Categorising Citations to Trace Research Impact

Stephen Hanney[*], Jonathan Grant[**], Teresa Jones[*] and Martin Buxton[*]

[*]*Stephen.Hanney@brunel.ac.uk, Teresa.Jones@brunel.ac.uk, Martin.Buxton@brunel.ac.uk*
Health Economics Research Group, Brunel University, Uxbridge, Middlesex UB8 3PH, (UK)

[**]*Jonathan_Grant@rand.org,*
RAND Europe – Cambridge, Grafton House, 64 Maids Causeway, Cambridge CB5 8DD (UK)

Introduction

Using citations as a 'radioactive tracer' of research impacts has been 'largely unutilized' (Kostoff, 1998) and citations' categorisations have rarely before been used to trace impacts through several generations of papers. Bowman (2001) previously described a method for the categorisation of proposals for a technology investment programme and his methods were considered adaptable for the assessment of citations. Clarification was considered necessary for the motives behind a citation, its strength of impact, location, and/or frequency of occurrence within an article. The method described here has been used as part of a wider study to identify the impact of diabetes research (Hanney, Home, Frame, et al (2005) and Hanney, Frame, Grant, et al (2003)).

Method

Existing literature on citations is extensive but as earlier reviews were carried out for different purposes this study began with a fresh literature review aimed at identifying previous studies with the greatest relevance. The aims of citation analyses and the similarity of methods, for example which citations were categorised and by whom, were considered. Studies that covered the factors that should be included in the categorisation and the location and frequency of a citation within a paper were also examined.

The findings from the literature search were used to develop a template to apply to articles citing the source, or first generation, papers. In particular, for motives we developed the list from Small (1982) and had 5 categories: develop; support; apply; refute; note/review only which are defined in Table 1. For strength of impact we used four categories drawn from Cano (1989); peripheral; limited; considerable; essential (See Table 2). The template was applied by one member of the team to the second generation papers that were accessed. A further five team members applied the template to a sample of the papers and inter-rater reliability was assessed by applying the kappa coefficient (Stata Press, 1999), with the standard six level classification scale of kappa results from *poor* to *almost perfect,* to the ratings of this sample by all six team members.

Table 1: Definitions of the reasons or motives for citing a paper applied to each citation occasion

Develop	The *citing* article is developing a concept or method previously described in the cited article.
Support	The *citing* article is supporting a concept or method previously described in the cited article.
Apply	The *citing* article uses a method (or methods) described in the cited article.
Refute	The *citing* article either claims that the cited article is incorrect or disputes the cited article but is unable to come to a firm conclusion.
Note/Review only	The *citing* article refers to the cited article as part of the relevant literature but it either serves no explicit role in the analysis (note) or is compared to other relevant literature (review).

Results

The 29 first generation diabetes papers published by Alberti and colleagues in 1981 were cited in 799 second generation papers. Of these, 623 were accessible for categorisation. 35% were categorised as only of peripheral importance, 56% as limited, 8% as considerable and only 1% as essential. Amongst those citing articles, most (80%) citations categorised as peripheral were for 'note or review', whereas amongst those categorised as essential the citations were mainly classified as 'supporting'. To test the inter-rater reliability of these categorisations, five additional raters categorised a sample of 62 of the second generation citations. Inter-rater reliability revealed a mixed picture with the level of agreement between all six assessors mainly *fair*, but ranging from *slight* for 'develop' as a motive to *almost perfect* for 'apply'.

Table 2: Definitions of the four categories of importance of the cited paper - applied to each citing paper

Peripheral	The work described in the cited article is of little importance to the *citing* article. Citation is simply background, an aside, for completeness or indeed irrelevant.
Limited	The work described in the cited article is of some limited importance to the *citing* article. It would be inappropriate to omit it, but it is not an important part of a central argument.
Considerable	The work described in the article is of considerable importance to the *citing* article. The work is one of a number central to the argument.
Essential	The work described in the cited article is of critical importance to the *citing* article, and central to the argument presented, and a key foundation for the paper.

Discussion/ Further Research

This approach has provided some useful information in terms of moving beyond mere counting of citations and has shown that only in a small minority of cases is the cited paper considered highly important to the citing paper. Due to the large numbers of citations, the method has not yet provided a pathway of important papers through several generations as had initially been envisaged. However, a number of component parts in the method have been identified which would benefit from strengthening, for example the definitions of the categories and the consistency of their application. The use of language ladders as discussed by Bowman (2001) may prove beneficial to the consistency of the application of the template.

Acknowledgments

We gratefully acknowledge funding from the Research and Development Directorate of the National Health Service Executive, London, and the Wellcome Trust. The Health Economics Research Group also receives a programme grant from the Department of Health's Policy Research Programme. We thank the project's Advisory Group and the study participants.

References

Bowman, C. W. (2001). Evaluating intellectual capital – III. Evaluating investment opportunities. *Canadian Chemical News,* March, 30-32.

Cano, V. (1989). Citation behaviour: Classification, utility and location. *Journal of the American Society of Information Science,* 40, 284-290.

Hanney, S., Frame, I., Grant, J., Green, P. & Buxton, M. (2003). From bench to bedside: Tracing the payback forwards from basic or early clinical research – A preliminary exercise and proposals for a future study. HERG Research Report No. 31. Uxbridge: Health Economics Research Group, Brunel University. [Available at: http://www.brunel.ac.uk/about/acad/herg/publications/hergreports/]

Hanney, S. R., Home, P. D., Frame, I., Grant, J., Green, P. & Buxton, M., (Forthcoming). Identifying the impact of diabetes research. *Diabetic Medicine.*

Kostoff, R. (1998). The use and misuse of citation analysis in research evaluation. *Scientometrics*, 43, 27-43.

Small, H., (1982). Citation context and content analysis. In B. Dervin, & M. Voight (Eds) *Progress in Communication Sciences*. Norwood, N. J: Ablex, vol 3, pp. 287-310.

Stata Press. (1999). *STATA Reference Manual: Release 7.0*. College Station, TX: Stata Press.