Published in final edited form as:

Stat Med. 2013 January 15; 32(1): 1–10. doi:10.1002/sim.5311.

Characteristics of Recent Biostatistical Methods Adopted by Researchers Publishing in General / Internal Medicine Journals

Paul J Nietert, PhD¹, Amy E Wahlquist, MS¹, and Teri Lynn Herbert, MS, MLIS²

¹Division of Biostatistics and Epidemiology, Department of Medicine

²Department of Library Science and Informatics Medical University of South Carolina

Abstract

Background—Novel statistical methods are constantly being developed within the context of biomedical research; however, the characteristics of biostatistics methods that have been adopted into the field of general / internal medicine (GIM) is unclear. This study highlights the statistical journal articles, the statistical journals, and the types of statistical methods that appear to be having the most direct impact on GIM research.

Methods—Descriptive techniques, including analyses of articles' keywords and controlled vocabulary terms, were used to characterize the articles published in statistics and probability journals that were subsequently referenced within GIM journal articles during a recent 10-year period (2000–2009).

Results—From the 45 statistics and probability journals of interest, a total of 989 unique articles were identified as being cited by 2,183 (out of a total of about 127,469) unique GIM journal articles. The most frequently cited statistical topics included general/other statistical methods, followed by randomized trials, epidemiologic methods, meta-analysis, generalized linear models, and computer simulation.

Conclusion—As statisticians continue to develop and refine techniques, the promotion and adoption of these methods should also be addressed so that their efforts spent in developing the methods are not done in vain.

Keywords

bibliometrics; biostatistical methods; general/internal medicine; journal impact factor

Introduction

Novel statistical methods are constantly being developed. The traditional routes of disseminating the new techniques typically involve presentations at local, regional, or national research meetings and publication in peer-reviewed statistical journals. However, it is unclear whether these communication strategies are effective at having measureable impact on real-world problems, such as in the field of general/internal medicine (GIM) research. In this paper, we quantify the extent to which recent biostatistics methods have been adopted into GIM research studies and provide some description of frequently cited biostatistics methods papers.

Please address all correspondence to: Paul J Nietert, PhD, Division of Biostatistics and Epidemiology, Department of Medicine, Medical University of South Carolina, 135 Cannon Street, Suite 303, PO Box 250835, Charleston, SC 29425, nieterpj@muscc.edu, P: 843-876-1204, F: 843-876-1126.

One way of measuring a journal article's influence is by counting how often it is cited in other publications. Although sometimes criticized for doing so [1–4], journals themselves use varieties of impact factors to measure their influence and to attract article submissions. These impact factors are based, in part, on citation frequency counts. Even though the journal impact factor is recognized as an imperfect measure, it is nonetheless considered one of the “best” ways to evaluate the “quality” or influence of a journal. For a number of reasons, prior research suggests that articles in statistical journals have slower rates of citation diffusion than are typically seen in other disciplines [5]. Additionally, within the biostatistics field, method adoption by clinical researchers and hence citation frequency may also be limited by the available software.

In this study, we were interested in determining which statistical articles, statistical journals, and statistical methods have the most direct impact on the field of GIM research. This was accomplished using descriptive techniques involving statistical journal articles and GIM articles published during a recent 10-year period (2000–2009).

