Commentary III
Validity and fairness of the impact factor – a comment on the article by Decker et al.

Dr. Dr. Kaltenborn is a lecturer at the Medical Center for Methodology and Health Research, Institute of Medical Informatics of the Philipps-University of Marburg

The article by Decker, Beutel and Brähler (2004) deals with a very important topic because the *impact factor* (IF) is meanwhile in widespread use as a scientometric parameter for the evaluation of scientists and medical science in Germany and other European countries. The authors select an innovative and original approach by evaluating the IF as an assessment procedure, highlighting its psychometric properties. While it is beyond the scope of a short comment to deal with all the manifold aspects of the IF as delineated in the article, I would like to focus on the validity and the resulting (un)fairness. As accurately outlined by the authors, the IF lacks validity: Its short-term calculation based on a time span of two years rewards research fields characterized by short publications and quotation lags, such as biochemistry and molecular biology, while penalizing disciplines known for a certain quotation delay but recording very lengthy half-lives of their scientific articles, such as social medicine or social sciences in general. Furthermore, research fields are differently represented in the SCI and publication as well as citation behavior is dependent on the specific research field (preference for publication in the national language when social aspects of medicine are delineated; different, field-based citation preferences for articles, books or Internet resources etc.; differences in the number of citations per article in different fields etc.). As a consequence, interdisciplinary comparison based on the IF would not only be unfair; it is in fact not feasible (see fairness and utility in the article). But it is this very comparison that is made if medical schools use the IF to allocate financial resources to their departments – with heterogeneous disciplines such as molecular genetics, surgery, dental medicine, medical informatics and social medicine being lumped together. Although the AWMF model reduces this interdisciplinary unfairness to some degree (and is therefore very important), we need to know much more about the consequences of the evaluation procedure based on the IF. This is especially true because the IF leads to a strategic adaptation of the people and institutions concerned, i.e., the IF opens up reflexivity concerning the behavior of the authors, the journals and the medical discipline, which ultimately affects knowledge transfer in society. Instead of one single quantitative parameter like the IF, more procedures based on qualitative aspects and aimed at multi-dimensional evaluation are discussed and applied (for different approaches see: Deutsche Forschungsgemeinschaft 1998; Giessler 2000; Kaltenborn & Kuhn 2002; Langenbeck 1998; Lewison 2000; 2002; Raspe 1999; Seglen 1997). Such a qualitative, multi-dimensional evaluation for an individual scientist might, for instance, rely on:

1. *Quantity and international visibility of scientific work* (quantity of articles and other media in national and international languages as well as visibility through inclusion in databases like Medline, Social Science Citation Index, Science Citation Index, Embase, Psychological Abstracts, etc.)
2. *Quality of scientific work* (reading and evaluating by experts of a limited number of relevant articles)
3. *Contribution of scientific work to medical knowledge and impact on healthcare improvements* (sensible use of citations including citations in the national language; prizes and honors; contribution to patents, to guidelines and medical practice; contribution to knowledge-based arti-

Soz.- Präventivmed. 49 (2004) 23–24
0303-8408/04/010023–2
DOI 10.1007/s00038-003-0035-0

facts such as medical instruments, hospital information systems, etc.; inclusion in legal commentaries and use for social policy; gaining of additional research funds etc.)

4. *Quality of medical and further scientific education* (guidance for dissertations, doctoral theses, postdoctoral theses and qualifications etc.)

(Not all dimensions need to have the same weight; *Quantity and international visibility of scientific work*, for instance, could have a much less significant influence than the others.) However, in contrast to the USA, German medicine has long neglected research about production and transfer of medical information and knowledge, and it appears that the proper evaluation of scientific work will depend on much more experience being gained in these basic processes (Jansz 2000; Kaltenborn & Kuhn 2002).

Karl-Franz Kaltenborn

References

Decker O, Beutel ME, Brähler E (2004). Deep impact – evaluation in the sciences. Soz Praventiv Med *49*: 10–4.

Deutsche Forschungsgemeinschaft (1998). Empfehlungen der Kommission „Selbstkontrolle in der Wissenschaft". Vorschläge zur Sicherung guter wissenschaftlicher Praxis, Januar 1998 (deutsch/englisch). http://www.dfg. de/aktuelles_presse/reden_stellungnahmen/download/empfehlung_wiss_praxis_0198.pdf http://www.dfg.de/aktuelles_presse/reden_stellungnahmen/download/self_regulation_98.pdf

Giessler A (2000). Zur Evaluierung der Forschungsleistung eines Universitätsklinikums. Dtsche Med Wochenschr *125*: 979–83.

Jansz CNM (2000). Some thoughts on the interaction between scientometrics and science and technology policy. Scientometrics *47*: 253–64.

Kaltenborn K-F, Kuhn K (2002). Der Journal impact factor als Indikator für das Ranking von Fachzeitschriften und die Evaluation von Forschung: eine kritische Analyse. In: Ekkernkamp A, Scheibe O, eds. Qualitätsmanagement in der Medizin: Handbuch für Klinik und Praxis. Landsberg: ecomed: 1–29.

Langenbeck U (1998): Kann eine Verwendung des Impact Factors für die Evaluation von Forschungsleistungen rational begründet werden? Medizinische Genetik *10*: 504–8.

Lewison G (2000): Citations as a means to evaluate biomedical research. In: Cronin B, Atkins HB, eds. Web of knowledge: a festschrift in honor of Eugene Garfield. Medford: Information Today: 361–72.

Lewison G (2002): Researchers' and users' perceptions of the relative standing of biomedical papers in different journals. Scientometrics *53*: 229–40.

Raspe H (1999): Leistungs- und belastungsorientierte Festsetzung von Personal- und Sachmittelbudgets auf der Basis fakultätsinterner und -übergreifender Evaluationen. Modelle deutscher medizinischer Fakultäten und ein vereinheitlichender Vorschlag. Med Klinik *94*: 702–7.

Seglen PO (1997). Why the impact factor of journals should not be used for evaluating research. BMJ *314*: 498–502.

Address for correspondence

Dr. Dr. Karl-Franz Kaltenborn
Institut für Medizinische Informatik
Universität Marburg
Bunsenstrasse 3
D-35037 Marburg
E-mail: kaltenbo@med.uni-marburg.de