Relation between online “hit counts” and subsequent citations: prospective study of research papers in the BMJ

Thomas V Perneger

Evaluation of published medical research remains a challenge. Two classic yardsticks are the citation count (the number of times a paper is cited by others) and the impact factor of the journal that published the paper (which reflects the average number of citations per article). However, the citation count can be assessed only several years after publication, and the impact factor is not paper specific and is thus virtually meaningless in assessing any paper’s influence. Another measure, which can be obtained rapidly and is paper specific, is the “hit count” (the number of times a paper is accessed online). Whether this count predicts citations is unknown. I examined this issue prospectively in a cohort of papers published in the BMJ.

Methods and results

The study used articles published in volume 318 of the BMJ (1999) in sections titled Papers, General Practice, and Information in Practice. The hit counts (full text articles, HTML version) for the main body of each article within a week of publication were provided by a BMJ staff member because the “hit parade” posted on the journal website was found to be unreliable for 1999. I obtained the number of citations on 24 May 2004 from the ISI Web of Science, an internet service to which the local medical library has a subscription. I also recorded for each paper the study design and the number of pages.

Nine papers were excluded because they did not report research (but reported discussions of, for example, NHS management and statistics methods). The remaining 135 papers comprised 29 randomised trials,

![Graph showing the relationship between citations and internet hits for 153 papers in volume 318 of the BMJ (1999)]
The value of a research study is traditionally assessed through citation counts or by the impact factor of the journal that published the study. Citation counts can be obtained only years after publication, and the impact factor is not paper specific. 

For a cohort of papers published in the BMJ in 1999, the hit count on the website in the week after online publication predicted the number of citations in subsequent years; the hit count is a potentially useful measure of the scientific value of a research paper.

When the hit count was included as predictor, however, the association between hits and citations became non-significant after adjustment for study design. Online readers judge the scientific value of an article from the title and the abstract, and if this assessment is favourable, they access the full paper. The paper’s scientific value also leads to citation by other researchers. This hypothesis is supported by the greater frequency of both hits and citations for papers that used the most scientifically rigorous study designs, such as randomised trials.

The number of early hits is a potentially useful measure of the scientific value of published medical research papers. Publication of hit counts by online journals should be encouraged.

Daniel Berhane from the BMJ provided valid hit counts for the journal’s website.

Contributor: TVP is the sole contributor.

Funding: None.

Competing interests: TVP is the editor of the International Journal for Quality in Health Care.

Ethical approval: Not required.

3 Seglen PO. Why the impact factor of journals should not be used for evaluating research. BMJ 1997;314:498-502.

(Accepted 10 July 2004)

Corrections and clarifications

Dr Foster’s case notes: How often are adverse events reported in English hospital statistics?

This article by Paul Aylin and colleagues contains errors that escaped the notice of both the BMJ and the authors during the editorial process. The first sentence of the opening box should read: “It has been suggested that an estimated 850 000 medical errors occur in NHS hospitals every year resulting in 40 000 deaths.” In the “basic figures” section, the first sentence should read: “On average 2.2% of all episodes (about 275 000 [not 27 500] per year) included a code for an adverse event.” And we initially posted an incomplete version of table A on bmj.com. It has now been updated to provide full ICD-10 codes. We apologise for our lapses.

Operative vaginal delivery and neonatal and infant adverse outcomes: population based retrospective analysis

We have been alerted to some errors in this paper by Kitaw Demissie and colleagues (3 July, pp 24-6). In the Participants section of the abstract and the opening sentence of the Results section, the published numbers for singleton live births were wrong. The correct figures are 11 463 829 (instead of 11 639 388) for the United States and 374 873 (instead of 556 597) for New Jersey. The authors state that these revisions do not affect any of the results in the table or the conclusion of the paper.

Neurocardiogenic syncope

Some referencing errors crept into this Clinical Review by Carol Chen-Scarabelli and Tiziano M Scarabelli (7 August, pp 336-41). In the Table summarising clinical trials, Takata et al should be reference w3 (not w2). In the “extra: further information” on tilt testing protocols on bmj.com, all references to w2 should in fact be to w3.

3 Seglen PO. Why the impact factor of journals should not be used for evaluating research. BMJ 1997;314:498-502.

(Accepted 10 July 2004)