Methods

The first step was to identify the journals in the “Statistics and Probability” category so defined in the Journal Citation Reports Science Edition from Thomson Reuters, Philadelphia, PA. According to the Thomson Reuters web site [6], this category “covers resources concerned with methods of obtaining, analyzing, summarizing, and interpreting numerical or quantitative data. Resources on the study of the mathematical structures and constructions used to analyze the probability of a given set of events from a family of outcomes are also covered.” Sorted according to their 2009 calculated Impact Factor, the top 40 journals (whose impact factor ranged from 1.13 to 4.00) were then selected for this study. In addition, we augmented this list with 5 other journals (*Canadian Journal of Statistics*, *Journal of Multivariate Analysis*, *Journal of Statistical Planning and Inference*, *Scandinavian Journal of Statistics*, and *Statistica Sinica*) that are focused in the area of biostatistical methodology that were believed (*a priori*) to contain methods possibly cited frequently in the GIM literature. For each of the 45 individual journals (Table 1), we searched the ISI Web of Science (ISIWoS) Science Citation Index Expanded Database to identify published articles in the ISIWoS category labeled “Medicine, General and Internal” that cited one or more articles in the statistics/probability journal of interest. See the Appendix for a listing of these journals. For the purposes of this study, both the *citing* article and the *cited* article must have been English language manuscripts published during the years 2000 through 2009. We also counted the total number of journal articles (including original and review articles and any published corrections) categorized as “Medicine, General and Internal” during the study time period. Because of the natural lag time for any citations to occur, biostatistical articles published later in the decade would be less likely to have been cited; thus we stratified our findings by 2 time periods, 2000–2006 and 2007–2009.

The processes described above resulted in two lists, one of articles in the GIM category that cited one or more articles in the statistics/probability category, and the other a list of articles in the statistics/probability category cited by one or more articles in the GIM category. The lists were managed in MS Excel 2007 (Redmond, WA), Thomson Reuters EndNote X3 (Philadelphia, PA), and SAS v9.3 (Cary, NC). For each of the statistics/probability articles, we counted the number of times each was cited by articles in the GIM listing.

For each of the cited articles, we also retrieved the National Library of Medicine's controlled vocabulary Medical Subject Headings (MeSH) and/or author-assigned keywords, when available. In order to obtain some type of understanding of the nature of the articles being cited, these MeSH and author-assigned keywords were mapped to one or more of the

following broader methodological categories: analysis of correlated data, Bayesian methods, bias, bioinformatics methods, computer simulation, diagnostic testing, econometrics, epidemiology, general/other statistical terms, generalized linear models, meta-analysis, missing data, multiple comparisons, non/semi-parametrics, psychometrics, randomized trials, reproducibility of results, sample size / power estimation, statistical software, study design, survival analysis, and non-statistical terms (including demographics, medical terms, etc.). The mapping was independently reviewed by two biostatisticians (PJN and AEW), and a consensus was reached on each term's / keyword's category mapping for any discrepancies.

Using the techniques described, several characteristics of the cited articles were examined, including the most frequently cited journal articles and journals. In addition, we were able to characterize the most frequently cited biostatistics methods using a simple descriptive analysis of the MeSH / keyword mappings.

Results

From 2000 through 2009, the ISIWoS search resulted in the identification of a total of 127,469 unique English language journal articles published in the GIM field. Of these articles, 2,183 (1.7%) included a citation of at least one of the 45 statistics and probability journals of interest published during the assigned time frame. Some of the GIM articles cited more than one statistics and probability article, for a total of 3,018 citations of 989 unique statistics and probability articles. Among individual cited articles, 60.8% of the articles were cited only once. Table 2 lists the top 10 articles (with 4 tied for 10th place) most frequently cited statistics and probability journal articles from 2000 through 2006, which together accounted for 31.6% of all the citations during that time period. Of these articles, the focal topics that emerged included meta-analysis, publication bias, missing data, epidemiologic methods, and prognostic models. Table 3 highlights the top 5 most frequently cited statistics and probability journal articles from 2006 through 2009, which accounted for 24.9% of the citations during that time period. Interestingly, several authors (DG Altman, RB D'Agostino Jr., RB D'Agostino Sr., S Duval, JP Higgins, MJ Pencina, P Royston, SG Thompson, R Tweedie, and RS Vasan) have more than 1 article in these Tables, and it appears that different articles by the same author may be addressing common methodologic issues.

