J Med Syst (2017) 41:150
DOI 10.1007/s10916-017-0794-4

EDUCATION & TRAINING

The Top 100 Articles in the Medical Informatics: a Bibliometric Analysis

Hamed Nadri[1,2] · **Bahlol Rahimi**[1] · **Toomas Timpka**[3,4] · **Shahram Sedghi**[5]

Received: 1 April 2017 / Accepted: 30 July 2017

Abstract The number of citations that a research paper receives can be used as a measure of its scientific impact. The objective of this study was to identify and to examine the characteristics of top 100 cited articles in the field of Medical Informatics based on data acquired from the Thomson Reuters' Web of Science (WOS) in October, 2016. The data was collected using two procedures: first we included articles published in the 24 journals listed in the "Medical Informatics" category; second, we retrieved articles using the key words: "informatics", "medical informatics", "biomedical informatics", "clinical informatics" and "health informatics". After removing duplicate records, articles were ranked by the number of citations they received. When the 100 top cited articles had been identified, we collected the following information for each record: all WOS database citations, year of publication, journal, author names, authors' affiliation, country of origin and topics indexed for each record. Citations for the top 100 articles ranged from 346 to 7875, and citations per year ranged from 11.12 to 525. The majority of articles were published in the 2000s (n=43) and 1990s (n=38). Articles were published across 10 journals, most commonly Statistics in medicine (n=71) and Medical decision making (n=28). The articles had an average of 2.47 authors. Statistics and biostatistics modeling was the most common topic (n=71), followed by artificial intelligence (n=12), and medical errors (n=3), other topics included data mining, diagnosis, bioinformatics, information retrieval, and medical imaging. Our bibliometric analysis illustrated a historical perspective on the progress of scientific research on Medical Informatics. Moreover, the findings of the current study provide an insight on the frequency of citations for top cited articles published in Medical Informatics as well as quality of the works, journals, and the trends steering Medical Informatics.

Keywords Medical informatics · Articles · Bibliometric · Top-cited

This article is part of the Topical Collection on *Education & Training*

Electronic supplementary material The online version of this article (doi:10.1007/s10916-017-0794-4) contains supplementary material, which is available to authorized users.

✉ Bahlol Rahimi
bahlol.rahimi@gmail.com

[1] Department of Health Information Technology, Urmia University of Medical Sciences, Urmia, Iran

[2] Student Research Committee, Urmia University of Medical Sciences, Urmia, Iran

[3] Department of computer and information Sciences, Linköping University, Linköping, Sweden

[4] Department of Medical and Health Sciences, Linköping University, Linköping, Sweden

[5] Department of Librarianship and Medical Information Science, Iran University of Medical Sciences, Tehran, Iran

Introduction

The first professional organization for practitioners working with medical information management was founded by Gustav Wagner in Germany back in 1949; while, the first appearance of the specific term Medical Informatics as a denotation of a research and practice area occurred in France in 1960s. Specialized academic departments and courses for Medical Informatics were established in European countries (1960s) and in the United States (1970s). In 2016, the International Medical Informatics Association (IMIA) had more than 40 national association members. To quantify the

Published online: 19 August 2017

impact from research in a sub-area of medicine such as Medical Informatics, the number of citations that a publication receives can be used as an objective indicator. Citation analysis is an approach that uses citation data, productivity, and evaluation based on the number of references that an article receives over the time [1, 2]. The pertinence of published article to a specific area is echoed in the amount of citations from peers it obtains. Citation analysis examines a network of published articles to assess an individual article's impact on its field [3]. Analysis of the most frequently cited articles is used to identify research trends within specific topics and to pinpoint the most frequently occurring authors, journals, and institutions [4]. The number of citations is hugely important for journals since the Impact Factor (IF) of a journal is reliant on the number of citations it receives [5–8], Nevertheless, this indicates the demands for certain articles brought out by researchers and the influence of the articles in generating changes in practice, controversies, discussions, or further researches. The number of citations is considered as a direct measure of the recognition that an article warranted in its field [9–11]. As the journal IF, it can be used as a proxy for scientific quality and originality [12]. The Institute for Scientific Information (ISI) has collected the citations data since 1945 and publicized these data electronically since 1979. Recently, ISI introduced the newest journal citation system called ": Science Citation Index® (SCI) Expanded," and it is one of the databases available in the Web of Science portal. Using the citation data retrieved from Web of Sciences databases, researchers have identified and analyzed highly cited articles in different areas of medicine such as Radiology [13, 14], Orthopedic [15], Pathology [16], Neurosurgery [17], Urology [18, 19], Emergency [20], and Dermatology [21].

In Medical Informatics, bibliometric studies have previously been used to characterize subdomains such as, modeling [22], computer-based medical records [23], as categorized and indexed in Medical Subject Headings of national Library of medicine(referred to as "MESH" terms) [24]. Andrews used a co-citation analyses method to visualize scholarly communication in the field, as well as identify the most productive and prominent authors [25, 26]. Studies have also used citation analysis to develop a core set of Medical Informatics serials [27, 28]. A study by Morris and colleagues found that Medical Informatics is a maturing interdisciplinary field when it identified a relatively small core literature [27]. However, to the best of our knowledge, there was not a comprehensive study focusing on top-cited articles in the field of Medical Informatics.

The aim of this study was to identify the 100 top-cited articles published in Medical Informatics and to examine their main characteristics to gain insight into the types of publications influential attributes in this field. The purpose is to expand the understanding of the role that Medical Informatics plays in medical research and practice.

