

Original Paper

# Two h-Index Benchmarks for Evaluating the Publication Performance of Medical Informatics Researchers

Khaled El Emam<sup>1,2</sup>, PhD; Luk Arbuckle<sup>1</sup>, MSc; Elizabeth Jonker<sup>1</sup>, BA; Kevin Anderson<sup>1</sup>, BA

<sup>1</sup>Electronic Health Information Laboratory, Children's Hospital of Eastern Ontario Research Institute, Ottawa, ON, Canada

<sup>2</sup>Department of Pediatrics, Faculty of Medicine, University of Ottawa, Ottawa, ON, Canada

**Corresponding Author:**

Khaled El Emam, PhD

Electronic Health Information Laboratory

Children's Hospital of Eastern Ontario Research Institute

401 Smyth Road

Ottawa, ON, K1H 8L1

Canada

Phone: 1 613 737 7600 ext 4181

Fax: 1 613 731 1374

Email: [kelemam@uottawa.ca](mailto:kelemam@uottawa.ca)

## Abstract

**Background:** The *h*-index is a commonly used metric for evaluating the publication performance of researchers. However, in a multidisciplinary field such as medical informatics, interpreting the *h*-index is a challenge because researchers tend to have diverse home disciplines, ranging from clinical areas to computer science, basic science, and the social sciences, each with different publication performance profiles.

**Objective:** To construct a reference standard for interpreting the *h*-index of medical informatics researchers based on the performance of their peers.

**Methods:** Using a sample of authors with articles published over the 5-year period 2006–2011 in the 2 top journals in medical informatics (as determined by impact factor), we computed their *h*-index using the Scopus database. Percentiles were computed to create a 6-level benchmark, similar in scheme to one used by the US National Science Foundation, and a 10-level benchmark.

**Results:** The 2 benchmarks can be used to place medical informatics researchers in an ordered category based on the performance of their peers. A validation exercise mapped the benchmark levels to the ranks of medical informatics academic faculty in the United States. The 10-level benchmark tracked academic rank better (with no ties) and is therefore more suitable for practical use.

**Conclusions:** Our 10-level benchmark provides an objective basis to evaluate and compare the publication performance of medical informatics researchers with that of their peers using the *h*-index.

(*J Med Internet Res* 2012;14(5):e144) doi: [10.2196/jmir.2177](https://doi.org/10.2196/jmir.2177)

**KEYWORDS**

*h*-Index; medical informatics; bibliometrics; evaluation; research output

## Introduction

Publication metrics, such as the impact factor of journals and the number of citations to papers, are often used directly or indirectly to evaluate the performance of researchers for hiring, promotion, and funding decisions [1-7]. For example, the US National Institutes of Health has developed an electronic Scientific Portfolio Assistant linked to publication metrics [8,9] (such as impact factor and number of citations) and is used by National Institutes of Health staff to “make close-call funding

decisions on individual grants” [10]. Similarly, some Wellcome Trust panels have used impact factor and applicant citation data to make grant funding decisions [7]. Publication metrics are also used to evaluate research institutions [11-13] and assess the impact of biomedical research funding policies and programs [9,10,14-16].

Direct comparisons of researchers from different backgrounds and disciplines on publication metrics can be quite misleading [17-21]. This can be a challenge for medical informatics in that it is generally considered a multidisciplinary field [22-26]. For

example, one analysis of the US National Library of Medicine's Medical Subject Headings (MeSH) that were used for medical informatics articles identified clusters indexed by terms related to the science and art of medicine, molecular genetics, statistical analysis, immunology, and biochemical communications [25]. A citation analysis of medical informatics articles found that work in general medicine journals was often cited [22].

The comparability problem is demonstrated in Table 1, which shows the average number of citations per paper over a 10-year period for a variety of disciplines [18]. There is an almost 10-fold difference in the average number of citations per paper

between a researcher in molecular biology and genetics, and a researcher in computer science. Consider a computer scientist who, with a mean of 5 citations to her papers, would be considered an above-average performer on that metric (for a computer scientist) but, when compared with a basic scientist with average performance she would be assessed quite poorly. Given that both a computer and a basic scientist can be medical informatics researchers and possibly affiliated with the same institution or department, there is a need for methods to evaluate and interpret their publication metrics that allow fair and meaningful comparisons with their medical informatics peers.

**Table 1.** Average number of citations per paper between 1995 and 2005 by discipline [18].

Discipline	Average citations per paper
Clinical medicine	10.58
Computer science	2.49
Economics and business	4.17
Engineering	3.17
Mathematics	2.66
Molecular biology and genetics	24.57
Neuroscience and behavior	16.41
Pharmacology and toxicology	9.4
Psychiatry and psychology	8.24
Social sciences, general	3.46

## The *h*-Index

One of the more-commonly used metrics to evaluate the publication performance of researchers is the *h*-index [27]. This was first proposed and defined by Hirsch in 2005 as follows: "A scientist has an index *h* if *h* of his or her *N<sub>p</sub>* papers have at least *h* citations each and the other (*N<sub>p</sub>* - *h*) papers have  $\leq h$  citations each" [27]. Hirsch designed the *h*-index to avoid the problems of other common bibliometrics, such as the total number of papers, total number of citations, number of citations per paper, number of significant papers with  $> y$  citations (*y* is determined by the evaluator), and number of citations to each of the *q* most-cited papers (*q* is determined by the evaluator). The *h*-index measures the impact of an individual's output rather than the volume, controls for the effect of a small number of highly cited papers, rewards consistent output, and is less arbitrary than measures for number of significant papers or number of citations to the *q* most-cited papers [27]. Its ease of use is also a benefit, as it is a single number that is simple to calculate using readily available databases that provide citation counts. Another advantage is that the *h*-index has been shown to predict the impact a researcher will make in the future. In a 2007 study by Hirsch, the predictive power of *h* was compared with that of 3 other bibliometrics: total number of papers, total number of citations, and mean number of citations per paper [28]. It was found that authors' *h*-index scores after the first 12 years of publication were best able to predict performance in the subsequent 12-year period, as well as cumulative achievement over the entire 24-year period. A high correlation

has also been found between an individual's *h*-index and his or her receipt of academic awards, appointments, and funding [17,29]. A person's *h*-index has also been found to be relatively unaffected by normal citation record errors—a 2010 review reported that *h*-scores before and after the correction of errors remained stable [29].

Many variations of and alternatives to the *h*-index have been proposed since 2005 [17,29-31], for example, to give more weight to highly cited papers [30], incorporating the variables of total number and age of citations [29] and allowing comparability across disciplines [31].

However, many of the subsequent variations proposed have been shown to be highly correlated with the *h*-index, and hence do not provide much additional information [32], and each variation increases the complexity and computational difficulty of the metric. As noted in a recent review, "many *h*-index variations, although being designed to overcome some of its supposed limitations, do indeed correlate quite heavily. This fact has made some researchers think that there is probably no need to introduce more *h*-index variations if it is not possible to prove that they are not redundant in real examples" [17]. Aided by the inclusion of automated *h*-index calculators in popular interdisciplinary databases, use of the *h*-index continues to grow [17].

A more-detailed critical review of the *h*-index and its measurement is provided in Multimedia Appendix 1.

## Uses and Interpretation of the h-Index

In the basic, natural, and applied sciences, there has been a trend toward objective performance evaluations of researchers for hiring, promotion, and funding decisions using bibliometrics, including the *h*-index [1-6]. In the health sciences, individuals, departments, and institutions have been compared using their *h*-index scores [21,33-39]. The *h*-index has also been used in medical informatics to evaluate the quality of panel sessions at the annual American Medical Informatics Association symposium [40] and to evaluate the national influence of medical informatics research [41].

Traditionally, subjective peer evaluations have been used as the main method to evaluate researcher performance. There is evidence that the *h*-index scores correlate well with peer assessments [42-45]. However, a case can be made for using the *h*-index to inform the peer-review decision-making process, which can arguably enhance interreviewer reliability (see [46]).

Proper interpretation of publication metrics requires a robust reference standard [47], and there is none for the *h*-index in the area of medical informatics. Given the relative advantages of the *h*-index as an objective measure of scientific output, such a standard is needed for the *h*-index to be used effectively. A defensible standard to interpret the *h*-index can help accelerate its adoption in medical informatics and allow objective, repeatable, and fair evaluations and comparisons of researchers.

