© *Basic & Clinical Pharmacology & Toxicology* 2005, **97**, 261–275.
Printed in Denmark . All rights reserved

Pharmacology & Toxicology
Copyright ©
ISSN 1742-7835

MiniReview

Bibliometric Methods: Pitfalls and Possibilities

Johan A. Wallin

University Library of Southern Denmark, University of Southern Denmark, DK-5230 Odense M, Denmark

(Received January 20, 2005; Accepted May 20, 2005)

Abstract: Bibliometric studies are increasingly being used for research assessment. Bibliometric indicators are strongly methodology-dependent but for all of them, various types of data normalization are an indispensable requirement. Bibliometric studies have many pitfalls; technical skill, critical sense and a precise knowledge about the examined scientific domain are required to carry out and interpret bibliometric investigations correctly.

Bibliometric indicators are increasingly being used as a tool for research performance evaluation. These indicators are based on bibliographic databases, which are designed primarily for information retrieval purposes so informetric studies represent only a secondary use of the systems (Hood & Wilson 2003). This causes many technical and interpretative problems, including methodological considerations: *One of the most crucial objectives in bibliometric analysis is to arrive at a consistent and standardised set of indicators* (van Raan 2004). If the necessary normalization is not undertaken, there is a risk of discrediting research that may be good enough by the standards of its own scientific discipline. At the same time there is always a considerable risk of ignoring important differences in the societal impact of a research programme, because this can not be captured using bibliometric methods (Council for Medical Sciences 2002). Bibliometrics and peer review can only comment with certainty on a research programme's short-term effects, whereas it is doubtful whether these methods make any predictions about the research programme's long-term effects (Kostoff 1998).

Bibliometric methods are quantitative by nature, but are used to make pronouncements about qualitative features. This is, in fact, the major purpose of all sorts of bibliometric exercises, to transform something intangible (scientific quality) into a manageable entity. Compared with peer review, which has a limited area of investigation, it is easy to use bibliometric methods to examine unlimited quantities of publications. Bibliometrics has given us a tool that can easily be scaled from micro (institute) to macro (world) level. But in terms of convincing scientific evidence, we do not know enough about the connection between these quantitative bibliometric objectives and their assertions about research quality. This is paradoxical in the light of the widespread and often uncritical use of bibliometric indicators (including Journal Impact Factor, JIF) for various assessment and resource allocation purposes, in spite of the warnings in a comprehensive bibliometric expert literature (in *Scientometrics,* etc.).

Citation analysis is a good example of this, since this type of bibliometric analysis is the most often used to couple a quantitative parameter to an evaluation of research performance. This is because it is thought to provide a simple, quick impression of the quality of the research: *Citation analyses generate relatively short-term quantifiable items, they have the appearance of short-term research impacts, and are therefore attractive candidates as short-term proxies for research impact and perhaps quality* (Kostoff 1998). This coupling is based on a theoretical assumption of a simple linear relationship between scientific quality and citation counts. But citation patterns vary greatly between disciplines, publication types and authors, as well as being dependent on the type of research and its long term significance. For example, reviews and methodology articles are cited conspicuously often, "bad" scientific work (negational citation) is cited more often than work of mediocre quality, and well-known fundamental work is not cited to the extent it deserves. One could add that scientific quality cannot be reduced to a few numerical parameters.

If one looks at references in a particular paper, many peculiarities may found, such as missing references to specifically important papers, or to the work of authors who have generally made essential contributions to the field, or an exaggerated attention to a specific author (...) *As soon as authors refer, already to a small extent ''reasonably'', i.e. not based on a 100% random ''reference generator'', valid patterns in citations will be detected if sufficiently large number of papers is used for analysis. Furthermore, it is statistically very improbable that all researchers in a field share the same distinct*

Author for correspondence: Johan A. Wallin, University Library of Southern Denmark, University of Southern Denmark, Odense Campus, Campusvej 55, DK-5230 Odense M, Denmark (fax +45 65 50 26 01, e-mail wallin@bib.sdu.dk).

reference-biases (for example, all authors cite deliberately earlier papers which did not contribute whatsoever in their field) (van Raan 1998a).

This sort of minimalistic citation theory is the starting point for the following considerations, and it has not yet been demonstrated satisfactorily that a more convincing general citation theory can be put forward (Cronin 1998; Leydesdorff 1998). This relates to the fact that bibliometric science is spread over many different scientific domains (Cronin 2000). The heart of the matter is that *High rates of citation may indicate a useful or provocative paper in a field of wide interest; low rates of citation may simply indicate a narrow field and cannot be constructed as prima facie evidence of a poor quality* (Chew & Relyea-Chew 1988). It has to be emphasized, therefore, that *the number of times this body of literature is cited world-wide, can be regarded as a measure of the impact or the international visibility of the research* (van Raan & Van Leeuwen 2002). This is underlined by the fact that *citation counts indicate impact rather than quality* (Moed *et al.* 1985b) and *citation-based indicators point to one specific, but important quality aspect referred to as international influence or impact (van Raan & Van Leeuwen 2002).* By their nature, citation studies can only reflect the importance of the publications in a contemporary perspective (Kostoff 1998), because technical and scientific knowledge becomes obsolete – much faster, in fact, within some scientific disciplines than others. Several investigations have consistently shown that the obsolescence time-frame for health science literature is just under 50 years (Hall & Platell 1997; Poynard *et al.* 2002). In addition, as already mentioned, generally accepted knowledge is absorbed into the existing universe of knowledge and is, therefore, mentioned but not explicitly cited – a phenomenon called *obliteration by incorporation* (Murugesan & Moravcsik 1978; MacRoberts & MacRoberts 1986).

The spectrum of bibliometric methods includes publication patterns studies, bibliographing, bibliographic coupling (co-citation and co-occurrence), and citation analysis (scientific papers and patents). The last three are most suitable for the task of evaluation, as most of them are based on a single type of publication, i.e. publications in scientific journals. All these methods require several forms of standardisation and normalization if they are to be correctly interpreted (Narin & Hamilton 1996). The use of publication patterns studies is restricted by the fact that this method assumes familiarity with the institutions' total publications lists as found in annual reports, etc. As it is much easier to establish a selective publication platform on the basis of the Science Citation Index Expanded Database, many investigators often ignore publication patterns studies, even though smaller or larger amounts of the publications will be lost. Publication loss is strongly dependent on the main discipline involved (engineering, science, humanities and social sciences, etc.) and ranges from a few % to more than 50% (Bourke & Butler 1996; Cronin *et al.* 1997; Council for Medical Sciences 2002). Publication loss varies within the health science disciplines too, being biggest within the humanities and social science health research (Porta 1996) and smallest within the more experimentally based disciplines such as clinical biochemistry and immunology. Citation studies are, therefore, valid within most, but not all of the health science disciplines (Wallin 2004).

Publication analysis

Publication patterns analysis (the types and numbers of publications) allows for an exhaustive division into different kinds of publications: research papers (theses and dissertations), books, chapters in books, contributions to anthologies, various kinds of articles in scientific journals, patents, etc.). Analysis of publication patterns permits precise comparisons between institutions in terms of how international their publication patterns are, how often they publish in publications with quality control (e.g. peer review), and how frequently they publish in journals that are considered to be flagships of the discipline, and how their publications are distributed between scientific and high-level synthetic (secondary) literature (reviews, chapters from textbooks, etc.). Publication patterns studies are especially used within the humanities and the social science disciplines (Nederhof *et al.* 1989; Nederhof and Zwaan 1991; Nederhof *et al.* 2001), where the expected spectrum of document types is much larger than within medical research and therefore presents significant problems for citation analysis (Glänzel & Schoepflin 1999). But publication patterns analysis could be used to a greater extent within health science disciplines, as this method represents a much more broadly-based statement about the usefulness and the quality of institutional publications, than if the analysis is restricted to citation analysis alone (Wallin 1999 & 2004). A publication patterns analysis is a necessary starting point for an analysis of the extent to which institutions publish in national compared with international journals, or in journals with no, low, middle or high ISI Journal Impact Factor relative to typical values for that discipline.

Bibliographing

Bibliographing (examination of the number of bibliographies indexing a publication) is another infrequently used bibliometric method. A journal's bibliographic count consists of the number of abstracting and indexing (A&I) bibliographies or databases that currently register the contents of the journal in question. This method also assumes that the institutions' publication basis is very well established. In principle, all forms of publications can be investigated bibliographically, but to put the institutions on a comparable basis it is more appropriate to limit the technique to scientific publications in journals, series and the like. The biggest current register of journals etc., **Ulrich's Periodicals Directory,** registered 716 A&I bibliographies or databases in the 39th edition. Bibliographing could be used to a larger extent within the health science disciplines, as also this method gives a more broadly based estimate of the quality of institutional publications than citations analysis alone (Wallin 2004). Bibliographing of journal publications represents an important statement about the availability and visibility of these journals for the international research society, a factor which furthermore cannot be controlled by the journal's editorial team (Yue & Wilson 2004). Bibliographing can also be considered as an expression of a publication's degree of internationalisation. Interpretation is somewhat restricted, however, by the fact that it is not possible to normalize bibliographic counts across the disciplines, because the average values have not yet been determined. Bibliographing is thus one of several important statements about the quality of a publication. No significant correlation can be proved conclusively between bibliographing and the size of the Journal Impact Factor in a pool of journals from all disciplines (Wallin 2004), whereas "journal visibility" (bibliographing) has a significant influence on the journal citation impact based on neurological journals (Yue & Wilson 2004). These two journal evaluation factors (i.e. bibliographing and journal impact factors) thus seem to reflect different aspects of the quality of journals.