Materials and methods

We performed a bibliometric analysis of the most highly cited original research articles (excluding Reviews, Proceeding Papers, Editorial Material, etc.) in Medical Informatics using data obtained from the WOS database in October 2016. No time limitations were implemented on the investigation. It is widely known that the citation counts from resources such as Google Scholar, Scopus, and WOS vary [2, 13, 29]. We chose the WOS since we found that it was the database providing the highest scientific quality. It also was compliant with our methods relying on search terms [13], the ability to classify articles to the "Medical Informatics" category as well as WOS has been shown to be the most robust scientific database resource for medicine [2]. The meaning of "Medical Informatics" term for this study was defined as "informatics applied to medical knowledge, practice, management, report, education, and research"[30]; Informatics was here used to denote the study of information and ways to process and handle it. We applied two procedures to identify research articles in the field of medical informatics. These were:

i) Articles published in the 24 journals listed under the Web of Knowledge subject category of "medical informatics", to include all the papers in the field of Medical Informatics.
ii) Articles indexed using following keywords: "informatics", "medical informatics", "biomedical informatics"," clinical informatics" and "health informatics" were identified by searching the WOS database.

Because of the multidisciplinary nature of the Medical Informatics field, it can be difficult to determine whether a scientific article belongs to the field per se. We therefore collected all articles classified using the subject category "Medical Informatics" in WOS, and articles that evidently were published in another primary field was excluded from the study.

All articles retrieved in both procedures were imported into a Microsoft Excel 2010 worksheet. Then we removed the duplicate papers from the list. The articles were sorted according to their citation number and listed in descending order. It was assumed that the least cited article in the list would have 100 or more citations. Articles with less than 100 citations were therefore excluded.

Each article was examined regarding following attributes: publication date, journal name, first and senior authors, year of publication, geographic origin, total number of citations, and citation density (total citations/article age). Continuous variables were summarized with descriptive statistics such as range, mean, and median. Categorical variables were

Table 1 Journals in which the 100 top-cited articles in the Medical Informatics were published

Rank	Journal	Frequency (n)	Impact factor
1	STATISTICS IN MEDICINE	71	1.533
2	MEDICAL DECISION MAKING	10	2.908
3	JOURNAL OF THE AMERICAN MEDICAL INFORMATICS ASSOCIATION	6	2.363
4	COMPUTER METHODS AND PROGRAMS IN BIOMEDICINE	4	1.862
5	JOURNAL OF BIOMEDICAL INFORMATICS	3	2.447
6	MEDICAL & BIOLOGICAL ENGINEERING & COMPUTING	2	1.797
7	IEEE TRANSACTIONS ON INFORMATION TECHNOLOGY IN BIOMEDICINE	1	2.493
8	JOURNAL OF MEDICAL INTERNET RESEARCH	1	4.532
9	METHODS OF INFORMATION IN MEDICINE	1	2.248
10	STATISTICAL METHODS IN MEDICAL RESEARCH	1	4.634

expressed as frequency and percentage. Finally the data were analyzed using HistCite[1] and SPSS 20.[2]

Results

Sources and citations

The 100 most cited articles were published in 10 journals (Table 1). The articles in the list had been cited between 346 and 7875 times (Table 2). The citation density for the articles ranged from 11.12 to 525 citations/year; mean 47.65 citations/year and median 31.26 citations/year. The top three articles based on citation density were "*Quantifying heterogeneity in a meta-analysis*" (525 citations/year), "*Research electronic data capture (REDCap)-A metadata-driven methodology and workflow process for providing translational research informatics suppor*t" [31] (305.25 citations/year), and "*Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond*" (78.33 citations/year). The years 1998 and 2004 were the years with the greatest number of top-cited articles (n = 8), followed by 2002 (n = 7), 2000 and 1999 (n = 6). That articles related to year 2002 had more Global Citation Score (shows the total number of citations to a paper) with (GCS = 11,104). In addition, 4 top-cited articles per year were published in 2002, 1984, 1998, and 1993.

The impact factors for journals with the top 100 cited articles ranged from 1.5 to 4.6. The higher immediacy index also belonged to the journal with the lowest impact factor. Specialized Medical Informatics journal such as the JAMIA, only had 6 articles to the list despite its high impact factor.

[1] HistCite™ Inc. Released 2008. Garfield, E: Thomson Reuters.
[2] SPSS Inc. Released 2007. SPSS for Windows, Version 20.0. Chicago, SPSS Inc.

Author affiliations and numbers

The country of origin, number of manuscripts per institution and type of manuscript are described in Table 2. The two most common departmental affiliations of first authors were Statistics and Biostatics (n = 71) and Artificial Intelligence/ Medical Decision Analysis (n = 12). Other affiliations of first authors included Cancer Research, Nursing, Public Health, Engineering, Bioinformatics and Nutrition Research. The

Table 2 Descriptors of the 100 top-cited manuscripts in the Medical Informatics including country of origin, number of manuscripts per institution and type of manuscript

	Descriptor	Frequency (percentage)
Country of origin (n = 100)	USA	52 (52%)
	UK	25 (25%)
	Canada	11 (11%)
	Australia	4 (4%)
	Netherlands	5 (5%)
	Finland	2 (2%)
	Germany	2 (2%)
	Italy	2 (2%)
	Sweden	2 (2%)
	Other: Austria, Israel, Switzerland and Thailand	1 each (1% each)
Number of Institutions (n = 103)	1	12 (12%)
	2	5 (5%)
	3	5 (5%)
	4	5 (5%)
	5–15	36 (36%)
	>15	40 (40%)
Type of manuscript (n = 100)	Original journal article	92 (92%)
	Conference proceedings paper	8 (8%)

average number of authors for the 100 most cited works was 2.47 authors per paper. The number of authors per paper ranged from 1 to 10, and 96% were co-written by six authors or fewer. Individual authors contributed anywhere from 1 to 20 articles on the list. Twenty-two authors contributed more than one article to the top 100 list.

Year of publication

The articles included in this study were published between 1983 and 2011. The average number of years since publication was 29 years and most of the articles were published in the 2000s. Figure 1 is a graphical representation of the distribution of the 100 top cited articles by decade of publication (Table 3).