In the past, different types of reference standards for the *h*-index have been constructed in other disciplines. Examples include using the mean for multiple scientific disciplines [18], by rank of radiology academics [33], by rank of neurosurgery academics [21], by comparison with chairs of medical departments in US medical schools [34], as a median for academic anesthesiologists [38,39], by rank of academic anesthesiologists [37], and by rank for academic urologists [48].

Because citation and publication distributions are known to be heavily skewed [49,50], reference standards based on percentiles have been recommended [51]. In this study we developed 2 percentile-based benchmarks for interpreting the value of the *h*-index for researchers who publish in medical informatics journals and validated the benchmarks using an independent measure of performance.

## Methods

Our objective was to develop appropriate *h*-index benchmarks for researchers who publish in medical informatics journals.

### Requirements for Benchmarks

We considered the following as requirements to maximize the utility of the benchmarks. We therefore used them to guide the methodological decisions made during their construction:

R1. The benchmarks should allow for the evaluation of researchers' performance at multiple stages in their careers. This means that the benchmarks should have sufficient granularity and variation to reflect the performance of early career investigators as well as more-established researchers.

R2. The benchmarks need to be contemporary, reflecting the performance of researchers at this point of time rather than serving as a historical analytical tool.

R3. The benchmarks should reflect the performance of researchers who publish influential work rather than work that does not get cited often.

We describe below how we made the tradeoffs necessary to meet these requirements.

### Author Sampling Frame

We intended the benchmarks to apply to individuals who are considered *medical informatics researchers*. One approach to identifying medical informatics researchers is to use a subset of prominent individuals in the field, such as American College of Medical Informatics (ACMI) fellows. This approach has an important disadvantage in that ACMI fellows are not necessarily representative of the rest of the medical informatics researchers and would therefore not meet requirement R1 above because (1) they have higher *h*-index values than other researchers, and (2) they constitute a very small fraction of all medical informatics researchers.

While constructing benchmarks based only on ACMI fellows would meet requirement R3, this group has higher *h*-index values and would not be a good representation of all medical informatics researchers. To confirm this, we compared the *h*-index values for fellows with those who were not fellows in a simple random sample of 430 authors from all authors in the top (by impact factor) 2 medical informatics journals as classified by *Journal Citation Reports: the Journal of Medical Internet Research* (JMIR) and the *Journal of the American Medical Informatics Association* (JAMIA), with 5-year impact factors of 4.98 and 3.94, respectively (according to the Web of Knowledge). Fellows had a mean *h*-index value of 16.5 versus 8.8 for the nonfellows. This difference was statistically significant ( $P < .05$  for a 2-tailed *t* test, and  $P < .05$  for a Mann-Whitney nonparametric test).

A very small proportion of authors in medical informatics journals are ACMI fellows. This is because there are only 338 fellows, they do not all publish in medical informatics journals, and many of the fellows are now retired or deceased and are no longer actively publishing. Table 2 shows the percentage of authors in some of the *Journal Citation Reports* medical informatics journals who were ACMI fellows during the period 2006-2011. The first number in the table is the maximum possible and assumes that all ACMI fellows publish in that journal. The second number is based on an exact match of the names of fellows and authors in the journal, and the third number is based on an approximate match of the names using the Jaro-Winkler string comparison distance [52] (with a cut-off of 0.9). The exact match rate is likely to have a higher false-negative rate and the latter, with the approximate matching, a higher false-positive rate. Therefore, the correct match rate is expected to be within these two values. Across all journals, the fellows account for 0.5%-0.68% (using exact and approximate matching, respectively) of all authors.

**Table 2.** Percentage of American College of Medical Informatics fellows who published in some of the *Journal Citation Reports* medical informatics journals over the period 2006–2011.

Journal name	Maximum match (%)	Exact match (%)	Approximate match (%)
<i>Journal of the American Medical Informatics Association (JAMIA)</i>	12.8	4.63	5.62
<i>Journal of Medical Internet Research (JMIR)</i>	23.5	0.9	1.32
<i>IEEE Engineering in Medicine and Biology</i>	37.3	0	0.22
<i>Artificial Intelligence in Medicine</i>	28.8	0.94	1.45
<i>BMC Medical Informatics and Decision Making</i>	21.5	1.4	2.1
<i>Computers, Informatics, Nursing</i>	34.5	1.43	2.14
<i>Computer Methods and Programs in Biomedicine</i>	11.5	0.24	0.37
<i>IEEE Transactions on Information Technology in Biomedicine</i>	12.7	0.07	0.22
<i>International Journal of Medical Informatics (IJMI)</i>	15.5	2.65	3.52
<i>International Journal of Technology Assessment in Health Care</i>	21.5	0.06	0.25
<i>Journal of Biomedical Informatics (JBI)</i>	16.9	3.4	4.35
<i>Journal of Medical Systems</i>	14.8	0.13	0.17
<i>Medical &amp; Biological Engineering &amp; Computing</i>	11.3	0	0.1
<i>Medical Decision Making</i>	21.4	0.63	1
<i>Methods of Information in Medicine</i>	18.7	0.33	2.77
<i>Statistics in Medicine</i>	8.1	0.17	0.38
<i>Statistical Methods in Medical Research</i>	45.3	0.13	0.4

Another approach to constructing a sampling frame is to identify all authors who publish in medical informatics journals over a specified time period and consider them to be medical informatics researchers. Various approaches have been used in the literature to identify a core set of medical informatics journals, which we review below.

A bottom-up method uses index terms in article databases to identify journals [22,23,53,54]. For example, some studies used MeSH terms for medical informatics concepts. However, a recent analysis found that the journals that published the majority of papers classified in this way had central concepts outside of medical informatics and were “not typically identified as medical informatics-specific journals,” such as physics, imaging, and engineering journals [23]. Therefore, this approach would not have strong face validity and is unlikely to be intuitively convincing.

A variant of that approach is to identify the journals with the relatively most-cited articles that are indexed or classified under medical informatics [22]. However, many of the journals with the most-cited articles were general medicine or general science journals, since these journals tend to have quite high average citations per article. A survey of ACMI fellows found that general medicine and general science journals ranked lower in terms of readership than did journals more typically associated with medical informatics [55]. Again, our benchmark would not pass the face validity test if it were based on publications that have a low readership among the most experienced members of the community.

Some authors subjectively decide on an initial set of medical informatics journals to study [24–26,55,56] or ask ACMI fellows

to rank or rate journals [55,57]. Others use existing classifications of journals, such as the *Journal Citation Reports* [58]. Sometimes multiple approaches are used [57]. The journals in which prominent members of the community publish have also been used as the core set of medical informatics journals, such as the most-cited ACMI fellows [58].

Our approach to identifying source journals for selecting medical informatics researchers was informed by the methods used in the literature. We selected the list of journals in the *Journal Citation Reports* in the medical informatics category. This is consistent with previous approaches [57,58]. We identified the top-2 ranked journals by impact factor at the time of writing: JMIR and JAMIA. If we had considered other definitions of “core medical informatics journals” [24,57], these 2 journals would still have had the highest impact factors among journals in those sets.

We considered all authors who published more than 1 article over the 2006–2011 period in any of the 17 journals in Table 2. Approximately 70%–77% (using exact and approximate matching, respectively) of these JMIR and JAMIA authors had also published at least one paper in one of the other 15 journals. While the choice of JMIR and JAMIA authors seemingly limited our analysis to those who published the articles that were most cited, in fact there is still significant community overlap with other journals.

By defining the sampling frame to consist of all authors in JMIR and JAMIA, plus the overlap in authorship with other journals, we met requirement R1. Requirement R3 was also met because these 2 journals have the highest impact factors.



## Sampling Period

We considered the  $h$ -index of authors who had published in the 5-year period 2006-2011 in these 2 journals: JMIR 2006;8(1) to 2011;13(2) and JAMIA 2006;13(1) to 2011;18(5). We chose the 5-year period because it is quite common for studies evaluating scholars, institutions, and programs to examine the previous 5 years' data on publications [59-68], and previous studies assessing the structure of the medical informatics literature and community have often used 5-year intervals [26,58]. In addition, a longer period would likely include more researchers who may no longer be as active in the field, hence reducing the benchmarks' representativeness of current researchers, and would therefore not meet requirement R2 above.