Journal Impact Factor

Another journal evaluation factor, the ISI Journal Impact Factor (JIF), has on the other hand attracted conspicuous interest. *The Journal Impact Factor is calculated by dividing the number of current citations to items published in the two previous years by the total number of articles & reviews published in the two previous years* (ISI def.). JIF is calculated annually by the Institute for Scientific Information (ISI), and the results are published in the **Journal Citation Reports (JCR)**. JIF was originally only envisaged as an aid for scientific libraries for the evaluation of their choice of scientific journals. JIF is today used by anyone who analyses and evaluates research with regard to assessment, prioritising the allocation of funds, etc. This has led to a widespread misuse of JIF. In the quest for scientific quality, many consider JIF the "deus ex machina", which can solve all problems. But it is naive reductionism to substitute a qualitative concept like research quality with one single quantitative parameter. JIF is calculated on the basis of a 2-year, current time window, which means that JIF favourizes journals with a steep citation curve, i.e. journals which are cited intensively for a very short period after their publication, but which after a few years are not cited any longer. Journals with a flatter citation curve, however, which reach the same citation counts over a much longer period, get a much lower JIF due to the standard calculation method. JIF thus favourizes journals within disciplines with a fast distribution of knowledge or, said in another way, journals which quickly become obsolete (Chew & Relyea-Chew 1988; Vinkler 1991; Glänzel & Schoepflin 1995). It is possible to study a journal's ageing distribution in the form of ISI Cited Half-Life. *The Cited Half-Life is the number of journal publication years going back from the current year, which accounts for 50% of the total citations received, by the cited journal in the current year* (ISI def.). This factor is used much more infrequently than JIF in bibliometric analyses, just as there are only few investigations that operate with a JIF based on a longer time window. Cited Half-Life can be used to calculate the so-called Cited Half-Life Impact Factor, which appears to be more correct than JIF (Sombatsompop *et al.* 2004).

In addition to the calculation method with 2-year time windows there is another even more serious problem affecting the size of the JIF. Different disciplines have very different citation patterns, which are correlated with the size of the pool of citeable literature: *The probability of an article in a certain field being cited two years after publication-date, is proportional to the average number of two-year old references per article in that field* (Moed *et al.* 1985a). The disciplinary citation patterns are also related to different citation traditions, the length of articles and the number of references in these, the average amount of articles per author and variations in the use of indirect citations (see later). Journals from disciplines with great attention from other disciplines will attract more citations than journals from the more closed (isolated) disciplines, without this in any way reflecting a difference in quality (Moed *et al.* 1985a; Vinkler 1991; Schubert & Braun 1996). Mathematics for example has a completely different citation pattern from pharmacology. But there is also an obvious difference between the medical disciplines both with regard to attention from other disciplines and citation patterns. These conditions result in very great differences between the disciplines with regard to the size of their JIF. Among the journals with the highest counts of JIF are the general medical journals Nature Medicine, JAMA, New England Journal of Medicine, Lancet and Annals of Internal Medicine, although the articles in these are not necessarily of a higher quality than articles in the discipline's most important (sub)specialised journals, which typically have a much lower JIF. One must strongly caution against evaluations which assume that journals with a higher JIF are always better than journals with a lower JIF.

JIF is also strongly influenced by formal factors in journals, since there is a direct linearity between JIF and the number of articles (Rousseau & Van Hooydonk 1996; Tsay & Ma 2003). JIF is also affected by the composition of the contents (articles, reviews, letters etc.) (Moed & Van Leeuwen 1995), as well as the formal alteration of content (Van Leeuwen *et al.* 1999) such as the presence or the absence of supplementum-numbers (Zetterström 2002). There is, furthermore, a significant relationship between journal accessibility (i.e. circulation or subscription, language, online versions) and journal citation impact (Yue & Wilson 2004). All these factors can be manipulated by the journal's editorial team. Since certain types of articles attract large numbers of citations, and if editors are driven to maximize JIF, as seems to be the case in more and more journals, they will be motivated to publish only certain types of articles (e.g multicenter-trials and clinical guidelines), and reject others. This could result in articles important to segments of the scientific community not being published because they are not highly citeable. The standard calculation of JIF is incorrect (Van Leeuwen *et al.* 1999) and arguments have been made for another and more correct calculation method, without this, however, having had any effect on the originator, ISI (Moed & Van Leeuwen 1995).

The JIF calculated and published by ISI is not, therefore, a simple expression of scientific quality, but on the contrary reflects the very different publication and citation patterns within the scientific disciplines (Moed *et al.* 1985a & 1985b), including the interdisciplinary and the specific and partly manipulable conditions applying to the journals (Yue & Wilson 2004). It is therefore scientifically incorrect to use (aggregate) JIF without taking the disciplines in consideration. The widespread misuse of JIF has received several sharp comments from bibliometricists, health science researchers and editors of scientific journals (Meenen 1997; Seglen 1997; Gisvold 1999; Whitehouse 2002; Lawrence 2003). Hecht *et al.* (1998) express the essence of this criticism: *We conclude that the ''impact factor'' is not a measure of true impact. Granted the ''impact factor'' is very appealing because it is a simple quantitative measure. The trouble is that it is a quantitative measure of a quality that cannot be quantified*. The criticism, which is based on a survey of numerous methodologically incorrect studies with erroneous conclusions, is unfortunately justified.

JIF can be an excellent bibliometric tool, however, but this requires the use of various kinds of normalization (Van Leeuwen *et al.* 1999; Glänzel & Moed 2002; Van Leeuwen & Moed 2002). Firstly, allowance must be made for the types of publication in the investigated journals (articles, reviews, letters etc.), as these publications do not have the same average citation counts. Secondly, intra- and interdisciplinary normalizations are necessary regarding the journals in the investigated discipline(s) (Van Leeuwen & Moed 2002; Vinkler 2002); this includes the "Journal to Field Impact Score", which is field-normalized. This important normalization means that a journal's JIF is compared to the citation average in the field(s) it covers (Van Leeuwen & Moed 2002). Thirdly, the so-called "relative impact" (Van Hooydonk 1998) must be found by calculating the relationship between the expected citation counts (i.e. the JIF) and the counts actually attained. The latter two types of JIF normalization are absolutely decisive, because they will have the effect of making the JIFs comparable across the disciplines, but they are still ignored by many clinicians and administrators in research committees, etc. Intradisciplinary normalization can be finely tuned by using variable time windows, because also journals in the same discipline can have differing citation sequences over time (as expressed in Cited Half-Life).

The use of normalized (standardised) JIF (sJIF) is a unique statement about publication performance, which tells us how good research units are at positioning their publications in journals with the highest impact within the discipline(s) in which they do their research. This statement gives an impression of the strength of a particular research environment by reflecting a great number of factors, such as the researcher's performance, the effectiveness of research management, capacity for fruitful co-operation in the organization, abilities in international co-operation, etc. It also indicates the position of the institute in question within its own discipline (e.g. JCR Pharmacology and Pharmacy) and in relation to the institutions within other disciplines. Such a calculated and normalized sJIF gives a methodologically correct picture of how good researchers are at getting their publications included in journals with the highest JIF, not only within the discipline but also between disciplines. JIF further represents a highly restricted estimate of the publications' subsequent citation history (see below). But unfortunately the picture is blurred by the fact that the prestige of the journals and the size of their JIF are bound together to a growing extent, because: *one expects scientists to publish in the top-journals available in their fields of science (...) However the perception of these journals being ''top-journals'' is partly created by their position in rankings based on, for example, Journal Impact Factors* (Van Leeuwen *et al.* 2003).

The fight to get publications accepted for journals with high prestige and/or JIF is decided mainly by the quality of the submitted manuscripts. The "quality" of manuscripts is, however, ambiguous, which reflects both immediate objective conditions (including scientific methodology) and subjective conditions including both the cognitive plan (the scientific ideas) and aesthetic qualities (Moed *et al.* 1985b). The quality of manuscripts is, however, also influenced by their potential importance ("impact"), which is difficult to assess, as usefulness for the scientific society depends on the subsequent scientific development. The quality of manuscripts can, therefore, never be completely assessed by a few experts. This is why it is not surprising that many articles even in the most prestigious journals remain uncited, just as, on the other hand, articles in more humble (low prestige) journals can attract many citations. It is only after the articles have been exposed to the global research society's "peer review" that the original reviewers' evaluation of the manuscript's importance becomes irrelevant.

With the starting point in the expert evaluation of the quality of manuscripts and their potential importance, it is especially remarkable that the curve between the number of citations and the number of articles in scientific journals shows a very skewed distribution, so that a very few articles are cited many times, a minority a few times and the majority of articles are not cited at all. The most cited 15% of the articles account for 50% of the citations, and the most cited 50% of the articles account for 90% of the citations (Seglen 1997). This extremely important circumstance, known under the expression "the skewness of science" (Seglen 1992), is evident in both specialised and general journals (Chew & Relyea-Chew 1988; Seglen 1992 & 1994; Opthof *et al.* 2004). This has the vital consequence, *that the impact factor (*JIF) *of a scientific journal is not a totum pro parte for its individual papers* (Opthof *et al.* 2004). In theory a sufficiently large, random sample of journal articles will correlate with the corresponding average of journal impact factors (JIF), but institutional publications (i.e. from separate institutions) will never represent a random sample, on the contrary they must be expected to reflect significant differences between institutions in research performance (and citation counts) and accordingly in the international impact of the institutions (Seglen 1994). At macro level this phenomenon means that *the average citation rates in national subfields are to a large extent determined by only a few highly cited papers* (Aksnes & Sivertsen 2004), which should be taken into consideration when interpreting bibliometric comparisons which do not correct for this phenomenon.