Discussion

Given the multidisciplinary nature of Medical Informatics, we used the WOS subject category 'Medical Informatics' to retrieve articles. This had as consequence that for some well-known articles in the core field of Medical Informatics and bioinformatics, the number of citations was not sufficient to enter the Top 100 list. We noted that the 1st and 2nd articles in the list with the highest number of citations were published in the biostatistics sub-area. "Statistics in Medicine" was also the Medical Informatics journal that accumulated the most publications (71 articles). However, the "Medical Decision Making" journal was the second with 28 articles and the "Journal of the American Medical Informatics Association" was the third with 6 articles. "Computer Methods and Programs in Biomedicine" was the 4th journal in rank and there was no other journal contributing more than three articles to the list. It is our understanding that the high number of articles belonging to the biostatistics sub-area in the top 100 list is due to that these papers to a high degree also was cited by authors from other medical disciplines than Medical Informatics. Correspondingly, further investigations revealed that 71 of the articles had statistics and biostatics as the subject area, while medical artificial intelligence/decision analysis was the focus of 12 articles. Journals specialized in the core topics of Medical Informatics contributed only a small portion of 100 top cited articles in comparison to non-specialized journals addressing a larger community of medical researchers and practitioners.

Our bibliometric analysis shows that the most cited article in Medical Informatics is the 2002 paper by Higgins JPT and et al., "Quantifying heterogeneity in a meta-analysis," in *Statistics in Medicine*. In addition, this was the article with the most citations per year since publication. Citations per year is a metric which allows us to evaluate which articles are presently being the most widely read and cited, effectively correcting the total citations for time since publication. Figure 2 is a graphical representation of the distribution of the 100 top cited articles by Citations per year. 61 of the 100 most cited articles were published before 2000, biasing the list in favor of articles that have had longer periods of time since publication to accumulate citations. Regarding research institutions, the largest number of records (12 articles) originated from Harvard University, followed by Boston University, and Toronto University with 5 articles each. The number of institutional subdivisions was 149, with Harvard University Medicine School contributing 8 articles, Harvard University School of Public Health 6 articles, and Cambridge Institute of Public Health/MRC Biostatistics Unit (BSU) 4 articles.

We also demonstrate that the majority (n = 43) of the most highly cited articles were published in 2000s. This is contrary to the most of the other bibliometric analyses, which generally have reported that the peak period for citations was between 1980 and 1995 [32–34]. However, elapsed time is required for the articles to accrue citations and gain significant coverage. This lack of elapsed time can explain the relatively few top cited articles published during 2000–2016. The distribution of citations in Fig. 2 with a significant number of citations over the last few years demonstrate the dynamics of the Medical

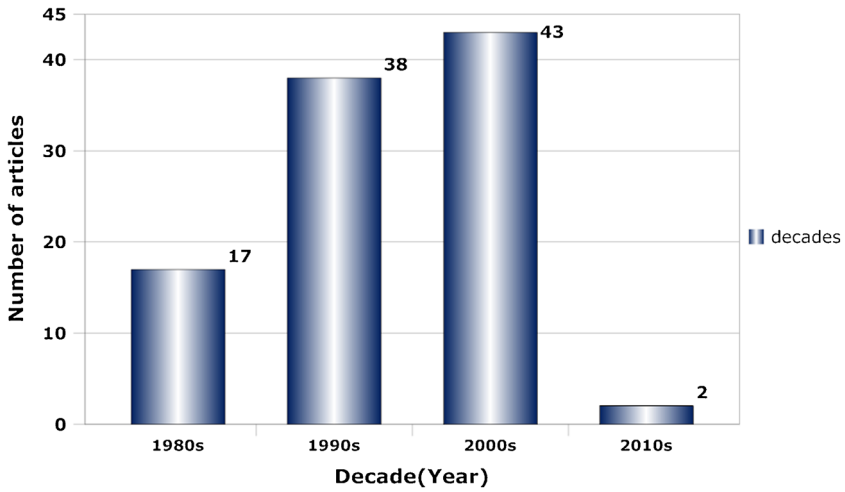

Fig. 1 Distribution of the 100 top-cited articles in the Medical Informatics by decade of publication

Table 3 The 100 top-cited articles in Medical Informatics ranked in descending order of number of citations

Rank	Article	Citations	Citation per year	Rank by citation per year
1	Higgins, Julian, and Simon G. Thompson. "Quantifying heterogeneity in a meta-analysis." Statistics in medicine 21.11 (2002)	7875	525	1
2	Harrell, Frank E., et al. "Regression modelling strategies for improved prognostic prediction." Statistics in medicine 3.2 (1984)	3462	164.86	5
3	d'Agostino, Ralph B. "Tutorial in biostatistics: propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group." Statistics in medicine 17.19 (1998)	2515	132.37	4
4	Pencina, Michael J., and Ralph B. D'Agostino. "Overall C as a measure of discrimination in survival analysis: model specific population value and confidence interval estimation." Statistics in medicine 23.13 (2004)	2505	278.33	2
5	Harris, Paul A., et al. "Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support." Journal of biomedical informatics 42.2 (2009)	2442	305.25	18
6	Wang, Lihong, Steven L. Jacques, and Liqiong Zheng. "MCML—Monte Carlo modeling of light transport in multi-layered tissues." Computer methods and programs in biomedicine 47.2 (1995)	1750	79.55	3
7	Parmar, Mahesh KB, Valter Torri, and Lesley Stewart. "Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints." Statistics in medicine 17.24 (1998)	1707	89.84	34
8	Newcombe, Robert G. "Interval estimation for the difference between independent proportions: comparison of eleven methods." Statistics in medicine 17.8 (1998)	1676	88.21	7
9	Gooley, Ted A., et al. "Estimation of failure probabilities in the presence of competing risks: new representations of old estimators." Statistics in medicine 18.6 (1999)	1433	79.61	8
10	Sonnenberg, Frank A., and J. Robert Beck. "Markov models in medical decision making a practical guide." Medical decision making 13.4 (1993)	1231	51.29	29
11	Kim, Hyune-Ju, et al. "Permutation tests for joinpoint regression with applications to cancer rates." Statistics in medicine 19.3 (2000)	1165	68.53	9
12	Hochberg, Yosef, and Yoav Benjamini. "More powerful procedures for multiple significance testing." Statistics in medicine 9.7 (1990)	1081	40.04	6
13	Cole, Timothy J., and Pamela J. Green. "Smoothing reference centile curves: the LMS method and penalized likelihood." Statistics in medicine 11.10 (1992)	1068	42.72	11
14	Prentice, Ross L. "Surrogate endpoints in clinical trials: definition and operational criteria." Statistics in medicine 8.4 (1989)	998	35.64	15
15	Thompson, Simon G., and Julian Higgins. "How should meta-regression analyses be undertaken and interpreted?." Statistics in medicine 21.11 (2002)	972	64.8	42
16	Van Buuren, Stef, Hendriek C. Boshuizen, and Dick L. Knook. "Multiple imputation of missing blood pressure covariates in survival analysis." Statistics in medicine 18.6 (1999)	902	50.11	32
17	Durrleman, Sylvain, and Richard Simon. "Flexible regression models with cubic splines." Statistics in medicine 8.5 (1989)	891	31.82	35
18	White, Ian R., Patrick Royston, and Angela M. Wood. "Multiple imputation using chained equations: issues and guidance for practice." Statistics in medicine 30.4 (2011)	838	139.67	10
19	Cuzick, Jack. "A wilcoxon-type test for trend." Statistics in medicine 4.4 (1985)	834	26.06	45
20	Moses, Lincoln E., David Shapiro, and Benjamin Littenberg. "Combining independent studies of a	807	33.62	16

