

# The journal impact factor as a parameter for the evaluation of researchers and research\*

K.-F. Kaltenborn and K. Kuhn

*Institut für Medizinische Informatik. Fachbereich Medizin und Klinikum. Philipps-Universität Marburg. Marburg, Germany*

## ABSTRACT

The journal impact factor (IF), which is published annually by the *Institute for Scientific Information*<sup>®</sup> (USA), is meanwhile in widespread use as a scientometric parameter for the evaluation of research and researchers in Germany and other European countries. The present article subjects the IF to critical analysis. It first deals with processes of production, transfer, and use of medical knowledge, because the IF intervenes in these processes on account of its reflexivity. Secondary effects of the IF resulting from its reflexivity are discussed with the focus on the level of the author, the journal and the medical discipline as well as on social knowledge processes in society. In addition, the extent to which the IF is appropriate for evaluating the quality of a specific article, of a journal or of individual and collective research achievements is discussed. The present article calls for a) research evaluation in accordance with the recommendations of the *Deutsche Forschungsgemeinschaft* (German Research Council, DFG) and of the *Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften* (Association of the Scientific Medical Societies, AWMF); and b) for more intensive occupation with and organization of medical knowledge processes.

**Key words:** Impact factor. Research. Evaluation. Bibliometrics. Communication media.

Received: 07-01-04.

Accepted: 12-01-04.

*Correspondence:* PD Dr. Dr. Karl-Franz Kaltenborn. Institut für Medizinische Informatik. Fachbereich Medizin und Klinikum. Philipps-Universität Marburg. Bunsenstrasse 3. D-35037 Marburg (Germany). Phone: (+49)/(0) 6421/2863683). e-mail: kaltenbo@mail.uni-marburg.de

\*Institute of Medical Informatics, Medical Faculty, University of Marburg, Germany. This article was first published in German in the journal *Medizinische Klinik* (2003, 98, 153-169). The authors are most grateful to the publisher, Urban & Vogel (Munich), for having kindly given permission for the English version to be printed here. Translation: Susan Griesbach in cooperation with Karl-Franz Kaltenborn.

## RESUMEN

El factor de impacto de las revistas (*journal impact factor* – IF), que publica anualmente el *Institute for Scientific Information*<sup>®</sup> (EE.UU.), se emplea generalizadamente como parámetro cuantitativo para evaluar las investigaciones y a los investigadores en Alemania y otros países europeos. El presente artículo somete al IF a un análisis crítico. Primero trata de los procesos de producción, comunicación y empleo del conocimiento médico, pues el IF interviene en estos procesos en virtud de su reflexividad. Se exponen los efectos secundarios del IF, que surgen de esta reflexividad, centrándose en el nivel del autor, de la revista y de la disciplina médica, así como en los procesos de conocimiento social de la sociedad. Además, se discute el grado en que el IF es adecuado para evaluar la calidad de un artículo específico, de una revista o de los logros científicos individuales o colectivos. El presente artículo demanda a) una evaluación de las investigaciones de acuerdo con las recomendaciones del *Deutsche Forschungsgemeinschaft* (Consejo de Investigaciones Científicas alemán, DFG) y de la *Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften* (Asociación de Sociedades Médicas Científicas, AWMF); y b) una dedicación más intensiva a los procesos del conocimiento médico y una mayor organización de los mismos.

**Palabras clave:** Factor de impacto. Investigación. Evaluación. Bibliometría. Medios de comunicación.

---

Kaltenborn K-F, Kuhn K. The journal impact factor as a parameter for the evaluation of researchers and research. *Rev Esp Enferm Dig* 2004; 96: 460-476.

---

## INTRODUCTION

According to a survey carried out at Germany's medical faculties, the journal impact factor (IF) published by the *Institute for Scientific Information*<sup>®</sup> (ISI<sup>®</sup>) is in common use as a scientometric parameter for the evaluation of research and researchers. In more than two-thirds of all faculties the IF served, together with

supplementary funding, as a criterion for the awarding of university funds; in the majority of faculties this indicator played a role in the decision-making process underlying university appointments and led to a postdoctoral qualification (habilitation) tailback of up to two years among young academics (1-4). In Finland, IF-based research evaluation has already attained judicial status, and in the corresponding allocation of funds, one IF-point is equivalent to approximately US \$ 7000 (5).

In Germany and in Europe in general, the IF is of greater significance than in the USA despite being published by a private American company (ISI®) and being primarily US-oriented. Zach Hall, vice-chancellor for research at USCF and former director of the *National Institute of Neurological Disorders and Stroke*, is quoted in a *Nature Neuroscience* editorial (6) as saying that it is the leading universities and research institutions in the USA that make least use of the IF as an evaluation instrument. In a survey of US, Canadian and European anesthesiologists, the IF was considered by a far higher proportion of the Europeans than of the North Americans (81 vs 38%) to be of importance to their academic career; in addition, the proportion of institutions using the IF as a basis for allocating funds was greater in Europe than in North America (56 vs 13%) (7). At the same time, however, there are numerous reports from the USA and from other countries worldwide of the IF being misused (6,8-23).

It is the discrepancy between value and use of the IF and its relevance to specific scientific cultures that have provided the motivation to deal with this topic in the present article. Prior to the IF being defined and its mode of calculation being presented, medical knowledge processes are to be briefly discussed, for this will allow the status of the IF in the world of medical information to be accurately presented and more critically appraised.

### KNOWLEDGE PROCESSES IN THE MEDICAL WORLD AND IN SOCIETY: THE GENERATION, TRANSFER AND UTILIZATION OF MEDICAL KNOWLEDGE

The complex structure of medical knowledge and its social construction are to be seen as essential medico-immanent and medico-transcendent specifics and determinants in the production of medical knowledge. The outcome of the complexity of medical knowledge is that the description and explanation of medical matters, e.g. diseases, are subject to reference being made to a wide range of scientific contents. These scientific contents are located in a hierarchical knowledge structure extending from a lower, atomic level to the highest level covering the patient as a whole (24) (Table I). Because of its complex knowledge structure, the field of medicine is confronted with a pronounced individuality and a diversity of pathologic phenomena, and with numerous transitional and borderline states between health and

**Table I. Hierarchical levels of medical descriptions (after 24)**

Level	0	Patient as a whole
Level	-1	Major patient part: e.g. chest, abdomen, head
Level	-2	Physiologic system: e.g. cardiovascular system, respiratory system, digestive system
Level	-3	System part, or organ: e.g. heart, major vessels, lungs, liver
Level	-4	Organ part, or tissue: e.g. myocardium, bone marrow, cerebellum
Level	-5	Cell: e.g. epithelial cell, fibroblast, lymphocyte
Level	-6	Cell part: e.g. cell membrane, organelles, nucleus
Level	-7	Macromolecule: e.g. enzyme, structural protein, nucleic acid
Level	-8	Micromolecule: e.g. glucose, ascorbic acid, vitamin B
Level	-9	Atoms or ions: e.g. sodium, iron

illness (24-27). The numerous scientific levels can be classified into three larger areas, according to which medical research is organized: basic research (lower levels), disease-oriented research (medium levels) and patient-oriented research (top level)<sup>1</sup>.

The premises of medical knowledge production are currently undergoing a crucial change: knowledge production in basic and disease-oriented research is becoming increasingly subject to an economic perspective. The outcome is that medical knowledge is losing its traditional character as public property and is tending to become a commodity. Patents granted in the field of genetic technology for the protection of *intellectual property* are an expression of the commercialization of medical knowledge: economically prized, yet socially disputed (28-33). By contrast, patient-oriented knowledge cannot be protected and marketed to the same extent as *intellectual property* (although proposals to this end have been submitted) and continues to function as public property (31,32). It is in this context that the social construction of medical knowledge too is manifested: Basic research and disease-oriented research represent the preferred scientific research fields, while patient-oriented research tends to be sidelined. Because of the shortage of patient-oriented research, many basic research findings can be put to only limited clinical use, if any; to this extent, this disparity among research areas proves to be a limiting factor to medical progress (31,32,34). With respect to the transfer of knowledge, the following aspects are of significance: knowledge generated within the framework of biomedical research does not make a direct contribution towards better clinical care or improved health status but is first evaluated, synthesized and aggregated in a complex scientific process (25-27,35-38) (Fig. 1). Publication forms

<sup>1</sup>Blois (24) has already pointed out that the medical knowledge structure proposed by him would have to be expanded by additional levels in order to cover social and societal aspects, too.

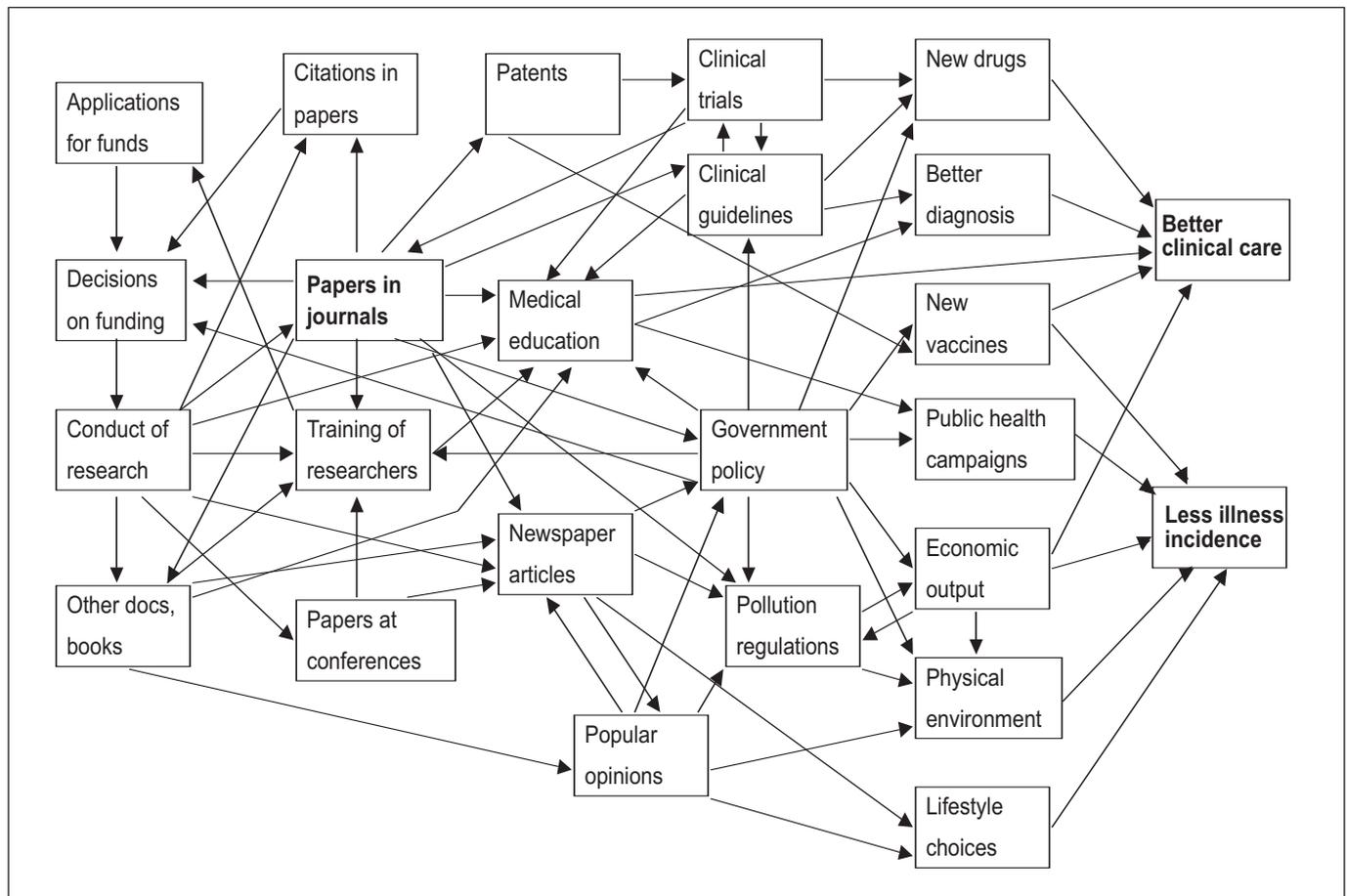


Fig. 1.- Processes of knowledge generation and knowledge transfer: from research to knowledge application in curative and preventive medicine (by kind permission of Prof. Dr. G. Lewison [37]).

such as articles in specialist journals or monographs, while playing an important role in the aggregation and transfer of medical knowledge, have a different function in the knowledge process: basis-oriented biomedical research institutions may conclude a substantial part of their curiosity-driven research activities with the presentation of their findings in journals. Application-oriented and clinically oriented institutions take recourse to this published stockpile of knowledge in order to forge ahead with the knowledge process, for example by performing clinical studies, formulating guidelines, or developing specific diagnostic or therapeutic techniques (16,37,39, 40). Special attention should be paid to four aspects:

1. There is a non-reciprocal citation relationship between basic research disciplines and application-oriented clinical disciplines: application-oriented clinical disciplines take recourse in their quotations to basic knowledge – but basic disciplines do not to application-oriented and clinical knowledge contents.