## Author Order

In addition to constructing a benchmark based on all authors within the sampling frame and sampling period, we could construct benchmarks based on first authors only. However, there is a lack of consistency across disciplines in how authors are ordered. For example, in some cases authors are ordered alphabetically or by the extent of contribution. This makes it difficult to determine when a first-author benchmark should be used. Furthermore, there is evidence of a strong correlation between the  $h$ -index values based on being a first author and those ignoring author order [69]. We therefore constructed benchmarks ignoring author order.

## Benchmark Levels

In general, a reference standard for publication metrics using percentiles has been recommended [70], and specifically one based on deciles [71]. We refer to a decile benchmark as PR10. A PR10 benchmark provides a 10-level scale based on deciles, where a level 10 means that a researcher is in the top 10% among his or her peers in terms of his or her  $h$ -index value, and a level 1 means that a researcher is in the bottom 10% among his or her peers. We deemed benchmarks with fewer levels, such as 5 levels based on quintiles, to be too coarse, as they would provide inadequate granularity to assess researchers' publication output as they move through their career stages and would therefore not meet requirement R1.

Another evaluation scheme used in the Science and Engineering Indicators by the US National Science Foundation has only 6 percentile levels, which we refer to as PR6 [72]. PR6 focuses on authors with  $h$ -index values that are higher than the median and on benchmarking more-established researchers.

## Calculation of h-Index

For our computation of the  $h$ -index we used Scopus. We manually checked the identity of all authors in Scopus by confirming their affiliation and comparing their listed publications with personal or academic websites. Scopus also uses specific algorithms [73] to combine all references from each author under a unique ID, which means that the time needed to manually match references to author names is reduced.

Furthermore, Scopus has been shown to have a more-accurate automated calculator for the  $h$ -index (vs Web of Science) [74], to include more peer-reviewed conference proceedings than

Web of Science [75] and to avoid the problems of duplicate records and false-positives associated with Google Scholar [74-76].

The  $h$ -index not only is computed from articles published in the 2 top medical informatics journals (JMIR and JAMIA) from which we sampled authors, but also covers all publications by these researchers in all journals and conferences that are indexed, going back to when indexing started. In the case of Scopus, indexing started from 1996. For example, if selected medical informatics researchers also published papers in general medicine journals in 2000, their  $h$ -index would include their publications in the general medicine journals from that year.

## Sample Size Calculation

We manually computed the  $h$ -index for medical informatics researchers using Scopus; therefore, rather than computing it for all authors, we decided to estimate it from a random sample. To ensure we had a representative sample of the underlying population, we decided to use a strict quantitative definition of representativeness. This definition also provides insight into the confidence intervals we can expect, based on the uncertainty we chose in sampling, therefore estimating the true unknown distribution of the  $h$ -index. We will use the PR10 benchmark to describe the method.

Following Sobel and Huyett [77], we selected a random sample of authors based on a nonparametric definition of representativeness. Namely, our sample of authors should be simultaneously representative of the true unknown cumulative distribution  $F$  of the  $h$ -index for deciles. That is, we divide the  $h$ -index into 10 pairwise disjoint subsets that we denote  $C_1, \dots, C_{10}$ . These subsets are unknown but have probability under  $F$  greater than zero, and in the case of equiprobable deciles are given by  $F(C_i) = 0.10$ , or  $i = 1, \dots, 10$  (where  $F$  as used here is a probability measure).

For an observed cumulative sample distribution  $F_n^*$  based on  $n$  observations, we say that a sample is representative relative to the fixed disjoint subsets  $C_1, \dots, C_{10}$  to within the common allowance  $\beta^*$  if we have  $|F_n^*(C_i) - F(C_i)| \leq \beta^*$  simultaneously for  $i = 1, \dots, 10$ . The degree of representativeness of  $F$  under this setup is subsequently defined as  $d_g^* = 1 - 10\beta^*$  for deciles. In particular, we consider a common allowance  $\beta^* = 0.05/2 = 0.025$ , as it is a standard threshold used in 1-sided definitions, resulting in a degree of representativeness of  $d_g^* = 0.75$  for deciles.

Next we choose a probability  $P^*$  that the sample will at least have the degree of representativeness of  $F$  that we selected. That is,  $P(|F_n^*(C_i) - F(C_i)| \leq \beta^* \text{ for all } i = 1, \dots, 10) \geq P^*$ , where we seek the sample size  $n$  needed to satisfy this equation. Note that larger sample sizes may not satisfy the inequality, and it is therefore necessary to treat this as an approximation only. For the sample sizes we consider, however, the duplicates in the original author list should decrease the author count such that the sample is inflated enough to meet the probability threshold we desire.

Determining the sample size  $n$  involves using a multinomial distribution for an infinite population and a hypergeometric distribution for a finite population. Sobel and Huyett [77] provide tables that give the sample size required under a variety of circumstances. Note that for a fixed number of pairwise disjoint subsets  $C_i$  and fixed allowance  $\beta^*$ , the greater the probability  $P^*$ , the greater the required sample size. Let  $N$  be the population size,  $n_\infty$  the sample size for an infinite population, and  $n_N$  the sample size for a finite population. Then, given  $n_\infty$  from the tables, the required sample size can be adjusted using the approximation  $n_\infty \approx n_N(N - 1) / (N - n_N)$ .

We identified  $N = 3220$  total authors (regardless of position in the author list) and chose a probability of simultaneous representativeness  $P^* = .75$  for deciles ( $\beta^* = 0.025$ ,  $d_g^* = 0.75$ ), resulting in a minimum sample size of  $n_{3220} = 430$  (where  $n_\infty = 500$  for an infinite population).

### Sample of Authors

We extracted article names, author names, and journal edition for JMIR and JAMIA over the 5 years of journal issues. We drew a simple random sample of 430 from the 3220 authors regardless of position in the author list, excluding correspondences or letters, editorials, errata, highlights, and articles without a designated author.

## Results

The distribution of  $h$ -index values from our sample is shown in Figure 1. We fitted a kernel density estimate with a normal kernel function and a bandwidth that minimizes the approximate mean integrated square error between the density estimator and the true density. As can be seen, the distributions have a long

tail. This means that the maximum  $h$ -index value will be significantly larger than the 90th percentile value (level 10 in PR10) and the 99th percentile in PR6. For example, the highest value for an author in our sample was 74, but the point estimate of the 90th percentile for our benchmarks was approximately 23. Therefore, authors with values significantly larger than the 90th percentile were all in the top 10% of medical informatics researchers. Since percentiles are a ranking, the actual  $h$ -index value for a level-10 author can be much higher than the 90th percentile value.

Percentile estimates of the  $h$ -index of authors who published in JAMIA or JMIR over the 5 years we examined are given in Table 3. Confidence intervals were calculated using the Woodruff method, which inverts the confidence intervals for the estimated distribution function (first proposed Woodruff [78] and further justified by Francisco and Fuller [79], Dorfman and Valliant [80], Sitter and Wu [81] and Chatterjee [82]).

We will consider an example to illustrate how to interpret Table 3. A medical informatics researcher with an  $h$ -index of 21 could be said to be in the 90th percentile of his or her peers in medical informatics, since that value falls right within the confidence interval. Any  $h$ -index value as high as 25 would still be in the 90th percentile. A researcher with a value in the range 17–20 is above the 80th percentile (since 17 is larger than the upper confidence limit for the 80th percentile), but not at the 90th percentile. To move beyond the 90th percentile, that researcher would need an  $h$ -index value of 26 or higher.

Table 4 provides the  $h$ -index benchmark values that would indicate statistically significant values for each of the levels. These can be used to directly determine the level of researchers based on their  $h$ -index values.

**Table 3.**  $h$ -Index percentile estimates for authors published in the *Journal of the American Medical Informatics Association* (JAMIA) or *Journal of Medical Internet Research* (JMIR) over the 5-year period 2006–2011.