Table 4 lists the top 10 most frequently cited statistics and probability journals from the GIM journal articles. Of the 3,018 citations, *Statistics in Medicine* accounted for n=2,051, or 68.0%, of the total, followed by *Biometrics* (n=193, 6.4%) and *Statistical Methods in Medical Research* (n=147, 4.9%). Some of the 45 journals were never cited in the GIM literature during this study time frame.

In Table 5, each of the keyword mapping categories is displayed, along with the most frequently occurring individual keyword terms included in the categories and the citation frequencies associated with the categories. Aside from non-statistical terms, general/other statistical keywords were cited most often (n=4,806 citations), followed by terms classified under the headings of randomized trials (n=2,003), epidemiology (n=1,463), meta-analysis (n=1,223), generalized linear models (n=1,043), and computer simulation (n=619). There are more citations than articles simply due to the fact that each citing article may cite multiple statistics and probability articles, which, in turn, often include multiple keywords.

Discussion

This study highlights relative trends from 2000–2009 in the applications of statistics and probability within the GIM research literature. Less than 2% of the GIM journal articles

cited an article published in any statistics and probability journal during this time period. Among 989 cited statistics and probability articles, the most frequently cited articles address topics including meta-analysis, publication bias, methods for handling missing data, epidemiologic methods, randomized trial networks, and prognostic models.

Three journals (*Statistics in Medicine*, *Biometrics*, and *Statistical Methods in Medical Research*) accounted for 79.2% of all the citations within 2,183 unique GIM journal articles. We speculate that this is due to several factors, including belief by biostatistical authors that these journals provide an excellent means to disseminate their work and the editorial screening process by which editors of these journals select only the articles that are most likely to have an impact on medical research.

Overall, the most frequently cited statistical keywords were classified as general/other statistical terms, followed by randomized trials, epidemiology, meta-analysis, generalized linear models, and computer simulation. In many respects, the results of the keyword analysis in this citation study are not entirely surprising. GIM research is essentially focused on identifying risk factors for disease, identifying optimal treatments, and finding ways to prevent disease, which typically rely on epidemiologic techniques and randomized clinical trials. Meta-analysis, which require specific attention to relevantly novel statistical methods, provides a means for even the most junior GIM investigator to publish highly relevant research, which may help explain the abundance of meta-analytic methods citations. If we had conducted a similar study from an earlier decade (e.g. 1980–1989 or 1990–1999), the topics identified may have been quite similar to the ones we highlight in this paper. In fact, a paper in 2005 featured the then top 25 most-cited statistical papers of all time[7], and the top 10 papers addressed topics including survival analysis[8–11], multiple comparisons[12], least squares estimation[13], measurement agreement[14], randomized trials[10, 15], bootstrapping[16], epidemiological methods[17].

From our study it is clear that GIM researchers definitely recognize the need for sound biostatistical methods related to meta-analysis, handling missing data within studies, using computer simulation, analysis of correlated data, and understanding the influence of various biases. In some respects however, our study may illustrate a different problem - the fact that relatively few GIM research articles are referencing novel statistical methods. In many areas of medicine, it can take years (perhaps 1 or 2 decades) for research to translate into routine clinical practice [18], so perhaps we should not be too surprised at the relatively low uptake of novel biostatistical methods into the GIM field. The biostatistical community may need to consider studying better ways of disseminating their findings. Van Nierop (2009) argues using diffusion curve analysis that it simply takes longer for articles in statistical journals to become frequently cited [5]. There is an entire branch of translational research dealing with the science of dissemination (typically of evidence-based treatments)[19], and there may be lessons learned in that arena which could help biostatisticians promote better methodological practices. Using an analogy involving drug development, if resources spent on the successful discovery of a new drug do not result in the drug being used to treat/prevent disease, then resources have been wasted. Similarly, if resources are used to develop novel/improved statistical methods that are never utilized, then resources have again been wasted. As with a new treatment that has been developed, successful communication and dissemination of novel biostatistical methods will likely need to involve statisticians' becoming more aware of relevant methodological gaps, promoting the novel ideas and methods, adapting them to different settings, and revising the methods to make them better.