2. An evaluation and selection of objects of knowledge takes place in the knowledge process; i.e., not all objects of knowledge are taken up and integrated into the extensive medical knowledge hierarchy for subsequent use in curative or preventive medicine.

3. For application-oriented and clinical disciplines, it is not only the presentation of knowledge in scientific papers but also “knowledge-incorporating artifacts” such as hospital information systems (medical informatics), surgical techniques with a corresponding instrumentarium (surgery), or filling techniques and materials (dental medicine) that may be end products of their work. These “artifacts” can undoubtedly be presented in papers and followed up bibliometrically, but such “artifacts” tend to varying degrees to have an intrinsic value not reflected completely in bibliometric indices, so that application-oriented progress is characterized to a certain extent by bibliometric imperceptibility (16,39-41). The dispersal of knowledge based on such “artifacts” and its implementation demand not only its presentation in publications but also additional means of imparting knowledge, such as extended periods of practical experience, concrete work in the department concerned, acquisition of intuition, etc.

4. Biomedical basic knowledge is of international relevance, whereas application-oriented clinical knowledge is directed rather towards national medical cultures. Clinical medicine is largely –as stated by Lindner (38)– a national medicine. In other words, the significance of knowledge relevant to national medical cultures increases with the

**Table II. Computation of the IF for the journal *Revista Española de Enfermedades Digestivas* (source: JCR 2002). IF: 0.594**

Cites in 2002 to articles published in:	2001 = 29	Number of articles published in:	2001 = 47		=	<b>0.594</b>
	2000 = 31		2000 = 54			
	1 + 0 = 60		1 + 0 = 101			
	Calculation: $\frac{\text{Cites to recent articles}}{\text{Number of recent articles}}$		$\frac{60}{101}$			

application orientation and exploitation of knowledge in curative and preventive medicine (1,42-47) (Fig. 1).

## BASIS, DEFINITION AND CALCULATION OF THE IMPACT FACTOR

The *Science Citation Index Expanded*<sup>TM</sup> (SCI), *Social Sciences Citation Index*<sup>®</sup> (SSCI) and *Arts & Humanities Citation Index*<sup>®</sup> form the basic databases from which the ISI<sup>®</sup> creates further chargeable information media on a secondary basis. These include the *Journal Citation Reports*<sup>®</sup> (JCR) with the IF as well as, for example, the well known *Current Contents Connect*<sup>®</sup> (CC Connect<sup>®</sup>) and further products (2,48)<sup>2</sup>.

The IF, which is re-calculated each year, is defined as the quotient from the citations that a journal has received for the preceding 2 years, divided by the number of citable documents<sup>3</sup> in that journal in the same period (Table II). The IF thus shows the relative citation frequency of an average article in a journal in relation to a limited 2-year time window.

## EVALUATION OF THE IMPACT FACTOR IN THE MEDICO-SCIENTIFIC LITERATURE

### Positive evaluation of the impact factor

In the medico-scientific literature there are only few examples where IF-based researcher evaluation is explicitly advocated, even if the fundamental desirability of an objective evaluation standard for the reliable and valid rating of scientific achievements is not denied.

Calza & Garbisa (49), Brody (50), and de Marchi & Rocchi (51) have the most positive attitude towards the IF being used. On the basis of their empirical study, de Marchi & Rocchi (51) recommend the IF as an instrument for monitoring as well as for evaluating research achievements. These recommendations are, however, in contrast with the limitations of the study, whose design and extremely low response rate hardly constitute a suitable foundation for these far-reaching interpretative conclusions. Despite his

positive fundamental attitude towards the IF as a bibliometric standard, Garfield (52) calls for the impact factor to be used "wisely" and gives an urgent warning in numerous publications against an uncritical approach.

A survey carried out among anesthesiologists showed them to have a less negative attitude towards the IF than suggested in the generally critical literature on the IF; nevertheless, the majority of the surveyed anesthesiologists called for an improvement in the IF (7).

### Critical evaluation of the impact factor

As the IF is constructed mathematically from the data material of the SCI and the SSCI, some criticisms apply equally to the two databases and the IF, but others specifically to the IF.

### Criticism of the citation databases and the impact factor

*Criticism of the lack of transparency in the selection process of source journals.* The abundance of scientific literature, which is on the increase worldwide and currently comprises some 100,000 to 126,000 journals, gives rise in the production of information databases to the fundamental problem of what literature to select (8,20-23,42-44,53). The journals selected and analyzed for the ISI<sup>®</sup> databases, entailing high prestige and many other advantages, are referred to as *source journals*, while the non-analyzed journals, the *cited-only journals*, are neglected. However, many scientific journals are not evaluated for the SCI/SSCI, although their publication profile and standard are fully in line with the basic requirements published by the ISI<sup>®</sup> and are moreover of scientific relevance (54-58). For this reason the lack of transparency in the selection process for the source journals of the SCI/SSCI is criticized by numerous authors (1,20-23,59).

*Criticism of the language- and discipline-specific bias in the selection of source journals.* In the ISI<sup>®</sup> databases and thus in the IF too, it is the English language that prevails, with more than 95% of all citations originating from English-language papers (10,60). The criticism of the unconstrained predominance of the English language in the ISI<sup>®</sup> databases and in the IF is of a thoroughly universal nature, hardly confined to specific scientific cultures or specialties and disciplines (1,2,8,10,17,20-23,42-44,58,60-

<sup>2</sup>http://www.isinet.com/isi/products/index.html (17.12.2001).

<sup>3</sup>Cf. criticisms below.

67). As a consequence of the monopoly of English, the scientific cultures of the ancient world and of the Third World are being ignored, no matter whether in South America, Africa, India, southeast Asia, China, Japan or Russia (43). Braun et al. (68) reported that “a group of major publishers are clearly over-represented in the SCI-JCR journal base, some of them extremely heavily indeed”. This applies in particular to Elsevier. In other words, the *global players* in (medical) publishing dominate the source journals and thus the IF with their English-language publications, whereas journals published in other languages by non-multinational publishing houses are under-represented (cf. 10). The lack of informative value and validity of the IF for scientific institutions publishing in other than the English language has been empirically verified by van Leeuwen et al. (67, cf. 58).

With reference to the level of representation of medical journals issued by German publishers, which are, of course, not necessarily confined to the German language, Braun et al. (68) found no far-reaching prejudicial effect. However, when the focus is placed on journals in the German language rather than from German publishers, a clear-cut decline in German-language papers in the SCI from 5.7 to 1.4% can be observed between 1977 and 1997 (10,67). This may apply analogously, of course, to other languages (e.g. Spanish, Portuguese, French).

Another point open to criticism is the varying degree to which the different disciplines are represented in the databases and the marginalization of specific disciplines (16,20-23,69,70).

*Criticism of the focus on journals and of the neglect of other media.* Despite the small number of monographs evaluated, the two *citation indices* are to be regarded purely as journal databases. This restriction of the literature published in journal form provokes criticism of those very specialties that submit a significant proportion of their publications in monographic form, as this publication format does greater justice to their knowledge structure than publication in journals. Pantel & Mundt (3), for instance, claim that psychiatry, being at the interface between the natural sciences, humanities, and social sciences, is confronted with a complex knowledge structure that lends itself to – or even demands – the publication of research findings in the form of monographs and contributions to books. The criticism of there being a prejudice towards monographic literature, one deriving primarily but not exclusively from psychiatry, is to be taken seriously, because although monographs are less frequently published in the fields of the natural sciences and medicine, they are nonetheless of importance to the transfer of knowledge and may also record very high numbers of citations. In other specialties too – for instance in (medical) informatics – scientific communication is based only to a limited extent on the classic printed journals, with congress reports and publications being of greater significance (2,8,16,20-23,40,42,65,71-73).

### *Criticism of the impact factor*

*Criticism of the inadequate definition of citable documents in the impact factor calculation algorithm.* Although the IF algorithm set out above may seem mathematically simple and unproblematic, the calculation is subject to substantial problems resulting from an inadequate definition of citable documents. When citations are counted in the JCR, no distinction is made between citations from letters, editorials, reviews or original articles with research findings (all of which are included in the numerator of the IF calculation algorithm); however, when the articles are counted, it is only original and review articles that are taken into account (i.e. only these are included in the denominator of the IF calculation algorithm). This may give rise to discrepancies between the number of citations in the numerator and the concrete number of articles in the denominator, so that the IF may consequently be substantially distorted. If, for example, a journal publishes a large number of letters and thus acquires allowable citations, the IF rises because citations appear in the numerator, but letters are disregarded in the denominator.

Numerous critical objections in the literature are consequently leveled too at the IF calculation algorithm (5, 6,8,13,20-23,43,74-78). For the variation in the scope of citations in editorials, letters and other documents among individual journals gives rise overall to significant distortions in the IF comparison (74,75,77,79).

*Criticism of the 2-year calculation period for the impact factor.* Further criticism is leveled at the 2-year time window, as the different scientific disciplines are characterized by varying citation conventions linked with the dynamics of the production and ageing of knowledge. On the one hand, for instance, dynamic scientific fields such as biochemistry and molecular biology very rapidly produce new knowledge that is integrated with no great time lag into new knowledge structures and thus results in citations; on the other hand, knowledge contents also become more rapidly obsolete or at least obsolescent in dynamic research fields of this kind (8,20-23,60,64,72,77, 78,80).

Some empirical studies suggest that the 2-year period allows the long-term trend too to be recorded, provided it is investigated within the same discipline: Although clear-cut differences between the 2-year, 7-year and 15-year impact are revealed when the cumulative IF is calculated over a number of years, there is no great change in the intra-disciplinary ranking (10,11,18,57,76,81). However, a comparison of journals representing different disciplines such as sociology and medicine (*American Sociological Journal* vs *Lancet*) reveals that the *Lancet* has the higher IF in the short term, but the *American Sociological Journal* in the longer term (82). With respect to the *Zeitschrift für Klinische Psychologie, Psychiatrie und Psychotherapie*, Decker & Brähler (72) found that the bulk of the cited literature (93.9%) had been published prior to the 2-year time window.