PR6 <sup>a</sup>			PR10 <sup>b</sup>		
Percentile	Estimate	95% CI <sup>c</sup>	Percentile	Estimate	95% CI
<50%			<10%		
50%	4.9	4.2–5.6	10%	0.6	0.5–0.6
75%	12.6	10.9–14.3	20%	1.4	1.1–1.6
90%	22.9	20.2–25.6	30%	2.3	1.9–2.7
95%	28.9	26–31.7	40%	3.4	2.9–3.9
99%	48.5	32.2–64.9	50%	4.9	4.2–5.6
			60%	7.5	6.3–8.7
			70%	10.8	9.3–12.3
			80%	14.8	13.1–16.5
			90%	22.9	20.2–25.6

<sup>a</sup> 6 Percentile-level benchmark.

<sup>b</sup> Decile benchmark.

<sup>c</sup> Confidence interval.

**Table 4.** *h*-Index benchmarks for authors published in the *Journal of the American Medical Informatics Association* (JAMIA) or *Journal of Medical Internet Research* (JMIR) over the 5-year period 2006–2011.

Level	Benchmark	
	PR6 <sup>a</sup>	PR10 <sup>b</sup>
1	0–5	0
2	6–14	1
3	15–25	2
4	26–31	3
5	32–64	4–5
6	≥65	6–8
7		9–12
8		13–16
9		17–25
10		≥26

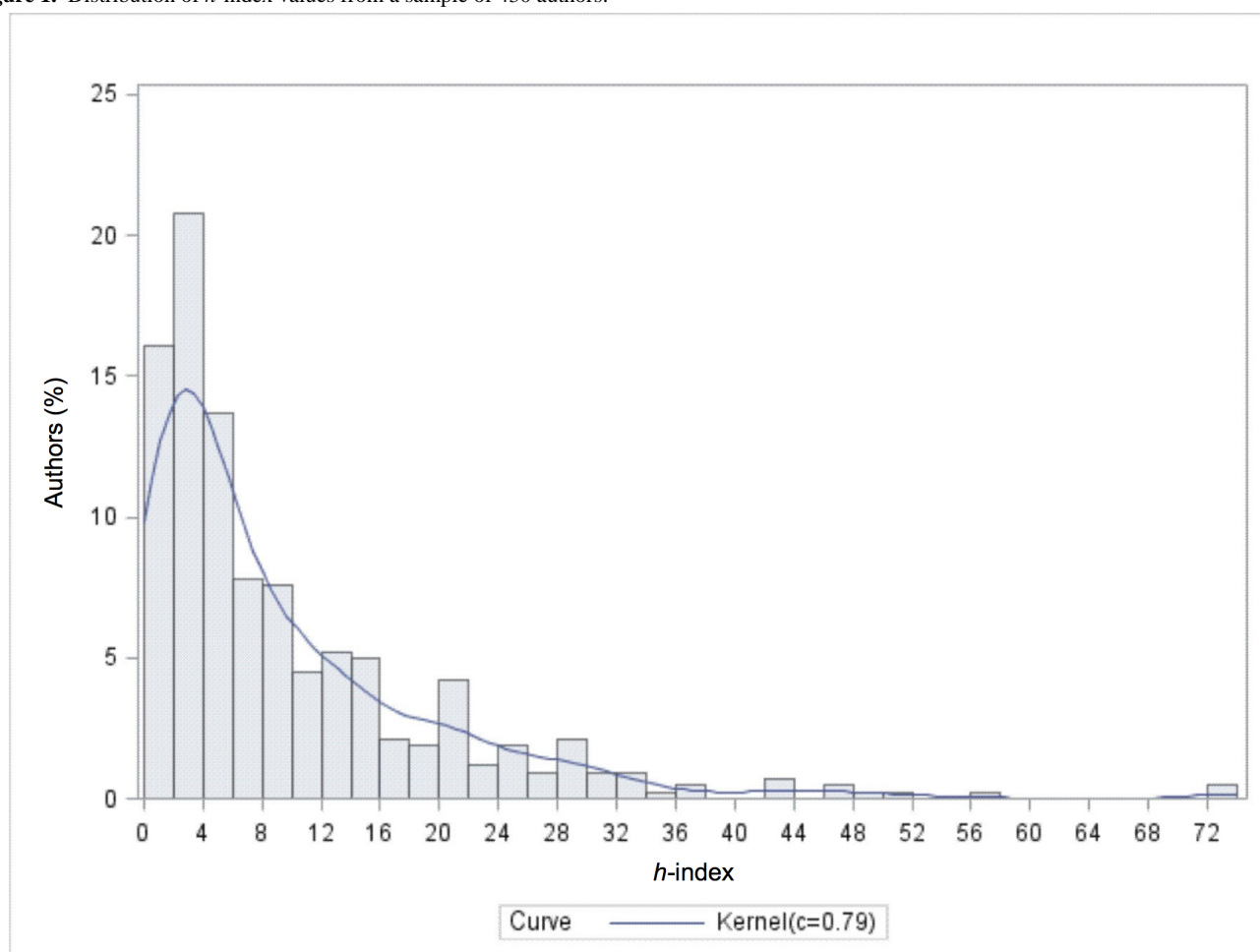
<sup>a</sup> 6 Percentile-level benchmark.

<sup>b</sup> Decile benchmark.

To explain how Table 4 was constructed, we take as an example the 50th percentile. Here the upper confidence limit is 5.6. Keeping in mind that *h*-index values can only be integers, any *h*-index value that is 5 or less will be in the lower 50% of all authors. Similarly, if we take the 75th percentile in PR6, any

*h*-index value that is 14 or less will be in the bottom 75% of all authors. Consequently, any value that is higher than 5 and equal to or less than 14 will be in the percentile range greater than 50% and less than or equal to 75% (in the third quartile). This is the 6–14 range shown in Table 4.

**Figure 1.** Distribution of *h*-index values from a sample of 430 authors.



## Discussion

### Summary

The objective of this study was to construct empirical benchmarks for interpreting the *h*-index for medical informatics researchers. We had specified three requirements to ensure the utility of the benchmarks to a large segment of the community: (1) they allow the tracking of career progress for researchers, (2) the benchmarks should be contemporary, reflecting current work in the field, and (3) the benchmarks should be based on researchers doing work that is cited often. The values we computed represent all publications by authors going back to 1996.

The benchmarks allow for the evaluation of researchers relative to their peers. These peers come from a mix of other disciplines, but they do represent the contemporary medical informatics researcher community.

More precisely, we have provided 2 empirical benchmarks that are slightly different. The first, PR6, uses a National Science Foundation scheme that has 6 percentile levels, and the second, PR10, is a broader 10-level benchmark based on deciles. The PR10 benchmark allows for the evaluation of performance of researchers at early and late stages of their careers. The PR6 benchmark is focused on the top half of performers.

### Validation

To validate the benchmarks, we examined the relationship between the *h*-index and some independent measure or ranking of researcher performance. One option was to rank scientists based on the number of recognitions and awards they have received (eg, the Nobel Prize) and the number of degrees they have received, and then determine whether our benchmarks reflect that ordering [83]. A proxy for achievements of researchers is their academic faculty rank.

We examined whether our benchmarks track the mean *h*-index value of medical informatics faculty in the United States. Several studies have used the mean (or median) *h*-index of faculty ranks

to characterize performance levels within a discipline [21,33,34,37-39,48].

We identified medical informatics departments in the United States using the list of medical informatics departments funded by the National Library of Medicine under the University-based Biomedical Informatics Research Training Programs [84], augmented with medical informatics departments listed in the Open Directory Project [85]. For each department we manually identified all faculty at any of the three ranks listed on their websites: assistant professor, associate professor, and full professor, for a total of 463 individuals. We then selected using simple random sampling 50 from each rank and computed their mean *h*-index and the confidence interval for the mean. The results are shown in Figure 2. There is greater variation in performance as the rank increases.

Given that the confidence intervals for the three ranks do not overlap, the differences in the mean *h*-index are statistically significant. Furthermore, our PR10 benchmark levels track the mean values by faculty rank well as seen in Table 5. This provides validation that the PR10 benchmark can be a useful tool for assessing the scientific performance of medical informatics faculty. For example, full professors in level 10 (top decile) on our PR10 benchmark would be above the average for US medical informatics faculty. On the other hand, a researcher with an *h*-index of 4 would be on level 5 of the PR10 benchmark and would therefore be below the mean for an assistant professor. However, such a level could be considered an acceptable target for someone completing a postdoctoral fellowship, for instance.

