ELSEVIER

Cardiovascular Research 61 (2004) 201–203

Cardiovascular Research

www.elsevier.com/locate/cardiores

Editorial

Impact factors: no *totum pro parte* by skewness of citation

Tobias Opthof [a,*], Ruben Coronel [b], Hans Michael Piper [c]

[a] *Department of Medical Physiology, University Medical Center Utrecht, PO Box 85060, 3508 AB Utrecht, The Netherlands*
[b] *Molecular and Experimental Cardiology Group, Academic Medical Center, Amsterdam, The Netherlands*
[c] *Insitute of Physiology, Justus Liebig University, Giessen, Germany*

Received 10 October 2003; received in revised form 15 November 2003; accepted 21 November 2003

Citation of the various papers published in one and the same journal is highly skewed. Journals with a high impact factor obtain this high value by frequent citation of only a limited number of their papers and, on the other hand, journals with low impact factors publish many papers that remain uncited [1]. Thus, mere publication of a paper in a given journal cannot be regarded as a quality marker of that particular paper [2], it just means that the authors have 'succeeded in surviving' the review process of that journal. Seglen [3] has analyzed that 50% of the obtained citations are accumulated by only 15% of the contents of a journal. In addition, the most frequently cited 50% of the contents obtain almost all citations (90%). These numbers were based on an analysis of three biochemical journals (*Biochimica Biophysica Acta*, *Biochemical Journal* and *Journal of Biochemical Chemistry*) and have later been confirmed for two cardiovascular journals (*Circulation Research* and *Cardiovascular Research*) [4]. Both analyses applied to recently published papers [3,4]. It is not known whether these data vary over time and whether or not they depend on impact itself.

1. Skewness of citation

We have analyzed the original research papers of *Cardiovascular Research* published between 1992 and 2000 ($n = 1886$). We have taken advantage from the fact that the impact factor of *Cardiovascular Research* increased from 1.47 in 1991 to 4.69 in 2002. This allowed us to relate the skewness of citations both to the time since publication and to impact. Thus, for year 1 (the calendar year of publication itself), we had nine sets of data (contents of 1992 as cited in 1992 through contents of 2000 as cited in 2000). For year 2 (the calendar year following the year of publication), we also had nine sets of data (contents of 1992 as cited in 1993 through contents of 2000 as cited in 2001). The same applied to year 3. Starting with year 4 the sets of data decreased by one per year for obvious reasons. In year 11 we only had one set of data (contents of 1992 as cited in 2002). In total we analyzed 63 sets of data.

Fig. 1 shows that for all 63 sets of data (varying with respect to years since publication and impact), 50% of citations are obtained by $14.0 \pm 4.8\%$ (S.D.) of papers (triangle). The most frequently cited half of the papers obtains $90.3 \pm 6.1\%$ (S.D.) of all citations (Fig. 1, circle). Finally, only $65.7\% \pm 19.5$ (S.D.) of the papers is cited in a particular year and thus 34.3% of papers remains uncited (Fig. 1, square). The solid line of identity has been included in Fig. 1 as an aid to appreciate the skewness of citation of individual papers. If all papers would be cited in line with the impact factor, this line would give the relationship between the percentage of papers and the percentage of obtained citations.

2. Skewness of citation and years since publication

Fig. 2 shows the relationship of the parameters described in Fig. 1 with the years since publication. Thus, the most frequently cited 50% of the papers (open circles) accumulate on average 90.3% of the citations (as shown in Fig. 1). Also, 50% of the citations are obtained by only 14.0% of the papers (triangles). About two thirds of the papers (67.3%) are less frequently cited than the average paper (data not shown). These parameters did not show a substantial dependence on time since publication. In contrast, there was a significant relationship between the years since publication and the percentage of uncited papers (squares). This figure is about 20% during years 3 and 4 since publication but is substantially higher during earlier and later years (Fig. 2).

* Corresponding author. Tel.: +31-30-2538900; fax: +31-30-2539036.
E-mail address: t.opthof@med.uu.nl (T. Opthof).

