Journal impact factors: a 'bioequivalence' issue?

A. Rostami-Hodjegan & G.T. Tucker

University of Sheffield, Section of Molecular Pharmacology and Pharmacogenetics, Clinical Sciences Division, The Royal Hallamshire Hospital, Sheffield S102JF, UK

Aims Journal impact factors (IMFs) are used increasingly by institutions as performance indicators of the quality of 'individual research output'. Although the need for discretion when using the numbers has been emphasized, there has been little formal analysis of the issues. We therefore investigated citation profiles for three clinical pharmacology journals to assess the validity of using IMF as a measure of 'individual research'.

Methods We compared the pattern of individual citations for random samples of 120 papers published in *Clin Pharmacol Ther* (CPT), *Br J Clin Pharmacol* (BJCP) and *Eur J Clin Pharmacol* (EJCP) in 1981, 1991, 1995 and 1996. Using an analogy between citation-time profiles of papers and concentration-time profiles of drugs, it was possible to define 'lag-time', $C_{\max}$, $t_{\max}$, $t_{1/2}$ and $AUC(t)$, and to investigate 'bioequivalence'.

Results Citation distributions for individual publications were widely variable and skewed (skewness = 1.47, 2.16 and 1.37 for CPT, BJCP and EJCP, respectively). The 90% CI values for the IMF of a publication in each journal (i.e. 90% CI for an observation as opposed to 90% CI for the mean) were 0.24–16.94, 0.08–10.3 and 0.09–5.68.

Conclusions IMF does not represent the impact of an individual paper. Furthermore, if the comparison of journals is treated as a bioequivalence issue, the citation data should be log transformed prior to calculating IMF such that they represent the likelihood of citation for the median article. After such transformation, absolute differences between the IMF of clinical pharmacology journals become much smaller.

Keywords: clinical pharmacology, journal impact factor, research assessment

Introduction

Last year the IMF of the BJCP, as reported by the Institute for Scientific Information (ISI), increased by 38% from 1.846 (1998) to 2.545 (1999), sustaining the Journal within the top 20% of journals (Figure 1). So what? For many journals such an increase is accompanied by a celebratory mood and editorial comment on achievement and relative ranking (e.g. [1–4]). However, many readers and also those within institutions who use IMF to assess performance and promotion may not be aware of the appropriate meaning and correct interpretation of IMF. We use this

opportunity to comment on some aspects of IMF which may be of concern.

IMF is a measure of the frequency at which the 'average article' in a journal has been cited in a particular year (the Institute for Scientific Information, ISI [5]). It is calculated by dividing 'the number of current citations to articles published in the two previous years' by 'the total number of articles published in the same time period'. For example, the IMFs of five 'clinical pharmacology journals'¹ for 1994–99 are shown in Figure 2. Data used by ISI to derive the numbers are presented in Table 1.

IMFs are used increasingly by institutions as an index of research excellence and the quality of 'individual research output' (or researcher). They are also used by libraries with respect to ranking and purchase of journals. Although

¹Although many journals publish research in the field of clinical pharmacology, amongst the 'top 100' journals within the 'pharmacology and pharmacy' discipline, these five journals are the only ones that contain 'clinical pharmacology' in their title.

Correspondence: Professor G.T. Tucker, Section of Molecular Pharmacology and Pharmacogenetics, Floor L, The Royal Hallamshire Hospital, Sheffield, S10 2JF. Tel.: 0114 2712052; Fax: 0114 2720275; E-mail: g.t.tucker@sheffield.ac.uk

Received 28 August 2000, accepted 27 November 2000.

Figure 1 Ranking of the *Br J Clin Pharmacol* amongst all journals in the ISI database (●) and amongst journals in the ISI subgroup of 'Pharmacology and Pharmacy' (○).