Our study is limited by several factors. There are novel and improved statistical methods that are published in journals that do not focus on statistical methods (e.g. clinical journals, bioinformatics journals) and would not have met the criteria for being included in this study;

however, identifying such articles would have required more resources (e.g. having to search and review perhaps tens of thousands of journal articles) than were available for this study. Also, since our time frame for the cited articles coincided with that of the citing articles, it is possible that there may already exist published novel statistical methods, not identified in this study, that may ultimately be extremely influential in future GIM research. In other words, a highly innovative and useful statistical method published in 2009 may not start being cited frequently for a few years outside of this study's time frame. There is often a significant lag between statistical methods development and acceptance in other areas of medical research. For example, in 1904 Karl Pearson is credited to have published the first meta-analysis [20]. However, the term 'meta-analysis' was not officially used until 1976 [21], and only afterwards did applications of the technique become routine [22]. Thus it is possible that a relatively obscure biostatistical paper published between 2000 and 2009 could be one of the most highly cited 10 or 20 years from now. Additionally, some GIM researchers may not feel comfortable using state-of-the-art statistical methods until they have been tested in circumstances beyond those anticipated in the original methods publication, resulting in relatively slow translation (uptake) of novel methods. Also, since some traditional and common statistical methods (e.g. t-tests, Cox regression) are not typically cited, it is possible that some GIM researchers may not deem it necessary to cite biostatistical methods at all, which could account for the relatively small number of cited articles identified. Unless a biostatistician is involved with the study, the GIM researchers may not be aware of more appropriate and powerful approaches to address their research questions. Had we performed a similar analysis in a different field of medical research (e.g. cancer), we would likely have obtained vastly different results, both in the quantity and types of statistical articles cited. Also, although our categorization of key words was performed by 2 biostatisticians, these classifications are limited to some extent by the nature of keywords themselves; keywords tend to be broad descriptions of the paper's contents and thus may not capture the novel aspect of the paper's contribution to the literature. Finally, there may be some types of biostatistical papers (e.g. ones that rely on more advanced theoretical mathematics) that may tend to be overlooked by clinical investigators but which are highly influential among biostatisticians, who are able to adapt the methods and publish the adaptations in journals that are more likely to be read and appreciated by clinical investigators. This type of indirect influence by more technical publications warrants further study.

The field of developing statistical methods has a bright future. With emerging emphases from the National Institutes of Health on topics such as translational research, comparative effectiveness research, community-based participatory research, and personalized medicine, it will be imperative that statisticians continue to develop and refine techniques to ensure accurate analyses and efficient uses of resources. However, they should also be mindful of finding optimal ways to promote and market their methods so that the effort for having developed the methods is not done in vain. Such techniques might include writing easily adapted computer programs (e.g. SAS ® macros, R packages), presenting methods within clinical scientific settings (from local journal clubs to international gatherings), or providing more user-friendly descriptions of the techniques on internet sites and/or via social media applications.

Acknowledgments

Grant Support

The project described was supported by Award Number UL1RR029882 from the National Center for Research Resources. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Center for Research Resources or the National Institutes of Health.

