

### What is already known on this topic

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

### What this study adds

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

11 systematic reviews, 41 prospective studies, 8 case-control studies, 41 cross sectional surveys, 6 qualitative studies, and 17 other designs (such as economic analyses or case reports).

The average hit count for the papers in the first week after publication was 685 (SD 410; 25th, 50th, and 75th centiles 437, 578, and 795 respectively; range 175 to 3181); the average number of citations in the five years after publication was 32.5 (SD 37.5; 25th, 50th, and 75th centiles 9.5, 22, and 42.5 respectively; range 0 to 291). Only one paper was never cited. The hit count was associated with the number of subsequent citations (Pearson correlation coefficient: 0.50,  $P < 0.001$ ). The result was similar for logarithms of the counts ( $r = 0.54$ ,  $P < 0.001$ ) (figure). For every 100 additional hits, 4.4 additional citations (95% confidence interval 3.1 to 5.7) accrued over the five years.

The average hit count for randomised trials or systematic reviews was 832, for prospective or case-control studies was 747, and for cross sectional, qualitative, and other studies was 545 hits ( $P = 0.001$ ). Longer papers attracted more hits than short papers (an extra 54.4 hits per page,  $P = 0.004$ ), but this association became non-significant after adjustment for study design.

Citations were predicted by paper length (an extra 9.3 citations per page,  $P < 0.001$ ) and study design (randomised trials and systematic reviews yielded 46.0 citations, prospective and case-control studies 38.9 citations, and other designs 19.3 citations ( $P = 0.001$ ). When the hit count was included as predictor, however, the effect of study design became non-significant; only page length (an extra 7.3 citations per page,  $P < 0.001$ ) and the hit counts (an extra 3.7 citations per 100 hits,  $P < 0.001$ ) remained as independent predictors. These variables explained 33% of variance in citation counts.

### Comment

Papers that attracted the most hits on the *BMJ* website in the first week after publication were subsequently cited more often than less frequently accessed papers. Thus early hit counts capture at least to some extent the qualities that eventually lead to citation in the scientific literature.

My hypothesis is that "scientific value" explains the association between hits and citations. 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.<sup>4</sup> 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.

1 Adam D. The counting house. *Nature* 2002;415:726-9.

2 Walter G, Bloch S, Hunt G, Fisher K. Counting on citations: a flawed way to measure quality. *Med J Aust* 2003;178:280-1.

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

4 Lee KP, Schotland M, Bacchetti P, Bero LA. Association of journal quality indicators with methodological quality of clinical research articles. *JAMA* 2002;287:2805-8.

(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 11 463 823 (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 w6.