Table 3 (continued)

Rank	Article	Citations	Citation per year	Rank by citation per year
	diagnostic test into a summary ROC curve: data-analytic approaches and some additional considerations." Statistics in medicine 12.14 (1993)			
21	Thompson, Simon G., and Stephen J. Sharp. "Explaining heterogeneity in meta-analysis: a comparison of methods." Statistics in medicine 18.20 (1999)	764	42.44	84
22	Newcombe, Robert G. "Two-sided confidence intervals for the single proportion: comparison of seven methods." Statistics in medicine 17.8 (1998)	738	38.84	31
23	Altman, Douglas G., and Patrick Royston. "What do we mean by validating a prognostic model?." Statistics in medicine 19.4 (2000)	724	42.59	33
24	O'Connor, Annette M. "Validation of a decisional conflict scale." Medical decision making 15.1 (1995)	697	31.68	27
25	Harrell, Frank E., Kerry L. Lee, and Daniel B. Mark. "Tutorial in biostatistics multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors." Statistics in medicine 15 (1996)	696	21.09	13
26	Hosmer DW, Hosmer T, leCessie S, Lemeshow S. "A comparison of goodness-of-fit tests for the logistic re-gression model." Statistic in medicine. 16.9 (1997).	694	34.7	23
27	Van Houwelingen, Hans C., Lidia R. Arends, and Theo Stijnen. "Advanced methods in meta-analysis: multivariate approach and meta-regression." Statistics in medicine 21.4 (2002)	688	45.87	21
28	Hämäläinen, Matti S., and Risto J. Ilmoniemi. "Interpreting magnetic fields of the brain: minimum norm estimates." Medical and biological engineering and computing 32.1 (1994)	666	28.96	88
29	Lunn, David, et al. "The BUGS project: Evolution, critique and future directions." Statistics in medicine 28.25 (2009)	662	82.75	56
30	Ash, Joan S., Marc Berg, and Enrico Coiera. "Some unintended consequences of information technology in health care: the nature of patient care information system-related errors." Journal of the American Medical Informatics Association 11.2 (2004)	627	36	48
31	Royston, Patrick, Douglas G. Altman, and Willi Sauerbrei. "Dichotomizing continuous predictors in multiple regression: a bad idea." Statistics in medicine 25.1 (2006)	622	48.23	77
32	Lu, Guobing, and A. E. Ades. "Combination of direct and indirect evidence in mixed treatment comparisons." Statistics in medicine 23.20 (2004)	619	56.55	12
33	Dweep, Harsh, et al. "miRWalk–database: prediction of possible miRNA binding sites by "walking" the genes of three genomes." Journal of biomedical informatics 44.5 (2011)	613	47.62	71
34	Harbord, Roger M., Matthias Egger, and Jonathan AC Sterne. "A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints." Statistics in medicine 25.20 (2006)	610	102.17	55
35	Fryback, Dennis G., and John R. Thornbury. "The efficacy of diagnostic imaging." Medical decision making 11.2 (1991)	604	55.45	22
36	Walter, S. D., M. Eliasziw, and A. Donner. "Sample size and optimal designs for reliability studies." Statistics in medicine 17.1 (1998)	599	31.53	39
37	Kulldorff, Martin, and Neville Nagarwalla. "Spatial disease clusters: detection and inference." Statistics in medicine 14.8 (1995)	599	23.23	78
38	Van Essen, David C., et al. "An integrated software suite for surface-based analyses of cerebral cortex." Journal of the American Medical Informatics Association 8.5 (2001)	595	27.23	95
39	Cole, Tim J., Jenny V. Freeman, and Michael A. Preece. "British 1990 growth reference centiles for weight,	589	37.19	30

Table 3 (continued)