References

1. O'Connor SJ. Citations, impact factors and shady publication practices: how should the lasting clinical and social value of research really be measured? *European Journal of Cancer Care*. 2010; 19:141–143. [PubMed: 20486987]
2. Hernán MA. Epidemiologists (of all people) should question journal impact factors. *Epidemiology*. 2008; 19:366–368. [PubMed: 18414080]
3. Reedijk J, Moed F. Is the impact of journal impact factors decreasing? *Journal of Documentation*. 2008; 64:183–192.
4. Feller SM. Beyond journal impact factors? *Cell Commun Signal*. 2010; 8:4. [PubMed: 20219121]
5. van Nierop E. Why do statistics journals have low impact factors? *Statistica Neerlandica*. 2009; 63:52–62.
6. Thomson Reuters. [Accessed August 5, 2011] Scope Notes: 2011 Science Citation Index Expanded. http://science.thomsonreuters.com/mjl/scope/scope_scie/
7. Ryan TP, Woodall WH. The most-cited statistical papers. *Journal of Applied Statistics*. 2005; 32:461–474.
8. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*. 1958; 53:457–481.
9. Cox DR. Regression models and life tables. *Journal of the Royal Statistical Society, Series B*. 1972; 34:187–220.
10. Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV, Mantel N, McPherson K, Peto J, Smith PG. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. analysis and examples. *British Journal of Cancer*. 1977; 35:1–39. [PubMed: 831755]
11. Mantel N. Evaluation of survival data and two new rank order statistics arising in its consideration. *Cancer Chemotherapy Reports*. 1966; 50:163–170.
12. Duncan DB. Multiple range and multiple F-tests. *Biometrics*. 1955; 11:1–42.
13. Marquardt DW. An algorithm for least squares estimation of non-linear parameters. *Journal of the Society for Industrial and Applied Mathematics*. 1963; 2:431–441.
14. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986; 1:307–310. [PubMed: 2868172]
15. Litchfield JT Jr, Wilcoxon F. A simplified method of evaluating dose-effect experiments. *The Journal of Pharmacology and Experimental Therapeutics*. 1949; 96:99–113. [PubMed: 18152921]
16. Felsenstein J. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*. 1985; 39:783–791.
17. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute*. 1959; 22:719–748. [PubMed: 13655060]
18. Agency for Healthcare Research and Quality. [Accessed August 8, 2011] Translating Research Into Practice (TRIP)-II Fact Sheet. <http://www.ahrq.gov/research/trip2fac.htm>
19. Dearing JW. Evolution of diffusion and dissemination theory. *Journal of Public Health Management and Practice*. 2008; 14:99–108. [PubMed: 18287914]
20. Pearson K. Report on certain enteric fever inoculation statistics. *British Medical Journal*. 1904; 2:1243–1246. [PubMed: 20761760]
21. Glass GV. Primary, secondary and meta-analysis of research. *Education Researcher*. 1976; 10:3–8.
22. O'Rourke K. An historical perspective on meta-analysis: dealing quantitatively with varying study results. *Journal of the Royal Society of Medicine*. 2007; 100:579–582. [PubMed: 18065712]

Table 1

Journals used in the analysis (in alphabetical order).

Journal Title	2009 Impact Factor [†]
<i>American Statistician</i>	1.252
<i>Annals of Applied Probability</i>	1.130
<i>Annals of Applied Statistics</i>	2.571
<i>Annals of Probability</i>	1.260
<i>Annals of Statistics</i>	3.185
<i>Biometrical Journal</i>	1.208
<i>Biometrics</i>	1.867
<i>Biometrika</i>	1.933
<i>Biostatistics</i>	3.246
<i>Canadian Journal of Statistics</i> *	0.905
<i>Chemometrics and Intelligent Laboratory Systems</i>	2.111
<i>Computational Statistics and Data Analysis</i>	1.228
<i>Econometric Reviews</i>	1.745
<i>Econometrica</i>	4.000
<i>Environmental and Ecological Statistics</i>	1.180
<i>Finance and Stochastics</i>	1.240
<i>Fuzzy Sets and Systems</i>	2.138
<i>IEEE/ACM Transactions on Computational Biology and Bioinformatics</i>	2.246
<i>Journal of Business and Economic Statistics</i>	1.562
<i>Journal of Chemometrics</i>	1.291
<i>Journal of Computational and Graphical Statistics</i>	1.258
<i>Journal of Computational Biology</i>	1.694
<i>Journal of Multivariate Analysis</i> *	1.017
<i>Journal of Quality Technology</i>	1.500
<i>Journal of Statistical Planning and Inference</i> *	0.725
<i>Journal of Statistical Software</i>	2.320
<i>Journal of the American Statistical Association</i>	2.322
<i>Journal of the Royal Statistical Society (A: Statistics in Society)</i>	1.690
<i>Journal of the Royal Statistical Society (B: Statistical Methodology)</i>	3.473
<i>Multivariate Behavioral Research</i>	2.328
<i>Pharmaceutical Statistics</i>	1.957
<i>Probabilistic Engineering Mechanics</i>	1.221
<i>Probability Theory and Related Fields</i>	1.373
<i>Scandinavian Journal of Statistics</i> *	1.022
<i>Stata Journal</i>	1.851
<i>Statistica Sinica</i> *	0.945
<i>Statistical Applications in Genetics and Molecular Biology</i>	2.247
<i>Statistics and Computing</i>	1.821