In view of the fact that knowledge contents from biomedicine as well as from the social sciences and social medicine are of general medical relevance, this difference in respect of the ageing of knowledge runs like a thread through the entire sphere of medicine, and the 2-year time window favors or discriminates against individual medical disciplines with no plausible explanation. Clinical medicine too is disadvantaged by the 2-year calculation mode, for clinical studies require a longer preparatory period, so that subsequent studies linking up with this state of knowledge are not completed nor published until later. This means that a response to previous publications in the form of citations is subject to an inevitable time lag in the field of clinical medicine (1,3,83,84). Difficult though it may be to answer the complex question concerning the ageing of knowledge and the degree to which its recording is distorted in the JCR, there will certainly be general agreement with Garfield's summary (11): "*The impact factors [...] may not provide a complete enough picture for slower moving fields with longer half-lives.*"

*Further criticism of the impact factor.* Self-citations by source journals may help raise the level of the IF quite considerably; conversely, their absence entails a lower IF. The fact that such self-citations by journals are not corrected in the IF is a point frequently raised against the IF (e.g. 3,20-23,60,78,85).

One further criticism relates to the fact that review articles have a higher citation rate, resulting in review journals or journals with a high proportion of review articles being favored by this citation characteristic and being rewarded with a higher IF (13,20-23,60,77,82,86).

Citation networks based on reciprocal citation display a strong *national bias*. As Anglo-American literature is predominant in the ISI® database system, the citation practice of US academics is in the foreground and emerges as a fencing-off mentality on the part of US researchers against non-US research<sup>4</sup>. The effects of this citation practice in terms of the IF are that, yet again, irrespective of quality characteristics, North American journals are rewarded (unjustifiably) with a higher IF (1, 3,20-23,37,45,47,61,87,88) (Fig. 2).

Further distortions of the IF can be interpreted as effects of the respective research field. As basic research is more often cited than clinical research on account of the outlined knowledge processes, journals covering basic research tend to achieve a higher IF (14,20-23,45,60).

Expanding or stagnating scientific fields also display different citation characteristics: rapidly expanding research fields produce numerous new publications which,

however, have recourse to only a limited supply of citable material. The outcome is that the numerous citations raise the IF of cited journals. In stagnating research fields, the converse process is observed (20-23). In addition, the individual scientific fields are found to vary in their citation practice with respect to the average number of citations per article; this too has consequences on the level of the IF (20-23).

The availability of a journal in electronic form on the Internet enhances *per se* its IF, thus distorting the IF ranking (16,89). Further criticism is leveled at the classification of specialist categories in the JCR and at the flawed data collection in the ISI® databases, which often has an impact on specific countries (1,8,20-23,80,82,90). Another point to be censured is that the IFs fail to give an accurate reflection of the true impact of a journal with very few published but frequently cited articles; in this case the IF overrates the influence of a journal (1,12,91,92).

The listed deficits of the IF are classified as being not contingent but rather systematic in their effects, with crucial consequences on the status of science worldwide. Seglen (22) expresses this aspect most trenchantly: "*Thus, both the apparent quality lead of American science and the values of the various impact factors are, to an important extent, determined by the large volume, the self citations, and the national citation bias of American science, in combination with the short term index used by the Science Citation Index for calculating journal impact factors*".

## THE REFLEXIVITY OF THE IMPACT FACTOR

If a doctor takes a patient's temperature to obtain information for diagnostic purposes or to monitor the course of therapy, this measuring procedure will normally have no effect on the patient's temperature. There is no interaction between the measuring procedure and the parameter to be measured. The situation is quite different in the case of the IF: the measuring of the IF and its use as a scientometric indicator for research evaluation purposes not only reflect on the measuring instrument but also influence the very area that the IF claims to measure. In sociological terminology, the IF entails reflexivity. The significance and scope of this reflexivity of the IF, the "impact of the impact factor" (13,93) as it were, is hardly to be overestimated, and its dynamics vary, depending on whether the focus is on Anglo-American or non-Anglo-American authors and journals. This reactivity of the IF can be observed at three different levels: at the level: a) of the author; b) of the journal and the discipline; and c) of society.

### The reflexivity of the impact factor at the level of the author

The increased focusing of attention on the IF, its perception and establishment as a crucial criterion of evalua-

<sup>4</sup>The perceived national citation bias may be due to more profound factors that should be taken seriously and are of significance to global medical knowledge transfer. For the reasons for the national citation bias may be linked with the problem of confidence in medical knowledge produced by foreign researchers and institutions unknown to the respective authors (cf. 27,46).

ISI JOURNAL CITATION reports<sup>®</sup> Powered by ISI Web of Knowledge<sup>™</sup>

HOME HELP

2002 JCR Science Edition

## JOURNAL SUMMARY LIST

Selection: **GASTROENTEROLOGY & HEPATOLOGY**

Sorted by: Impact Factor

Journals 1 - 20 (of 45) Page 1 of 3

Ranking is based on your journal and sort selections.

Mark	Rank	Abbreviated Journal Title (linked to full journal information)	ISSN	2002 Total Cites	Impact Factor	Immediacy Index	2002 Articles	Cited Half-life
<input type="checkbox"/>	1	<a href="#">GASTROENTEROLOGY</a>	0016-5085	44966	13.440	2.429	361	6.8
<input type="checkbox"/>	2	<a href="#">HEPATOLOGY</a>	0270-9139	30310	9.825	1.212	363	5.3
<input type="checkbox"/>	3	<a href="#">GUT</a>	0017-5749	19722	6.323	0.969	352	6.4
<input type="checkbox"/>	4	<a href="#">SEMIN LIVER DIS</a>	0272-8087	2232	5.954	0.385	39	4.5
<input type="checkbox"/>	5	<a href="#">J HEPATOL</a>	0168-8278	10558	4.974	0.880	208	4.6
<input type="checkbox"/>	6	<a href="#">AM J GASTROENTEROL</a>	0002-9270	16245	3.953	0.572	428	5.1
<input type="checkbox"/>	7	<a href="#">LIVER TRANSPLANT</a>	1527-6465	2178	3.786	0.661	165	2.8
<input type="checkbox"/>	8	<a href="#">AM J PHYSIOL-GASTR L</a>	0193-1857	10957	3.346	0.537	287	5.5
<input type="checkbox"/>	9	<a href="#">INFLAMM BOWEL DIS</a>	1078-0998	1083	3.175	0.377	61	3.5
<input type="checkbox"/>	41	<a href="#">DIGEST SURG</a>	0253-4886	538	0.645	0.044	90	3.9
<input type="checkbox"/>	42	<a href="#">ACTA GASTRO-ENT BELG</a>	0001-5644	335	0.636	0.111	36	4.8
<input type="checkbox"/>	43	<a href="#">DIS ESOPHAGUS</a>	1120-8694	373	0.603	0.044	68	4.9
<input type="checkbox"/>	44	<a href="#">REV ESP ENFERM DIG</a>	1130-0108	360	0.594	0.341	41	5.7
<input type="checkbox"/>	45	<a href="#">CHIR GASTROENTEROL</a>	0177-9990	74	0.120	0.067	60	

Fig. 2.- Journal summary list Gastroenterology & Hepatology of the 2002 JCR Science Edition, sorted by impact factor. The total number of journals in the category is 45. The journals ranking 1<sup>st</sup> to 9<sup>th</sup> and 41<sup>st</sup> to 45<sup>th</sup> are presented. Revista Española de Enfermedades Digestivas ranks 44<sup>th</sup>.

tion and achievement, and the implications for professional aspirations and academic careers have resulted in authors (involuntarily) adapting their publication behavior to a maximization of their personal IF. The outcome on the one hand is that manuscripts are submitted preferentially to high IF journals, even if those journals do not represent the most appropriate outlet on the basis of the work itself and of the intended readership.

Whereas Seglen (20) reported in the early 1990s that researchers interviewed by him did not align their publication strategy with the IF, indeed were hardly aware of the IF, and whereas surveyed NIH researchers indicated journal preferences not consistent with the IF (94), this attitude seems to have undergone a fundamental change in recent years, with the journal impact factor being promoted to the status of preferred characteristic of a journal selected for

publication –albeit subject to the respective scientific culture (6-8,10-12,20-23,62,69,79, 95-97).

However, pharmaceutical companies also appear to show some interest in where a paper is published and may exert pressure to ensure that studies sponsored by them are published where they think fit (95).

Even in the past, the compulsion to publish vast quantities has not only led to a deluge of severely redundant literature but has also entailed an inherent risk of the requirements of good scientific practice being ignored (98,99). Furthermore, the use of publication indices like the IF to measure achievement has added momentum to the drawback-ridden *publish or perish* publication strategy, compelling scientists to boost their personal IF by means of mini-papers, often referred to as the “least publishable unit” or “salami publications”, and of unfounded twofold publications.

Through their behavior-controlling effects, evaluation measures founded on publication indices, and funding based on them, may also result in researchers from technical fields giving a different weight to the focal points of their work and starting to publish more in terms of strategic adaptation, even though they themselves consider this to be dysfunctional because it is in conflict with their actual working objectives and is hardly conducive to the transfer of specialist knowledge (39). As the field of medicine, as stated above, also comprises a large number of subdisciplines whose knowledge is localized in the applications sphere, the findings of Jansz (39) and of Jansz & le Pair (41) are noteworthy. Although they may not be directly transferable, undesirable secondary effects of this kind are certainly more likely to occur in application-oriented than in basic research fields.

### Reflexivity of the impact factor at the level of the journal and of the discipline

It is, however, not only to the authors that a high IF is important. It is also of central significance to the journal, as the IF –as with the authors– is a (co-) determinant of the success or failure and thus of the fate of the journal: a high IF attracts authors to a journal and thus ensures that good-quality studies are submitted; a high IF moreover forms the basis for successful publicity campaigns, an increasing circulation, and higher advertising revenue (6,8,13,37,93). The relevance of the IF to the journal is responsible for a number of consequential problems:

The algorithm by which the IF is calculated offers scope for a wide range of fairly manipulative interventions that permit the IF to be raised artificially. Taking recourse to more trendy topics encouraging citations, and publishing more frequently cited reviews instead of original papers as well as documents not counted in the calculation algorithm but nonetheless providing the coveted citations, can boost the IF (3,6,8,13,20-23,43,59,75,78, 86,93,100; occasionally with examples from prominent journals such as *Lancet*, *Nature*, and *FASEB*). Self-citations by journals are also a good means of boosting the IF and may consequently be misused for this purpose. Discreet references can be made in the editorial of a journal or in correspondence with authors to the value of self-citations. Sometimes this is done in a less subtle manner: Smith (101) reported that a letter to authors from the journal *Leukemia* contained the following advice: “Consequently, we kindly ask you to add references of articles published in *Leukemia* to your present article”. *Leukemia* was, however, then accused of manipulating the IF (101). In view of the widespread practice of IF-doping, reactions in the scientific community may well be tainted with a touch of hypocrisy (43). In the context of self-citations and of the 2-year IF calculation period, it is moreover advisable from maximization aspects to speed up the manuscript review process in order to publish articles as quickly as possible, creating an opportunity for (self-) citations (43,100,102).