Figure 2 Impact factors of five clinical pharmacology journals from 1994 to 1998. Key: ■ = *Clin Pharmacol Ther*, ○ = *Br J Clin Pharmacol*, ◆ = *Fund Clin Pharmacol*, △ = *Eur J Clin Pharmacol*, ● = *J Clin Pharmacol*, Solid line = impact factor for median of 100 top journals in the ISI group of 'Pharmacology and Pharmacy'.

Table 1 Data used to derive IMFs reported in Figure 2.

	Year	CPT	BJCP	EJCP
Number of citations of 1999 papers to articles published in 1997/1998	1997	569	522	325
	1998	509	405	214
Number of research/review articles published in 1997/1998	1997	119	190	151
	1998	122	186	159
Calculation of IMF		$\frac{(569+509)}{(119+122)}$	$\frac{(522+405)}{(190+186)}$	$\frac{(325+214)}{(151+159)}$
IMF		4.846	2.545	1.770

much has been written about the deficiencies and limitations of IMF as an indicator of research excellence, particularly across different specialities [5–26], and ISI clearly indicates the need for discretion when using the numbers [6–12], formal analysis of the issues are rare [13, 14] and may not be visible to many scientists outside of 'information science and scientometrics'. For clinical pharmacologists, as the target audience of *BJCP*, pharmacokinetic terms and parameters are well known. Thus, we have used an analogy to pharmacokinetics in analysing the time-course of citations and have applied the principles of bioequivalence testing to consider:

- 1) The relationship between journal IMF and the impact of individual articles.
- 2) The appropriate calculation and interpretation of IMF.

Methods

The patterns of article-by-article citations of 120 randomly chosen papers published in *CPT*, *BJCP* and *EJCP* in 1981, 1991, 1995 and 1996 ($n=30$ per journal per year)

were compared. Citations of these articles were counted individually using 'cited reference search' within the Science Citation Index of the ISI data base through access to Bath Information and Data Services (BIDS). By analogy to plasma drug concentration–times profiles, it was possible to estimate citation lag-time, $C_{\max}$, $t_{\max}$, $t_{1/2}$ and $AUC(t)$ values from plots of number of citations *vs* time for individual papers.

The most commonly used IMF is based on 2 year citation. Thus, 1999 IMFs are related to the number of citations that 1997 papers received 2 years after publication and the number of citations that 1998 papers received 1 year after publication (Table 1). Therefore, it is reasonable to assume that the cumulative citations ($AUC(2 \text{ year})$) for 2 successive years indicate, on average, the IMF for the subsequent year. For example, average cumulative 2 year citations for 1995 and 1996 should resemble the ISI reported IMF for 1997. We used this approach to test whether randomly selected papers in this study were representative of overall publications in each journal.

To investigate the link between short-term and long-term impact, the correlation between cumulative 2 year

citations (AUC(2 year)) of 1981 papers with their cumulative 17 years citations (AUC(17 year)) was analysed. Similarly, the relationship between cumulative two year citations (AUC(2 year)) of 1991 papers with their cumulative 7 years citations (AUC(7 year)) were studied using linear regression analysis. Descriptive statistics were used to characterize the distributions of cumulative two year citations for 1995 and 1996 papers.

The 'bioequivalence' of the three specific journals examined with respect to IMF in 1996–97 was estimated from the ratio of log transformed AUC(2 year) values of 60 papers published in 1995 and 1996 (30 per year).

Results

The distribution of 2 year cumulative citations to 60 publications from three journals is shown in Figure 3. The calculated IMFs for *CPT*, *BJCP* and *EJCP* based on these random samples taken from 1995 to 96 were 3.48, 1.93,

1.17, respectively (Figure 3). These values resembled those of ISI reported IMF values for 1997 (Figure 2); there was -27 , 6 , -4% bias for *CPT*, *BJCP* and *EJCP*, respectively. The latter indicates that the samples were adequately representative of the papers published in these journals.