These results also indicate that the PR6 benchmark does not track medical informatics faculty rank very well, since both associate and full professors would be within the same level. This is due to PR6 making finer distinctions at the top end of the distribution and coarser distinctions otherwise. This results in multiple faculty ranks grouped into the same performance level. Therefore, one can argue that for practical purposes the PR10 benchmark is a more-useful tool for assessing and tracking performance of medical informatics academic faculty.

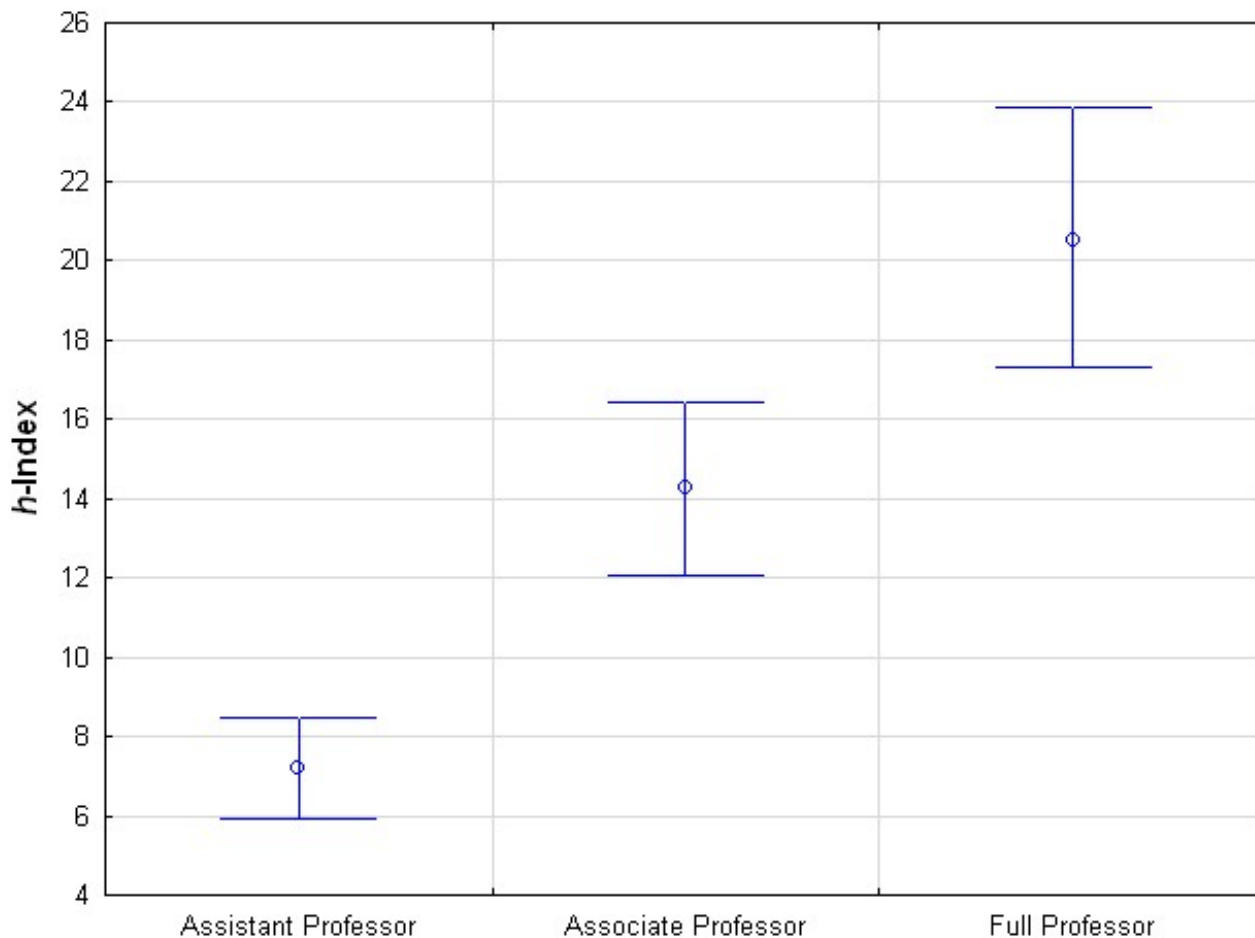
**Table 5.** Benchmark levels based on the mean *h*-index for the three US medical informatics academic faculty ranks.

Rank of faculty in US medical informatics departments	PR6 level <sup>a</sup>	PR10 level <sup>b</sup>
Assistant professor	2	6
Associate professor	3	8
Full professor	3	9

<sup>a</sup> 6 Percentile-level benchmark.

<sup>b</sup> Decile benchmark.



**Figure 2.** Mean  $h$ -index and 95% confidence interval for the three faculty ranks in US medical informatics departments.

### Interpretation and Use of the Benchmarks

Although medical informatics researchers come from multiple disciplines, the PR10 benchmark applies to the subset of individuals from these disparate disciplines who do medical informatics research. If the vast majority of medical informatics researchers were basic scientists, for example, then the benchmark would tilt more toward the performance of basic scientists. Similarly, if the vast majority of medical informatics researchers were computer scientists, then the benchmark would tilt toward the publication performance of that community.

The benchmark would be meaningful only for researchers whose area of research is clearly medical informatics. For example, a clinical researcher may have a high ranking on PR10, but this would not be relevant unless that individual has medical informatics as a primary area of work.

We suggest three scenarios where the PR10 benchmark can be useful. Researchers at early stages of their careers would be expected to be at the lower levels on the benchmark and to progress to higher levels over time. Therefore, the benchmark can be used as a yardstick to track performance over time. For research leaders, administrators, or funding agencies evaluating researchers at the same career stage competing for positions, funding, or promotions, the PR10 benchmark can be used to assess their relative standing in terms of scientific output. For instance, researchers with scores of 6 and 7 are in the same

decile and may be considered equal on the  $h$ -index metric. Finally, the PR10 benchmark can be used to set objective gates (see Table 5), such as for hiring or promotion decisions.

In practice, the PR10 benchmark could replace or augment metrics such as the number of published papers or number of citations when assessing performance. It should not be used as the sole method for evaluating the publication performance of medical informatics researchers, but can serve as another useful input for such an evaluation. Furthermore, the PR10 benchmark would need to be updated on a regular basis to ensure that it reflects contemporary performance levels in the field.

### Limitations

An underlying assumption of our method is that medical informatics researchers will at some point publish in medical informatics journals (they only need to publish once in a top medical informatics journal to be in our sampling frame). For example, an author who has published medical informatics papers only in general medicine or general science journals, or who has published in conferences only but never in a top medical informatics journal, would not be in our sampling frame. Such an individual, however, would also likely not be considered to have medical informatics as a primary area of his or her research.

Our results are limited by the journals we selected from which to sample researchers. It is possible that a different set of

journals would have produced different values for the benchmarks because they would have included a different group of researchers. However, we have argued that our choice of journals balances representativeness of the community and covers authors who publish influential work in the field.

While we used Scopus to compute our benchmark, one can argue that the use of another tool, such as Web of Science, may

have produced different results. For example, Scopus indexes publications only since 1996. This would not account for citations to earlier research articles. On the other hand, medical informatics is a recent discipline, with JAMIA starting publication in 1994 and JMIR in 1999. Furthermore, there is evidence that Web of Science and Scopus produce very similar citation counts [86], which would also mean very similar *h*-index values.

## Acknowledgments

We wish to thank Brian Dewar for his assistance with data collection. We also would like to thank the reviewers for their thoughtful feedback on an earlier version of this paper.

## Conflicts of Interest

None declared.

## Multimedia Appendix 1

A critical review of the *h*-index and its measurement.