Rank	Article	Citations	Citation per year	Rank by citation per year
	height, body mass index and head circumference fitted by maximum penalized likelihood." Statistics in medicine 17.4 (1998)			
40	Metz, Charles E., Benjamin A. Herman, and Jong-Her Shen. "Maximum likelihood estimation of receiver operating characteristic (ROC) curves from continuously-distributed data." Statistics in medicine 17.9 (1998)	588	31	14
41	Fryback, Dennis G., et al. "The Beaver Dam Health Outcomes Study: initial catalog of health-state quality factors." Medical Decision Making 13.2 (1993)	576	30.95	26
42	Putter, Hein, M. Fiocco, and R. B. Geskus. "Tutorial in biostatistics: competing risks and multi-state models." Statistics in medicine 26.11 (2007)	570	57	20
43	Rubin, Donald B., and Nathaniel Schenker. "Multiple imputation in health-are databases: An overview and some applications." Statistics in medicine 10.4 (1991)	570	24	62
44	Jiang, Hangyi, et al. "DtiStudio: resource program for diffusion tensor computation and fiber bundle tracking." Computer methods and programs in biomedicine 81.2 (2006)	560	21.92	72
45	Lindberg, Donald AB, Betsy L. Humphreys, and Alexa T. McCray. "The unified medical language system." IMIA Yearbook (1993)	557	50.91	49
46	Jonsson, E. Niclas, and Mats O. Karlsson. "Xpose—an S-PLUS based population pharmacokinetic/pharmacodynamic model building aid for NONMEM." Computer methods and programs in biomedicine 58.1 (1998)	553	23.21	73
47	Pencina, Michael J., Ralph B. D'Agostino, and Ramachandran S. Vasan. "Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond." Statistics in medicine 27.2 (2008)	535	30.72	65
48	Lipkus, Isaac M., Greg Samsa, and Barbara K. Rimer. "General performance on a numeracy scale among highly educated samples." Medical decision making 21.1 (2001)	530	41.15	17
49	Normand, S. L. T. "Meta-analysis: formulating, evaluating, combining and reporting." Statistics in medicine (1999).	529	33.12	24
50	Beck, J. Robert, and Stephen G. Pauker. "The Markov process in medical prognosis." Medical Decision Making 3.4 (1983)	526	29.39	36
51	Clayton, D., and E. Schifflers. "Models for temporal variation in cancer rates. II: age–period–cohort models." Statistics in medicine 6.4 (1987)	525	15.47	40
52	Carpenter, James, and John Bithell. "Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians." Statistics in medicine 19.9 (2000)	516	17.5	41
53	Pacini, Giovanni, and Richard N. Bergman. "MINMOD: a computer program to calculate insulin sensitivity and pancreatic responsivity from the frequently sampled intravenous glucose tolerance test." Computer methods and programs in biomedicine 23.2 (1986)	513	30.35	47
54	J Sweeting, Michael, Alexander J Sutton, and Paul C Lambert. "What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data." Statistics in medicine 23.9 (2004)	511	16.55	53
55	Cope, M., and David T. Delpy. "System for long-term measurement of cerebral blood and tissue oxygenation on newborn infants by near infra-red transillumination." Medical and Biological Engineering and Computing 26.3 (1988)	506	39.31	50
56	Eysenbach, Gunther. "The law of attrition." Journal of medical Internet research 7.1 (2005)	506	42.17	99

Table 3 (continued)

Rank	Article	Citations	Citation per year	Rank by citation per year
57	Cnaan, Avital, N. M. Laird, and Peter Slasor. "Tutorial in biostatistics: using the general linear mixed model to analyse unbalanced repeated measures and longitudinal data." Statistics in medicine 16 (1997)	499	17.45	63
58	Whitehead, Anne, and John Whitehead. "A general parametric approach to the meta-analysis of randomized clinical trials." Statistics in medicine 10.11 (1991)	492	24.95	28
59	D'Agostino, Ralph B., et al. "Relation of pooled logistic regression to time dependent Cox regression analysis: the Framingham Heart Study." Statistics in medicine 9.12 (1990)	481	18.92	96
60	Miettinen, Olli, and Markku Nurminen. "Comparative analysis of two rates." Statistics in medicine 4.2 (1985)	474	17.81	86
61	Bates, David W., et al. "The impact of computerized physician order entry on medication error prevention." Journal of the American Medical Informatics Association 6.4 (1999)	468	14.81	38
62	Bates, David W., et al. "Ten commandments for effective clinical decision support: making the practice of evidence-based medicine a reality." Journal of the American Medical Informatics Association 10.6 (2003)	468	33.43	98
63	Macaskill, Petra, Stephen D. Walter, and Les Irwig. "A comparison of methods to detect publication bias in meta-analysis." Statistics in medicine 20.4 (2001)	466	29.12	64
64	Hirth, Richard A., et al. "Willingness to pay for a quality-adjusted life year in search of a standard." Medical Decision Making 20.3 (2000)	458	26.94	79
65	Rosse, Cornelius, and José LV Mejino. "A reference ontology for biomedical informatics: the Foundational Model of Anatomy." Journal of biomedical informatics 36.6 (2003)	455	32.5	80
66	Clayton, D., and E. Schifflers. "Models for temporal variation in cancer rates. I: age–period and age–cohort models." Statistics in medicine 6.4 (1987)	453	15.1	67
67	Littell, Ramon C., Jane Pendergast, and Ranjini Natarajan. "Tutorial in biostatistics: modelling covariance structure in the analysis of repeated measures data." Statistics in medicine 19.1793 (2000)	451	26.53	81
68	Tibshirani, Robert. "The lasso method for variable selection in the Cox model." Statistics in medicine 16.4 (1997)	445	22.25	19
69	Sankoh, Abdul J., Mohammad F. Huque, and Satya D. Dubey. "Some comments on frequently used multiple endpoint adjustment methods in clinical trials." Statistics in medicine 16.22 (1997)	440	22	58
70	Yusuf, Salim, Rory Collins, and Richard Peto. "Why do we need some large, simple randomized trials?." Statistics in medicine 3.4 (1984)	438	13.27	93
71	Vickers, Andrew J., and Elena B. Elkin. "Decision curve analysis: a novel method for evaluating prediction models." Medical Decision Making 26.6 (2006)	437	39.73	76
72	Dickman, Paul W., et al. "Regression models for relative survival." Statistics in medicine 23.1 (2004)	432	33.23	43
73	Higgins, Julian, and Simon G. Thompson. "Controlling the risk of spurious findings from meta-regression." Statistics in medicine 23.11 (2004)	424	32.62	92
74	Stinnett, Aaron A., and John Mullahy. "Net health benefits a new framework for the analysis of uncertainty in cost-effectiveness analysis." Medical decision making 18.2 suppl (1998)	415	21.84	37
75	DeMets, David L. "Methods for combining randomized clinical trials: strengths and limitations." Statistics in medicine 6.3 (1987)	414	13.8	46
76	Chinn, Susan. "A simple method for converting an odds ratio to effect size for use in meta-analysis." Statistics in medicine 19.22 (2000)	413	24.29	68