Distributions of cumulative 2 year citations for individual publications (1995–96) were highly variable and markedly skewed ('skewness' = 1.47, 2.16 and 1.37 for *CPT*, *BJCP* and *EJCP*, respectively) (Figure 3). This is in agreement with other reports on skewness of citations [13].

The 90% CIs for the IMF of a publication (i.e. the probability of an individual paper having a particular number of citations; *not* the same as the CI of the journal IMF, which is the 90% CI for the mean citation per article) were 0.24–16.94, 0.08–10.3 and 0.09–5.68, respectively (Figure 4), with considerable overlap between all three journals.

Figure 3 Distributions of total 2 year citations for 60 randomly chosen papers published in three clinical pharmacology journals in 1995 and 1996 (30 per year per journal). Solid square and bars are mean and 90% CI; open square and bars are median and geometric 90% CI.

Figure 4 Bioequivalence of IMF for three clinical pharmacology journals (a) as opposed to bioequivalence of individual publications in these journals (b).

Figure 5 Relationship between citations over 17 years and over 2 years for 30 randomly selected papers published in 1981 in each of three clinical pharmacology journals.

Logarithmic transformation prior to calculation of journal IMF decreased the apparent differences between journals. Whilst IMFs without transformation were 3.48, 1.93, 1.17, they were 2.01, 0.93 and 0.72 with transformation for *CPT*, *BJCP* and *EJCP*, respectively. This indicates, for example, that the article representing the modality of citation in *CPT* in 1995–96 received 1 more rather than 1.5 more citations than the article representing the modality of citations in *BJCP*.

The time-profiles of citations of individual papers were highly variable (Figure 5) with respect to maximum number of citations ($C_{\max}$), time to reach $C_{\max}$ ($t_{\max}$), lag-time before observing any citations (t_{lag}) and decay in the number of citations over time (characterized by $t_{1/2}$, which was calculated from log-transformed post $t_{\max}$) (data not shown). Median $t_{\max}$ values were the same for all three journals (3 years) but median $C_{\max}$ values were different (5, 3, 2 for *CPT*, *BJCP* and *EJCP*, respectively). Some articles had very long lag times (similar to enteric-coated pharmaceutical formulations!), perhaps indicating some difficulty in digestion.

There were significant correlations between short-term (2 years) and long-term (7 or 17 years) citations of individual papers (Figure 6; data for 2 years citation *vs* 7 years citations are not shown). The latter is in agreement with other observations of relatively consistent ranking for journals based on short-term or long-term IMF, particularly in the same field of studies [27]. Nevertheless, individual articles with similar short-term citation could have widely variable long-term citation (Figures 5 and 6).

Comments and conclusions

Over the last 10 years, in parallel with an increase in perception of the IMF as a metaphor of research excellence, there have been many publications warning against the limitations of IMF and advocating its prudent use. It is worth emphasizing that some of these warnings, which are largely ignored (!), have come from the founder

Figure 6 Time-profiles of citations of randomly selected papers published in *CPT*, *BJCP* and *EJCP* in 1981.

and chairman emeritus of ISI, Dr Eugene Garfield [6–12]. We have summarized some of the problems associated with the use of IMF in Table 2 and refer the interested reader to appropriate references. Although many of the issues itemized in Table 2 are of general concern, we have concentrated on the use of IMF as an indicator of research performance.

The first question that we wanted to answer was whether short-term citation (e.g. 2 years) is representative of long-term citation. Such a relationship will clearly depend on the consistency of half-life. This echoes the issue of using truncated AUC as a measure of AUC(infinity) in bioequivalence testing [28]. The average citation

half-life in the field of clinical pharmacology, as assessed for three journals, was found to be fairly consistent (5.5–5.7 years for the citations to 1981 papers). Therefore, 2 years IMF is a reasonable predictor of long-term IMF as long as comparison is made within the field of clinical pharmacology.