0008-6363/$ - see front matter
doi:10.1016/j.cardiores.2003.11.023

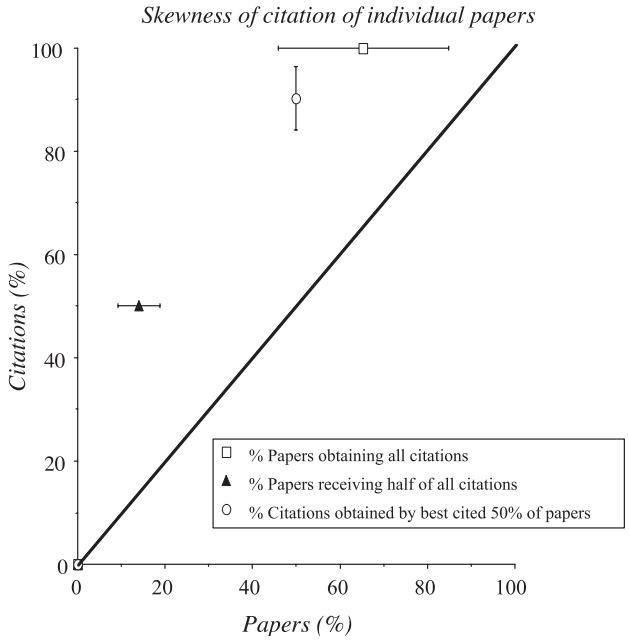

Fig. 1. The contents of *Cardiovascular Research* as published between 1992 and 2000 (1886 original papers only) were analyzed concerning citation from 1992 till 2002. This led to 63 sets of data. For each set, the percentage of papers needed to accumulate 50% of the citations was scored (triangles). Also, the percentage of citations obtained by the most frequently cited 50% of the papers was calculated (circles). Finally, the percentage of papers accounting for all citations in a given year was determined (squares). The error bars indicate standard deviations. The solid line of identity depicts equal citation of all published papers.

3. Skewness of citation and impact

Fig. 3 shows that the same parameters as tested in Fig. 2 have a significant correlation with the impact of the underlying scientific material. Impact was simply defined as the number of citations obtained in all years since publication divided by the number of published papers and was based on original papers only. 'Impact' is thus not the same as 'impact factor'. The number of uncited papers decreases substantially when the impact is higher, which would be expected, although higher impact might also result from more frequent citation of a limited number of papers. The amount of citations obtained by the 'best 50%' of the papers is also negatively correlated with impact. In contrast, the percentage of papers that accumulates 50% of the citations is positively correlated with impact. This suggests that for journals with a higher impact factor, (i) the square in Fig. 1 (indicating the percentage of cited vs. uncited papers) tends to shift to the right, (ii) the circle in Fig. 1 (indicating the percentage of citations obtained by the best cited 50% of papers) shifts to a lower value along a vertical line, and (iii) the triangle in Fig. 1 (indicating the percentage of papers receiving half of all citations) tends to shift to the right. Thus, although skewness of citation is substantial for journals with a high impact factor (such as *Cardiovascular Research* or *Circulation Research* [4]), it will probably be even more skewed for journals with a low impact factor. This is also suggested by comparison of the citations obtained in 1994 by the papers of *Cardiovascular Research* and *Circulation Research* published in 1992 and 1993. The

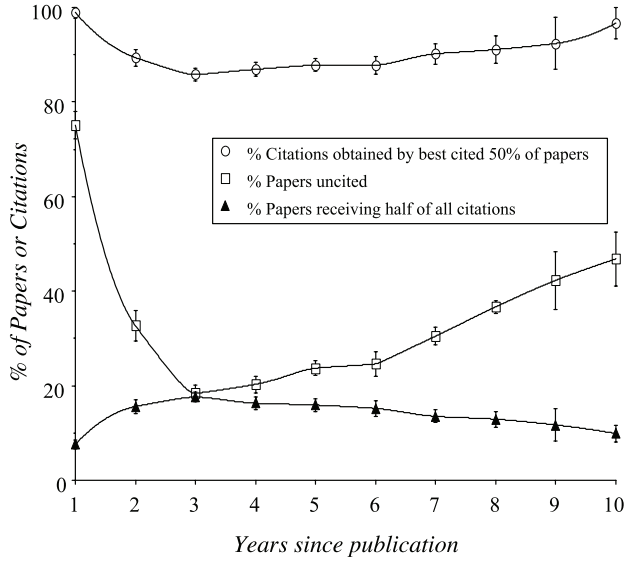

Fig. 2. Within the same 63 data sets as described in Fig. 1, we analyzed whether the percentage of papers required to accumulate 50% of the citations (triangles), the percentage of citations obtained by the most frequently cited 50% of the papers (circles), and the percentage of uncited papers (squares) were correlated with the years since publication. The latter parameter varied substantially over the years, whereas the first two did not. Polynomial curve fits are shown. Error bars indicate standard error of the mean.

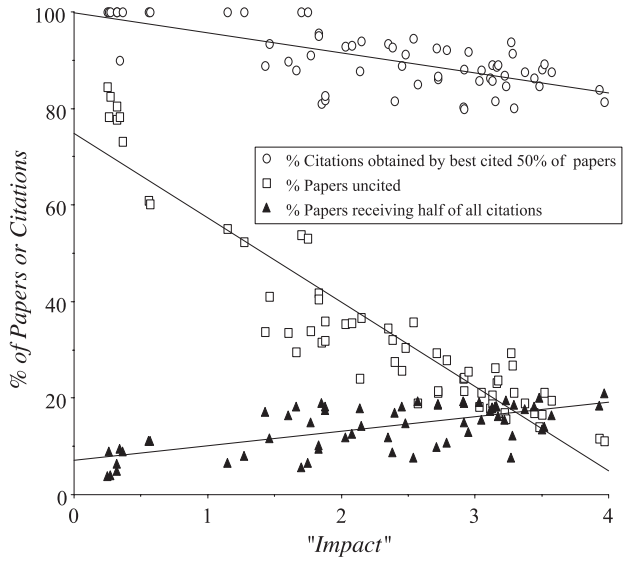

Fig. 3. Within the same 63 data sets as described in Fig. 1, we analyzed whether the percentage of papers required to accumulate 50% of the citations (triangles), the percentage of citations obtained by the most frequently cited 50% of the papers (circles), and the percentage of uncited papers (squares) were correlated with impact. This was the case for all linear regression lines. Correlation coefficients were 0.654, −0.697, −0.471 and −0.932, respectively. All fulfilled $P<0.0005$. Error bars indicate standard error of the mean.