Table 3 (continued)

Rank	Article	Citations	Citation per year	Rank by citation per year
77	Van Buuren, Stef. "Multiple imputation of discrete and continuous data by fully conditional specification." Statistical methods in medical research 16.3 (2007)	406	40.6	69
78	Karantonis, Dean M., et al. "Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring." IEEE transactions on information technology in biomedicine 10.1 (2006)	406	36.91	44
79	Pocock, Stuart J., et al. "Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practiceand problems." Statistics in medicine 21.19 (2002)	404	26.93	74
80	Heinze, Georg, and Michael Schemper. "A solution to the problem of separation in logistic regression." Statistics in medicine 21.16 (2002)	400	26.67	25
81	Lumley, Thomas. "Network meta-analysis for indirect treatment comparisons." Statistics in medicine 21.16 (2002)	395	26.33	59
82	Haynes, R. Brian, et al. "Developing optimal search strategies for detecting clinically sound studies in MEDLINE." Journal of the American Medical Informatics Association 1.6 (1994)	395	17.17	60
83	Zeger, Scott L., and Kung-Yee Liang. "An overview of methods for the analysis of longitudinal data." Statistics in medicine 11.14–15 (1992)	395	15.8	85
84	Austin, Peter C. "Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples." Statistics in medicine 28.25 (2009)	390	48.75	52
85	Jaro, Matthew A. "Probabilistic linkage of large public health data files." Statistics in medicine 14.5–7 (1995)	388	17.64	57
86	Muggeo, Vito MR. "Estimating regression models with unknown break-points." Statistics in medicine 22.19 (2003)	382	27.29	82
87	Rosner, B., W. C. Willett, and D. Spiegelman. "Correction of logistic regression relative risk estimates and confidence intervals for systematic within-person measurement error." Statistics in medicine 8.9 (1989)	381	13.61	90
88	Austin, Peter C. "A critical appraisal of propensity-score matching in the medical literature between 1996 and 2003." Statistics in medicine 27.12 (2008)	380	42.22	54
89	Freedman, Laurence S., Barry I. Graubard, and Arthur Schatzkin. "Statistical validation of intermediate endpoints for chronic diseases." Statistics in medicine 11.2 (1992)	380	15.2	83
90	Berkey, Catherine S., et al. "A random-effects regression model for meta-analysis." Statistics in medicine 14.4 (1995)	377	17.14	51
91	Donner, Allan, and Michael Eliasziw. "Sample size requirements for reliability studies." Statistics in medicine 6.4 (1987)	377	12.57	89
92	Bates, David W., et al. "Reducing the frequency of errors in medicine using information technology." Journal of the American Medical Informatics Association 8.4 (2001)	373	23.31	66
93	Gauderman, W. James. "Sample size requirements for matched case-control studies of gene–environment interaction." Statistics in medicine 21.1 (2002)	370	24.67	61
94	McNeil, Barbara J., and James A. Hanley. "Statistical approaches to the analysis of receiver operating characteristic (ROC) curves." Medical decision making 4.2 (1984)	367	11.12	75
95	Austin, Peter C., Paul Grootendorst, and Geoffrey M. Anderson. "A comparison of the ability of different propensity score models to balance measured variables between treated and untreated subjects: a Monte Carlo study." Statistics in medicine 26.4 (2007)	363	36.3	87

Table 3 (continued)

Rank	Article	Citations	Citation per year	Rank by citation per year
96	Sullivan, Lisa M., Joseph M. Massaro, and Ralph B. D'Agostino. "Presentation of multivariate data for clinical use: The Framingham Study risk score functions." Statistics in medicine 23.10 (2004)	363	27.92	70
97	Laird, Nan M. "Missing data in longitudinal studies." Statistics in medicine 7.1–2 (1988)	361	12.45	91
98	Lunceford, Jared K., and Marie Davidian. "Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study." Statistics in medicine 23.19 (2004)	354	27.23	97
99	Thakkinstian, Ammarin, et al. "A method for meta-analysis of molecular association studies." Statistics in medicine 24.9 (2005)	350	29.17	100
100	Begg, Colin B. "Biases in the assessment of diagnostic tests." Statistics in medicine 6.4 (1987)	346	11.53	94

Informatics field, i.e. that the body of literature has flourished in recent years and researchers tend to rely on the latest guidelines. A review of citations in Fig. 1 shows that the authors of the Medical Informatics articles use more up-to-date resources than the forty papers originally published in the 2000s. Noteworthy, all papers in the list of Top cited 100 articles that were retrieved by keyword (ii procedure) were also retrieved through journals related to Medical Informatics searches (i procedure). In other words, all the articles that were retrieved using key words could also be retrieved through the journals search.