The hypothetical citing (expressed as the average JIF figure) is thus an extremely poor estimate of the individual articles' actual citation history. The correlation between the publications' JIF and their actual citation history is actually so bad that the cumulated JIF can only account for a quarter (27%) of the citations of the articles in a pool (Larsen 1999). So the use of JIF must at the same time be accompanied by actual citation counts, not just to evaluate the publication strength, but also to evaluate the publications' actual importance (citation impact). Used in this way, the simultaneous use of JIF scoring and the investigation of the individual publications' citation history will give a broader assessment of a research group's quality than citation history used alone. The conclusion of the discussion about use (and misuse) of JIF is, therefore, that JIF is not a straightforward substitute for scientific quality, but when correctly used is an excellent tool in the bibliometric toolbox, which can throw light on certain aspects of a research group's performance with regard to the publication process.

Citation processes

Science Citation Index (SCI) was started in 1963. The use of citation indexing was intended to be a supplement to literature searching in the classical bibliographies **Medline**, **Chemical Abstracts** etc., and only to a very limited extent a bibliometric research tool. This pattern, however, changed completely when SCI became available electronically in 1974. Since then the use of citation analysis has developed explosively and is no longer confined to informetric journals, but has spread to the scientific and health science journals. Unfortunately, the applied literature is full of misapplications of methods and mistaken interpretations of results indicating ignorance of the major development in methods, including standardisation of indicators (Glänzel 1996; van Raan 2004), which has taken place in fundamental bibliometric research over the past 25 years.

SCI is based on scientific citation practice but a number of investigations has shown that there can be very many different reasons for citing older literature (Moravcsik & Murugesan 1975; Brooks 1985 & 1986; Cano 1989; White & Wang 1997; Case & Higgins 2000; Kim 2004). It has to be concluded that the citation styles of researchers cover such a broad spectrum of motives that *indiscriminate use of citation counts for evaluative purposes* has to be dismissed (Maricic *et al.* 1998). Citing styles also vary between researchers from the same discipline (Cronin & Shaw 2002), because every researcher, so to speak, establishes his own "*citation identity*" (White 2001 & 2004). It is obvious that more knowledge about citing styles would be desirable, not least on how they vary between the journals of different disciplines or countries (Murugesan & Moravcsik 1978). The investigation's spectrum of results emphasizes that citing results cannot be transferred between different scientific disciplines (or countries) without normalization. There are also differences between the disciplines with regard to the use of indirect citations, a phenomenon called "indirect-collective referencing" (ICR). Usage of indirect referencing is about citing a whole series of references just by citing one specifically referenced publication, and adding a phrase such as "and references cited therein". ICR is found, for example, in 17,2% of the articles in physics journals (Szava-Kovats 2001). ICR can result in a considerable loss of data in citation analyses, as this kind of reference is not normally registered in citation databases.

If a relationship between citation frequency and research quality does exist, this relationship is not likely to be linear. The relationship between research quality and citation frequency probably takes the form of a J-shaped curve, with exceedingly bad research cited more frequently than mediocre research (Bornstein 1991). This approximately J-shaped distribution of citedness has been confirmed as regards the relationship between peer evaluations reflected in scholarly book reviews and the citation frequencies of reviewed books, as caused by a skewed allocation of negative citations (Nicolaisen 2002). It is a fully accepted practice to deliberately cite "poor" work (Kostoff 1998); "poor" is put in quotation marks because this quality evaluation is typically done by the citing author, even though the evaluation may well have been shared with other authors in a certain subject relationship (MacRoberts & MacRoberts 1984). "Negational" citing is, however, especially related to scientific disciplines with a more critical discourse.

Several high-level but contradictory theories have been put forward about citation styles, which all seem to contain important descriptions of researchers' motives for citing literature. They reflect different views of scientific research. The widespread normative citation model, which arises from a conception of science's observations and objective description of the laws of nature as absolute truth, is in the purest sense based on the view that *such a norm might be the expectation that authors acknowledge prior work in an accurate manner and true to the original author's intentions* (Small 2004). The normative model becomes a problem, however, because the scientific process does not normally allow for objective, documented principles for systematically correct searches and critical selection (or deselection) and use of the older scientific literature. This means that the character of the argumentation, personal preferences and random selection can easily play an important part. This attitude is equivalent to the social constructivist model. *Social constructivism maintains that scientific knowledge is soci-*

ally constituted, that 'facts' are made by us. Thus it challenges the objectivity of knowledge (Def.: Routledge Encyclopedia of Philosophy). In its fullest consequence this theory leads to a total rejection of the use of citation analysis for research evaluation (MacRoberts & MacRoberts 1986 & 1996). The social constructivist model sees referencing as "tools of persuasion", because *one can therefore argue that the scientific 'norm' that one should cite the research on which one's work depends, may not be a product of a pervasive concern to acknowledge 'property rights', but rather may arise from scientist's interest in persuading their colleagues by using all the resources available to them, including those respected papers which can be cited to bolster their own arguments* (Gilbert 1977). The normative model is, however, supported by the fact that *citations to very famous names are roughly balanced by citations to obscure ones, and most citations go to authors of middling reputation* (White 2004), and furthermore that a *significant positive effect of cited article cognitive content and cited article quality* (Baldi 1998) as well as the fact that *in basic science the percentage of 'authoritative' references decreases as bibliographies become shorter* (Moed & Garfield 2004). But even though this model must be presumed to represent the best description, there are good reasons to maintain that *both normative and constructivist perspectives have positions in citation analysis* (Yue & Wilson 2004).

It is a well-known observation that the same errors (among others the spelling of author names and journal references, especially volume, issue and page numbers), often repeat themselves in different authors' lists of references (Broadus 1983; Moed & Vriens 1989; Abt 1992). If this is investigated sufficiently closely, the conclusion seems to be that a very great deal (70–90%) of the cited literature has not been read at all by the citers but that the citations have just been moved on from one scientific publication to another (Simkin & Roychowdhury 2003 & 2005). These results confirm the Matthew effect ("cumulative advantage process") discussed below, but also places a decisive question mark against the connection between quality and citation counts, by suggesting that *simple mathematical probability, not genius, can explain why some papers are cited a lot more than others* (Simkin & Roychowdhury 2003) and that *a large amount of citations can be a result of a stochastic process rather than a consequence of intrinsic merit* (Simkin & Roychowdhury 2005) An alternative explanation of how the same error appears in different lists of references is, however, that researchers fail to find their original copy of an article they wish to cite, and therefore re-use the first and the best reference to it in another relevant article.

A certain amount of empiric evidence exists for a correlation between citation parameters and various measurable expressions of scientific quality: Firstly, there is expert assessment of the articles (Virgo 1977; McAllister *et al.* 1980; Lawani and Bayer 1983; Lawani 1986; Abt 2000a). Secondly, there is expert assessment of the researchers (Small 1977; Meho & Sonnenwald 2000). Thirdly, there are the British Research Assessment Exercise ratings within genetics, anatomy, archaeology, information science respectively (Oppenheim 1995 & 1997). Fourthly, peer reviews have been made of research groups in economy (Nederhof & van Raan 1993) or research programs in condensed matter physics (Rinia *et al.* 1998). Fifthly, studies have been made of state subsidies within academic chemical research (Moed & Hesselink 1996). Sixthly, expert assessments have been made of journals within health care administration (Dame & Wolinsky 1993). However, it can not be denied that all these studies in principal are flawed by a fundamental problem of methodology: *The problem with these correlations is that the two parameters (peer review and number of citations) are probably not independent* (Opthof 1997). That the relationship is much more complex is evident from other, empirical investigations, which cannot confirm such a relationship between citation parameters and measurements of quality (Lewison 2002; Nisonger 2002; West & McIlwaine 2002; Lee *et al.* 2003; Gupta *et al.* 2004).

The conclusion must therefore be that there is no unambiguous relationship between citation parameters and scientific importance and/or quality. If we then assume that there must after all be some sort of relationship, an explanation for these clearly conflicting investigations must therefore be that the relationship is so complex that we have difficulty in capturing it with the tools available to us. This reflects the fundamental problem of reducing the concept of research quality to one objective, manipulable size. The only clear relationships are found between objective parameters, i.e. a significant correlation between the number of citations and the length of the articles (Abt 2000b) and a significant difference between cited and uncited articles, as uncited articles have respectively a lower average of authors, and a lower number of references (Stern 1990). There is, however, one indispensable requirement before citation data can be used for research quality evaluation: *All bibliometric indicators of the quality of research produced are normally based on either direct citation counts, or various citation surrogates such as journal influence or Journal Impact Factor, and all are subject to one extremely important constraint: the data must be normalized for differences in field, subfield, and sometimes specialty parameters* (Narin & Hamilton 1996). How this is done will be examined below.

Citation mining

Identifying the full scope of impacts produced by scientific publications must include both the directly identifiable research impacts (e.g. citation counts) and the secondary or indirect impacts on the scientific user community. This can be achieved by means of citation mining, which has been developed by Kostoff *et al.* (2001). The user community is characterized by the articles in the SCIE that cite the original research articles and that cite the succeeding generations of these articles as well. The original set of articles may be publications by one particular author, by a research unit, an institution, or some other well-defined entity. Text mining, which selects relevant information from the chosen generations of citing articles by means of computational linguistics, is an essential starting point for citation mining. Text mining of the SCIE free (e.g. title and abstract) and non-free (e.g. address) text fields illuminates the trans-citational thematic relationships between all the analysed publications. Citation mining thus combines the features of citation bibliometrics and text mining to track and document the impact of research on the larger scientific community across many generations of publications, but this elegant method has not yet come into widespread use owing to its technical difficulty (Kostoff *et al.* 2001; del Rio *et al.* 2002).