[\[PDF File \(Adobe PDF File\), 70KB-Multimedia Appendix 1\]](#)

## References

1. Abbott A, Cyranoski D, Jones N, Maher B, Schiermeier Q, Van Noorden R. Metrics: Do metrics matter? *Nature* 2010 Jun 17;465(7300):860-862. [doi: [10.1038/465860a](https://doi.org/10.1038/465860a)] [Medline: [20559361](https://pubmed.ncbi.nlm.nih.gov/20559361/)]
2. Académie des Sciences. Annex 2: Evaluation of individual researchers in theoretical and experimental sciences. In: *On the Proper Use of Bibliometrics to Evaluate Individual Researchers*. Paris: Académie des Sciences, Institut de France; 2011:31-36.
3. Académie des Sciences. Annex 3: Bibliometric practices by discipline. In: *On the Proper Use of Bibliometrics to Evaluate Individual Researchers*. Paris: Académie des Sciences, Institut de France; 2011:37-47.
4. European Research Council. 2011. Supporting Top Researchers From Anywhere in the World URL: <http://erc.europa.eu/> [accessed 2012-10-07] [WebCite Cache ID 6BFC6fUjn]
5. European Commission. *Assessing Europe's University-Based Research*. Luxembourg: European Commission; 2010.
6. Patel VM, Ashrafian H, Ahmed K, Arora S, Jiwan S, Nicholson JK, et al. How has healthcare research performance been assessed?: a systematic review. *J R Soc Med* 2011 Jun;104(6):251-261. [doi: [10.1258/jrsm.2011.110005](https://doi.org/10.1258/jrsm.2011.110005)] [Medline: [21659400](https://pubmed.ncbi.nlm.nih.gov/21659400/)]
7. Lewison G, Cottrell R, Dixon D. Bibliometric indicators to assist the peer review process in grant decisions. *Res Eval* 1999;8(1):47-52. [doi: [10.3152/147154499781777621](https://doi.org/10.3152/147154499781777621)]
8. Fitzpatrick S. Thomson Reuters. 2009 Aug 3. National Institutes of Health Chooses Web Of Science® to Power Electronic Scientific Portfolio Assistant (e-SPA) URL: [http://thomsonreuters.com/content/press\\_room/science/468364](http://thomsonreuters.com/content/press_room/science/468364) [accessed 2012-10-07] [WebCite Cache ID 6BFD5G6yD]
9. Bertuzzi S. *Behind Medicine: Science and Policy. A View From the US National Institutes of Health on How Return on Investment Analysis Can Inform Science Policy*. Brussels: European Parliament; 2010.
10. Wadman M. Study says middle sized labs do best. *Nature* 2010 Nov 18;468(7322):356-357. [doi: [10.1038/468356a](https://doi.org/10.1038/468356a)] [Medline: [21085145](https://pubmed.ncbi.nlm.nih.gov/21085145/)]
11. Hendrix D. An analysis of bibliometric indicators, National Institutes of Health funding, and faculty size at Association of American Medical Colleges medical schools, 1997-2007. *J Med Libr Assoc* 2008 Oct;96(4):324-334. [Medline: [18979684](https://pubmed.ncbi.nlm.nih.gov/18979684/)]
12. National Research Council (US) Committee on an Assessment of Research Doctorate Programs; Ostriker J, Holland P, Kuh C, Voytuk J, editors. In: Ostriker JP, Kuh CV, Voytuk JA, editors. *A Data-Based Assessment of Research-Doctorate Programs in the United States (National Research Council)*. Washington, DC: National Academies Press; 2011.
13. Oppenheim C. The correlation between citation counts and the 1992 research assessment exercise ratings for British research in genetics, anatomy and archaeology. *J Doc* 1997;53(5):477-487. [doi: [10.1108/EUM0000000007207](https://doi.org/10.1108/EUM0000000007207)]
14. Boyack KW, Jordan P. Metrics associated with NIH funding: a high-level view. *J Am Med Inform Assoc* 2011;18(4):423-431 [FREE Full text] [doi: [10.1136/amiajnl-2011-000213](https://doi.org/10.1136/amiajnl-2011-000213)] [Medline: [21527408](https://pubmed.ncbi.nlm.nih.gov/21527408/)]
15. Lyubarova R, Itagaki BK, Itagaki MW. The impact of National Institutes of Health funding on U.S. cardiovascular disease research. *PLoS ONE* 2009;4(7):e6425 [FREE Full text] [doi: [10.1371/journal.pone.0006425](https://doi.org/10.1371/journal.pone.0006425)] [Medline: [19641617](https://pubmed.ncbi.nlm.nih.gov/19641617/)]
16. US National Science Foundation. *Research Performance Progress Report (RPPR)* URL: <http://www.nsf.gov/bfa/dias/policy/rppr/> [accessed 2012-09-28] [WebCite Cache ID 6B1O4gYWb]
17. Alonso S, Cabrerizo FJ, Herrera-Viedma E, Herrera F. *h*-Index: a review focused in its variants, computation and standardization for different scientific fields. *J Informetr* 2009;3:273-289. [doi: [10.1016/j.joi.2009.04.001](https://doi.org/10.1016/j.joi.2009.04.001)]