Limitations

As with all bibliometric analyses, our study has limitations. The primary limitation is the use of search terms in the creation of the study database. Articles that did not contain our query terms were not retrieved and included in our analysis. However, we collected all journals that published articles in Medical Informatics according to the WOS "Medical Informatics" category. Another limitation is that Medical Informatics is a multidisciplinary field, and Medical

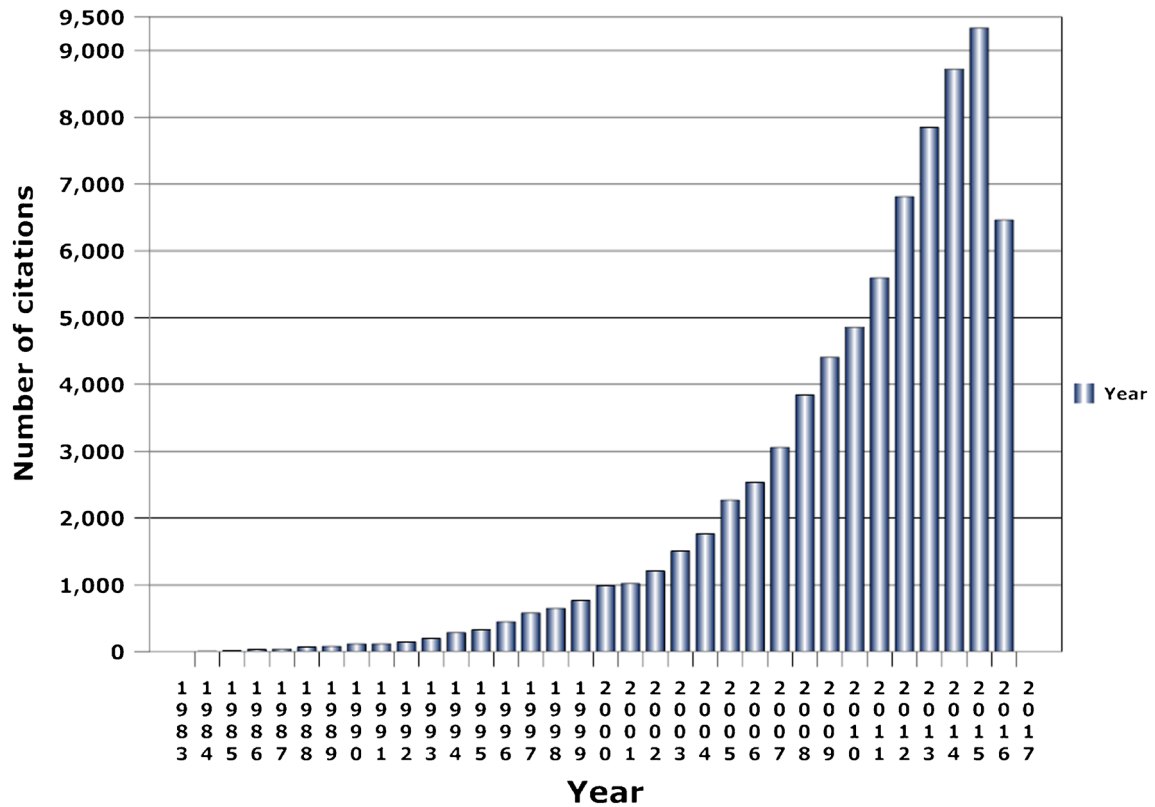

Fig. 2 The distribution of the 100 top cited articles by Citations per year, Web of Knowledge

Informatics articles may be published in non-specialized journals. WOS may not always index them using the Medical Informatics category, and too, recognize their articles which are absolutely related to Medical Informatics are difficult. Moreover, like other bibliometric analyses, our top 100 most cited articles are biased in favor of older publications. However, we also offer the alternative citations per year measure to identify articles with the most impact, regardless of publication year. Finally, articles published in languages other than English might have received unfair citation counts because of bias and poor recognition in the field. Despite these limitations, our investigation provides some insights into the most read and cited articles in the field of Medical Informatics.

Conclusion

This study highlights the role of Medical Informatics in medicine. We found that articles in the statistics and biostatics subarea dominated the Top 100 cited list, followed by articles in the medical artificial intelligence/decision analysis sub-area. We find that our subspecialty bibliometric analyses has revealed the characteristics of highly cited papers, which have implications for Medical Informatics specialists, librarians, researchers, editors, and reviewers. The most cited articles are continuously changing, meaning that the present study provides a snapshot of the most influential articles of the current time, while also showing trends in the literature.

Acknowledgements The authors are grateful to the Research Committee of Urmia University of Medical Sciences, for their valuable supporting of this research.

Authors' Contributions Bahlol Rahimi and Hamed Nadri were the responsible for the study design, data analysis, interpretation of results, and drafting the manuscript.

Toomas Timpka contributed to the design of the study, data analysis, and revision of the manuscript.

Shahram Sedghi contributed to the design of the study and revision of the manuscript.

All authors read and approved the final manuscript.

Compliance with Ethical Standards

Funding There was no funding associated with this research.

Conflict of Interest The authors declare that they have no conflicts of interests.

Research Involving Human Participants and/or Animals This article does not contain any studies with human participants performed by any of the authors.

Informed Consent Not applicable.

