

ARE MEDICAL ARTICLES HIGHLIGHTING DETAILED STATISTICS MORE CITED?

*Mike Thelwall**

School of Mathematics and Computer Science. University of Wolverhampton.

Resumen: Al llevar a cabo una revisión de la literatura científica, es natural buscar artículos y leer sus resúmenes con el fin de seleccionar los documentos que posteriormente serán leídos completamente. Por esta razón, los resúmenes informativos son importantes para asegurar que la investigación se lee. La descripción de los métodos seguidos en la elaboración de un artículo puede ayudar a aumentar la confianza sobre la calidad del mismo. Este artículo evalúa si los artículos médicos que mencionan tres métodos estadísticos, cada uno de los cuales es sin duda indicativo de un análisis estadístico más detallado que el promedio, son más altamente citados. Los resultados muestran que los artículos médicos que mencionan correcciones de Bonferroni, *Bootstrapping* y del Tamaño del Efecto tienden a ser un 7 %, 8 % y 15 % más altamente citadas que la media, respectivamente. Aunque esto es consistente con la hipótesis de que mencionar técnicas estadísticas más detalladas genera una investigación más altamente citada, estas técnicas pueden también tender a ser utilizadas en las áreas más altamente citadas de la Medicina.

Palabras clave: Análisis de citas, Bootstrap, corrección de Bonferroni, Tamaño del efecto.

Título: ¿SON MÁS CITADOS LOS ARTÍCULOS MÉDICOS QUE INCLUYEN ESTADÍSTICAS DETALLADAS?

Abstract: When conducting a literature review, it is natural to search for articles and read their abstracts in order to select papers to read fully. Hence, informative abstracts are important to ensure that research is read. The description of a paper's methods may help to give confidence that a study is of high quality. This article assesses whether medical articles that mention three statistical methods, each of which is arguably indicative of a more detailed statistical analysis than average, are more highly cited. The results show that medical articles mentioning Bonferroni corrections, bootstrapping and effect size tend to be 7%, 8% and 15% more highly ranked for citations than average, respectively. Although this is consistent with the hypothesis that mentioning more detailed statistical techniques generate more highly cited research, these techniques may also tend to be used in more highly cited areas of Medicine.

Keywords: Citation analysis, Bootstrap, Bonferroni correction, Effect Size.

1. INTRODUCTION

Paper abstracts are presumably read much more often than the full text of papers as researchers scan them for clues about whether the paper is relevant and important enough to be properly read (Cronin, Ryan & Coughlan, 2008), (Dijkers, 2003). It is therefore important for authors to ensure that their abstracts clearly indicate the topic of a paper but also perhaps to make the case that the paper is important enough to fully read. One way in which a paper can be impressive is through the use of detailed statistical methods and some authors include such details in their abstracts. Although it is known that various properties of abstracts are important (Hartley & Benjamin, 1998), no research has assessed the importance of the level of detail given to the statistical methods. This paper applies a scientometric technique (Thelwall & Wilson, in press) to assess whether medical articles mentioning a detailed statistical technique tend to be more cited than average. Although this is a limited approach, it has the advantage of being applicable on a large scale and hence being perhaps able to identify trends that are hidden at the level of individual articles. The new scientometric technique is demonstrated with an examination of three statistical methods used in medical research.

To give some context to the study, the importance of high quality statistical methods is widely acknowledged in science, both explicitly and implicitly. Nevertheless, applications of statistics vary in quality many studies have shown that published refereed journal articles do not always include essential information and sometimes make mistakes (Schreiber, Nora, Stage, Barlow, & King, 2006). This article takes a different approach by showing that articles using more detailed statistical methods tend to be more highly cited. Whilst citations in general are not good indicators of the quality of individual articles, citation counts can be used when aggregated on a large scale (van Raan, 1998). Moreover, in general, citation counts have significant positive correlations with peer judgements of article quality in some fields. For example, one study found a Spearman correlation of 0.56 in medicine between peer review judgments and citation counts for 2369 articles in an Italian research assessment exercise (Franceschet & Costantini, 2011).

* m.thelwall@wlv.ac.uk

Recibido: 15-04-2015; aceptado: 15-06-2015.