Journal Title	2009 Impact Factor [†]
<i>Statistical Methods in Medical Research</i>	2.569
<i>Statistical Science</i>	3.523
<i>Statistics in Medicine</i>	1.990
<i>Stochastic Environmental and Research Risk Assessment</i>	1.419
<i>Stochastic Processes and their Applications</i>	1.543
<i>Technometrics</i>	1.711
<i>TEST</i>	1.241

* Not one of the top 40 journals as ranked by the 2009 Science Citation Index impact factor

[†] Impact factors were obtained from the ISI Web of Knowledge Journal Citation Reports

Table 2

Statistical journal articles most frequently cited within general / internal medicine research articles, 2000–2006.

Full Citation	Citation Count
Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. <i>Statistics in Medicine</i> , 2002, 21(11): 1539–1558.	361
Donner A, Klar N. Issues in the meta-analysis of cluster randomized trials. <i>Statistics in Medicine</i> , 2002, 21(19): 2971–2980.	89
Sweeting MJ, Sutton AJ, Lambert PC. What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data. <i>Statistics in Medicine</i> , 2004, 23(9): 1351–1375.	57
Macaskill P, Walter SD, Irwig L. A comparison of methods to detect publication bias in meta-analysis. <i>Statistics in Medicine</i> , 2001, 20(4): 641–654.	56
Duval S, Tweedie R. Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. <i>Biometrics</i> , 2000, 56(2): 455–463.	53
Thompson SG, Higgins JP. How should meta-regression analyses be undertaken and interpreted? <i>Statistics in Medicine</i> , 2002, 21(11): 1559–1573.	53
Lumley T. Network meta-analysis for indirect treatment comparisons. <i>Statistics in Medicine</i> , 2002, 21(16): 2313–2324.	28
Altman DG, Royston P. What do we mean by validating a prognostic model? <i>Statistics in Medicine</i> , 2000, 19(4): 453–473.	27
Kim HJ, Fay MP, Feuer EJ, Midthune DN. Permutation tests for joinpoint regression with applications to cancer rates. <i>Statistics in Medicine</i> , 2000, 19(3): 335–351.	26
Curtin F, Altman DG, Elbourne D. Meta-analysis combining parallel and cross-over clinical trials. I: Continuous outcomes. <i>Statistics in Medicine</i> , 2002, 21(15): 2131–2144.	24
Duval S, Tweedie R. A nonparametric “trim and fill” method of accounting for publication bias in meta-analysis. <i>Journal of the American Statistical Association</i> , 2000, 95(449): 89–98.	24
Royston P. Multiple imputation of missing values. <i>Stata Journal</i> , 2004, 4(3): 227–241.	24
Williams RL. A note on robust variance estimation for cluster-correlated data. <i>Biometrics</i> , 2000, 56(2): 645–646.	24

Table 3

Statistical journal articles most frequently cited within general / internal medicine research articles, 2007–2009.