Citation analysis

Citation analyses can be carried out at different aggregation levels, i.e. **1**. from lists of publications at institutional level, **2**. using institutional names (intermediary level) or **3**. at country level. The former is precise and exhaustive, as it is based on the knowledge of official institutional publication lists. This also applies in principle to the latter, as citation databases operate with a consistent control of countries ("Geolocation"). This control cannot be complete, however, since approximately 17% of the records in Science Citation Index Expanded (SCIE) are not provided with complete author addresses, because the authors do not give this information in their publications. This defect is especially frequent in articles with multiple authors (Wallin 2004). On the other hand, citation databases have no institutional name control at all, which leads to a particularly important source of error in many bibliometric investigations at the intermediate aggregation level. Certain of the secondary citation products, such as **ISI Essential Science Indicators**, admittedly operate with some institutional name control, but its mechanism is obscure and/or often fails. Many citation analyses take place at the intermediate aggregation level, i.e. by searching at the institution's address in SCIE, which, based on experience, is often very problematical, because it is impossible to ensure in practice that all of an institution's possible names have been checked, regardless whether the institution has established fixed name control or not. Citation analyses at the intermediate aggregation level may therefore ignore a small pool of highly cited publications and thus give a false picture of an institution's true citation performance.

The practical implementation of citation analyses requires appreciable expertise, as there are many serious errors in citation databases which complicate the analyses (Moed & Vriens 1989; Hood & Wilson 2003). These errors have for example been documented by a chemist with a starting point in his own authorship (Reedijk 1998). Among the most important are: **1.** errors in author name control, including errors and inconsistencies in corporate authors (authors that act both independently and as a group), which amongst other things often has the effect of causing underrating of citations to corporate authors (MacKinnon & Clarke 2002). **2.** errors in journal data (volume, issue and page numbers), and **3**. errors in journal series, which include supplements (supplementum).

The error rate in citation databases is estimated to be about 7–9%. Unfortunately, these errors do not turn up randomly but vary in a systematic way, which means that ignorance of them can distort citation analyses to such a degree as actually to invalidate them (Moed & Vriens 1989; Moed 2002).

In principle, citation analyses can be applied as soon as a publication has been registered in SCIE, but a robust estimate requires at least two to three years of observations dependent on the discipline's ageing pattern. The analyses can take place both synchronously and diachronously or in other words they can be retrospective and/or prospective (Ingwersen *et al.* 2000). The prospective (diachronous) method is the most appropriate to use: *Since ageing of scientific literature has to be considered a real "process" (not only in the mathematical sense) with maturing and decay, this process can best be reflected by a measure of the use of (scientific) information ... through the change of citedness in time* (Glänzel 2004). Citation analyses provide citation frequency as well as the absolute number of citations for a specific number of articles (citation counts). Citation frequency (whether publications are cited or not) gives an indication of the overall attention given to a body of publications. Together with the citation impact (citation counts divided by number of publications), therefore, this gives a somewhat broader picture of citation conditions than citation impact alone. A scientifically based use of these citation parameters, however, invariably requires several types of normalization.

Patent citation statistics

Patents can be viewed upon as materializations of technologies (Tijssen 2001), and patent statistics are therefore an important tool for technological performance assessments: *the use of patent counting, clustering, and citation analysis in the evaluation of corporate, industry-wide, and national technological activity* (Narin *et al.* 1984). Patent statistics are also used in financial assessment of industrial enterprises, normally on the basis of their patent portfolios. These uses of patent statistics are regarded today as basic tools for evaluating the quality and economic value of patented technologies, for monitoring science and technology portfolios, for investment analyses etc. An excellent and up to date survey of methods and applications is found in "Handbook of quantitative science and technology research. The use of publication and patent statistics in studies of S&T systems" Eds. Moed, H.F., Glänzel, W., Schmock, U. Kluwer Academic Publishers 2004 (ISBN 1-4020-2702-8).

The underlying assumption in patent citation analysis is that a highly cited patent (a patent which is referred to by many subsequently issued patents) is likely to contain technological advances of particular importance that has led to numerous subsequent technological improvements. It follows *that a company (or institution) whose patent portfolio contains a large number of highly cited patents is generating high quality technology* (Narin *et al.* 2004).

Patent citation analysis is based on the examiner references that appear on the front page of patents (Michel & Bettels 2001). However, the patent applicant may also cite references within the body of the patent, and even though these may show great similarity to those on the front page (Narin *et al.* 1997), it is not certain that excluding them will not affect the outcome of the analysis (Meyer 2000a). Examiner references consist both of references to other patents and references to scientific literature (i.e. patent and non-patent references). The non-patent references are a wide variety of references to journal papers, meetings, books, and many non-scientific sources (e.g. industrial standards). Selection of these references is performed in a standardised way in the various national patent organisations (e.g. EPO, JPO and USPTO), but there are significant differences between these patent organisations in their use of patent and non-patent references, so that direct comparisons of citation data from different patent organisations are not possible (Michel & Bettels 2001).

The selection of patent references for the front page of a patent is done by experts in a centralised and consistent process, and therefore, as might be expected, *numerous validation studies have shown the existence of a strong positive relationship between patent citations and technological importance* (Narin *et al.* 2004). Important similarities can be seen between literature bibliometrics and patent bibliometrics. There is one decisive similarity between citation patterns in scientific literature and patent literature, namely that a relatively small number of patents are cited very frequently, whereas the majority are rarely cited: *a relatively small number of patents receiving many citations, and the majority being very lightly cited, if cited at all ... This skewness of the citation distribution is a common characteristic of papers and patents* (Narin 1994). This highly skewed patent citation distribution corresponds perfectly with the highly skewed citation of scientific papers ("the skewness of science") (Seglen 1992; Narin & Hamilton 1996).

An important aspect of patent references is their role in linking science with technology. Patents contain a steadily growing number of citations to the scientific literature, indicating an increasing linkage between technology and public science, a good example being pharmaceutical patents (Narin & Olivastro 1992; Narin *et al.* 1997). Patent citation analysis is therefore increasingly being used to analyse national and international knowledge and technology transfer: *the citations from articles to articles, from patents to patents, and from patents to articles provide indicators of intellectual linkages between the organizations that are producing the patents and articles, and knowledge linkage between their subject areas* (Narin *et al.* 1994). However, a significant national component may be detected in the patent science linkage, since patents from a given country cite literature from the inventor's own country more frequently than literature from other countries (Narin & Olivastro 1992; Narin *et al.* 1997).

Narin's interpretations, envisaging a one-way relationship between technology and public science ("linkage bibliometric"), has been criticized, amongst other things, for being based solely on the front page references, for restricting the citation analyses to references covered by SCIE, and especially because the use of references in patents must be assumed to have additional purposes to their use in scientific literature (Meyer 2000a & 2000b; Tijssen 2001). When using non-patent references to analyse the flow of knowledge between scientific literature and patents, therefore, there are good reasons for assuming: *that the citation links do not indicate science-dependence of technology, but should be taken as an indication of the multifacetted interplay between science and technology* (Meyer 2000a).

Last but not least, the correct use of patent citation analysis presupposes a normalization of the data: *A major advantage inherent in quantitative technological performance assessment is that these techniques allow for cross-disciplinary normalization. With these techniques it is possible to explicitly allow for the differences in patent citation habits with different parts of the patent literature ... With cross disciplinary normalization is it possible to fully account for the variations within each patent class* (Narin *et al.* 1984).

Normalization of citation data

The following paragraphs present the various normalization techniques used for scientific literature.

Firstly there is author citation standardization of publications with 2 or more authors (Harsanyi 1993). The choices can be:

1. To give the first author all the citations: only the first of the N authors is given full credit for the multiauthored article (first author counting) or
2. To give all the authors the same number of citations: each of the N authors is given full credit for the multiauthored article (normal counting).

   Both these methods are often used. In order to obtain a larger degree of fairness one can also choose
3. To give the authors a proportion of the number of citations corresponding to the number of authors (fractional counting) or
4. To give the authors a proportion of the number of citations so that the share falls according to the position in the list of authors (proportional counting). These forms of author citation standardisation with variants can obviously give very different results (Van Hooydonk 1997; Egghe *et al.* 2000; Trueba & Guerrero 2004). The correctness of these methods are continually discussed but even though there is much in favour for using them, a consensus has not yet been reached on any one of them, when the substantial formal and technical problems are taken into consideration (Lange 2001). This is connected with the fact that although the first author

will usually have the greatest responsibility for the publication, there may be a number of reasons for the order of listing of the authors, and thus to the weighting due to the individual authors, because the intellectual author contribution does not necessarily fall linearly with the position in the list of authors but may vary systematically within it: *The mean contribution percentages decreased greatly from first to second to last to middle authors* (Hwang *et al.* 2003). Whatever the method chosen, it is important to be aware of the principles behind them.

The next normalization method concerns document types. As already stated, there is a great difference in citation impact between the different types of document: monographs, journal articles, etc. (Bourke & Butler 1996; Cronin *et al.* 1997; Glänzel & Schoepflin 1999). Citation analyses can, of course, be performed on all kinds of publications, whether they are monographical (dissertations, books, anthologies, etc.) or periodical (journals) but as the citation patterns will vary greatly between them (Cronin *et al.* 1997), comparable analyses can only be made using the 5.900 journals that are indexed for SCIE. The varying citation patterns arise from the differences both in the choice and use of citations in monographic and periodic literature and in the extent to which journals cite themselves. All journals cite themselves, the so-called “self-citing rate” (Billesbølle *et al.* 1988; Porta 1996; Meenen 1997), and even though only 20% of all journals have a self-citing rate of more than 20% (McVeigh 2002) the citation counts for articles in journals which are not indexed for citation databases will always be too low. Comparable citation analyses can thus only be performed on publications from journals indexed for citation databases. Completely correct citation analyses ought also to distinguish between formal document types: editorial material, original articles, reviews, letters, meeting abstracts, etc., as these do not have the same citation rate (citation impact) (Van Leeuwen *et al.* 2003). Only the document types article, letter, note, review and proceedings can be regarded as important conveyors of relevant scientific information. Since meeting abstracts have proved not to be citeable items, this type should normally be omitted from bibliometric studies (Glänzel 1996). In principle one ought also to distinguish between the different subtypes of journal articles, as there can be a colossal difference between the citing of methodology articles and ordinary articles (Van Leeuwen *et al.* 1999). Such differences can also appear between articles with differing research design; thus the international multicenter-trials seem to have a conspicuously high citation impact (Wallin 2004). This presumably reflects the existence of a significant correlation between numbers of authors per article and the expert-assessed quality of articles within oncology (Lawani 1986). It is technically possible to distinguish between the different document types in databases, but in practice it is impossible to distinguish between most subtypes of journal articles, so the disciplines which for example are influenced by fast methodology development and, therefore, have many methodology articles, must necessarily dominate in terms of the size of the average citation rate over the disciplines without such an influence.