There are a number of traditional contribution categories of journals which are not beneficial to the IF and which tend to be avoided if the IF is to be raised. For instance, case reports, which may be reflected in the IF calculation algorithm depending on layout but which yield only few citations and are therefore detrimental to the IF. They have consequently come into the firing line: Bloch & Walter (8) reported that editors were exposed to criticism and pressure on the part of authors for not boosting the IF sufficiently. Although this pressure was initially resisted, a decision was eventually taken to reduce the number of clinical case reports, partly because of the IF, although case reports were considered to make “a useful contribution to the advancement of psychiatry and the education of psychiatrists” (103, cf. 104). Short reports on congresses and postgraduate education courses on offer are becoming more and more difficult to place in journals, irrespective of their informative value, because they suppress the IF on account of the calculation algorithm and consequently tend to have a negative effect on a journal’s reputation and prestige (12,64)<sup>5</sup>.

The outlined trends can be summarized as follows: the long-term outcome of the reflexivity of the IF is a shift in the journal’s structure and content. This shift is targeted at maximizing the IF, not at optimizing the transfer of medical information and knowledge.

The higher IFs of Anglo-American journals suggest that part of the literature – and possibly the most interesting part – is no longer submitted by authors to journals published in their native language but to those published in English. Now the *lingua franca* of the sciences today is English, and publication in that language is not only legitimate but, in view of the international nature of the target readership, also indicated in many cases. However, the trend towards English-language journals appears to be so pronounced that “the publication of original papers in German-language journals is losing ground to an ever-increasing extent” and some scientific publishing houses are complaining of an impoverished supply of original papers in the German language (105). If journals published in the German language, for example, were devoid of an attractive IF as a motivation for manuscript submission, this would imply in the long term not only that the German language was defunct as a scientific language but also that the journals themselves were largely “superfluous” (87,106). For if authors publish only in high IF journals, national journals lose their scientific qualitative and economic basis (12,14,106). On the one hand, the loss of the German language as a medium for scientific publication is being accelerated, and the ability of German surgical research, for example, to hold its own is being impeded by the IF (106). On the other hand, a “surgical journal in the

<sup>5</sup>“The JIF is thereby dramatically reduced. In spite of the educational value of these supplements, they hurt the journal’s reputation” (12).

*German language is indispensable if relevant research findings and clinical data are to be made widely available and implemented in our linguistic region*" (45). This applies analogously, of course, to other languages (e.g. Spanish, Portuguese, French).

However, it is not only in the relationship between German-language and English-language journals that dysfunctional shifts occur as a result of the IF orientation but also, with reference to broader geographic dimensions, on a transatlantic scale between Europe and the USA. Attention is drawn to this in British and Scandinavian editorials and manuscripts in particular. Cavalli (95) reported how, as publisher of a journal, he made vain attempts to appeal to the "European conscience", with the aim of attracting good manuscripts. In view of the threat of "a gradual decline of European journals by shifting of submissions to North American journals with higher IMF", Rostami-Hodjegan & Tucker (18) drew attention to the negative consequence for medical associations and societies as publishers of journals. Baddeley (107), Finzen (42) and Gisvold (12) saw European research jeopardized by the IF, urging that European research traditions and qualities should be preserved: "It cannot possibly be in anyone's interest that the United States should become more and more the single scientific superpower taking over all control of high quality scientific publishing. It is not only a question of where you publish, it is also a question of preserving the creativity of European scientists and institutions" (12). "Finally, we should value our journals and try to ensure that they continue to be able to compete in terms of originality and quality with those of our more overloaded North American friends" (107). Frömter et al. (108) too considered that the criticism that the IF was essentially an instrument sponsoring established (Anglo-American) journals was undeniably accurate.

South America, Africa and China, for example, are naturally even more severely affected by this development. Yet they have set up their own information systems, which ensures some degree of autonomy and provides an opportunity for the medical knowledge processes of their own continent or their own country to be promoted (56,109,110).

The reflexivity of the IF is also of significance to the content orientation of medical disciplines, and perhaps even to the individual competing paradigms in the field of medicine (3,42-44,46,111,112). Consideration of psychiatry, for example, reveals that a major part of its research is bound specifically to language and scientific culture, so that it cannot be readily expressed in different languages. Its complex knowledge structure is moreover conducive to the publication of research findings in the form of monographs and chapters of books (3). Furthermore, under the dictates of the IF, biologic-psychiatric research approaches have far better chances than philosophic-historic approaches of scoring high cumulative IFs within the field of psychiatry. To this extent, IF orien-

tion in the field of psychiatry gives biologic-psychiatric research an advantage over philosophic-historic approaches (3,111). The IF orientation –as stated by Finzen & Hoffmann-Richter (43)– therefore becomes a disaster, especially for the behavioral and cultural sciences within the field of medicine, i.e. for social medicine, social psychiatry, forensic psychiatry, and medical philosophy and history. However, there is nothing to say that developmental trends of this kind have to be confined to psychiatry<sup>6</sup>. When planning the expansion of university surgical research departments, with the option of either a "quality of life" or a "molecular genetics" research group, for example, the decision might go in favor of the latter in view of the IF too: for molecular-genetic basic research is more rewarding in terms of the IF. In other words: instead of medical or social requirements, it might be the IF (and the associated funding) that played a major role in the selection of research fields and topics (114). Baddeley (107), Gillis (115) and Mattern (16) point out that the pressure resulting from the need to publish in high IF journals is more likely to result in controversial scientific concepts being suppressed, in the eagerness for scientific innovation being spoilt, and in conservative scientific attitudes gaining influence.

Autonomous research traditions specific to national medical cultures and the plurality of medical research approaches and paradigms, aspects of special importance to clinical medicine in view of the complexity of medical knowledge (24), cannot be cultivated and implemented unless journals and other media have a sense of commitment towards these research traditions and approaches and are available for publications and discourses. Excessive orientation towards the IF and thus towards foreign, English-language publications with a high IF but with different research traditions and preferences is detrimental to this desideratum of a multiplicity of theoretical and methodological approaches.

### **Reflexivity of the impact factor and knowledge processes in society**

In today's postindustrial individualized society, higher demands than ever before are made on the individual to take responsibility for organizing the course of his or her own life. Since tradition and religion have lost influence as means of orientation, information and knowledge are gaining greater significance as resources on which action is based (116). The general use of information and knowledge in coping with everyday and professional demands corresponds to very concrete information-based actions on the part of patients in coping with illness and in obtaining adequate health care, with the Internet play-

<sup>6</sup>Lorenz (34) drew attention to the overweighting of biomedical aspects of American surgery while appealing for social psychology to be integrated into surgical studies in Germany and Europe (cf. 113).

ing a very central role as an information medium (35, 117-121). In the accompanying social knowledge processes, scientific and everyday knowledge, specialized and general knowledge are in a complex interrelationship (25-27,37,43,46,118,122-124).

As the President of the *Deutsche Forschungsgemeinschaft* (German Research Council, DFG), E. L. Winnacker (125) stressed, this is why it is “so important for science to be opened up to the general public”. To this effect it is crucial for medical research and doctors to “take a step towards repayment” in terms of providing information to the layman by making scientific findings generally accessible in a differentiated and adequate form (43,126). Scientific knowledge is moreover dependent upon mediation, not only within the sciences but also by society, if it is to be applied in curative and preventive medicine (Fig. 1).

However, if the population is to be provided with fundamental medical knowledge and if medical knowledge is to find its way into the social discourse, then medico-scientific knowledge must (also) be available in the native language. This is no longer the case if, under the pressure imposed on medical researchers by the IF, publishing in native-language journals is no longer worthwhile and native-language scientific journals consequently vanish from the market. The ensuing impoverishment of the social discourse in terms of medical knowledge impedes knowledge processes and inhibits the use of medical research findings. Furthermore, confidence in knowledge and in (national) science as well as the willingness to show solidarity in funding national research are likely to be affected (43).

The reported reflexive effects of the IF are, however, not like laws of nature and are thus open to change and to corrective measures (see below).

## CORRELATION BETWEEN IMPACT FACTOR AND QUALITY CHARACTERISTICS

The described distortions of the ISI® databases and of the IF notwithstanding, it still has to be investigated in relation to the issues of interest here whether: a) the IF is a valid indicator of the quality of an article in that it demonstrates its inherent citation potential and thus permits a statement on quality in this context; b) the IF reflects the quality of a journal; and c) the IF can serve as an indicator of research achievement.

### The impact factor as an indicator of the quality of the individual article?

Seglen (20-23) investigated the extent to which the IF can be taken as an indicator of the quality of an individual article in a journal. Based on empirical journal analyses, he reported that a mere 15% of the most frequently cited articles accounted for 50% of the citations of a journal,

and 50% for as much as 90% of all citations. The same results were reported by other authors too (92,127). The IF, which gives the average citation rate of an article published in a journal, is thus not an indicator of the frequency with which the individual article is cited. Fundamentally, a large number of articles that are cited extremely rarely, if at all, benefit from the high citation rates of a small number of citation classics: scientific standing and reputation are transferred from successful to less successful papers. From this aspect, the IF cannot be regarded as a good indicator of the scientific quality of an individual article (6,10,12,14,18,20-23,57,59,60,78,80, 83,84,92,106,126-128).

Pittler et al. (129), who used a different methodological approach to investigate areas of alternative medicine, found that the overall methodological quality of articles was very low; however, it was no better in high IF journals than in those with no IF. A clear-cut publication bias was revealed in respect of positive outcome: positive results led to publication despite a poor underlying methodological quality. Even in high IF journals, articles reporting positive findings were given preference at the expense of methodological quality. Kiroff (130), too, showed in his own investigation that studies with positive findings were more likely to be published; studies with a high level of evidence were given publication preference by journals with a higher IF (cf. 96). Kiroff (130) moreover reported a study (Helsinki Heart Study) where a primary prevention component with positive results was promptly published in a high IF journal, whereas the secondary prevention arm with less favorable results was published with a considerable delay, and then in a low IF journal. The primary component with positive results was cited disproportionately more often than the second. These findings, too, suggest that the quality of an individual article cannot be assessed with reference to the IF of the journal; the fact is rather that complex relationships exist in both IF and non-IF journals between the quality and results of a study and publication bias.

Another question of special importance to authors is whether publication in a high IF journal gives the specific article a “free ride”, as it were. Seglen (20-23) and Rostami-Hodjegan & Tucker (18) refute this on the basis of empirical findings, claiming that articles are cited irrespective of the IF of the journal in which they are published. Empirical analyses by Opthof (128) suggest that the quality of an article has a greater impact on the citation rate than the IF of the journal; on the other hand, the exposure accorded to an article through publication in a highly reputed journal does contribute to its citation success (interpreted by Seglen (21) as a *national bias*; cf. 57). According to van Dalen & Henkens (131), the citation rate of an article is influenced to varying degrees by numerous factors: publication in *core journals* covering a broad range of topics (expressed in the journal having a higher IF), language (English preferred), location of the article (greater exposure of the first article in a journal),

the Matthew effect with respect to the author (cf. 132), country of origin, co-authors, circulation of the journal, and topics covered. Although further studies are needed on this aspect, the citation rate can be assumed to be determined multifactorially. While the quality of the article will play a central role in the hierarchy of influencing factors, the IF of the journal concerned may well have a supporting function (with the IF not being an independent value but –as shown above– being influenced by language, nationality, etc.).

