual. Original AS values

---

**Fig. 1** Scatter plots illustrating relationship between original variables and increase in number of citations over time

Relationship between a) original Altmetric score (AS) and interval citation gain (iCG) ($R^2 = 0.024$, $P = 0.121$), b) original AS and interval citation gain iCG with outliers (marked by arrows in a) removed ($R^2 = 0.199$, $P < 0.001$), c) original citation count and iCG ($R^2 = 0.797$, $P < 0.001$), d) original citation rate index (CRI) and iCG ($R^2 = 0.908$, $P < 0.001$), and e) original 5-year impact factor (IF) and iCG ($R^2 = 0.011$, $P = 0.311$).
Fig. 2 Receiver operating characteristic curves to identify critical thresholds for variables

a Original Altmetric score (AS), b original citation count, c original citation rate index (CRI), d original 5-year impact factor (IF), and e Oxford Centre for Evidence-Based Medicine level. Arrows indicate the point taken to establish critical thresholds.
were not recorded at identical follow-up times; nonetheless, AS and CRI were associated with a higher iCG, and were therefore irrespective of time lapsed. These bibliometric markers were strongly associated with citation accrual. All articles predate the founding of Altmetric scoring (2011) and the majority (89.0 per cent) predate the creation of Twitter. As such, conclusions regarding the impact of social media on AS cannot be established for this cohort, and future work should focus on prospectively collected serial measures of bibliometrics and AS over time using current articles. The 5-year IFs used in this study were acquired via the Web of Science database during the original data collection period in 2017. Some of these values may have fluctuated over time, having been higher or lower at the time of each paper’s release. The 100 most cited articles related to surgery were the focus, and the findings must therefore be interpreted with caution, as these findings may not be applicable to scientific articles published in other research arenas, or articles published in lower Q-score-rated journals given the preponderance of Q1 journals in this cohort. Citations carry inherent bias in the form of institutional or self-citation that cannot be identified from an online database search. The study also has a number of strengths. Article citations and AS values were recorded simultaneously and from a reputable database—Web of Science. The bibliometric analysis has statistical power, limiting confounding variables, and bias relating to time from publication.

Table 2 Relationship between bibliometric variables, interval citation gain, and Altmetric score in predicting gains in citations.

<table>
<thead>
<tr>
<th>Score</th>
<th>AUC</th>
<th>P</th>
<th>Threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original AS</td>
<td>0 (0–5; 0–53)</td>
<td>0.65 (0.55, 0.76)</td>
<td>0.008</td>
</tr>
<tr>
<td>Original citation count</td>
<td>573.5 (488.8–70.7; 446–5746)</td>
<td>0.65 (0.54, 0.75)</td>
<td>0.012</td>
</tr>
<tr>
<td>Original CRI</td>
<td>34.7 (26.3–45.9; 8.0–442)</td>
<td>0.86 (0.79, 0.93)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Original 5-year IF</td>
<td>4.9 (3.3–8.7; 1.4–8.7)</td>
<td>0.52 (0.41, 0.64)</td>
<td>0.680</td>
</tr>
<tr>
<td>OCEBM level</td>
<td>n.a.</td>
<td>0.46 (0.35, 0.58)</td>
<td>0.545</td>
</tr>
</tbody>
</table>

Values are median (i.q.r.; range). *Data assumptions for the Youden index were violated, so dichotomization points associated with original impact factor (IF) and quality of evidence were not considered. AUC, area under the curve; AS, Altmetric score; CRI, citation rate index; n.a., not available; OCEBM, Oxford Centre for Evidence-Based Medicine.

Table 3 Univariable and multivariable binary logistic regression analysis of factors associated with higher interval citation gain

<table>
<thead>
<tr>
<th>n</th>
<th>Univariable analysis</th>
<th>Multivariable analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Odds ratio</td>
<td>P</td>
</tr>
<tr>
<td>Median</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RCT (no/yes)</td>
<td>87/13</td>
<td>1.19</td>
</tr>
<tr>
<td>Multicentre study (no/yes)</td>
<td>65/35</td>
<td>2.70</td>
</tr>
<tr>
<td>Primary institute appears more than once in top 100 (no/yes)</td>
<td>73/27</td>
<td>0.85</td>
</tr>
<tr>
<td>First or last author appears more than once in top 100 (no/yes)</td>
<td>82/18</td>
<td>3.16</td>
</tr>
<tr>
<td>OCEBM level (1/2/3/4/5)</td>
<td>13/19/34/46/15</td>
<td>0.84</td>
</tr>
<tr>
<td>Original AS ≥ 2* (low/high)</td>
<td>65/35</td>
<td>4.94</td>
</tr>
<tr>
<td>Original citation count ≥ 521* (low/high)</td>
<td>37/63</td>
<td>16.44</td>
</tr>
<tr>
<td>Original CRI ≥ 34.4* (low/high)</td>
<td>49/51</td>
<td>16.44</td>
</tr>
<tr>
<td>Upper quartile</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RCT (no/yes)</td>
<td>87/13</td>
<td>2.09</td>
</tr>
<tr>
<td>Multicentre study (no/yes)</td>
<td>65/35</td>
<td>1.67</td>
</tr>
<tr>
<td>Primary institute appears more than once in top 100 (no/yes)</td>
<td>73/27</td>
<td>1.34</td>
</tr>
<tr>
<td>First or last author appears more than once in top 100 (no/yes)</td>
<td>82/18</td>
<td>3.06</td>
</tr>
<tr>
<td>OCEBM level (1/2/3/4/5)</td>
<td>13/19/34/46/15</td>
<td>0.90</td>
</tr>
<tr>
<td>Original AS ≥ 2* (low/high)</td>
<td>65/35</td>
<td>4.13</td>
</tr>
<tr>
<td>Original citation count ≥ 521* (low/high)</td>
<td>37/63</td>
<td>2.98</td>
</tr>
<tr>
<td>Original CRI ≥ 34.4* (low/high)</td>
<td>49/51</td>
<td>15.95</td>
</tr>
</tbody>
</table>

Values in parentheses are 95 per cent confidence intervals. *Identified by receiver operating characteristic (ROC) curve analysis. OCEBM, Oxford Centre for Evidence-Based Medicine level; AS, Altmetric score; CRI, Citation Rate Index.
Despite the perceived reciprocal association and attraction between high-quality research articles and journals with a higher IF, it appears that article fit, social media promotion, and author power are also predictors of prospective article impact and citation accrual. If any journal’s editorial strategy is to publish articles expected to acquire high citation numbers, early upload to one of the emerging repository websites may well develop as a preferred plan before official periodical submission. Such tactics would allow a period of attention, scrutiny, reflection, AS attainment, and early citation accrual. Such a strategy may be an improvement on the current peer review system before publication in a journal.

**Funding**
Royal College of Surgeons England
Health Education and Improvement Wales

**Acknowledgements**
D.B.T.R. and A.G.M.T.P. are joint first authors. O.P.J. was supported by a Joint Surgical Research Fellowship from the Royal College of Surgeons England and Health Education and Improvement Wales.

Disclosure. The authors declare no conflict of interest.

**Supplementary material**
Supplementary material is available at BJS Open online.

**References**
13. Norgren L, Hiatt WR, Bell K, Nehler MR, Harris KA, Fowkes FGR et al. Inter-society consensus for the management of peripheral arterial disease (TASC II). *Eur J Vasc Endovasc Surg* 2007; 33: S1–S75