References

1. Garfield, E., Citation analysis as a tool in journal evaluation. *Science*. 178:471–479, 1972.
2. Moed, H.F., New developments in the use of citation analysis in research evaluation. *Arch. Immunol. Ther. Exp.* 57:13–18, 2009. doi:10.1007/s00005-009-0001-5.
3. Choudhri, A.F., Siddiqui, A., Khan, N.R., and Cohen, H.L., Understanding bibliometric parameters and analysis. *Radiographics*, A review publication of the Radiological Society of North America, Inc. 35:736–746, 2015. doi:10.1148/rg.2015140036.
4. Hirsch, J.E. An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America*. 102:16569–16572, 2005. doi:10.1073/pnas.0507655102.
5. Cheek, J., Garnham, B., and Quan, J., What's in a number? Issues in providing evidence of impact and quality of research(ers). *Qual. Health Res.* 16:423–435, 2006. doi:10.1177/1049732305285701.
6. Garfield, E., 100 citation classics from the journal of the American Medical Association. *JAMA*. 257:52–59, 1987.
7. Garfield, E., The impact factor and its rightful use. *Anaesthesist*. 47:439, 1998. doi:10.1007/s001010050581.
8. Garfield, E., Journal impact factor: A brief review. *Can. Med. Assoc. J.* 161:979–980, 1999.
9. Glanville, J., Kendrick, T., McNally, R., Campbell, J., and Hobbs, F.R., Research output on primary care in Australia, Canada, Germany, the Netherlands, the United Kingdom, and the United States: Bibliometric analysis. *BMJ*. 342:d1028, 2011.
10. Gu, W., Yuan, Y., Yang, H., Qi, G., Jin, X., and Yan, J., A bibliometric analysis of the 100 most influential papers on COPD. *Int. J. Chron. Obstruct. Pulmon. Dis.* 10:667-676, 2015. doi:10.2147/COPD.S74911.
11. Ye, S., Xing, R., Liu, J., and Xing, F., Bibliometric analysis of Nobelists' awards and landmark papers in physiology or medicine during 1983–2012. *Ann. Med.* 45:532–538, 2013.
12. Joyce, C.W., Sugrue, C.M., Joyce, K.M., Kelly, J.L., and Regan, P.J., 100 citation classics in the melanoma literature: A bibliometric analysis. *Dermatol. Surg.* 40:1284–1298, 2014.
13. Dolan, R.S., Hanna, T.N., Warraich, G.J., Johnson, J.O., and Khosa, F., The top 100 articles in the radiology of trauma: A bibliometric analysis. *Emerg. Radiol.* 22:667–675, 2015. doi:10.1007/s10140-015-1345-2.
14. Pagni, M., Khan, N.R., Cohen, H.L., and Choudhri, A.F., Highly cited works in radiology: The top 100 cited articles in radiologic journals. *Acad. Radiol.* 21:1056–1066, 2014. doi:10.1016/j.acra.2014.03.011.
15. Nayar, S.K., Dein, E.J., Spiker, A.M., Bernard, J.A., and Zikria, B.A., The top 100 cited articles in clinical orthopedic sports medicine. *Am J Orthop (Belle Mead NJ)*. 44:E252–E261, 2015.
16. Pathology International TOP 100 citation articles. *Pathology International*, 65:2–8, 2015. doi:10.1111/pin.12225.
17. Khan, N. R., Lee, S. L., Brown, M., Reding, J., Angotti, J., Lepard, J., Gabrick, K., Klimo, P., Jr., and Michael, L. M., Highly cited works in skull base neurosurgery. *World Neurosurg.*. 83: 403–418, 2015. doi:10.1016/j.wneu.2014.12.005.
18. Hennessey, K., Afshar, K., and Macneily, A.E., The top 100 cited articles in urology. *Can Urol Assoc J*. 3:293–302, 2009.
19. Nason, G.J., Tareen, F., and Mortell, A., The top 100 cited articles in urology: An update. *Can. Urol. Assoc. J*. 7:16–24, 2013.
20. Shuaib, W., Acevedo, J.N., Khan, M.S., Santiago, L.J., and Gaeta, T.J., The top 100 cited articles published in emergency medicine journals. *Am. J. Emerg. Med.* 33:1066–1071, 2015. doi:10.1016/j.ajem.2015.04.047.

21. Wu, J.J., Choi, Y.M., and Marczynski, W., The 100 most cited psoriasis articles in clinical dermatologic journals, 1970 to 2012. *J. Clin. Aesthet. Dermatol*. 7:10–19, 2014.
22. Hasman, A., and Haux, R., Modeling in biomedical informatics–an exploratory analysis (part 1). *Methods Inf. Med.* 45:638–642, 2006.
23. Moorman, P.W., and van der Lei, J., An inventory of publications on computer-based medical records: An update. *Methods Inf. Med.* 42:199–202, 2003. doi:10.1267/meth03030199.
24. DeShazo, J.P., LaVallie, D.L., and Wolf, F.M., Publication trends in the medical informatics literature: 20 years of" Medical Informatics" in MeSH. *BMC Med. Inform. Decis. Mak*. 9:1, 2009.
25. Mendis, K., Health informatics research in Australia: Retrospective analysis using PubMed. *Inform. Prim. Care*. 15:17–23, 2007.
26. Otero, P., Pedernera, F., Montenegro, S., Borbolla, D., Garcia Marti, S., Luna, D., and de Quiros, F.G., Evolution of medical informatics in bibliographic databases. *Stud. Health Technol. Inform.* 107:301–305, 2004.
27. Morris, T.A., and McCain, K.W., The structure of medical informatics journal literature. *J. Am. Med. Inform. Assoc.* 5:448–466, 1998.
28. Sittig, D.F., and Kaalaas-Sittig, J., A citation analysis of medical informatics journals. *Medinfo. MEDINFO*. 8(Pt 2):1452–1456, 1995.
29. Yoon, D.Y., Yun, E.J., Ku, Y.J., Baek, S., Lim, K.J., Seo, Y.L., and Yie, M., Citation classics in radiology journals: The 100 top-cited articles, 1945–2012. *Am. J. Roentgenol.* 201:471–481, 2013.
30. Eysenbach, G., Consumer health informatics. *BMJ: British Medical Journal*. 320:1713–1716, 2000.
31. Harris, P.A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., and Conde, J.G., Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J. Biomed. Inform.* 42: 377–381, 2009.
32. Lefaivre K A, Shadgan B and O'Brien PJ (2011) *100 most cited articles in orthopaedic surgery.* Clinical Orthopaedics and related research® 469: 1487-1497.
33. Paladugu, R., Schein, M., Gardezi, S., and Wise, L., One hundred citation classics in general surgical journals. *World J. Surg.* 26: 1099–1105, 2002.
34. Ponce, F.A., and Lozano, A.M., Highly cited works in neurosurgery. Part I: The 100 top-cited papers in neurosurgical journals: A review. *J. Neurosurg.* 112:223–232, 2010.