### The impact factor as an indicator of the quality of journals?

In an attempt to determine the extent to which the IF indicates the quality of a journal, controversial empirical findings based on different methodological approaches have been presented. On the one hand, experts were questioned on the quality of specific journals and their assessments were then compared with the corresponding IFs. In this research tradition, Foster (94) reported on an empirical survey<sup>7</sup> showing that journal rankings by scientists from the National Institutes of Health (Bethesda, MA, USA) were by no means in accord with the ISI<sup>®</sup> ranking, and characterized the ISI<sup>®</sup> approach to journal evaluation as “*a poor indicator of prestige journals*”. Hansson (84), who performed an empirical evaluation of the relationship between the IF and discipline-specific citation preferences and traditions, also concluded on the basis of his findings: “*Quality plays a small part in determining IF. I propose that the impact factor be rejected as a guide to quality*”; he pointed out moreover that clinical journals are handicapped by the IF. Metcalfe (133) registered a dependence of the IF on the delay in publication and on the discipline-specific ageing tendency of scientific articles. He too concluded that it is largely these parameters rather than the quality of the papers that accounts for the differing IF rankings in an interdisciplinary comparison of journals. The author stated: “*We should ensure that funding bodies and others who use IFs are aware of the unscientific nature of their current measure of scientific excellence.*”

Lines of reasoning in the current discourse are often associated with these negative evaluations of the IF. It is argued that the IF may perhaps be an appropriate evaluation method for prestigious journals –whose outstanding quality is, however, recognizable and accepted even without the IF ranking– but not for the bulk of journals (60). It is moreover emphasized that the IF is at most an indicator of the assertive potential of a journal but not of its quality (13,82,83,94,133,134).

By contrast, Opthof (128), who analyzed two cardiovascular journals, concluded that “*the impact*

*factor indeed permits assessment of the quality of journals*”. Using a different line of reasoning, it is pointed out that the correlation between a high rejection rate of manuscripts submitted to journals and a high IF permits the IF to be interpreted as an indicator of the quality of a journal: more manuscripts are submitted to high IF journals, which can consequently select those of the highest quality and can afford a high rejection rate, thus attaining a high quality standard and even increasing it in the long term (51,65,93). For peer review practice seems also to confirm in some instances that thorough appraisal does indeed allow the most important articles to be identified –“skimming off the cream”, as it were–, as a means of raising the IF (135).

On the other hand, reasons given for rejection are not only verdicts on quality but may –as shown in the study by Ray et al. (136)– also be associated with the degree of specialization of the submitted manuscripts and the journal’s target readership. One finding of that study was that those manuscripts rejected by the *Annals of Internal Medicine*, a journal from the field of general medicine, were then published in journals that were more specialized, had a smaller target readership, and consequently, fully in line with the logic underlying IF calculation, had a lower IF; this subsequent shift to a subspecialty journal was rated by the authors as “desirable and rational”.

However, divergent standpoints with respect to the IF as a gauge of the quality of journals need not be quite so incompatible as initial impressions may suggest. With reference to the algorithm used for calculating the IF and its inherent potential for distortion (see above), it can be logically deduced that the IF permits a differentiating statement on journal quality provided that preconditions for such a comparison are fulfilled. In other words, the journals have to belong to one and the same discipline, to be comparably aligned with respect to the topics selected from the basic sciences and from the clinical field, to be published in the same language in one and the same country, and to be similar in their structure (original papers, reviews, letters, case reports, etc.). Only under these conditions is a comparison admissible (137).

To this extent, Garfield’s (10) statement that very little can be inferred from the IF about journals publishing local clinical studies is correct. And this implies too that the IF has very little bearing on the actual impact of non-English language journals on their national medical information and knowledge transfer (for Germany cf. 38,47, 60). In other words, the significance of native language journals in Spanish, Portuguese, French or German for the medical information and knowledge transfer in these countries is neither defined nor indicated by the IF.

### The impact factor as an indicator of individual and collective research achievements?

According to Garfield (9), the wary attitude to the journal impact factor results from its being misused in

<sup>7</sup>Study by Hooper in NIH Catalyst (not available to the present authors).

many situations, especially in evaluating individuals, e.g. during the habilitation process for the awarding of postgraduate degrees. He reported having noted in many European countries that the journal impact factor was used as an easy option instead of citation analysis for estimating the citation frequency of an academic, and stated that he had always warned against such a practice. Despite all his public advocating of ISI® products, Garfield has repeatedly called for meaningful, cautious application of the IF and has endeavored to prevent misuse, especially for the assessment of individual research achievement. Other authors also displayed a critical or negative attitude towards the IF as a gauge of research (1,3,6,13,14,20-23,78,80,128). This stance can moreover be corroborated empirically, despite the slight divergence of opinion with respect to larger institutions.

On the basis of empirical evaluations, Seglen (20-23) reported that, for publications by individual researchers, the correlation between the IF of the publishing journal and the citations received for the article concerned is often low; even when the analytic approach is expanded to cover entire nations, substantial differences are recorded between the IF and actual citations (22,138). To this extent the IF, as pointed out by Seglen (22), may be misleading both for countries and for individuals (cf. 67). Bauin & Rothman (61), who used a different bibliometric indicator from the IF in their empirical study, concluded that the use of approximation measures based on the impacts of journals is quite legitimate as long as the groups being evaluated have at least ca. 100 publications. Their reasoning was that such general methods are more economical in their procedure than the more complex citation analysis. A similar line of reasoning is put forward by Opthof (128) who, while rejecting the IF-based evaluation of individual academics, does consider the IF to be a reliable means of assessing research quality at large institutions such as universities. However, he also draws attention to the problems of IF application to research groups too, pointing out that, as a selection of articles by researchers is not randomized, preference should be given to citation analyses. A further empirical finding was reported by Moed (139), who demonstrated that the ranking of institutions is influenced by certain publication and management strategies which for their part are dependent on the developmental stage of the institution: differences in ranking cannot be construed solely in terms of the quality and significance of research but have to be seen and interpreted more comprehensively against the background of institutional and personnel development (139).

## THE AWMF MODEL AS A CORRECTIVE OF THE IMPACT FACTOR AND OTHER ALTERNATIVES

The bibliometry commission of the *Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften*

(*Association of the Scientific Medical Societies*, AWMF) turned its attention to IF usage and evolved an application procedure (108). The AWMF recommendations are as follows<sup>8</sup>:

1. The IFs listed in the SCI should be given a discipline-specific weighting.
2. The IFs listed in the SSCI should be applied unweighted.
3. The IF of German-language journals listed in the SCI and SSCI should be doubled.
4. Original papers from journals not listed in the SCI and SSCI should be given an "IF equivalent" of 0.2.
5. Like papers published in journals, those published in textbooks and manuals as well as monographs should be given an IF equivalent as follows: 5-20 pages = 0.5 IF points, 21-50 pages = 1.0 IF points, 51-200 pages = 2.0 IF points, and more than 200 pages = 4.0 IF points.

According to the procedure for discipline-specific weighting, the normalized IF allocated to an article in a specific journal is calculated by dividing it by the arithmetic mean of all impact factors (mIF) of the discipline under which the journal is listed in the SCI. The mIF of each specialist category is calculated with the exclusion of pure review journals and is published annually in the Internet by the AWMF (see given URL).

As a supplement to the AWMF model, attention is drawn here too to the model proposed by Kochen et al. (2) for general medicine, to the recommendations made by Langenbeck (140), and to the IF assessment procedures for non-SCI journals (57,141). A survey of alternative impact measures and corrective factors has been presented by Glänzel & Moed (82).

As a further alternative to the IF, citation data are occasionally recommended in the literature on research evaluation, for these reflect achievement –or so it is claimed– more individually and thus more justly in the individual case than the IF (in citation analysis, reference is made to the actual citations of an article rather than to the average citation rate as with the IF). In general, those advocating this system are by no means uncritical towards citation indices, for these do indeed have their weak points: distortion through self-citations and citation cartels, the Matthew effect (citation classics are comparatively "over"-cited), predominance of the English language, citation preference for review articles rather than original papers, discipline-specific citation customs, weighting problems with respect to multi-author papers, negative citations (flawed and unreasonable research approaches and results are also cited), occasional delays in the citing of innovative research in particular, etc. Recommendations for rectifying some of these deficiencies of citation analysis have been submitted (e.g. giving self-citations a

<sup>8</sup><http://www.awmf-online.de/> (11.02.2004).

reduced weight, taking citations in the native language into account).

But even if citation analysis is slightly more viable as an evaluation system than the IF, this relative advantage does not justify and permit its application without criticism. The ISI® citation indices, in their concept certainly a fascinating instrument for the analysis of communication structures and communication deficits of the scientific discourse, especially of the Anglo-American linguistic and cultural region<sup>9</sup>, can be applied to other issues only with expert insight and with their limitations being taken into account (for various experiences and a controversial discussion, see 3,9,12,16,40,63,65,72,83,92,93,99,127,128,131,142-144).

### CONCLUSION: THE EVALUATION OF INDIVIDUAL AND COLLECTIVE RESEARCH ACHIEVEMENTS

With respect to the evaluation of research achievements, our attention is confined here to those two fields that play a special role in the university medical setting: a) the evaluation of individual research achievements; and b) the evaluation of collective research achievements within the framework of an interdisciplinary comparison between the individual disciplines of a medical faculty for research funding purposes.

With respect to the evaluation of individual research achievements, the *Recommendations of the Commission on Professional Self Regulation in Science* of the *Deutsche Forschungsgemeinschaft* (German Research Council, cf. 99) in the *Proposals for Safeguarding Good Scientific Practice* envisage in recommendation 6: “Universities and research institutes shall always give originality and quality precedence before quantity in their criteria for performance evaluation. This applies to academic degrees, to career advancement, appointments and the allocation of resources”. The subsequent explanations go into further detail: “An adequate evaluation of the achievements of an individual or a small group, however, always requires qualitative criteria in the narrow sense: their publications must be read and critically compared to the relevant state of the art and to the contributions of other individuals and working groups”. Wherever achievement has to be evaluated, the evaluators must be encouraged to make explicit judgements of quality before all else. They should therefore “receive the smallest reasonable number of publications – selected by their authors as the best examples of their work according to the

criteria by which they are to be evaluated”. The DFG proposals for the evaluation of individual research achievements, which represent only one selective aspect of the overall recommendation, have attracted considerable attention abroad and have been explicitly welcomed as indicating that hopes of change in Germany’s scientific culture may not be forlorn (“encouraging [...] hopes for change”) (8); “a broader cultural change in German science”; W. Singer, appreciatively cited by an anonymous author (6). The AWMF too takes the same line, stating<sup>10</sup>: “The journal impact factors published by the Institute for Scientific Information in the annual *Journal Citation Report (JCR)* are acceptable as a measure of quality – if at all – only within close discipline-oriented limits and solely for the internal distribution of research funds in institutes/research groups, i.e., they are not suitable for an evaluation of individuals with respect to the granting of postgraduate degrees (habilitation) and to appointments.”

With qualifying performance evaluations of this kind, which recognize originality and quality in the different achievements, we feel that supplementary quantifying data based on citation analysis may be taken into account in order to define some of the qualifying data more closely and to substantiate them, though without succumbing to the suggestive force of indices (98,145).

Restricting publication to a few significant, rewarding articles is by no means inevitably detrimental to the author’s career prospects but may rather prove to be more beneficial than extensive publication<sup>11</sup>.

If the IF is to be used for the evaluation of collective research achievements, then the AWMF model is to be recommended in the present situation; attention must moreover be paid to the incidental consequences. Under these circumstances, taking account of the AWMF model is important to clinical medicine and absolutely vital to smaller, (socio-)medical disciplines with different publication traditions and target readerships (publication in German-language journals). It must also be borne in mind that smaller medical disciplines with their few scientists at varying stages in their studies are unable to ensure a continuous publication output in the same way as larger disciplines where the numerous ongoing research projects generally ensure that some are in the publication phase. The AWMF guidelines are also helpful in maintaining the tradition of German-language medical publishing, although

<sup>9</sup>Unfortunately, very little use can be made of this potential for analysis in the web version of the ISI® citation databases SCI and SSCI available in Germany.