Interdisciplinary normalization of citation data

This leads to the third and most important form of citation normalization, namely disciplinary normalization. There is an extraordinarily large difference in citation impact between disciplines (Kostoff 2002), but many bibliometric investigations nevertheless ignore this circumstance. There is, in fact, a spectacular difference between the citation count a publication can expect to get within the individual scientific disciplines, regardless whether this arises from intradisciplinary (van Raan 1996) or interdisciplinary conditions (Rinia *et al.* 2001). The disciplinary citation counts are published continuously in several products from ISI, e.g. in **ISI Essential Science Indicators**, which are all based on the classification of journals according to disciplines (“Subject Category”). The total average citation impact of 8.03 for 22 main scientific disciplines for the period 1993–2003 give the following rankings when citation impact is broken down according to disciplines: 1. Molecular Biology & Genetics (23.38), 2. Immunology (18.40), 3. Neuroscience & Behavior (15.14), 4. Biology & Biochemistry (14.42) and 5. Microbiology (12.08). The lowest positions (20 to 22) are held by Engineering (2.73), Mathematics (2.38) and Computer Science (2.20). It is clear that citation impact does not reflect the importance of these disciplines and still less the quality of research in them, but on the contrary it reveals important differences in the disciplinary citation behaviour as well as in the amount of citeable publications, the so-called Matthew effect (Merton 1968) or “cumulative advantage” principle (Price 1976). In the citation world this effect relates to the fact that citing a publication singles it out for other authors, which increases its chances of being cited again. The differences can also be observed in the proportion of uncitedness, where Computer Science lies at the top among the main disciplines with 44.52% uncited publications and Immunology at the bottom with just 13.17% uncited publications (*ISI Science Watch* January/February 1999).

If we look at **ISI Institutional Performance Indicators** for the period 1981–97 and limit the focus to 43 health science disciplines, the same marked differences can be observed in the citation impact, with values ranging from 3.23 (Orthopedics & Sports Medicine) to 21.79 (Immunology) (Wallin 1999). If we investigate **JCR 2003 Science Edition** there is an opportunity to rank the journals according to the size of their JIF, and the same extreme difference is observed between journals with regard to the distribution as well as the size of their JIF. Taking the top-100 journals in terms of JIF value, then some of the disciplines are overrepresented and a large number of the other disciplines are either underrepresented or completely absent. The five most frequently represented disciplines are: Biochemistry & Molecular Biology (14 journals), Cell Biology (10), Genetics & Heredity (8), Neurosciences (8) og Immunology (7). It is not scientifically tenable to claim that such distributions represent the "quality" of these journals, and it is therefore very clear that the results of citation analyses can only be valid within the investigated disciplines, and that the results must not be transferred between the disciplines. Every comparison of institutions with publication activity in several disciplines is, therefore, inconclusive unless the citation data has been normalised. This also applies to comparisons between countries (King 2004). At macro level (countries) important differences in the countries' publication patterns determine the publication activity in respectively international and nationally-oriented journals, a condition which distorts every comparison of the citation impacts of different countries, unless corrections are made for this phenomenon (Zitt *et al.* 2003).

The interdisciplinary normalization can be performed in several ways, all of which have advantages and disadvantages (Schubert & Braun 1993 & 1996). The following methods can be recommended:

1. The publishing journal as reference standard. This is done by normalizing the measured publications' citation rate with the average citation rate for all the publications in the journal for the same volume – also called diachronic analysis (Ingwersen *et al.* 2000). Ideally, it should be limited to publications of exactly the same type. This is a simple, effective and very widespread method which gives a precise normalization, although an objection to this method is that in principle it rewards researchers who publish noteworthy manuscripts in journals with low scientific prestige and/or low JIF (Lewison 2002). This objection is rather theoretical, however, as all researchers must be expected to aspire to publish in journals with as high scientific prestige (Bonitz & Scharnhorst 2001; Van Leeuwen *et al.* 2003) or size of JIF (Garfield 1998; Bordons *et al.* 2002) as possible. This is supported at macro level by the fact that all countries reaching a higher than expected citation rate (based on publishing journal as reference standard), have a citation rate above the world average (Schubert & Braun 1993). This method is, however, less suitable for multidisciplinary journals (e.g. Nature) and general medical journals (e.g. The Lancet), which include articles from several disciplines.
2. Subject specific normalization. This is done by normalizing the citation rate in proportion to a defined cluster of articles on the same subject. It can be carried out in practice by using Related Records in SCIE or Related Articles in PubMed. In principle the method gives a more precise normalization than the method 1, as many journals carry articles within a number of sub-specialisations, and the method focuses on specific subject relationships. However, defining the necessary cluster's size is problematic, and the method is difficult in practice. It is, therefore, most suitable for investigations of small pools of articles. Subject specific normalization can possibly be made more exact by using a pool of *thematically related papers* (Kostoff 2002), but this, too, is a technically very complex method.
3. A set of cited journals as reference standard. In opposition to the other methods this method builds on the authors' own selection, as it uses the journals appearing in the investigated publications' lists of literature as the standard. This method has not become particularly widespread, however.
4. Reference standards based on (sub)disciplinary journal classification. This very much used method is discussed in the section about JIF. However, the method is not an ideal solution either, because an overall classification heading (for example "Neurosciences") will draw together journals from sub-disciplines which may have different publication and citation patterns.

Besides the above-mentioned types of normalization, there could be an argument for making corrections for the proportion of self-citations, i.e. citations by the author to publications by the same author, so that these self-citations are not included. An estimate of author self-citations in a Danish medical faculty is 10,4% (Wallin 1999). The definition can be widened to include authors who appear in a collective of authors, i.e. the set of co-authors of the citing paper and that of the cited one are not disjoint, but share at least one author (van Raan 1998b). This correction is technically extremely difficult to perform.

The rate of self-citation varies between countries and between disciplines. Some substantiated figures, based on the widened definition, are 36% for Norwegian science (Aksnes 2003), around 40% for the world reference standard, 25% for the world standard for biomedical research, and 19–21% for clinical and experimental medicine, depending on the choice of medical specialties (Glänzel & Thijs 2004). The proportion of self-citations is much lower in journals with a high impact or visibility than in journals with a relatively low impact or visibility (Glänzel *et al.* 2004; Glänzel & Thijs 2004). However self-citations are inherently neither good nor bad, although they are typically assumed to be self-serving. If an author has done the seminal work in a field, then self-citing would be most appropriate and a perfect reason for the retention of the proportion of self-references. Differences in self-citation frequency can in fact be a bibliometric indicator by itself (Glänzel *et al.* 2004).

Citation analyses are usually made at the previously mentioned aggregation levels. They can be supplemented with advantage by other sampling methods at citation data level. Among the possibilities should be named:

1. Restriction to the highest cited publications from one country, for example 1% (Van Leeuwen *et al.* 2003; King 2004). This value can be related to all the publications from institutions or to the number of their researchers, the allocation of funds, etc. As the limitation is at a numeric level, the method is good for collecting all the

highly cited publications from institutions as long as they have identifiable address data.

2. How big a proportion the investigated countries have in the highest-cited publications distributed by discipline. This value can either be put in relation to all the publications from that country in the individual disciplines or in relation to the population, gross national earnings, research funding, etc. The method is robust, because citation databases operate with country control.
3. Analysis of the number of publications in journals with the highest JIF calculated for the individual disciplines, for general medical journals, and for multidisciplinary top journals. This value could be related to all the publications in the individual disciplines, number of research-employees, etc. Finally, one could distinguish between citations from the same journal and from all other journals, giving citations from other journals (and publications) higher values. It is also possible to analyse the proportion of the citations from high prestige journals, as this proportion does have a special power of statement of *scientific excellence* (Van Leeuwen *et al.* 2003).

Bibliographic coupling

Bibliographic coupling can be used to investigate and render visible the relationship between publications. It is based on the coupling of elements in bibliographic records, and the coupling strength is measured by the number of coupling units between them. Although bibliographic coupling was first studied by Kessler (1963), it was Fano who first formulated the concept (Fano 1956). The idea behind bibliographic coupling was later mathematically formalised (Sen & Gan 1983). An important new methodological approach was given by Glänzel & Czerwon (1996). Greatest attention is usually given to the coupling of citation data, where one can distinguish between 1. co-citation analyses and 2. co-reference analyses (Small 1974).

