What do impact factors tell us?

There is an increasing tendency, when considering appointment, tenure or promotion, or when comparing research groups, to look to impact factors (as defined according to the citation index values (Box 1) of the journals in which they publish. Although the 'research assessment exercise' process that determines UK government funding does not explicitly use impact rating, some institutions have set a standard of expecting that all biologists publish in journals with impact factors of at least 4, and those that fall below this number are likely to be excluded from the exercise. A concern is that such a criterion is unfair, especially in comparisons among fields of research.

Here, we show how this usually disadvantages evolutionary biologists and ecologists. One widespread belief is that the size of the field (measured as the annual number of published articles) largely determines the impact of the best papers within that field, and hence has a long publication time. Consequently, papers that provide the solution to the problem took a year and the author, Andrew Wiles, had previously devoted seven years to the problem. Indeed, the ten mathematical journals with the highest impact factors all have half-lives of at least 9.9 years, whereas none of the same subset of cell biology journals achieves a half-life of six years.

As an illustration of this difference, the highest ranking journal in mathematics has an impact factor equivalent to the 51st in cell biology. This low performance clearly does not indicate that the quality of mathematics as a science is poor. Instead, it probably arises from the tradition of mathematical research progressing on many questions simultaneously, where papers quote few references, but where each paper contains a large amount of relevant information and has a long publication time. Consequently, published information is useful for a long time.

This is exemplified by the proof of Fermat’s Last Theorem for which the original reference is a scribble in the margin of Diophantus’ Arithmetica (the margin text was published in 1670, after Fermat’s death), and the paper that provides the solution quotes 84 references of which only four are published in the previous two years – hardly surprising as the reference took a year and the author, Andrew Wiles, had previously devoted seven years to the problem. Indeed, the ten mathematical journals with the highest impact factors all have half-lives of at least 9.9 years, whereas none of the same subset of cell biology journals achieves a half-life of six years.

The number of papers published within a field does not affect the average citations at all (Box 2) and hence has nothing to do with the impact factor of a field. The size-independence can be understood by imagining an isolated subject with only one journal, which therefore always quotes itself. If each paper in this journal had 20 references to work published in the same journal in the previous year, the impact factor of this journal would equal 20 (assuming a time window of one year), regardless of the size of the field.

Box 1. Some widely used terms, and their definitions

Impact factor of a journal: the average number of times that articles published in a specific journal in the two previous years (e.g. 1997–1998) were cited in a particular year (e.g. 1999). The citing journals have to be included in the Science Citation Index.

Impact time window: number of years included in the calculation of the impact factor (usually two, as indicated above).

Impact factor of a field: the average number of times that articles published in journals of that field in the two previous years (e.g. 1997–1998) were cited in a particular year (e.g. 1999). This can be calculated by weighting journal impact factors by the number of articles they have published. The citing journals have to be included in the Science Citation Index.

Half-life of a journal: the number of years, going back from the current year, that account for 50% of the total citations received by the journal in the current year. If 1500 citations appear in year 1999 that refer to journal X and 500 of them refer to papers published in 1996 or later, the half-life of journal X equals three years. A long half-life implies that a small proportion of citations are included in the time window that influences impact factors.

Half-life of a field: the number of years, going back from the current year, that account for 50% of the total citations received by journals within that field in the current year.

Growth rate of a field or a journal: the number of papers appearing in a specific year divided by the number in the preceding year.
Box 2. A demographic model of impact factors

Assume that all science is divided into $F$ different ‘fields’, such as astronomy or virology, among which impact factors are to be compared. A field $j$ has produced $N_j(t)$ scientific articles 7 years ago ($T = 0$, marks current year), and the number of references in each of these papers averages $c_j(t)$. The citation index of journal $j$ which produces $n_j$ papers, can then be written as follows:

$$C_j = \frac{\sum_{t=0}^{T} n_j(t)c_j(t)}{\sum_{t=0}^{T} n_j(t)}$$

(1)

assuming the index is calculated using citations within 7 years typically ($T = 7$). The factor $p_j(t_1, t_2)$ gives the probability that a paper that appeared $t_1$ years ago in journal $j$ is cited in a randomly chosen article published $t_2$ years ago ($T = t_2 - t_1$) and belonging to field $i$. Assuming that the journal $j$ belongs to field $f(j)$, this factor can be further divided into the following components:

$$p_j(t_1, t_2) = \frac{N_{f(j)}(t_2, t_1, N_j(t_1)r_j - t_1)}{\sum_{f(j)} N_{f(j)}(t_2, t_1, N_j(t_1)r_j - t_1)}$$

(2)

$R_{i, f(j)}(t)$ is the average relevance of a paper belonging to field $f(j)$ to the development of field $i$ on a time scale of $t$ years, defined such that the numerator of the first term gives the probability that a randomly chosen citation in a paper of field $i$ refers to field $f(j)$ when the respective publication years differ by $t$. It is clear that $R_{i, f(j)}(t)$ typically peaks within the same field, thus when $t = 0$, and with a time difference of a few years; also, $R_{i, f(j)}(t)$ has a wider distribution in fields with long citation half-lives (Box 1). We then get:

$$C_j = r_j \cdot \frac{\sum_{t=0}^{T} N_j(t)c_j(t)}{\sum_{t=0}^{T} N_j(t)}$$

(3)