<sup>10</sup><http://www.awmf-online.de/> (11.02.2004).

<sup>11</sup>In the investigation by the renowned sociologist Merton (132) on the publication behavior of Nobel prizewinners, restriction to the publication of scientifically central and significant articles proves to be an essential precondition for scientific success. The following passage from that study serves as an example: “Seymour Benzer, for example, tells of how he was saved from going ‘down the biochemical drain’: ‘Delbrück saved me, when he wrote to my wife to tell me to stop writing so many papers. And I did stop.’”

it is currently not clear whether they can actually guarantee its survival<sup>12</sup>.

For the evaluation of collective research achievements, the AWMF has moreover suggested evaluating input (supplementary funding from various sources, etc.) in addition to output parameters (number and quality of scientific publications, patents, etc.)<sup>13</sup>. Langenbeck (140), a member of the AWMF working group, proposed elsewhere that evaluation of the “budget-worthiness” of scientific working groups should be undertaken on the basis of the following parameters: quality (evaluated supplementary funding, cooperation with industrial partners, stipendiaries), productivity (weighted or cumulated IF, contributions to books, patents), vitality (originality of concepts, method spectrum, dissertations, diplomas, congress papers with abstracts), and relevance (topicality, citations during the past 5 years, university cooperation, editorial boards, reviewer activity). On the basis of his survey, Raspe (4) likewise recommended a more complex model covering five different areas of achievement: a) acquisition of supplementary funding; b) publications (with discipline-specific weighting in line with the AWMF recommendation); c) teaching; d) dissertations and other academic qualifications; and e) transfer performance (knowledge transfer through scientific symposia, continuing education courses, etc.). To this effect, for example at the Lübeck campus of Schleswig-Holstein University Medical Faculty, a number of performance areas are registered by means of documentation sheets, the individual achievements are weighted (e.g. in the advancement of young academics, a special point system for post-doctoral and professorial theses, with account being taken of the grade awarded) and used for performance-related funding (personal communication from Prof. Dr. H. Raspe dated 15.01.2003; the registration system is currently undergoing further development) (cf. 4). The overall aim is to register a broad range of scientific productiv-

ity (doctoral and postdoctoral theses, supplementary funding, etc.) in terms of a multidimensional evaluation strategy (3,9,145,146). However, it is also reported that such multidimensional evaluations are by no means a simple undertaking; the experience gained by these authors suggests that corresponding application at the present time would be premature, and further testing is considered necessary: “*Because of the variety of problems we have identified with the approach we used and because most of these problems seem inherent to the measurement of the relative value of faculty activities, the application of such estimates of relative value-based production to budgetary decisions is premature*” (73).

In addition to the performance-evaluating funding approach, a performance-stimulating approach has been proposed and applied (see, for example) (147).

In general, the evaluation of research achievements in the fields of medicine and health sciences is a difficult, complex undertaking, as pointed out also by G. Gieselmann (148) from the Scientific Council<sup>14</sup> in her presentation at the “Research Achievement Evaluation” workshop: “*Clinical research is a highly complex activity. I would like to advise against apparently simple correlations. In my opinion, individual measuring parameters must not be accorded absolute status. Those wishing to improve performances in clinical research have to understand the structures and the processes taking place there in order to then change and optimize them.*”

## FUTURE PROSPECTS

The aim of medical research is to generate knowledge for the successful prevention of disease and for the medical care of the population, and to put it effectively into practice. For this purpose, basic knowledge, disease-oriented knowledge, and patient-oriented knowledge have to be functionally interlinked –to this effect, knowledge processes are like functioning ecosystems, with the neglect of one link jeopardizing the system as a whole. The AWMF’s important corrective proposal (108) aimed at damage limitation for both researcher and medical research should in our opinion be further developed as *targeted optimization of medical knowledge processes*. This requirement will not be met by the isolated IF or by citation counts; the fact is rather that, in a far more comprehensive sense, the contribution made by researchers to knowledge processes as a whole has to be the starting point for multidimensional performance-evaluating and performance-stimulating processes (cf. 82,97). The aim must be to prevent the unbalanced channeling of professional activities towards one (flawed) measuring param-

<sup>12</sup>With a view to obtaining an efficient medical publication system for Germany, Austria and the German-speaking region of Switzerland, it will be quite reasonable to discuss in this context whether specific credits (IF equivalents) should not be set interregionally for publication in selected German-language journals and publishing forms. If publishing in these journals can be made attractive, the quality of these journals will adjust itself automatically in terms of reflexivity. As has been shown, this very transformation from quantity to quality is taking place in a large number of US journals on account of this reflexivity. The AWMF model, by giving IF equivalents for scientific papers in German-language journals and books, is a step in the right direction. This applies analogously, of course, to other non-English languages (e.g. Spanish, Portuguese, French). Another point open to serious discussion would be a weighting of categories relevant to medicine in the SSCI, as it is there that a not insignificant proportion of the (socio-)medical literature is published (e.g. GERONTOLOGY; HEALTH POLICY & SERVICES; PSYCHOLOGY, CLINICAL; SOCIAL SCIENCES, BIOMEDICAL; SUBSTANCE ABUSE). Research into the consequences of applying the AWMF model is moreover urgently needed if inappropriate developments are to be detected in good time and corrected.

<sup>13</sup><http://www.awmf-online.de/> (11.02.2004).

<sup>14</sup>According to the data in the list of authors, head of the Medical Division, Scientific Council, Cologne office; later: ScienceConsult.

eter and the resulting negative consequences –as well as to ensure that the very purpose of medical science, to preserve health and to heal, does not slip out of the focus of attention.

## REFERENCES

1. Meenen NM. Der Impact-Faktor - ein zuverlässiger scientometrischer Parameter? Unfallchirurgie 1997; 23: 128-34.
2. Kochen MM, Fischer GC, Himmel W, Abholz HH. Bewertung wissenschaftlicher Veröffentlichungen: ein Vorschlag (nicht nur) für die Allgemeinmedizin. Gesundheitswesen 1998; 60: 113-9.
3. Pantel J, Mundt C. Über die Evaluation von Forschungsleistungen in der Psychiatrie. Möglichkeiten und Grenzen bibliometrischer Analysen. Nervenarzt 1999; 70: 281-7.
4. Raspe H. Leistungs- und belastungsorientierte Festsetzung von Personal- und Sachmittelbudgets auf der Basis fakultätsinterner und -übergreifender Evaluationen. Modelle deutscher medizinischer Fakultäten und ein vereinheitlichender Vorschlag. Med Klin 1999; 94: 702-7.
5. Adam D. The counting house. Nature 2002; 415: 726-9.
6. Anonymous. Citation data: the wrong impact? Nat Neurosci 1998; 1: 641-2.
7. Fassoulaki A, Sarantopoulos C, Papilas K, Patris K, Melemeni A. Academic anesthesiologists' views on the importance of the impact factor of scientific journals: a North American and European survey. Can J Anaesth 2001; 48: 953-7.
8. Bloch S, Walter G. The Impact Factor: time for change. Aust N Z J Psychiatry 2001; 35: 563-8.
9. Garfield E. Der Impact Faktor und seine richtige Anwendung. Unfallchirurg 1998; 101: 413-4.
10. Garfield E. The diverse roles of citation indexes in scientific research. Rev Invest Clin 1998; 50: 497-504.
11. Garfield E. Use of Journal Citation Reports and Journal Performance Indicators in measuring short and long term journal impact. Croat Med J 2000; 41: 368-74.
12. Gisvold SE. Citation analysis and journal impact factors - is the tail wagging the dog? Acta Anaesthesiol Scand 1999; 43: 971-3.
13. Hecht F, Hecht BK, Sandberg AA. The journal "impact factor": a misnamed, misleading, misused measure. Cancer Genet Cytogenet 1998; 104: 77-81.
14. Höök O. Scientific communications - history, electronic journals and impact factors. Scand J Rehabil Med 1999; 31: 3-7.
15. Lowy C. Impact factor limits funding. Lancet 1997; 350: 1035.
16. Mattern F. Zur Evaluation der Informatik mittels bibliometrischer Analyse. Informatik-Spektrum 2002; 25: 22-32.
17. Rey-Rocha J, Martín-Sempere MJ, Martínez-Frias J, López-Vera F. Some misuses of journal impact factor in research evaluation. Cortex 2001; 37: 595-7.
18. Rostami-Hodjegan A, Tucker GT. Journal impact factors: a 'bio-equivalence' issue? Br J Clin Pharmacol 2001; 51: 111-7.
19. Saiz-Salinas JJ. Failed professor. Nature 1996; 381: 186.
20. Seglen PO. Die Evaluierung von Wissenschaftlern anhand des 'journal impact'. In: Weingart P, Sehringer R, Winterhager M, eds. Indikatoren der Wissenschaft und Technik. Theorie, Methoden, Anwendungen. Frankfurt am Main: Campus, 1991. p. 72-90.
21. Seglen PO. Citations and journal impact factors: questionable indicators of research quality. Allergy 1997; 52: 1050-6.
22. Seglen PO. Why the impact factor of journals should not be used for evaluating research. BMJ 1997; 314: 498-502.
23. Seglen PO. Citation rates and journal impact factors are not suitable for evaluation of research. Acta Orthop Scand 1998; 69: 224-9.
24. Blois MS. Information and medicine. The nature of medical descriptions. Berkeley: University of California Press, 1984.
25. Fleck L. Erfahrung und Tatsache. Gesammelte Aufsätze. Frankfurt am Main: Suhrkamp, 1983.
26. Fleck L. Entstehung und Entwicklung einer wissenschaftlichen Tatsache. Einführung in die Lehre vom Denkstil und Denkkollektiv. Frankfurt am Main: Suhrkamp, 1994.
27. Kaltenborn K-F. Information und Wissen in der Medizin und im Gesundheitswesen. In: Kaltenborn K-F, ed. Informations- und Wissenstransfer in der Medizin und im Gesundheitswesen. Frankfurt am Main: Klostermann, 1999. p. 1-12.
28. Anonymous. Patent protection versus public health. Lancet 2001; 358: 1563.
29. Buss K-P, Wittke V. Wissen als Ware - Überlegungen zum Wandel der Modi gesellschaftlicher Wissensproduktion am Beispiel der Biotechnologie. In: Bender G, ed. Neue Formen der Wissenserzeugung. Frankfurt am Main: Campus, 2001. p. 123-46.
30. Cassier M. Private property, collective property, and public property in the age of genomics. International Social Science Journal 2002; 54: 83-98.
31. Goldstein JL, Brown MS. The clinical investigator: bewitched, bothered, and bewildered-but still beloved. J Clin Invest 1997; 99: 2803-12.
32. Rees J. Patents and intellectual property: a salvation for patient-oriented research? Lancet 2000; 356: 849-50.
33. Williamson AR. Gene patents: socially acceptable monopolies or an unnecessary hindrance to research? Trends Genet 2001; 17: 670-3.
34. Lorenz W. Klinimetrie (klinische Studien) - Bilanz. Chirurg 1998; 69: 325-32.
35. Hurrelmann K, Leppin A. Moderne Gesundheitskommunikation - eine Einführung. In: Hurrelmann K, Leppin A, eds. Moderne Gesundheitskommunikation: vom Aufklärungsgespräch zur E-Health. Bern: Huber, 2001. p. 9-21.
36. Koller M, Lorenz W. Methoden zur Aggregation von medizinischem Wissen. In: Kaltenborn K-F, ed. Informations- und Wissenstransfer in der Medizin und im Gesundheitswesen. Frankfurt am Main: Klostermann, 1999. p. 184-93.
37. Lewison G. Citations as a means to evaluate biomedical research. In: Cronin B, Atkins HB, eds. Web of Knowledge - a Festschrift in Honor of Eugene Garfield. Medford: Information Today, 2000. p. 361-372.
38. Lindner UK. Zur Funktionalität der Fachzeitschrift. In: Frömmel C, Heß D, eds. Workshop "Leistungsbewertung Forschung". Förderung der Forschung an Medizinischen Fakultäten und Hochschulen. Workshop 28. November 1997. Berlin: Charité, Medizinische Fakultät der Humboldt-Universität, u. Klarsicht Verlag, 1998. p. 75-85.
39. Jansz CNM. Some thoughts on the interaction between scientometrics and science and technology policy. Scientometrics 2000; 47: 253-64.
40. Nebel B. Ranking? Publikationen, Zitate, Drittmittelprojekte und Promotionen an deutschen Informatikfakultäten im Spiegel des WWW. Informatik-Spektrum 2001; 24: 234-49.
41. Jansz MCN, le Pair CK. Die bibliometrische Unsichtbarkeit des technischen Fortschritts. In: Weingart P, Sehringer R, Winterhager M eds. Indikatoren der Wissenschaft und Technik. Theorie, Methoden, Anwendungen. Frankfurt am Main: Campus, 1991. p. 209-23.
42. Finzen A. Der Impact Factor - die Veränderung der Wissenschaftskultur durch die Quantifizierung wissenschaftlicher Leistung. Med Welt 1998; 49: 128-34.
43. Finzen A, Hoffmann-Richter U. Publikation und Rezeption von Forschungsergebnissen der Sozialpsychiatrie - Annabelle und der Impact Factor. Sozialpsychiatrische Informationen 1997; 27: 10-5.
44. Finzen A, Hoffmann-Richter U, Dittmann V, Haug HJ. Deutsch lesen - Englisch schreiben: Fachzeitschriften zwischen Science Citation Index und Nulltarif. Psychiatr Prax 1996; 23: 1-3.
45. Herfarth C, Schürmann G. Deutsche klinische Zeitschriften und der Impact Factor. Chirurg 1996; 67: 297-9.
46. Kaltenborn K-F. Informations- und Wissenstransfer in der Informationsgesellschaft. In: Kaltenborn K-F, ed. Informations- und Wissenstransfer in der Medizin und im Gesundheitswesen. Frankfurt am Main: Klostermann, 1999. p. 534-63.
47. Niethard FU. Im Pakt mit dem Impact-Faktor? Z Orthop Ihre Grenzgeb 1996; 134: 293-4.
48. Schwarz I, Umstätter W. Medizinische Online-Dienste und CD-ROM-Datenbanken. In: Kaltenborn K-F, ed. Informations- und Wissenstransfer in der Medizin und im Gesundheitswesen. Frankfurt am Main: Klostermann, 1999. p. 207-91.
49. Calza L, Garbisa S. Italian professorships. Nature 1995; 374: 492.
50. Brody S. Impact factor as the best operational measure of medical journals. Lancet 1995; 346: 1300-1.
51. De Marchi M, Rocchi M. The editorial policies of scientific journals: testing an impact factor model. Scientometrics 2001; 51: 395-404.
52. Garfield E. The Impact Factor. (This essay was originally published in the Current Contents print editions June 20, 1994.)