1. Co-citation analyses are based on the frequency of simultaneous citation of two publications, on the principle that the more researchers who cite the same two publications, the bigger the probability that the double citation is not a chance event, but expresses a type of subject relationship between the cited publications. Thus, relationships within research areas and between scientific disciplines can be made visible, relationships that often give surprising results, as they cannot usually be studied in any other way. The relationships can with advantage be illustrated graphically by means of "bibliometric cartography". It may however take time before a 'critical mass' of papers has been created on a new research topic, sufficient to produce the highly cited publications on which the co-citation mapping is based. Consequently this method might fail when applied to young disciplines or new topics (Hicks 1987).
2. Co-reference analyses build on the commonality between publications in the form of two or more shared literature references. The number of shared literature references determines the precision of the commonality between publications. Commercial software exists that performs both co-citation and co-reference analyses. The use of co-citation analysis is nowadays the most dominant method for the investigation of the structure of scientific communication, despite the methodological discussions (Gmur 2003), but co-reference analyses have found a commercial application in the essential Related Records function in the SCIE database.
3. Co-occurrence analyses are about the coupling of other data like subject data and address data, including country ("Geolocation"). Knowledge discovery or data mining is the process of extracting patterns from bibliographic records. Text mining is data mining applied to natural language data, which can be applied to bibliometric problems and patent statistics. Text mining offers a variety of approaches for extracting information and knowledge from textual data (Leopold *et al.* 2004). Co-heading analysis has been developed by Todorov & Winterhager (1990) and Todorov (1992), and is based on the co-occurrence of subject headings. This approach shows some advantage as compared to other bibliometric coupling methods. Analyses of subject data are best performed in Medline and Embase, which have controlled classification, whereas country information can best be studied in SCIE (or PsycINFO), which contains all author addresses. Co-occurrence analyses are especially used for studying co-operation between institutions and countries, and co-occurrence analyses represent an important statement on the strength of international co-operation relations. Co-occurrence analyses have found practical use in the essential Related Articles function in PubMed. Web of Science 7.0 (SCIE etc.) contains certain possibilities for co-occurrence analyses.

Concluding remarks

Bibliometric methods make it possible to evaluate unlimited amounts of publications from institutions or countries. It is tempting to substitute these quantitative estimates with definitive, undifferentiated statements about scientific quality. However, a true assessment of scientific quality is not obtained just by analysing the publications' citation impact, but ought also to include peer review of the societal effects of research. If bibliometric methods are used to their full extent, then a soundly based statement can be achieved about the degree of internationalisation within research, the researchers' ability to publish in journals with high prestige, as well as about the publication type, visibility and subsequent impact of the publications in the scientific community. An indispensible requirement for all this, however, is a detailed knowledge of the methods' weaknesses and meticulous standardization of indicators and normalization of results. Unfortunately, this normalization of bibliometric data is all too frequently neglected in current biomedical use. Bibliometricians have recognised this problem, and have taken the initiative on the issue (Proceedings of the Workshop on "Bibliometric Standards" 1996; van Raan 2004).

Acknowledgements

The author wish to thank the two anonymous reviewers for their useful comments and suggestions for the improvement of the paper.

References

Abt, H. A.: What fraction of literature references are incorrect. *Publications of the Astronomical Society of the Pacific* 1992, **104**, 235–236.

Abt, H. A.: Do important papers produce high citation counts? *Scientometrics* 2000a, **48**, 65–70.

Abt, H. A.: The reference-frequency relation in the physical sciences. *Scientometrics* 2000b, **49**, 443–451.

Aksnes, D. W.: A macro study of self-citation. *Scientometrics* 2003, **56**, 235–246.

Aksnes, D. W. & G. Sivertsen: The effect of highly cited papers on national citation indicators. *Scientometrics* 2004, **59**, 213–224.

Baldi, S.: Normative versus social constructivist processes in the allocation of citations: A network-analytic model. *Amer. Sociol. Rev.* 1998, **63**, 829–846.

Billesbølle, P., P. J. Hindhede Blyme & O. Glerup: Citeringsfrekvens i Ugeskrift for Læger – Hvilke tidsskrifter blev citeret i 1986? *Ugeskrift for Læger* 1988, **150**, 2650–2652.

Bonitz, M. & A. Scharnhorst: Competition in science and the Matthew core journals. *Scientometrics* 2001, **51**, 37–54.

Bordons, M., M. T. Fernandez & I. Gomez: Advantages and limitations in the use of impact factor measures for the assessment of research performance in a peripheral country. *Scientometrics* 2002, **53**, 195–206.

Bornstein, R. F.: The predictive validity of peer review: A neglected issue. *Behav. Brain Sci.* 1991, **14**, 138.

Bourke, P. & L. Butler: Publication types, citation rates and evaluation. *Scientometrics* 1996, **37**, 473–494.

Broadus, R. N.: An investigation of the validity of bibliographic citations. *J. Amer. Soc. Inf. Sci.* 1983, **34**, 132–135.

Brooks, T. A.: Private acts and public objects: An investigation of citer motivations. *J. Amer. Soc. Inform. Sci.* 1985, **36**, 223–229.

Brooks, T. A.: Evidence of complex citer motivations. *J. Amer. Soc. Inform. Sci.* 1986, **37**, 34–36.

Cano, V.: Citation behavior: Classification, utility, and location. *J. Amer. Soc. Inform. Sci.* 1989, **40**, 284–290.

Case, D. O. & G. M. Higgins: How can we investigate citation behavior? A study of reasons for citing literature in communication. *J. Amer. Soc. Inform. Sci.* 2000, **51**, 635–645.

Chew, F.S. & A. Relyea-Chew: How research becomes knowledge in radiology: An analysis of citations to published papers. *Amer. J. Roentgenol.* 1988, **150**, 31–37.

Council for Medical Sciences of the Royal Netherlands Academy of Arts and Sciences. The societal impact of applied health research. Towards a quality assessment system. ISBN 90-6984-360-9, 5-49. Amsterdam, KNAW Koninklijke Nederlandse Akademie van Wetenschappen, 2002.

Cronin, B.: Metatheorizing citation – Comments on theories of citation? *Scientometrics* 1998, **43**, 45–55.

Cronin, B.: Semiotics and evaluative bibliometrics. *J. Document.* 2000, **56**, 440–453.

Cronin, B. & D. Shaw: Identity-creators and image-makers: Using citation analysis and thick description to put authors in their place. *Scientometrics* 2002, **54**, 31–49.

Cronin, B., H. Snyder & H. Atkins: Comparative citation rankings of authors in monographic and journal literature: A study of sociology. *J. Document.* 1997, **53**, 263–273.

Dame, M. A. & F. D. Wolinsky: Rating journals in health care administration. The use of bibliometric measures. *Med. Care* 1993, **31**, 520–524.

del Rio, J. A., R. N. Kostoff, E. O. Garcia, A. M. Ramirez & J. A. Humenik: Phemomenological approach to profile impact of scientific research: Citation mining. *Advances in Complex Systems* 2002, **5**, 19–42.

Egghe, L., R. Rousseau & G. Van Hooydonk: Methods for accrediting publications to authors or countries: Consequences for evaluation studies. *J. Amer. Soc. Inform. Sci.* 2000, **51**, 145–157.

Fano, R. M.: Information theory and the retrieval of recorded information. In: *Documentation in action. Based on the Proceedings of the 1956 Western Reserve Conference "Documentation in Action"*. Eds.: J. H. Shera, A. Kent and J. W. Perry. Reinhold Publishing Corporation. Chapman & Hall Ltd., New York, London, 1956, pp. 238–244.

Garfield, E.: From citation indexes to informetrics: Is the tail now wagging the dog? *Libri* 1998, **48**, 67–80.

Gilbert, G. N.: Referencing as persuasion. *Social Studies of Science* 1977, **7**, 113–122.

Gisvold, S. E.: Citation analysis and Journal Impact Factors – Is the tail wagging the dog? *Acta anaesthesiol. Scand.* 1999, **43**, 971–973.

Glänzel, W.: The need for standards in bibliometric research and Technology. *Scientometrics* 1996, **35**, 167–176.

Glänzel, W.: Towards a model for diachronous and synchronous citation analyses. *Scientometrics* 2004, **60**, 511–522.

Glänzel, W. & H. J. Czerwon: A new methodological approach to bibliographic coupling and its application to the national, regional and institutional level. *Scientometrics* 1996, **37**, 195–221.

Glänzel, W. & H. F. Moed: Journal impact measures in bibliometric research. *Scientometrics* 2002, **53**, 171–193.

Glänzel, W. & U. Schoepflin: A bibliometric study on ageing and reception processes of scientific literature. *J. Inform. Sci.* 1995, **21**, 37–53.

Glänzel, W. & U. Schoepflin: A bibliometric study of reference literature in the sciences and social sciences. *Information Processing & Management* 1999, **35**, 31–44.

Glänzel, W. & B. Thijs: World flash on basic research – the influence of author self-citations on bibliometric macro indicators. *Scientometrics* 2004, **59**, 281–310.

Glänzel, W., B. Thijs & B. Schlemmer: A bibliometric approach to the role of author self-citations in scientific communication. *Scientometrics* 2004, **59**, 63–77.

Gmur, M.: Co-citation analysis and the search for invisible colleges: A methodological evaluation. *Scientometrics* 2003, **57**, 27–57.

Gupta, A. K., K. Nicol & A. Johnson: Pityriasis versicolor: Quality of studies. *J. Dermatolog. Treat.* 2004, **15**, 40–45.

Hall, J.C. & C. Platell: Half-life of truth in surgical literature. *Lancet* 1997, **350**, 1752.

Harsanyi, M. A.: Multiple authors, multiple problems – Bibliometrics and the study of scholarly collaboration: A literature review. *Library & Information Science Research* 1993, **15**, 325–354.

Hecht, F., B. K. Hecht & A. A. Sandberg: The journal "Impact Factor": a misnamed, misleading, misused measure. *Cancer Genet. Cytogenet.* 1998, **104**, 77–81.