From this, it can be seen that growing numbers of citations per paper always increases impacts, but the effect of growth of journals or whole fields can vary. A growing subject may increase its impact factor since there are more current references citing a smaller pool of past papers (effect of $r_j$), and with a time difference of a few years; also, $R_{i, f(j)}(t)$ has a wider distribution in fields with long citation half-lives (Box 1). Instead, $r_j$, which gives the relative relevance of results published in journal $j$ compared with the average of its field, reflects the importance of a result for further work: if the relevance $r_j$ of one journal is twice that of another in the same field, results published in it are cited twice as often. To give correct probabilities in Eqn 2, the relevance values must scale according to $\sum_{i} r_j(t) = N_j(t)$. To account for growth in either the number of articles or citations in each field, we assign a growth rate $\lambda_i$ to the average number of references per paper in this field, and $\lambda_j$ to the growth of the focal journal (see Box 3 for definitions of growth). We then get:

$$C_j = r_j \cdot \frac{\sum_{t=0}^{T} N_j(t)c_j(t)}{\sum_{t=0}^{T} N_j(t)}$$

(4)

The parameter $r_j$, which is nearest the idea of ‘quality’ of research, is automatically limited by the scaling $\sum_{j} r_j(t) = N_j(t)$. In other words, the expected $r_j$ of a randomly chosen paper must equal one within every field, and increasing the relevance $r_j$ of all results within the field by the same factor will have no effect at all on the overall impact. Instead, if growth rates are small (each $\lambda_j = 1$), the impact factor of a field is completely determined by the field-specific average of the number of citations in each paper, $c_j$, scaled by the fraction of cited papers that were published in the last 7 years. This proves that impact comparisons among fields are meaningless unless the primary interest is to find a complicated way of expressing average number of citable references per paper combined with the proportion of citations that have appeared recently.

Unlike the size of a field, its growth rate has the potential to alter impact factors (Box 2). A high growth rate in the number of papers published annually within a field can increase its impact, but this effect can be reversed if citations tend to accumulate slowly (Box 2). Paradoxically, successful journals often expand, yet this may reduce their impact factor. This is because most papers are cited when they are at the old end of the time window – having a lower proportion of papers at this end would weight the impact factor unfavourably.

The problem of comparing impacts among fields might be reduced by the fact that comparisons often take place within fields; for example, when comparing the curriculum vitae of applicants for academic positions. We agree with this, but we fear that the temptation to use impacts as comparative tools also exists.
among fields that appear related, such as the different biological sciences in Fig. 1. Even within a field, there will be variation in the traditions of different subfields, again making comparisons difficult.

For example, within ecology and evolution, there are molecular ecology journals with a typically rapid turnover, and taxonomy journals that have much longer half-lives. Scientists publishing in journals of a subfield that get listed along with another subfield with a shorter half-life or a longer number of citations are not indicated by the impact factor, such biases could be reduced if half-lives were considered an equally important measure of quality as impact factors. However, further complications are found when among-field citations are possible. Citing papers from other fields reduces the impact of the quoting field while increasing the impact of other fields reduces the impact of the cited field (Eqn 3 in Box 2). It is equally clear that the impact factor of interdisciplinary journals gets exaggerated in fields that have lower impact factors, such as ecology, but is understated for those with otherwise higher impacts, such as developmental biology.

Comparisons using impact factors have serious implications for ecologists and evolutionary biologists. The constraints of the slow life cycles of many species can make rapid studies impossible, and often details of the species or study site demand reference to older or obscure studies. Improvements such as relative rankings or impact divided by mean impact for that field are unlikely to offer completely satisfactory solutions simply because fields are never totally isolated in their citation patterns. It seems that any proper comparison must be constrained to journals within a field, and, even then, more fine-scaled divisions are likely to be problematic: the same issues apply at a smaller scale because any field can be subdivided.

Similar problems arise from the practice of counting citations to an individual researcher’s work. Furthermore, because of time lags between publication and citation, this tends to underestimate young scientists’ work and might overestimate the contribution of individuals with declining productivity. A further concern is that the demand for high impact ratings will subvert science; for example, by placing increasing emphasis on areas where it is possible to complete research, rather than on areas where it is possible to conduct and carry out research very quickly. This is of considerable concern for ecologists and evolutionary biologists where long-term studies are a fundamental part of the science.

Acknowledgements
We thank Neil Metcalfe and an anonymous referee for very helpful comments, and the TMR programme of the European Commission for funding.

Hanna Kokko
Department of Zoology, University of Cambridge, Downing Street, Cambridge, UK CB2 3EJ
(h.kokko@zoo.cam.ac.uk)

References
4 Statteiner, B., Kush, V.H. and Kolotina, N.G. (1995) Scale effects on impact factors of scientific journals: ecology compared to other fields, Oikos 72, 446-448.