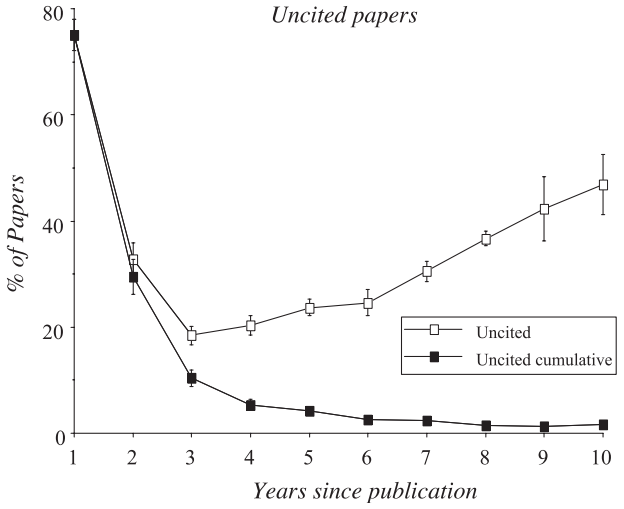

Fig. 4. The percentage of uncited papers as a function of the years since publication. Open squares: data per year. A paper qualified for this group if it remained uncited during that particular year, irrespective of its citations in all other years. Filled squares: Cumulative data per year, including all preceding years. A paper qualified for this group if it remained uncited during that particular year *and all* preceding years. Error bars indicate standard error of the mean.

percentage of papers receiving half of all citations was 16% for *Circulation Research* and 12% for *Cardiovascular Research* [4]. Indeed, the difference between the impact factors of those two journals was larger in 1994 (6.97 vs. 2.89) than at present. The large majority of scientific papers is published in journals with impact factors well below 2.00, at least in the field of cardiovascular science, underscoring the significance of the title of this editorial.

4. Uncited papers

In 1996 De Jong and Schaper [5] reported that in the field of clinical cardiology, 46% of a total of 137,019 papers published between 1981 and 1992 remained uncited during that same period. Papers published in 1981 scored in this category if they remained uncited during the full 12-year period between 1981 and 1992. Papers published in 1992 scored in this category if uncited in 1992. Thus, by and large, half of papers remained uncited for a period of 6 years after publication. We analyzed our 63 data sets, published between 1992 and 2000 and cited between 1992 and 2002, in the same way. Fig. 4 shows the results and focuses on the percentage of uncited papers *per year* (see also Fig. 2) and *per all years*. Although the percentage of uncited papers is between 30% and 50% between years 7 and 10 since publication (open squares), the percentage of papers that remains uncited during all *consecutive* years (filled squares) declines to below 4% after year 5 and below 2% during year 8 since publication. In order to 'qualify' for the papers depicted with the filled squares, a paper should—at any year along the abscissa—have been uncited during *all* preceding years as well. The observation of De Jong and Schaper [5] with respect to 46% of uncited papers over a period of 12 years since publication in the field of clinical cardiology, therefore, probably results from the inclusion of many journals with a low impact factor. Despite the fact that many papers remain uncited per year (see (Figs. 1, 2 and 4)), the number of papers that remains uncited during a sustained period is very low for a journal with a higher impact factor such as *Cardiovascular Research* (Fig. 4).

5. Conclusion

In conclusion, skewness of citation of papers is substantial. On the average, the 50% most frequently cited papers obtain about 90% of the citations. This percentage is relatively constant over the years after publication, but may be higher (almost 100%) for journals with a low impact factor and may be lower (but still 80%) for journals with a high impact factor. The percentage of papers that accumulate 50% of the citations is about 14%, but it may be below 10% for journals with a low impact factor and around 20% for journals with a high impact factor. Despite this correlation with impact, it is obvious that the impact factor of a scientific journal is not a *totum pro parte* for its individual papers [1–4]. Obviously, for the assessment of the quality of scientific papers there is no alternative to reading... as pointed out previously by Sidney Brenner (see Ref. [3] for source).

References

[1] Anonymous. Deciphering impact factors. Nat Neurosci 2003;8:783.
[2] Opthof T. Sense and nonsense about the impact factor. Cardiovasc Res 1997;33:1–7.
[3] Seglen PO. Why the impact factor of journals should not be used for evaluating research. BMJ 1997;314:498–502.
[4] Opthof T, Coronel R. The impact factor of leading cardiovascular journals: where is your paper best cited? Neth Heart J 2002;10:198–202.
[5] De Jong JW, Schaper W. The international rank order of clinical cardiology. Eur Heart J 1996;17:35–42.