Full Citation	Citation Count
Pencina MJ, D'Agostino RB, Sr., D'Agostino RB, Jr., Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. <i>Statistics in Medicine</i> , 2008, 27(2): 157–172.	33
Bradburn MJ, Deeks JJ, Berlin JA, Russell Localio A. Much ado about nothing: a comparison of the performance of meta-analytical methods with rare events. <i>Statistics in Medicine</i> , 2007, 26(1): 53–77.	22
Salanti G, Higgins JP, Ades AE, Ioannidis JP. Evaluation of networks of randomized trials. <i>Statistical Methods in Medical Research</i> , 2007, 17: 279–301.	11
Austin P C, Grootendorst P, Anderson GM. A comparison of the ability of different propensity score models to balance measured variables between treated and untreated subjects: a Monte Carlo study. <i>Statistics in Medicine</i> , 2007, 26: 734–753.	9
Pencina MJ, D'Agostino RB, Sr., D'Agostino RB, Jr., Vasan RS. Comments on 'Integrated discrimination and net reclassification improvements – practical advice'. <i>Statistics in Medicine</i> , 2008, 27(2): 207–212.	9

Table 4

Top 10 most frequently cited statistics and probability journals

Journal Title	Citation Frequency
<i>Statistics in Medicine</i>	2,051
<i>Biometrics</i>	193
<i>Statistical Methods in Medical Research</i>	147
<i>Stata Journal</i>	101
<i>Biostatistics</i>	87
<i>Journal of the American Statistical Association</i>	83
<i>Journal of the Royal Statistical Society (A: Statistics in Society)</i>	56
<i>American Statistician</i>	52
<i>Biometrical Journal</i>	27
<i>Pharmaceutical Statistics</i>	27

Table 5

Keyword mappings categories and citation frequencies

Keyword mapping category	Common individual keywords within mapping category	Number of times keyword mapping category was cited
General/other statistical	models, statistical; data interpretation, statistical; biometry	4,806
Randomized trials	randomized controlled trials as topic; clinical trials as topic; treatment outcome	2,003
Epidemiology	risk factors; risk; epidemiologic methods	1,463
Meta-analysis	meta-analysis; metaregression; inverse probability weighted estimation	1,223
Generalized linear models	analysis of variance; regression analysis; logistic models	1,043
Computer simulation	computer simulation; Monte Carlo method; Markov chains; bootstrap	619
Analysis of correlated data	cluster analysis; longitudinal studies; mixed models	610
Missing data	missing data; multiple imputation; missing at random	529
Bias	bias (epidemiology); publication bias; file drawer problem	451
Survival analysis	survival analysis; proportional hazards models; survival-time	435
Sample size / power estimation	sample size; power; statistical power	426
Diagnostic testing	sensitivity and specificity; ROC curve; predictive value of tests	355
Bayesian methods	Bayes' theorem; Winbugs; empirical Bayes	349
Study design	research design; design; 2-period crossover design	333
Bioinformatics methods	models, biological; oligonucleotide array sequence analysis; gene expression profiling	298
Statistical software	Software; st0067 (a Stata program for multiple imputation); databases, factual	194
Econometrics	cost-benefit analysis; cost analysis; costs and cost analysis	192
Non/Semi-parametrics	statistics, nonparametric; semiparametrics; nonparametric estimation	156
Reproducibility of results	reproducibility of results	104
Psychometrics	psychometrics; kappa statistic	64
Multiple comparisons	Bonferroni procedure; false discovery rate; multiple comparisons procedures	57
Non-statistical terminology	humans; female; male	16,097

Appendix

Journals categorized as “Medicine, General and Internal” by the ISI Web of Science (ISIWoS).ⁱ Note that some of these journals may not be published in English and would therefore have been excluded from our analyses.