- [http://sunweb.isinet.com/isi/hot/essays/journalcitationreports/7.html#muscat\\_highlighter\\_first\\_match](http://sunweb.isinet.com/isi/hot/essays/journalcitationreports/7.html#muscat_highlighter_first_match) (17.12.2001).
53. Mabe M, Amin M. Growth dynamics of scholarly and scientific journals. *Scientometrics* 2001; 51: 147-62.
  54. Garfield E. How ISI selects journals for coverage: quantitative and qualitative considerations. In: Garfield E, ed. *Journalology, KeyWords Plus, and other essays*. USA: ISI Press, 1991. p. 185-93.
  55. Testa J. The ISI® database: the journal selection process (Note: This essay was updated in November, 2001. The Editor; 2001) <http://sunweb.isinet.com/isi/hot/essays/selectionofmaterialforcoverage/199701.html> (17.12.2001)
  56. Nitzsche J. Informationsrealität im global village des Gesundheitswesens. Die Mißachtung des medizinischen Wissens der Entwicklungsländer. In: Kaltenborn K-F, ed. *Informations- und Wissenstransfer in der Medizin und im Gesundheitswesen*. Frankfurt am Main: Klostermann, 1999. p. 194-206.
  57. Stegmann J, Grohmann G. Citation rates, knowledge export and international visibility of dermatology journals listed and not listed in the Journal Citation Reports. *Scientometrics* 2001; 50: 483-502.
  58. Winkmann G, Schweim HG. Medizinisch-biowissenschaftliche Datenbanken und der Impact-Faktor. *Dtsch Med Wochenschr* 2000; 125: 1133-41.
  59. Lindner UK, Oehm V. Die Magie des Impact Faktors - Enttarnung eines Phänomens. *Unfallchirurg* 1997; 100: 253-4.
  60. Golder W. Der Impact Factor: Eine kritische Analyse. *Rofo Fortschr Geb Röntgenstr Neuen Bildgeb Verfahr* 1998; 169: 220-6.
  61. Bauin S, Rothman H. Der 'Impact' von Zeitschriften als Annäherungsmaß für Zitationsraten. In: Weingart P, Sehringer R, Winterhager M, eds. *Indikatoren der Wissenschaft und Technik. Theorie, Methoden, Anwendungen*. Frankfurt am Main: Campus, 1991. p. 91-111.
  62. Bordons M, Fernandez MT, Gomez I. 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.
  63. Bourdieu P. *Homo academicus*. Frankfurt am Main: Suhrkamp, 1988.
  64. Eich P, Wachinger M. Impact-Faktor - was er ist und wozu er nicht taugt. *Pneumologie* 1998; 52: 181-2.
  65. Lehl S. Der Impact-Faktor als Bewertungskriterium wissenschaftlicher Leistungen - das Recht auf Chancengleichheit. *Strahlenther Onkol* 1999; 175: 141-53.
  66. Marx W, Schier H, Wanitschek M. Citation analysis using online databases: feasibilities and shortcomings. *Scientometrics* 2001; 52: 59-82.
  67. Van Leeuwen TN, Moed HF, Tijssen RJW, Visser MS, van Raan AFJ. Language biases in the coverage of the Science Citation Index and its consequences for international comparisons of national research performance. *Scientometrics* 2001; 51: 335-46.
  68. Braun T, Glänzel W, Schubert A. How balanced is the Science Citation Index's journal coverage? - a preliminary overview of macrolevel statistical data. In: Cronin B, Atkins HB, eds. *Web of Knowledge - a Festschrift in Honor of Eugene Garfield*. Medford: Information Today, 2000. p. 251-77.
  69. Barnaby DP, Gallagher EJ. Alternative to the Science Citation Index impact factor as an assessment of emergency medicine's scientific contributions. *Ann Emerg Med* 1998; 31: 78-82.
  70. Gallagher EJ, Barnaby, D P. Evidence of methodologic bias in the derivation of the Science Citation Index impact factor. *Ann Emerg Med* 1998; 31: 83-6.
  71. Bourke P, Butler L. Publication types, citation rates and evaluation. *Scientometrics* 1996; 37: 473-94.
  72. Decker O, Brähler E. Von Büchern und Zeitschriften - Diskussion der Bewertung wissenschaftlicher Leistungen in den kultur- und sprachgebundenen Fächern in der Medizin. *Z Klin Psychol Psychiatr Psychother* 2001; 49: 235-46.
  73. Scheid DC, Hamm RM, Crawford SA. Measuring academic production. *Fam Med* 2002; 34: 34-44.
  74. Jacsó P. A deficiency in the algorithm for calculating the impact factor of scholarly journals: the journal impact factor. *Cortex* 2001; 37: 590-4.
  75. Moed HF, van Leeuwen TN. Impact factors can mislead. *Nature* 1996; 381: 186.
  76. Moed HF, van Leeuwen TN, Reedijk J. Towards appropriate indicators of journal impact. *Scientometrics* 1999; 46: 575-89.
  77. Van Leeuwen TN, Moed HF, Reedijk J. Critical comments on Institute for Scientific Information impact factors: a sample of inorganic molecular chemistry journals. *Journal of Information Science* 1999; 25: 489-98.
  78. Whitehouse GH. Citation rates and impact factors: should they matter? *Br J Radiol* 2001; 74: 1-3.
  79. Van Leeuwen TN, Moed HF. Development and application of journal impact measures in the Dutch science system. *Scientometrics* 2002; 53: 249-66.
  80. Brähler E, Beutel M, Decker O. Deep impact - evaluation in sciences. *Soz Praventivmed* 2004; 49: 10-4.
  81. Garfield E. Long-term vs. short-term journal impact: Does it matter? *Scientist* 1998; 12: 11-2 (February 2; Part I), 12: 12-3 (July 6; Part II). [http://www.the-scientist.com/yr1998/feb/research\\_980202.html](http://www.the-scientist.com/yr1998/feb/research_980202.html) (31.01.2002); [http://www.the-scientist.com/yr1998/july/research\\_980706.html](http://www.the-scientist.com/yr1998/july/research_980706.html) (31.01.2002).
  82. Glänzel W, Moed HF. Journal impact measures in bibliometric research. *Scientometrics* 2002; 53: 171-93.
  83. Haller U, Hepp H, Reinold E. Tötet der "Impact Factor" die deutsche Sprache? *Chirurg BDC* 1999; 38: 39-41.
  84. Hansson S. Impact factor as a misleading tool in evaluation of medical journals. *Lancet* 1995; 346: 906.
  85. Fassoulaki A, Paraskeva A, Papilas K, Karabinis G. Self-citations in six anaesthesia journals and their significance in determining the impact factor. *Br J Anaesth* 2000; 84: 266-9.
  86. Gowrishankar J, Divakar P. Sprucing up one's impact factor. *Nature* 1999; 401: 321-2.
  87. Beller FK. Der Zusammenhang zwischen Index Medicus, dem Impact factor und der deutschen Sprache. *Geburtshilfe Frauenheilkd* 1999; 59: M53-M56.
  88. Grange RI. National bias in citations in urology journals: parochialism or availability? *BJU Int* 1999; 84: 601-3.
  89. Curti M, Pistotti V, Gabutti G, Klersy C. Impact factor and electronic versions of biomedical scientific journals. *Haematologica* 2001; 86: 1015-20.
  90. Joseph KS, Hoey J. CMAJ's impact factor: room for recalculation. *CMAJ* 1999; 161: 977-8.
  91. Garfield E. How can impact factors be improved? *BMJ* 1996; 313: 411-3.
  92. Kaltenborn K-F, Kuhn K. Der Journal impact factor als Indikator für das Ranking von Fachzeitschriften und die Evaluation von Forschung: eine kritische Analyse. In: Ekkernkamp A, Scheibe O, eds. *Qualitätsmanagement in der Medizin: Handbuch für Klinik und Praxis*. Landsberg: ecomed, 2002. p. 1-29.
  93. Gannon F. The impact of the impact factor. *EMBO Rep* 2000; 1: 293.
  94. Foster WR. Impact factor as the best operational measure of medical journals. *Lancet* 1995; 346: 1301.
  95. Cavalli F. What has prevented Annals of Oncology from reaching an impact factor of 6? *Ann Oncol* 1999; 10: 1403-4.
  96. Egger M, Zellweger-Zahner T, Schneider M, Junker C, Lengeler C, Antes G. Language bias in randomized controlled trials published in English and German. *Lancet* 1997; 350: 326-9.
  97. Lewison G. Researchers' and users' perceptions of the relative standing of biomedical papers in different journals. *Scientometrics* 2002; 53: 229-40.
  98. Alewell K. Beurteilung der Leistungen von Hochschulen. In: Daniel H-D, Fisch R, eds. *Evaluation von Forschung. Methoden - Ergebnisse - Stellungnahmen*. Konstanz: Universitätsverlag, 1988. p. 41-58.
  99. Deutsche Forschungsgemeinschaft. Empfehlungen der Kommission "Selbstkontrolle in der Wissenschaft". Vorschläge zur Sicherung guter wissenschaftlicher Praxis, Januar 1998. [http://www.dfg.de/aktuelles\\_presse/reden\\_stellungnahmen/download/empfehlung\\_wiss\\_praxis\\_0198.pdf](http://www.dfg.de/aktuelles_presse/reden_stellungnahmen/download/empfehlung_wiss_praxis_0198.pdf) (24.01.2003, deutsche Fassung) [http://www.dfg.de/aktuelles\\_presse/reden\\_stellungnahmen/download/self\\_regulation\\_98.pdf](http://www.dfg.de/aktuelles_presse/reden_stellungnahmen/download/self_regulation_98.pdf) (12.02.2003, English version).
  100. Harder DR. Impact factors and the competitive nature of journal publishing. *Am J Physiol Heart Circ Physiol* 2000; 279: H457.
  101. Smith R. Journal accused of manipulating impact factor. *BMJ* 1997; 314: 463.
  102. Garfield E. Journal impact factor: a brief review. *CMAJ* 1999; 161: 979-80.
  103. Walter S, Rey JM, Dekker F. The humble case report. *Aust N Z J Psychiatry* 2001; 35: 240-5.
  104. Dixon AK. Impact factor. *Clin Radiol* 2000; 55: 414.
  105. Dick W. (2000). Forschung - Publikationen - Impact-Faktoren - "I-Mittel"-Korruption. Gedanken zu einem deutschen Phänomen. *Anaesthetist* 2000; 49: 371-2.