Two widely used statistical techniques and one statistical concept are assessed to ascertain whether medical articles that mention them are more highly cited than are comparable articles. Medical articles that use statistics typically contain explicit descriptions of their methods in abstracts, making it possible to identify the methods with simple queries. The first technique is bootstrapping, which indicates the use of statistics beyond standard testing. The second technique, Bonferroni correction (Hochberg, 1988), indicates that the familywise error rate for a set of statistical tests has been managed when conducting multiple hypotheses. This is a controversial technique, however, because some argue that it is misleading and that the issue it corrects for should be addressed in a different way (Nakagawa, 2004), (Perneger, 1998). The concept assessed is effect size. Including effect size goes beyond simple hypothesis tests to discuss the magnitude of an identified effect.

2. DATA AND METHODS

Medical articles mentioning the two techniques and one concept were identified by searching the Scopus citation database with the following queries. These exclude general journals as well as journals with a statistical or data focus in order to obtain lists of articles in medical journals that do not focus on statistics. Articles were ignored if they were published before 1996, when Scopus had much lower coverage of the academic literature, or after 2013, when the articles would have had little time to be cited.

- TITLE-ABS-KEY(bonferroni) AND SUBJAREA(medi) AND DOCTYPE(ar) AND NOT SRCTITLE(plos one) AND NOT SRCTITLE(statistics) AND NOT SRCTITLE(BIOMETRIC*) AND NOT SRCTITLE(BIOSTATISTICS) AND NOT SRCTITLE(DATA).
- TITLE-ABS-KEY(bootstrap) AND SUBJAREA(medi) AND DOCTYPE(ar) AND NOT SRCTITLE(plos one) AND NOT SRCTITLE(statistics) AND NOT SRCTITLE(BIOMETRIC*) AND NOT SRCTITLE(BIOSTATISTICS) AND NOT SRCTITLE(DATA).
- TITLE-ABS-KEY("effect size") AND SUBJAREA(medi) AND DOCTYPE(ar) AND NOT SRCTITLE(plos one) AND NOT SRCTITLE(statistics) AND NOT SRCTITLE(biometric*) AND NOT SRCTITLE(biostatistics) AND NOT SRCTITLE(data).

The average number of citations to articles depends upon the year in which they were published, with older articles having longer to attract citations, and also by field, with some fields tending to cite more than others (Moed, 2006). In order to assess whether the three sets of articles were more cited than comparable articles, for each article all other articles in the same journal, year, volume and issue were downloaded and then ranked in order of citation counts. The normalised citation rank of each article in its issue was recorded using $(rank-1)/(articles-1)$, a figure between 0 (fewer citations than any other article in the issue) and 1 (more citations than any other article in the issue). Ranks were used rather than means because citation distributions are highly skewed, even within a single field and year (Thelwall & Wilson, 2014). This normalised rank is adjusted for time because it depends only on articles published in the same issue and it is normalised for field since articles tend to be published in specialist journals, with the general journals having been excluded by the queries.

3. RESULTS AND DISCUSSION

From Table 1, medical articles mentioning Bonferroni correction tend to be ranked 7% higher, and medical articles mentioning bootstrapping tend to be ranked 8% higher than average for the issues in which they were published. This is consistent with the belief that using these techniques helps to produce better articles, although it could also be that they are more likely to be used in fields that have higher citation norms. Medical articles explicitly mentioning effect sizes (in their titles, abstracts or keywords) tend to be 15% higher than average for the issues in which they were published, which is a substantial difference. Again, this is consistent with the belief that using these techniques helps to produce better articles, although it could also be that effect sizes are more likely to be calculated or mentioned in fields that have higher citation norms.

Query	Articles	Mean rank	Lower bound	Upper bound
bonferroni	3720	0.571	0.563	0.580
bootstrap	2389	0.578	0.567	0.589
"effect size"	9544	0.653	0.648	0.659

Table 1. Normalised citation ranks and 95% confidence intervals (using the normal distribution) for Scopus-indexed articles (1996-2013) matching the three queries. For each article, the normalised citation rank is for the journal, year, volume and issue in which it was published.

The absences of statistical terms from article abstracts does not necessarily mean that the techniques were not used, just that they were not mentioned in abstracts. To estimate the proportion of articles mentioning each technique in the abstract out of all articles using the techniques, ScienceDirect Full Text searches suggest that less than 10% of medical articles mentioning these terms in their abstract (Table 2). Hence it is not possible to extrapolate the findings to cover all articles using the three techniques. Note, however, that an article can mention a technique outside its abstract in its literature review or to discount its use without applying it. Moreover, a technique may be used but referred to by a different name and so the figures in Table 2 (and Table 1) are estimates.