Hicks, D.: Limitations of cocitation analysis as a tool for science policy. *Social Studies of Science* 1987, **17**, 295–316.

Hood, W. W. & C. S. Wilson: Informetric studies using databases: opportunities and challenges. *Scientometrics* 2003, **58**, 587–608.

Hwang, S. S., H. H. Song, J. H. Baik, S. L. Jung, S. H. Park, K. H. Choi & Y. H. Park: Researcher contributions and fulfillment of ICMJE authorship criteria: Analysis of author contribution lists in research articles with multiple authors published in Radiology. International Committee of Medical journal Editors. *Radiology* 2003, **226**, 16–23.

Ingwersen, P., B. Larsen & I. Wormell: Applying diachronic citation analysis to research program evaluations. In: *The Web of Knowledge: A Festschrift in Honor of Eugene Garfield*. Eds.: B. Cronin & H. B. Atkins. Information Today & American Society for Information Science., Medford N.J., 2000, pp. 373–387.

Kessler, M. M.: Bibliographic coupling between scientific papers. *Amer. Document.* 1963, **14**, 10–25.

Kim, K.: The motivation for citing specific references by social scientists in Korea: The phenomenon of co-existing references. *Scientometrics* 2004, **59**, 79–93.

King, D. A.: The scientific impact of nations. *Nature* 2004, **430**, 311–316.

Kostoff, R. N.: The use and misuse of citation analysis in research evaluation – Comments on theories of citation? *Scientometrics* 1998, **43**, 27–43.

Kostoff, R. N.: Citation analysis of research performer quality. *Scientometrics* 2002, **53**, 49–71.

Kostoff, R. N., J. A. del Rio, J. A. Humenik, E. O. Garcia & A. M. Ramirez: Citation mining: Integrating text mining and bibliometrics for research user profiling. *J. Amer. Soc. Inform. Sci. Technol.* 2001, **52**, 1148–1156.

Lange, L. L.: Citation counts of multi-authored papers – First-named authors and further authors. *Scientometrics* 2001, **52**, 457–470.

Larsen, B.: Journal Impact Factors i forskningsevaluering. En undersøgelse af sammenhængen mellem Journal Impact Factor og antal modtagne citationer. CIS Rapport 10, 1–23. København, Center for Informetriske Studier. Danmarks Biblioteksskole, 1999.

Lawani, S. M.: Some bibliometric correlates of quality in scientific research. *Scientometrics* 1986, **9**, 13–25.

Lawani, S. M. & A. E. Bayer: Validity of citation criteria for assessing the influence of scientific publications: New evidence with peer assessment. *J. Amer. Soc. Inform. Sci.* 1983, **34**, 59–66.

Lawrence, P. A.: The politics of publication – Authors, reviewers and editors must act to protect the quality of research. *Nature* 2003, **422**, 259–261.

Lee, J. D., K. J. Vicente, A. Cassano & A. Shearer: Can scientific impact be judged prospectively? A bibliometric test of Simonton's model of creative productivity. *Scientometrics* 2003, **56**, 223–233.

Leopold, E., M. May & G. Paass: Data mining and text-mining for science & technology research. In: *Handbook of quantitative science and technology transfer. The use of publication and patent statistics in studies of S&T systems*. Eds.: H. F. Moed, W. Glanzel & U. Schmoch. Kluwer Acdaemic Publishers, Dordrecht/Boston/London, 2004, pp. 187–213.

Lewison, G.: Researchers' and users' perceptions of the relative standing of biomedical papers in different journals. *Scientometrics* 2002, **53**, 229–240.

Leydesdorff, L.: Theories of citation? *Scientometrics* 1998, **43**, 5–25.

MacKinnon, L. & M. Clarke: Citation of group-authored papers. *Lancet* 2002, 1513–1514.

MacRoberts, M. H. & B. R. MacRoberts: The negational reference: or the art of dissembling. *Social Studies of Science* 1984, **14**, 91–94.

MacRoberts, M. H. & B. R. MacRoberts: Quantitative masures of communication in science: A study of the formal level. *Social Studies of Science* 1986, **16**, 151–172.

MacRoberts, M. H. & B. R. MacRoberts: Problems of citation analysis. *Scientometrics* 1996, **36**, 435–444.

Maricic, S., J. Spaventi, L. Pavicic & G. Pifat-Mrzljak: Citation context versus the frequency counts of citation histories. *J. Amer. Soc. Inform. Sci.* 1998, **49**, 530–540.

McAllister, P. R., R. C. Anderson & F. Narin: Comparison of peer and citation assessment of the influence of scientific journals. *J. Amer. Soc. Inform. Sci.* 1980, **31**, 147–152.

McVeigh, M. E.: Journal self-citation in the Journal Citation Reports – Science edition (2002): A citation study from the Thomson Corporation. The Thomson Corporation, 2004.

Meenen, N. M.: Der Impact-Faktor – Ein zuverlässiger scientometrischer Parameter? *Unfallchirurgie* 1997, **23**, 128–134.

Meho, L. I. & D. H. Sonnenwald: Citation ranking versus peer evaluation of senior faculty research performance: A case study of Kurdish scholarship. *J. Amer. Soc. Inform. Sci.* 2000, **51**, 123–138.

Merton, R. K.: The Matthew effect in science. *Science* 1968, **159**, 56–63.

Meyer, M.: Does science push technology? Patents citing scientific literature. *Research Policy* 2000a, **29**, 409–434.

Meyer, M.: What is special about patent citations? Differences between scientific and patent citations. *Scientometrics* 2000b, **49**, 93–123.

Michel, J. & B. Bettels: Patent citation analysis – A closer look at the basic input data from patent search reports. *Scientometrics* 2001, **51**, 185–201.

Moed, H. F.: The Impact-Factors debate: the ISI's uses and limits. *Nature* 2002, **415**, 731–732.

Moed, H. F., W. J. M. Burger, J. G. Frankfort & A. F. J. van Raan: The application of bibliometric indicators: Important field- and time-dependent factors to be considered. *Scientometrics* 1985a, **8**, 177–203.

Moed, H. F., W. J. M. Burger, J. G. Frankfort & A. F. J. van Raan: The use of bibliometric data for the measurement of university research performance. *Research Policy* 1985b, **14**, 131–149.

Moed, H. F. & E. Garfield: In basic science the percentage of 'authoritative' references decreases as bibliographies become shorter. *Scientometrics* 2004, **60**, 295–303.

Moed, H. F. & F. T. Hesselink: The publication output and impact of academic chemistry research in the Netherlands during the 1980s: Bibliometric analyses and policy implications. *Research Policy* 1996, **25**, 819–836.

Moed, H. F. & T. N. Van Leeuwen: Improving the accuracy of Institute for Scientific Information's Journal Impact Factors. *J. Amer. Soc. Inform. Sci.* 1995, **46**, 461–467.

Moed, H. F. & M. Vriens: Possible inaccuracies occurring in citation analysis. *J. Inform. Sci.* 1989, **15**, 95–107.

Moravcsik, M. J. & P. Murugesan: Some results on the function and quality of citations. *Social Studies of Science* 1975, **5**, 86–92.

Murugesan, P. & M. J. Moravcsik: Variation of the nature of citation measures with journals and scientific specialties. *J. Amer. Soc. Inform. Sci.* 1978, **29**, 141–147.

Narin, F.: Patent bibliometrics. *Scientometrics* 1994, **30**, 147–155.

Narin, F., A. Breitzman & P. Thomas: Using patent citation indicators to manage a stock portfolio. In: *Handbook of quantitative science and technology research. The use of publication and patent statistics in studies of S&T systems*. Eds.: H. F. Moed, W. Glänzel & U. Schmock. Kluwer Academic Publishers, Dordrecht/Boston/London, 2004, pp. 553–568.

Narin, F., M. P. Carpenter & P. Woolf: Technological performance assessments based on patents and patent citations. *Ieee Transactions on Engineering Management* 1984, **31**, 172–183.

Narin, F. & K. S. Hamilton: Bibliometric performance measures. *Scientometrics* 1996, **36**, 293–310.

Narin, F., K. S. Hamilton & D. Olivastro: The increasing linkage between US technology and public science. *Research Policy* 1997, **26**, 317–330.

Narin, F. & D. Olivastro: Status report: Linkage between technology and science. *Research Policy* 1992, **21**, 237–249.

Narin, F., D. Olivastro & K. A. Stevens: Bibliometrics/Theory, practice and problems. *Evaluation Review* 1994, **18**, 65–76.

Nederhof, A. J., M. Luwel & H. F. Moed: Assessing the quality of scholarly journals in linguistics: An alternative to citation-based Journal Impact Factors. *Scientometrics* 2001, **51**, 241–265.

Nederhof, A. J. & A. F. J. van Raan: A bibliometric analysis of six economics research groups: A comparison with peer review. *Research Policy* 1993, **22**, 353–368.

Nederhof, A. J. & R. A. Zwaan: Quality judgments of journals as indicators of research performance in the humanities and the social and behavioral sciences. *J Amer. Soc. Inform. Sci.* 1991, **42**, 332–340.

Nederhof, A. J., R. A. Zwaan, R. E. De Bruin & P. J. Dekker: Assessing the usefulness of bibliometric indicators for the humanities and the social and behavioural sciences: A comparative Study. *Scientometrics* 1989, **15**, 423–435.

Nicolaisen, J.: The J-shaped distribution of citedness. *J. Document.* 2002, **58**, 383–395.

Nisonger, T. E.: The Relationship between international editorial board composition and citation measures in political science, business, and genetics journals. *Scientometrics* 2002, **54**, 257–268.

Oppenheim, C.: The correlation between citation counts and the 1992 Research Assessment Exercise ratings for British library and information science university departments. *J. Document.* 1995, **51**, 18–27.