The modality (highest likelihood) of citation for a published paper in a journal will only be associated with the mean of overall citations of papers in that journal if the distribution of individual citations is *normal*. Our study (in the field of clinical pharmacology) together with other reports (in other fields or in general) [13, 26] indicate that this assumption is not valid. In fact, the distributions of citations for individual publications are positively skewed necessitating log or even double log transformation for normalization. With the assumption of log-normality, the median (geometric mean) can represent the highest likelihood of the impact factor for an individual publication

Table 2 Known problems with IMF.

Problem	Reference
-Citation frequency may not be a valid indicator of scientific quality.	18
-IMF is not suitable to compare different fields or specialities within broad fields (e.g. 'Clinical Pharmacology vs Pharmaceutics' within the ISI determined field of 'Pharmacology and Pharmacy').	6–12
-IMF of review based journals are higher in general.	16
-Citations to any type of article (including letters, editorials, communications, meeting abstracts) are used to arrive at a number for total citations, but this is then divided by the number of normal articles and review articles only.	10,16
-IMF is heavily influenced by self-citation and the national bias of North American scientists to cite each other.	16
-Only a small proportion of highly cited articles determine journal IMF. Thus, journal IMFs are not representative of individual articles.	16
-Journal IMFs do not affect the citation of articles published in that journal.	15, 16

Figure 7 Relationship between citations of an identical paper that was published in four different journals and the IMF of those journals. Key: *Eur J Pharm Sci* = EJPS, *Pharm Res* = PR, *Eur J Drug Met Pharmacokin* = EJDMP, *J Pharm Sci* = JPS.

based on the journal IMF. Currently, ISI does not derive IMF based on individual citations but calculates them based on total citations and total number of papers. This calculation exaggerates the distance between likely citations that a paper may receive if it is published in one journal relative to another. It should be borne in mind that even after the corrections that have been mentioned, due to large variability in citations, it is not possible *a priori* to know that the citations of a paper accepted for one particular clinical pharmacology journal are going to be more than those for another paper to be published in another clinical pharmacology journal.

There has been debate on whether the journal has an impact on the number of citations that a paper may receive. A direct answer to this question is impossible without assessment of citations of the same paper when published in *different* journals at the *same time*. Clearly, repeat publication is not generally to be encouraged, and is never acceptable without disclosure to editors. However, legitimate examples of similar or repeat publication may occur occasionally. Seglen compared the performance of two research groups who were publishing in a similar set of journals (IMF from 0.5 to 8) and showed that the ratio of individual citations for the groups remained the same within each journal, while the expectation was to see a diminished difference as a function of high IMF [15]. We have evaluated a consensus report of a meeting on bioequivalence that was published in four different journals with varying IMF between September and December of 1993 [29–32]. No significant correlation was found between the citations and IMF (Figure 7), indicating that the journals did not significantly influence the citation rate. Arguably, there were too few data points to make such a conclusion. However, finding papers which are published simultaneously and in identical format in more than four journals will be a rare occurrence!

A current significant problem in academia, particularly in Europe and Australia (based on personal communications with clinical pharmacologists) is the crude use of IMF without any weighting or processing as a measure of the quality of individual research output or that of a research

Table 3 Weighted and nonweighted performance of eight professors in the University of Sheffield Faculty of Medicine based on ISI citations.

Professor	1	2	3	4	5	6	7	8
Average journal IMF ¹	3.1	2.7	3.5	6.5	5.5	4.9	3.5	5.9
Rank (Average journal IMF)	(7)	(8)	(5=)	(1)	(3)	(4)	(5=)	(2)
Cited papers	173	362	301	172	189	111	174	54
Citations	1813	2925	3110	3223	2656	1281	729	301
Individual impact	10.5	8.1	10.3	18.7	15.4	11.5	4.2	5.6
Rank (Individual impact)	(4)	(6)	(5)	(1)	(2)	(3)	(8)	(7)
Weighted impact ²	3.39	3.00	2.94	2.87	2.80	2.35	1.20	0.95
Rank (Average impact)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)

1-Average IMF of the latest 14 journals that the author had published in (or all journals if less than 14) 2-Weighted for the field/speciality effect based on individual 'Average IMF'.

group. As a consequence of such practices, researchers are forced to choose journals for publication based primarily on IMF; this despite indications that national (as opposed to institutional) assessments are alleged to be less proscriptive (e.g. the UK Research Assessment Exercise).