18. Iglesias JE, Pecharrromán C. Scaling the h-index for different scientific ISI fields. *Scientometrics* 2007;73:303-320. [doi: [10.1007/s11192-007-1805-x](https://doi.org/10.1007/s11192-007-1805-x)]
19. Radicchi F, Fortunato S, Castellano C. Universality of citation distributions: toward an objective measure of scientific impact. *Proc Natl Acad Sci U S A* 2008 Nov 11;105(45):17268-17272 [FREE Full text] [doi: [10.1073/pnas.0806977105](https://doi.org/10.1073/pnas.0806977105)] [Medline: [18978030](https://pubmed.ncbi.nlm.nih.gov/18978030/)]
20. Podlubny I. Comparison of scientific impact expressed by the number of citations in different fields of science. *Scientometrics* 2005;64:95-99. [doi: [10.1007/s11192-005-0240-0](https://doi.org/10.1007/s11192-005-0240-0)]
21. Lee J, Kraus KL, Couldwell WT. Use of the h index in neurosurgery. *Clinical article. J Neurosurg* 2009 Aug;111(2):387-392. [doi: [10.3171/2008.10.JNS08978](https://doi.org/10.3171/2008.10.JNS08978)] [Medline: [19392590](https://pubmed.ncbi.nlm.nih.gov/19392590/)]
22. Vishwanatham R. Citation analysis in journal rankings: medical informatics in the library and information science literature. *Bull Med Libr Assoc* 1998 Oct;86(4):518-522. [Medline: [9803294](https://pubmed.ncbi.nlm.nih.gov/9803294/)]
23. Deshazo JP, Lavallie DL, Wolf FM. Publication trends in the medical informatics literature: 20 years of "Medical Informatics" in MeSH. *BMC Med Inform Decis Mak* 2009;9:7 [FREE Full text] [doi: [10.1186/1472-6947-9-7](https://doi.org/10.1186/1472-6947-9-7)] [Medline: [19159472](https://pubmed.ncbi.nlm.nih.gov/19159472/)]
24. Morris TA, McCain KW. The structure of medical informatics journal literature. *J Am Med Inform Assoc* 1998;5(5):448-466 [FREE Full text] [Medline: [9760393](https://pubmed.ncbi.nlm.nih.gov/9760393/)]
25. Morris TA. Structural relationships within medical informatics. *Proc AMIA Symp* 2000:590-594. [Medline: [11079952](https://pubmed.ncbi.nlm.nih.gov/11079952/)]
26. Malin B, Carley K. A longitudinal social network analysis of the editorial boards of medical informatics and bioinformatics journals. *J Am Med Inform Assoc* 2007;14(3):340-348 [FREE Full text] [doi: [10.1197/jamia.M2228](https://doi.org/10.1197/jamia.M2228)] [Medline: [17329730](https://pubmed.ncbi.nlm.nih.gov/17329730/)]
27. Hirsch JE. An index to quantify an individual's scientific research output. *Proc Natl Acad Sci U S A* 2005 Nov 15;102(46):16569-16572 [FREE Full text] [doi: [10.1073/pnas.0507655102](https://doi.org/10.1073/pnas.0507655102)] [Medline: [16275915](https://pubmed.ncbi.nlm.nih.gov/16275915/)]
28. Hirsch JE. Does the H index have predictive power? *Proc Natl Acad Sci U S A* 2007 Dec 4;104(49):19193-19198 [FREE Full text] [doi: [10.1073/pnas.0707962104](https://doi.org/10.1073/pnas.0707962104)] [Medline: [18040045](https://pubmed.ncbi.nlm.nih.gov/18040045/)]
29. Norris M, Oppenheim C. The h-index: a broad review of a new bibliometric indicator. *J Doc* 2010;66(5):681-705. [doi: [10.1108/00220411011066790](https://doi.org/10.1108/00220411011066790)]
30. Egghe L. An improvement of the h-index: the g-index. *ISSI Newsl* 2006;2(1):8-9.
31. Antonakis J, Lalive R. Quantifying scholarly impact: IQp Versus the Hirsch h. *J Am Soc Inf Sci Technol* 2008;59(6):956-969. [doi: [10.1002/asi.v59:6](https://doi.org/10.1002/asi.v59:6)]
32. Bornmann L, Mutz R, Hug SE, Daniel HD. A multilevel meta-analysis of studies reporting correlations between the h index and 37 different h index variants. *J Informetr* 2011;5(3):346-359. [doi: [10.1016/j.joi.2011.01.006](https://doi.org/10.1016/j.joi.2011.01.006)]
33. Rad AE, Brinjikji W, Cloft HJ, Kallmes DF. The H-index in academic radiology. *Acad Radiol* 2010 Jul;17(7):817-821. [doi: [10.1016/j.acra.2010.03.011](https://doi.org/10.1016/j.acra.2010.03.011)] [Medline: [20471868](https://pubmed.ncbi.nlm.nih.gov/20471868/)]
34. Hedley-Whyte J, Milamed DR, Hoaglin DC. Chairpersons of pathology in the United States: limited benchmarks for publications. *Am J Clin Pathol* 2010 Aug;134(2):185-192 [FREE Full text] [doi: [10.1309/AJCPMVAYD5V8DLZN](https://doi.org/10.1309/AJCPMVAYD5V8DLZN)] [Medline: [20660319](https://pubmed.ncbi.nlm.nih.gov/20660319/)]
35. Sypsa V, Hatzakis A. Assessing the impact of biomedical research in academic institutions of disparate sizes. *BMC Med Res Methodol* 2009;9:33 [FREE Full text] [doi: [10.1186/1471-2288-9-33](https://doi.org/10.1186/1471-2288-9-33)] [Medline: [19480665](https://pubmed.ncbi.nlm.nih.gov/19480665/)]
36. Ponce FA, Lozano AM. Academic impact and rankings of American and Canadian neurosurgical departments as assessed using the h index. *J Neurosurg* 2010 Sep;113(3):447-457. [doi: [10.3171/2010.3.JNS1032](https://doi.org/10.3171/2010.3.JNS1032)] [Medline: [20380531](https://pubmed.ncbi.nlm.nih.gov/20380531/)]
37. Bould MD, Boet S, Sharma B, Shin E, Barrowman NJ, Grantcharov T. h-Indices in a university department of anaesthesia: an evaluation of their feasibility, reliability, and validity as an assessment of academic performance. *Br J Anaesth* 2011 Mar;106(3):325-330 [FREE Full text] [doi: [10.1093/bja/aeq403](https://doi.org/10.1093/bja/aeq403)] [Medline: [21258072](https://pubmed.ncbi.nlm.nih.gov/21258072/)]
38. Pagel PS, Hudetz JA. An analysis of scholarly productivity in United States academic anaesthesiologists by citation bibliometrics. *Anaesthesia* 2011 Oct;66(10):873-878. [doi: [10.1111/j.1365-2044.2011.06860.x](https://doi.org/10.1111/j.1365-2044.2011.06860.x)] [Medline: [21864299](https://pubmed.ncbi.nlm.nih.gov/21864299/)]
39. Pagel PS, Hudetz JA. H-index is a sensitive indicator of academic activity in highly productive anaesthesiologists: results of a bibliometric analysis. *Acta Anaesthesiol Scand* 2011 Oct;55(9):1085-1089. [doi: [10.1111/j.1399-6576.2011.02508.x](https://doi.org/10.1111/j.1399-6576.2011.02508.x)] [Medline: [22092205](https://pubmed.ncbi.nlm.nih.gov/22092205/)]
40. Kashyap R, Pickering B, Assudani J, Singh B, Ahmed A, Chute C, et al. H-index of panel session participants from AMIA 2010 annual meeting. 2011 Presented at: AMIA 2011 Annual Symposium; Oct 22-26, 2011; Washinton, DC. USA.
41. Shahbodaghi A, Sajjadi S. A scientometric investigation of the publication trends of Iranian medical informatics articles based on ISI Citation Databases. *J Paramed Sci* 2010;1(4):2-11.
42. Bornmann L, Daniel H-D. Does the h-index for ranking of scientists really work? *Scientometrics* 2005;65(3):391-392. [doi: [10.1007/s11192-005-0281-4](https://doi.org/10.1007/s11192-005-0281-4)]
43. Bornmann L, Wallon G, Ledin A. Is the h index related to (standard) bibliometric measures and to the assessments by peers? An investigation of the h index by using molecular life sciences data. *Res Eval* 2008;17(2):149-156. [doi: [10.3152/095820208X319166](https://doi.org/10.3152/095820208X319166)]
44. van Raan AFJ. arxiv. 2006. Comparison of the Hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups URL: <http://arxiv.org/ftp/physics/papers/0511/0511206.pdf> [accessed 2012-10-08] [WebCite Cache ID 6BGH6KOJW]