Acta Clinica Belgica
Acta clinica Croatica
Acta Medica Mediterranea
American Family Physician
The American Journal of Chinese Medicine
American Journal of Managed Care
The American Journal of Medicine
The American Journal of the Medical Sciences
American Journal of Preventive Medicine
Amyloid
Annals, Academy of Medicine, Singapore
The Annals of Family Medicine
Annals of Internal Medicine
Annals of Medicine
Annals of Saudi Medicine
Annual Review of Medicine
Archives of Internal Medicine
Archives of Iranian Medicine
Archives of Medical Science
Atención Primaria
Australian Family Physician
Aviation, Space, and Environmental Medicine
Bulletin de l'Academie Nationale de Medecine
BMC Family Practice
BMC Medicine
Bratislava Medical Journal
The British Journal of General Practice
British Journal of Hospital Medicine
British Medical Bulletin
British Medical Journal
Canadian Family Physician
Canadian Medical Association Journal
Central European Journal of Medicine
Chinese Medical Journal
Cleveland Clinic Journal of Medicine
Clinical Medicine
Clinics
Cochrane Database of Systematic Reviews
Croatian Medical Journal

Current Medical Research and Opinion
Danish Medical Bulletin
Deutsche Medizinische Wochenschrift
Disease-A-Month
Deutsches Ärzteblatt International
European Journal of Clinical Investigation
European Journal of Internal Medicine
Family Practice
Gaceta Medica De Mexico
Gender Medicine
Healthmed
Indian Journal of Medical Research
International Journal of Clinical Practice
International Journal of Osteopathic Medicine
International Medical Journal
Internal and Emergency Medicine
Internal Medicine
Internal Medicine Journal
Internist
Iranian Red Crescent Medical Journal
Irish Journal of Medical Science
Israel Medical Association Journal
Journal of the American Board of Family Medicine
Journal of Evaluation in Clinical Practice
Journal of Family Practice
Journal of the Formosan Medical Association
Journal of General Internal Medicine
Journal of Hospital Medicine
Journal of Internal Medicine
Journal of Investigative Medicine
Journal of the Korean Medical Association
Journal of Korean Medical Science
Journal of the National Medical Association
Journal of Pain and Symptom Management
Journal of Postgraduate Medicine
Journal of the Royal Society of Medicine
Journal of Travel Medicine
Journal of Urban Health-Bulletin of the New York Academy of Medicine
Journal of Women's Health
JAMA-Journal of the American Medical Association
JCPSP-Journal of the College of Physicians and Surgeons Pakistan
Lancet
Mayo Clinic Proceedings

Medical Clinics of North America
Medicina Clinica
Medicinski Glasnik
Medical Journal of Australia
Medizinische Klinik
Medicina-Lithuania
Medical Principles and Practice
Medical Problems of Performing Artists
Medicina Dello Sport
Medicina-Buenos Aires
Medicine
Minerva Medica
Mount Sinai Journal of Medicine
National Medical Journal of India
Netherlands Journal of Medicine
New England Journal of Medicine
Nobel Medicus
Pain Medicine
Pakistan Journal of Medical Sciences
Palliative Medicine
Panminerva Medica
PLOS Medicine
Postgraduate Medical Journal
Postgraduate Medicine
Presse Medicale
Preventive Medicine
Primary Care
Primary Care & Community Psychiatry
QJM-An International Journal of Medicine
Revista Clinica Espanola
Revista Da Associacao Medica Brasileira
Revista De Investigacion Clinica
Revista Medica De Chile
Revue De Medecine Interne
SAMJ South African Medical Journal
Sao Paulo Medical Journal
Saudi Medical Journal
Scandinavian Journal of Primary Health Care
Scottish Medical Journal
Southern Medical Journal
Swiss Medical Weekly
Terapevticheskii Arkhiv
Tohoku Journal of Experimental Medicine

Trakya Universitesi Tip Fakultesi Dergisi

Translational Research

Turkish Journal of Medical Sciences

Turkiye Klinikleri Tip Bilimleri Dergisi

Uppsala Journal of Medical Sciences

West Indian Medical Journal

Wiener Klinische Wochenschrift

Yonsei Medical Journal

ⁱThompson Reuters (2009). The ISI Web of Science (ISIWoS) Science Citation Index Expanded Database –journal listing for the “Medicine, General & Internal” category.