Query	Search matches in title, keywords or abstract	Search matches in full text
bonferroni	1102 (2.5%)	44700
bootstrap	616 (6.9%)	8877
"effect size"	2185 (9.0%)	24405

Table 2. ScienceDirect Full Text searches for statistical terms, with all searches restricted to journal articles 1996-2013 in the category Medicine and Dentistry.

4. CONCLUSIONS

On the basis of the above results, whilst it is clear that articles using the three statistical techniques investigated tend to be more highly cited than other articles, there are several possible reasons for this that the methods used cannot separate.

- The research is higher quality through the use of more detailed statistics.
- Research for which the more detailed statistics are important enough to be mentioned in the title, keywords or abstract of an article tends to be higher quality.
- Research for which the more detailed statistics are mentioned in the title, keywords or abstract of an article tends to attract more citations, perhaps because this highlighting convinces other researchers that the article is important enough to read.
- Journals in higher citation areas within medicine encourage authors to include more detailed statistical information within their abstracts.
- Research using detailed statistics tends to occur in topics with high citation rates, such as more quantitative areas.
- Research into statistical methods is within the sample and has higher citation rates than average for medicine.

Despite the multiple possible explanations for the results above, the first three listed explanations are plausible and so the findings should give an extra incentive for researchers to use more detailed statistics, when relevant to a study, and to report them in their titles, keywords or abstracts. This advice seems unlikely to be detrimental even if the use or reporting of more detailed statistics is not in fact a cause of higher citation rates for articles.

For future research it would be useful to try the techniques here for other statistical methods and for research methods more widely, as well as for other subject areas. It would also be useful to access a large collection of individual full-text article records in order to test whether the more important factor is using a statistical method or reporting it in the abstract.

5. REFERENCES

- CRONIN, P.; RYAN, F. & COUGHLAN, M. Undertaking a literature review: a step-by-step approach. *British Journal of Nursing*, 2008, vol. 17, num. 1, p. 38-43.
- DIJKERS, M.P.J.M. Searching the literature for information on traumatic spinal cord injury: the usefulness of abstracts. *Spinal cord*, 2003, vol. 41, num. 2, p. 76-84.
- FRANCESCHET, M. & COSTANTINI, A. The first Italian research assessment exercise: A bibliometric perspective. *Journal of Informetrics*, 2011, vol. 5, num. 2, p. 275-291.
- HARTLEY, J. & BENJAMIN, M. An evaluation of structured abstracts in journals published by the British Psychological Society. *British Journal of Educational Psychology*, 1998, vol. 68, num. 3, p. 443-456.
- HOCHBERG, Y. A sharper Bonferroni procedure for multiple tests of significance. *Biometrika*, 1998, vol. 75, num. 4, p. 800-802.

- MOED, H.F. *Citation analysis in research evaluation*. Berlin: Springer, 2006.
- NAKAGAWA, S. A farewell to Bonferroni: the problems of low statistical power and publication bias. *Behavioral Ecology*, 2004, vol. 15, num. 6, p. 1044-1045.
- PERNEGER, T. V. What's wrong with Bonferroni adjustments. *British Medical Journal*, 1998, vol. 316, num. 7139, p. 1236-1238.
- SCHREIBER, J.B.; NORA, A.; STAGE, F.K.; BARLOW, E.A. & KING, J. Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of Educational Research*, 2006, vol. 99, num. 6, p. 323-338.
- THELWALL, M. & WILSON, P. Distributions for cited articles from individual subjects and years. *Journal of Informetrics*, 2014, vol. 8, num. 4, p. 824-839. Available in: <<http://www.scit.wlv.ac.uk/~cm1993/papers/DistributionsForCitedArticlesPreprintx.pdf>>. Consultation: April 2, 2015.
- THELWALL, M. & WILSON, P. (in press). Does research with statistics have more impact? The citation rank advantage of structural equation modelling. *Journal of the Association for Information Science and Technology*.
- VAN RAAN, A.F. In matters of quantitative studies of science the fault of theorists is offering too little and asking too much. *Scientometrics*, 1998, vol. 43, num. 1, p. 129-139.