106. Peiper M, Zornig C. Chirurgische Forschung und Publikationen in Deutschland - eine Analyse von 1984-1994. *Langenbecks Arch Chir Suppl Kongressbd* 1997; 114: 361-6.
107. Baddeley A. So where should we publish? *Cortex* 2001; 37: 598-9.
108. Frömter E, Brähler E, Langenbeck U, Meenen NM, Usadel KH. Das AWMF-Modell zur Evaluierung publizierter Forschungsbeiträge in der Medizin. *Dtsch Med Wochenschr* 1999; 124: 910-5.
109. Nitzsche J. Medizinische Datenbanken aus Afrika, Asien und Lateinamerika. In: Kaltenborn K-F, ed. *Informations- und Wissenstransfer in der Medizin und im Gesundheitswesen*. Frankfurt am Main: Klostermann, 1999. p. 314-32.
110. Rousseau R, Jin BH, Yang NH, Liu XM. Observations concerning the two- and three-year synchronous impact factor, based on the Chinese Science Citation Database. *Journal of Documentation* 2001; 57: 349-57.
111. Adams KM. Impact factors: aiming at the wrong target. *Cortex* 2001; 37: 600-3.
112. Nitzsche J. Medizinimmanente und medizintranszendente Determinanten des Informationsbedarfs in der Medizin und im Gesundheitswesen. In: Kaltenborn K-F, ed. *Informations- und Wissenstransfer in der Medizin und im Gesundheitswesen*. Frankfurt am Main: Klostermann, 1999. p. 71-80.
113. Koller M, Lorenz W. Chirurgisches Entscheiden und Handeln. Erklärungen und Forschungsperspektiven der Sozialpsychologie. *Chirurg* 2002; 73: 846-54.
114. Schoonbaert D, Roelants G. Impact takes precedence over interest. *Nature* 1998; 391: 222.
115. Gillis JM. Overuse of impact factors suppresses controversial ideas. *Nat Neurosci* 1999; 2: 101.
116. Beck U, Giddens A, Lash S, eds. *Reflexive modernization. Politics, tradition and aesthetics in the modern social order*. Cambridge: Polity Press, 1994.
117. Eysenbach G. Neue Medien in Public Health, Prävention und Gesundheitsförderung. In: Hurrelmann K, Leppin A, eds. *Moderne Gesundheitskommunikation: vom Aufklärungsgespräch zur E-Health*. Bern: Huber, 2001. p. 205-10.
118. Kaltenborn K-F. Medizin- und gesundheitsrelevanter Wissenstransfer durch Medien. In: Hurrelmann K, Leppin A, eds. *Moderne Gesundheitskommunikation: vom Aufklärungsgespräch zur E-Health*. Bern: Huber, 2001. p. 36-69.
119. Lerch M. Gesundheitskommunikation über das Internet. In: Hurrelmann K, Leppin A, eds. *Moderne Gesundheitskommunikation: vom Aufklärungsgespräch zur E-Health*. Bern: Huber, 2001. p. 224-33.
120. Mühlbacher A, Wiest A, Schumacher N. E-Health: Informations- und Kommunikationstechniken im Gesundheitswesen. In: Hurrelmann K, Leppin A, eds. *Moderne Gesundheitskommunikation: vom Aufklärungsgespräch zur E-Health*. Bern: Huber, 2001. p. 211-23.
121. Stöbel A, Rohrwasser A. Die Rolle von Informations-Datenbanken. In: Hurrelmann K, Leppin A, eds. *Moderne Gesundheitskommunikation: vom Aufklärungsgespräch zur E-Health*. Bern: Huber, 2001. p. 234-46.
122. Kaltenborn K-F. Bedarf, Nutzung und Nutzen von Information und Wissen in der Medizin und im Gesundheitswesen. In: Kaltenborn K-F, ed. *Informations- und Wissenstransfer in der Medizin und im Gesundheitswesen*. Frankfurt am Main: Klostermann, 1999. p. 81-136.
123. Kaltenborn K-F. Ergebnisse einer explorativen Studie zu Bedarf und Nutzung von Information und Wissen in der Medizin und im Gesundheitswesen. In: Kaltenborn K-F, ed. *Informations- und Wissenstransfer in der Medizin und im Gesundheitswesen*. Frankfurt am Main: Klostermann, 1999. p. 137-83.
124. Nitzsche J. Thesen zu Bedarf und Nutzung medizinischer Information und Literatur. *Medizin – Bibliothek – Information* 2001; 1: 10-4, 23.
125. Winnacker EL. Wissenschaftsdialog, Impact-Faktor und Hierarchien. *Dtsch Med Wochenschr* 2000; 125: A18-20.
126. Gerok W. Aufgaben und Qualitätsmerkmale medizinischer Fachliteratur – zwischen Wissenschaft und Fortbildung. In: Creutzfeldt W, Gerok W, eds. *Medizinische Publizistik: Probleme und Zukunft*. Stuttgart: Thieme, 1997. p. 1-9.
127. Hofbauer R, Gmeiner B, Kaye K, Kaye AD, Frass M. "Wiener klinische Wochenschrift": publication patterns 1990-2000. *Wien Klin Wochenschr* 2001; 113: 610-5.
128. Opthof T. Sense and nonsense about the impact factor. *Cardiovasc Res* 1997; 33: 1-7.
129. Pittler MH, Abbot NC, Harkness EF, Ernst E. Location bias in controlled clinical trials of complementary/alternative therapies. *J Clin Epidemiol* 2000; 53: 485-9.
130. Kiroff GK. Publication bias in presentations to the Annual Scientific Congress. *ANZ J Surg* 2001; 71: 167-71.
131. Van Dalen HP, Henkens K. What makes a scientific article influential? The case of demographers. *Scientometrics* 2001; 50: 455-82.
132. Merton RK. The Matthew effect in science. In: Merton RK, ed. *The sociology of science. Theoretical and empirical investigations*. Chicago: University of Chicago Press, 1973. p. 439-59.
133. Metcalfe NB. Journal impact factors. *Nature* 1995; 376: 720.
134. Hitzemberger G. Der Impact-Faktor als Bewertungsgrundlage für medizinische Zeitschriften. *Wien Med Wochenschr* 1995; 145: 668-9.
135. Opthof T, Coronel R, Janse MJ. Submissions, impact factor, reviewer's recommendations and geographical bias within the peer review system (1997-2002): focus on Germany. *Cardiovasc Res* 2002; 55: 215-9.
136. Ray J, Berkwitz M, Davidoff F. The fate of manuscripts rejected by a general medical journal. *Am J Med* 2000; 109: 131-5.
137. Vinkler P. Subfield problems in applying the Garfield (Impact) Factors in practice. *Scientometrics* 2002; 53: 267-79.
138. Bonitz M, Scharnhorst A. Competition in science and the Matthew core journals. *Scientometrics* 2001; 51: 37-54.
139. Moed HF. Bibliometric indicators reflect publication and management strategies. *Scientometrics* 2000; 47: 323-46.
140. Langenbeck U. Kann eine Verwendung des Impact Factors für die Evaluation von Forschungsleistungen rational begründet werden? *Medizinische Genetik* 1998; 10: 504-8.
141. Stegmann J. How to evaluate journal impact factors. *Nature* 1997; 390: 550.
142. Abt HA. Do important papers produce high citation counts? *Scientometrics* 2000; 48: 65-70.
143. Lehl S. Der SII als Maß der Durchsetzung sowie der Forscherqualität und -kapazität von Wissenschaftlern. Nürnberg: Media Point, 1991.
144. Lehl S. Bewertung von Autoren und Co-Autoren durch Citation Index und Impact-Faktor. *Dtsch Med Wochenschr* 2000; 125: 1109-11.
145. Vinkler P. Some aspects of the evaluation of scientific and related performances of individuals. *Scientometrics* 1995; 32: 109-16.
146. Fisch R. Ein Rahmenkonzept zur Evaluation universitärer Leistungen. In: Daniel HD, Fisch R, eds. *Evaluation von Forschung. Methoden – Ergebnisse – Stellungnahmen*. Konstanz: Universitätsverlag, 1988. p. 13-31.
147. Frömmel C. Drittmittel, Impactfaktoren usw. - Kriterien für die Forschungsevaluation und der Zuordnung von Ressourcen? In: Frömmel C, Heß D, eds. *Workshop "Leistungsbewertung Forschung"*. Förderung der Forschung an Medizinischen Fakultäten und Hochschulen. Workshop 28. November 1997. Berlin: Charité, Medizinische Fakultät der Humboldt-Universität, Klarsicht Verlag, 1998. p. 122-44.
148. Gieselmann G. Qualitätskriterien in der klinischen Forschung. Erfahrungen aus der Praxis des Wissenschaftsrates. In: Frömmel C, Heß D, eds. *Workshop "Leistungsbewertung Forschung"*. Förderung der Forschung an Medizinischen Fakultäten und Hochschulen. Workshop 28. November 1997. Berlin: Charité, Medizinische Fakultät der Humboldt-Universität, Klarsicht Verlag, 1998. p. 116-9.