In a recent internal Research Assessment Exercise at the University of Sheffield School of Medicine, academic staff were asked to provide information on the IMF of the journals they had published in, as a primary part of the filter to determine who would be returned in the national exercise. These data were then used to rank their research performance regardless of their broad speciality/field of interest. Moreover, it was assumed that the journal IMF could be used as an indicator of the IMF of a particular publication in that journal. ISI's recommendation for evaluation of 'research groups' and 'individual scientists' is to use 'Expected Citation Rates' (ECR) as part of 'Personal Citation Reports' (PCR) [7]. Use of these indices comprises a very similar approach to that which we have taken in this study to follow up individual citations of papers (AUC(*t*)) and weight them by the expected citation in their field. However, ECR and PCR are rarely used as they need to be calculated separately for each individual, and they cannot be applied to recent publications unless some extrapolation (e.g. using half-life) is applied. In Table 3 we have tried to show the different interpretations that arise when using crude IMFs of the journals as opposed to individual impact and weighted individual impact. Eight professors in the Faculty of Medicine of the University of Sheffield with different primary specialty interests were ranked based on the aforementioned measures. Clearly, each measure of performance gives a totally different ranking for each professor. Number 4 would be ranked first when using the journal IMF of the papers that he published in, or according to his individual impact. However, by weighting his individual impact factor by the general impact of his research area, his rank decreases to fourth. Professor number 8 was second in rank

based on average journal IMF, while his individual impact and also his weighted impact put him at rank 7 and 8, respectively. Thus, for their field of studies professors 1–3 are doing better than the others despite the fact that their journal IMFs rank them at 5–8.

In this commentary, we have attempted to clarify the assumptions of using IMF for assessing individual research by using an analogy to pharmacokinetics that is understandable to clinical pharmacologists, and to show that the individual impact factor of a publication cannot be assessed using the journal IMF. As emphasized by Hamilton [26], researchers should be warned against 'publishing by and for IMF numbers'. The aim always should be to target the report to the most suitable audience. Research organizations that use journal IMF to assess the scientific merit of research reports are ignorant of the meaning of IMF and they do not appreciate the potentially negative effects of their practice on diverting information to less appropriate audiences. Obviously, another negative aspect of such practice might be a gradual decline of European journals by shifting of submissions to North American journals with higher IMF. This could have detrimental effects on many scientific societies and communities in Europe and elsewhere, especially those that are linked closely to particular journals.

References

- Howard L, Wilkinson G. Impact factor of psychiatric journals- the British Journal of Psychiatry now has the highest impact factor of all psychiatric journals outside Europe. *Br J Psychiatry* 1998; **172**: 457.
- Lee VHL. Pharmaceutical Research: a quality journal on a mission. *Pharm Res* 2000; **17**: 251.
- Schoonbaert D, Van der Stuft P, Roelants G. Has TM & IH entered the top 3 of the tropical medicine journals? Reflections on the journals's unexpected first impact factor. *Trop Med Int Health* 1998; **3**: 935.
- Peter TJ. Impact factors: believe them or not. *Addiction Biol* 1999; **4**: 109–110.