45. Lovegrove BG, Johnson SD. Assessment of research performance in biology: how well do peer review and bibliometry correlate? *BioScience* 2008 Feb 1;58(2):160-164.
46. Bornmann L. Scientific peer review. *Annu Rev Inf Sci Technol* 2011;45:199-245.
47. Bornmann L, Mutz R, Neuhaus C, Daniel H-D. Citation counts for research evaluation: standards of good practice for analyzing bibliometric data and presenting and interpreting results. *Ethics Sci Environ Polit* 2008;8:93-102. [doi: [10.3354/esep00084](https://doi.org/10.3354/esep00084)]
48. Benway BM, Kalidas P, Cabello JM, Bhayani SB. Does citation analysis reveal association between h-index and academic rank in urology? *Urology* 2009 Jul;74(1):30-33. [doi: [10.1016/j.urology.2008.10.045](https://doi.org/10.1016/j.urology.2008.10.045)] [Medline: [19567283](https://pubmed.ncbi.nlm.nih.gov/19567283/)]
49. Seglen PO. The skewness of science. *J Am Soc Inf Sci* 1992;43(9):628-638. [doi: [10.1002/\(SICI\)1097-4571\(199210\)43:9<628](https://doi.org/10.1002/(SICI)1097-4571(199210)43:9<628)]
50. Seglen PO. Why the impact factor of journals should not be used for evaluating research. *BMJ* 1997 Feb 15;314(7079):498-502. [Medline: [9056804](https://pubmed.ncbi.nlm.nih.gov/9056804/)]
51. Leydesdorff L, Bornmann L, Mutz R, Ophhof T. Turning the tables on citation analysis one more time: principles for comparing sets of documents. *J Am Soc Inf Sci Technol* 2011;62(7):1370-1381. [doi: [10.1002/Asi.21534](https://doi.org/10.1002/Asi.21534)]
52. Winkler W. String comparator metrics and enhanced decision rules in the Fellegi-Sunter model of record linkage. In: *Proceedings of the Survey Research Methods Section*. Alexandria, VA: American Statistical Association; 1990.
53. Otero P, Pedernera F, Montenegro S, Borbolla D, Garcia Marti S, Luna D, et al. Evolution of medical informatics in bibliographic databases. *Stud Health Technol Inform* 2004;107(Pt 1):301-305. [Medline: [15360823](https://pubmed.ncbi.nlm.nih.gov/15360823/)]
54. Sittig DF. Identifying a core set of medical informatics serials: an analysis using the MEDLINE database. *Bull Med Libr Assoc* 1996 Apr;84(2):200-204. [Medline: [8826624](https://pubmed.ncbi.nlm.nih.gov/8826624/)]
55. Greenes RA, Siegel ER. Characterization of an emerging field: approaches to defining the literature and disciplinary boundaries of medical informatics. *Proc Annu Symp Comput Appl Med Care* 1987 Nov 4:411-415.
56. Raghupathi W, Nerur S. Research themes and trends in health information systems. *Methods Inf Med* 2008;47(5):435-442. [Medline: [18852917](https://pubmed.ncbi.nlm.nih.gov/18852917/)]
57. Sittig DF, Kaalaas-Sittig J. A quantitative ranking of the Biomedical Informatics serials. *Methods Inf Med* 1995 Sep;34(4):397-340. [Medline: [7476471](https://pubmed.ncbi.nlm.nih.gov/7476471/)]
58. Andrews J. An author co-citation analysis of medical informatics. *J Med Libr Assoc* 2003 Jan;91(1):47-56. [Medline: [12568157](https://pubmed.ncbi.nlm.nih.gov/12568157/)]
59. Glass RL. An assessment of systems and software engineering scholars and institutions. *J Syst Softw* 1994;27(1):63-67.
60. Glass RL. An assessment of systems and software engineering scholars and institutions (1993-1997). *J Syst Softw* 1998;43(1):59-64. [doi: [10.1016/S0164-1212\(98\)10037-7](https://doi.org/10.1016/S0164-1212(98)10037-7)]
61. Wong WE, Tse TH, Glass RL, Basili VR, Chen TY. An assessment of systems and software engineering scholars and institutions (2003-2007 and 2004-2008). *J Syst Softw* 2011;84(1):162-168. [doi: [10.1016/j.jss.2010.09.036](https://doi.org/10.1016/j.jss.2010.09.036)]
62. Ren J, Taylor RN. Automatic and versatile publications ranking for research institutions and scholars. *Commun ACM* 2007;50(6):81-85. [doi: [10.1145/1247001.1247010](https://doi.org/10.1145/1247001.1247010)]
63. Geist R, Chetuparambil M, Hedetniemi S, Turner AJ. Computing research programs in the U.S. *Commun ACM* 1996;39(12):96-99. [doi: [10.1145/240483.240505](https://doi.org/10.1145/240483.240505)]
64. Glass RL, Ramesh V, Vessey I. An analysis of research in computing disciplines. *Communications of the ACM* 2004;47(6):89-94. [doi: [10.1145/990680.990686](https://doi.org/10.1145/990680.990686)]
65. Bornmann L, Daniel H-D. Selecting scientific excellence through committee peer review: a citation analysis of publications previously published to approval or rejection of post-doctoral research fellowship applicants. *Scientometrics* 2006;68:427-440. [doi: [10.1007/s11192-006-0121-1](https://doi.org/10.1007/s11192-006-0121-1)]
66. Endler NS. Where the "stars" are: the 25 most cited psychologists in Canada (1972-1976). *Can Psychol Rev* 1979;20(1):12-21. [doi: [10.1037/h0081489](https://doi.org/10.1037/h0081489)]
67. Jokic M. Scientometric evaluation of the projects in biology funded by the Ministry of Science and Technology, Republic of Croatia, in the 1991-1996 period. *Period Biol* 2000;102:129-142.
68. Tutarel O. Geographical distribution of publications in the field of medical education. *BMC Med Educ* 2002 Apr 30;2:3 [FREE Full text] [Medline: [12031092](https://pubmed.ncbi.nlm.nih.gov/12031092/)]
69. Ophhof T, Wilde AA. The Hirsch-index: a simple, new tool for the assessment of scientific output of individual scientists: The case of Dutch professors in clinical cardiology. *Neth Heart J* 2009 Apr;17(4):145-154. [Medline: [19421360](https://pubmed.ncbi.nlm.nih.gov/19421360/)]
70. Garfield E. *Citation Indexing: Its Theory and Applications in Science, Technology, and Humanities*. New York, NY: Wiley; 1979.
71. Schubert A, Braun T. Cross-field normalizations of scientometric indicators. *Scientometrics* 1996;36(3):311-324. [doi: [10.1007/BF02129597](https://doi.org/10.1007/BF02129597)]
72. National Science Foundation. 2012 Jan. Science and Engineering Indicators 2012 URL: <http://www.nsf.gov/statistics/seind12/appendix.htm> [accessed 2012-10-08] [WebCite Cache ID 6BGaKExEu]
73. Scopus. 2009 Mar. Frequently Asked Questions: Author Identifier URL: [http://www.info.sciverse.com/documents/files/scopus-training/resourcelibrary/pdf/FAQ\\_Author\\_Identifier\\_09.pdf](http://www.info.sciverse.com/documents/files/scopus-training/resourcelibrary/pdf/FAQ_Author_Identifier_09.pdf) [accessed 2012-10-08] [WebCite Cache ID 6BGaL0yc6]



74. Jacso P. Testing the Calculation of a Realistic h-index in Google Scholar, Scopus, Web of Science for F. W. Lancaster. *Libr Trends* 2008;56(4):784-815. [doi: [10.1353/lib.0.0011](https://doi.org/10.1353/lib.0.0011)]
75. Meho LI, Yang K. Impact of data sources on citation counts and rankings of LIS faculty: Web of Science versus Scopus and Google Scholar. *J Am Soc Inf Sci Technol* 2007;58(13):2105-2125. [doi: [10.1002/asi.v58:13](https://doi.org/10.1002/asi.v58:13)]
76. Jacso P. The pros and cons of computing the h-index using Google Scholar. *Online Inf Rev* 2008;32(3):437-452.
77. Sobel M, Huyett MJ. Nonparametric definition of the representativeness of a sample: with tables. *Bell Syst Tech J* 1958;37.
78. Woodruff RS. Confidence intervals for medians and other position measures. *J Am Stat Assoc* 1952;47(260):635-646.
79. Francisco CA, Fuller WA. Quantile estimation with a complex survey design. *Ann Stat* 1991;19(1):454-469.
80. Dorfman A, Valliant R. *Quantile Variance Estimators in Complex Surveys*. Alexandria, VA: American Statistical Association; 1993.
81. Sitter RR, Wu C. A note on Woodruff confidence intervals for quantiles. *Stat Probab Lett* 2001;52(4):353-358.
82. Chatterjee A. Asymptotic properties of sample quantiles from a finite population. *Ann Inst Stat Math* 2008;63:157-179.
83. Kosmulski M. Calibration against a reference set: a quantitative approach to assessment of the methods of assessment of scientific output. *J Informetr* 2012;6(3):451-456. [doi: [10.1016/j.joi.2012.03.004](https://doi.org/10.1016/j.joi.2012.03.004)]
84. US National Library of Medicine. 2012 Apr 13. NLM's University-Based Biomedical Informatics Research Training Programs URL: <http://www.nlm.nih.gov/ep/GrantTrainInstitute.html> [accessed 2012-10-08] [[WebCite Cache ID 6BGHoDbwm](#)]
85. Wikipedia. 2012 Aug 31. Open Directory Project URL: [http://en.wikipedia.org/wiki/Open\\_Directory\\_Project](http://en.wikipedia.org/wiki/Open_Directory_Project) [accessed 2012-10-08] [[WebCite Cache ID 6BGHpoStc](#)]
86. Bornmann L, Marx W, Schier H, Rahm E, Thor A, Daniel HD. Convergent validity of bibliometric Google Scholar data in the field of chemistry: citation counts for papers that were accepted by *Angewandte Chemie International Edition* or rejected but published elsewhere, using Google Scholar, Science Citation Index, Scopus, and Chemical Abstracts. *J Informetr* 2009;3(1):27-35. [doi: [10.1016/j.joi.2008.11.001](https://doi.org/10.1016/j.joi.2008.11.001)]

## Abbreviations

- ACMI:** American College of Medical Informatics  
**JAMIA:** Journal of the American Medical Informatics Association  
**JMIR:** Journal of Medical Internet Research  
**MeSH:** Medical Subject Headings  
**PR6:** 6 percentile-level benchmark  
**PR10:** decile benchmark

*Edited by G Eysenbach; submitted 17.05.12; peer-reviewed by R Scherer, L Bornmann; comments to author 13.06.12; revised version received 18.08.12; accepted 23.09.12; published 18.10.12*

*Please cite as:*

El Emam K, Arbuckle L, Jonker E, Anderson K

*Two h-Index Benchmarks for Evaluating the Publication Performance of Medical Informatics Researchers*

*J Med Internet Res* 2012;14(5):e144

URL: <http://www.jmir.org/2012/5/e144/>

doi: [10.2196/jmir.2177](https://doi.org/10.2196/jmir.2177)

PMID: [23079075](https://pubmed.ncbi.nlm.nih.gov/23079075/)

©Khaled El Emam, Luk Arbuckle, Elizabeth Jonker, Kevin Anderson. Originally published in the *Journal of Medical Internet Research* (<http://www.jmir.org>), 18.10.2012. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the *Journal of Medical Internet Research*, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.