Oppenheim, C.: The correlation between citation counts and the 1992 Research Assessment Exercise ratings for British research in genetics, anatomy and archaeology. *J. Document.* 1997, **53**, 477–487.

Opthof, T.: Sense and nonsense about the Impact Factor. *Cardiovasc. Res.* 1997, **33**, 1–7.

Opthof, T., R. Coronel & H. M. Piper: Impact Factors: No totum pro parte by skewness of citation. *Cardiovasc. Res.* 2004, **61**, 201–203.

Porta, M.: The bibliographic "Impact Factor" of the Institute for Scientific Information: How relevant is it really for public health journals? *J. Epidemiol. Community Health* 1996, **50**, 606–610.

Poynard, T., M. Munteanu, V. Ratzlu, Y. Benhamou, V. Di Martino, J. Taieb & P. Opolon: Truth survival in clinical research: an evidence-based requiem? *Ann. Int. Med.* 2002, **136**, 888–895.

Price, D. J. D.: A general theory of bibliometric and other cumulative advantage processes. *J. Amer. Soc. Inform. Sci.* 1976, **27**, 292–306.

Proceedings of the Workshop on "Bibliometric Standards". Satellite event to the 5th International Conference on Scientometrics and Informetrics. Rosary College, River Forest, Illinois (USA) Sunday, June 11, 1995. W. Glänzel, J. S. Katz & H. Moed. Scientometrics, Volume 35, Issue 2, 1996 , 166–290. 1996. Budapest, Akademiai Kiado.

Reedijk, J.: Sense and nonsense of Science Citation analyses: Comments on the monopoly position of ISI and citation inaccuracies. Risks of possible misuse and biased citation and impact data. *New J. Chem.* 1998, **22**, 767–770.

Rinia, E. J., T. N. Van Leeuwen, E. E. W. Bruins, H. G. van Vuren & A. F. J. van Raan: Citation delay in interdisciplinary knowledge exchange. *Scientometrics* 2001, **51**, 293–309.

Rinia, E. J., T. N., Van Leeuwen, H. G. van Vuren & A. F. J. van Raan: Comparative analysis of a set of bibliometric indicators and central peer review criteria – Evaluation of condensed matter physics in the Netherlands. *Research Policy* 1998, **27**, 95–107.

Rousseau, R. & G. Van Hooydonk: Journal production and journal Impact Factors. *J. Amer. Soc. Inform. Sci.* 1996, **47**, 775–780.

Schubert, A. & T. Braun: Reference standards for citation based ASsessments. *Scientometrics* 1993, **26**, 21–35.

Schubert, A. & T. Braun: Cross-field normalization of scientometric indicators. *Scientometrics* 1996, **36**, 311–324.

Seglen, P. O.: The skewness of science. *J. Amer. Soc. Inform. Sci.* 1992, **43**, 628–638.

Seglen, P. O.: Causal relationship between article citedness and journal impact. *J. Amer. Soc. Inform. Sci.* 1994, **45**, 1–11.

Seglen, P. O.: Why the Impact Factor of journals should not be used for evaluating research. *Brit. Med. J.* 1997, **314**, 498–502.

Sen, S. K. & S. K. Gan: A mathematical extension of the idea of bibliographic coupling and its applications. *Ann. Library Sci. Documentation* 1983, **30**, 78–82.

Simkin, M. V. & V. P. Roychowdhury: Copied citations create renowned papers? *Condensed Matter 0305150 (to appear in Annals of Improbable Research)* 2003, **http://arxiv.org/abs/cond-mat/10305150**.

Simkin, M. V. & V. P. Roychowdhury: Stochastic modeling of citation slips. *Scientometrics* 2005, **62**, 367–384.

Small, H.: Co-citation in the scientific literature – A new measure of the relationship between two documents. *Current Contents* 1974, 7–10.

Small, H.: A Co-citation model of a scientific specialty: A longitudinal study of collagen research. *Social Studies of Science* 1977, **7**, 139–166.

Small, H.: On the shoulders of Robert Merton: Towards a normative theory of citation. *Scientometrics* 2004, **60**, 71–79.

Sombatsompop, N., T. Markpin & N. Premkamolnetr: A modified method for calculating the impact factors of journals in ISI Journal Citation Reports: Polymer science category in 1997–2001. *Scientometrics* 2004, **60**, 217–235.

Stern, R. E.: Uncitedness in the biomedical literature. *J. Amer. Soc. Inform. Sci.* 1990, **41**, 193–196.

Szava-Kovats, E.: Indirect-collective referencing (ICR) in the elite journal Literature of physics. I. A literature science study on the journal level. *J. Amer. Soc. Inform. Sci. Technol.* 2001, **52**, 201–211.

Tijssen, R. J. W.: Global and domestic utilization of industrial relevant science: Patent citation analysis of science-technology interactions and knowledge flows. *Research Policy* 2001, **30**, 35–54.

Todorov, R.: Displaying content of scientific journals – A co-heading analysis. *Scientometrics* 1992, **23**, 319–334.

Todorov, R. & M. Winterhager: Mapping Australian geophysics – A co-heading analysis. *Scientometrics* 1990, **19**, 35–56.

Trueba, F. J. & H. Guerrero: A robust formula to credit authors for their publications. *Scientometrics* 2004, **60**, 181–204.

Tsay, M. Y. & S. S. Ma: The nature and relationship between the productivity of journals and their citations in semiconductor literature. *Scientometrics* 2003, **56**, 201–222.

Van Hooydonk, G.: Fractional counting of multiauthored publications: Consequences for the impact of authors. *J. Amer. Soc. Inform. Sci.* 1997, **48**, 944–945.

Van Hooydonk, G.: Standardizing relative impacts: Estimating the quality of research from citation counts. *J. Amer. Soc. Inform. Sci.* 1998, **49**, 932–941.

Van Leeuwen, T. N. & Moed, H. F.: Development and application of journal impact measures in the Dutch science system. *Scientometrics* 2002, **53**, 249–266.

Van Leeuwen, T. N., H. F. Moed & J. Reedijk: Critical comments on Institute for Scientific Information Impact Factors: A sample of inorganic molecular chemistry journals. *J. Inform. Sci.* 1999, **25**, 489–498.

Van Leeuwen, T. N., M. S. Visser, H. F. Moed, T. J. Nederhof & A. F. J. van Raan: The holy grail of science policy: Exploring and combining bibliometric tools in search of scientific excellence. *Scientometrics* 2003, **57**, 257–280.

van Raan, A. F. J.: Advanced bibliometric methods as quantitative core of peer review based evaluation and foresight exercises. *Scientometrics* 1996, **36**, 397–420.

van Raan, A. F. J.: In matters of quantitative studies of science the fault of theorists is offering too little and asking too much – Comments on theories of citation? *Scientometrics* 1998a, **43**, 129–139.

van Raan, A. F. J.: The influence of international collaboration on the impact of research results – Some simple mathematical considerations concerning the role of self-citations. *Scientometrics* 1998b, **42**, 423–428.

van Raan A. F. J.: Measuring science. Capita selecta of current main issues. In: *Handbook of quantitative science and technology research. The use of publication and patent statistics in studies of S&T systems*. Eds.: Moed, H. F., W. Glänzel & U. Schmock. Kluwer Academic Publishers, Dordrecht/Boston/London, 2004, pp. 19–50.

van Raan, A. F. J. & T. N. Van Leeuwen: Assessment of the scientific basis of interdisciplinary, applied research – Application of bibliometric methods in nutrition and food research. *Research Policy* 2002, **31**, 611–632.

Vinkler, P.: Possible causes of differences in information impact of journals from different subfields. *Scientometrics* 1991, **20**, 145–161.

Vinkler, P.: Subfield problems in applying the Garfield (Impact) Factors in practice. *Scientometrics* 2002, **53**, 267–279.

Virgo, J. A.: A statistical procedure for evaluating the importance of scientific papers. *Library Quarterly* 1977, **47**, 415–430.

Wallin, J. A.: Forskningen ved Sundhedsvidenskab i Odense 1966–1997. En bibliometrisk undersøgelse. Det Sundhedsvidenskabelige Fakultet. Syddansk Universitet, Odense, Denmark. 1999.

Wallin, J. A.: Forskningen ved Det Sundhedsvidenskabelige Fakultet Syddansk Universitet 1998–2000. En bibliometrisk undersøgelse. Det Sundhedsvidenskabelige Fakultet. Syddansk Universitet, Odense, Denmark, 2004.

West, R. & A. McIlwaine: What do citation counts count for in the field of addiction? An empirical evaluation of citation counts and their link with peer ratings of quality. *Addiction* 2002, **97**, 501–504.

White, H. D.: Authors as citers over time. *J. Amer. Soc. Inform. Sci. Technol.* 2001, **52**, 87–108.

White, H. D.: Reward, persuasion, and the Sokal hoax: A study in citation identities. *Scientometrics* 2004, **60**, 93–120.

White, M. D. & P. L. Wang: A qualitative study of citing behavior: contributions, criteria, and metalevel documentation concerns. *Library Quarterly* 1997, **67**, 122–154.

Whitehouse, G. H.: Impact Factors: facts and myths. *Eur. Radiol.* 2002, **12**, 715–717.

Yue, W. P. & C. S. Wilson: Measuring the citation impact of research journals in clinical neurology: A structural equation modelling analysis. *Scientometrics* 2004, **60**, 317–332.

Zetterström, R.: Bibliometric data: A disaster for many non-american biomedical journals. *Acta Paediatrica* 2002, **91**, 1020–1024.

Zitt, M., S. Ramanana-Rahary & E. Bassecoulard: Correcting glasses help fair comparisons in international science landscape: Country indicators as a function of ISI database delineation. *Scientometrics* 2003, **56**, 259–282.