- 5 Garfield E. The impact factor.
http://isinet.com/isi/hot/essays/7.html 1994.
- 6 Garfield E. Using impact factor.,
http://isinet.com/isi/hot/essays/8.html 1994.
- 7 Garfield E. Expected citation rates, half-life, and impact ratios: comparing apples to apples in evaluation of research.,
http://isinet.com/isi/hot/essays/10.html 1994.
- 8 Garfield E. How can impact factors be improved? *Br Med J* 1996; **313**: 411–413.
- 9 Garfield E. The Impact Factor and its correct usage. *Anaesth J* 1998; **47**: 439–440.
- 10 Garfield E. Journal impact factor: a brief review. *Can Med Assoc J* 1999; **161**: 979–980.
- 11 Garfield E. The diverse roles of citation indexes in scientific research. *Rev Invest Clin* 1998; **6**: 497–504.
- 12 Garfield E. The multiple meanings of impact factors. *J Am Soc Inf Sci* 1998; **49**: 768.
- 13 Seglen PO. Skewness of science. *J Am Soc Inf Sci* 1992; **43**: 628–638.
- 14 Seglen PO. Quantification of scientific article contents. *Scientometrics* 1996; **35**: 358–366.
- 15 Seglen PO. Casual relationship between article citedness and journal impact. *J Am Soc Inf Sci* 1994; **45**: 1–11.
- 16 Seglen PO. Why the impact factor of journals should not be used for evaluating research. *Br Med J* 1997; **314**: 498–502.
- 17 Seglen PO. Citations and journal impact factors: questionable indicators of research quality. *Allergy* 1997; **52**: 1050–1056.
- 18 Seglen PO. Citations frequency and journal impact: valid indicator of scientific quality. *J Intern Med* 1991; **229**: 109–111.
- 19 Cronin B, Overfelt K. Citation-based auditing of academic-performance. *J Am Soc Inf Sci* 1994; **45**: 61–72.
- 20 Forrest MC. Impact factor abuse. *J Chromatogr* 1997; **9**: 3–4.
- 21 Harter SP, Nisonger TE. ISI's impact factor as misnomer: a proposed new measure to assess journal impact. *J Am Soc Inf Sci* 1997; **48**: 11460–11468.
- 22 Hecht F, Hecht BK, Sandberg AA. The journal impact factor: a misnamed, misleading, misused measure. *Cancer Genet Cytogenet* 1998; **104**: 77–81.
- 23 Linde A. On the pitfalls of journal ranking by impact factor. *Eur J Oral Sci* 1998; **106**: 525–526.
- 24 Stoelinga PJW. The rise and fall of the impact factor. *Int J Oral Maxillofac Surg* 1997; **26**: 81.
- 25 Opthof T. Sense and nonsense about the impact factor. *Cardiovascular Res* 1997; **33**: 1–7.
- 26 Hamilton DP. Research papers: who's uncited now. *Science* 1991; **251**: 25.
- 27 Garfield E. Long-term *vs* short-term journal impact: does it matter? *Scientist* 1998; **12**: 11–13.
- 28 Midha KK, Hubbard JW, Rawson MJ. Retrospective evaluation of relative extent of absorption by the use of partial areas under plasma concentration versus time curves in bioequivalence studies on conventional release products. *Eur J Pharm Sci* 1996; **4**: 381–384.
- 29 Blume HH, Midha KK. Report of consensus meeting. Bio-International 92, conference on bioavailability, bioequivalence, and pharmacokinetic studies. *J Pharm Sci* 1993; **82**: 1186–1189.
- 30 Blume HH, Midha KK. Report of consensus meeting. Bio-International 92, conference on bioavailability, bioequivalence, and pharmacokinetic studies. *Eur J Pharm Sci* 1993; **1**: 163–171.
- 31 Blume HH, Midha KK. Report of consensus meeting. Bio-International 92, conference on bioavailability, bioequivalence, and pharmacokinetic studies. *Pharm Res* 1993; **10**: 1806–1811.
- 32 Blume HH, Midha KK. Report of consensus meeting. Bio-International 92, conference on bioavailability, bioequivalence, and pharmacokinetic studies. *Eur J Drug Metab Pharmacokin* 1993; **18**: 225–232.