Assessing Prevention Research Impact
A Bibliometric Analysis

Adele L. Franks, MD, Eduardo J. Simoes, MD, MSc, MPH, Rajdeep Singh, MBBS, MS, Barbara Sajor Gray, MIA, MLn

Background: This study was undertaken to explore a bibliometric approach to assessing the impact of selected prevention research center (PRC) peer-reviewed publications.

Methods: The 25 eligible PRCs were asked to submit 15 papers that they considered the most important to be published in the decade 1994-2004. Journal articles (n =227) were verified in 2004 and categorized: 73% were research reports, 10% discussion articles, 9% dissemination articles, and 7% review articles.

Results: Only 189 articles (83%) were searchable via the Institute of Scientific Information (ISI) Web of Science databases for citation tracking in 2004. These 189 articles were published in 76 distinct journals and subsequently cited 4628 times (range 0 to 1523) in 1013 journals. Articles published before 2001 were cited a median of 14 times each. Publishing journals had a median ISI impact factor of 2.6, and ISI half-life of 7.2. No suitable benchmarks were available for comparison. The PRC influence factor (number of PRCs that considered a journal highly influential) was only weakly correlated with the ISI impact factor and was not correlated with half-life.

Conclusions: Conventional bibliometric analysis to assess the scientific impact of public health prevention research is feasible, but of limited utility because of omissions from ISI's databases, and because citation benchmarks for prevention research have not been established: these problems can and should be addressed. Assessment of impact on public health practice, policy, or on the health of populations, will require more than a bibliometric approach.
(Am J Prev Med 2006;30(3):211–216)

Introduction

The Prevention Research Centers (PRC) program is the largest extramural research program supported by the Centers for Disease Control and Prevention (CDC). Our purpose was to characterize selected publications emerging from this program over the past decade and explore the feasibility of assessing their scientific impact using a bibliometric approach.

The PRC program was authorized by Congress in 1984 to support innovative applied research to improve disease prevention and health promotion in communities.[1] This program was created to bring academic institutions into partnerships with disadvantaged communities and public health professionals in an effort to identify feasible ways to improve health in difficult real-world settings such as urban poverty. The program grew from three initial centers funded in 1986 to 28 centers in 2003. In 2003, CDC funding exceeded $40 million dollars ($27 million in core spending and about $17 million in special projects). Over the past 20 years, PRC researchers have built coalitions with diverse communities around the United States and have implemented numerous projects relevant to these communities. To some public health practitioners, community members, and researchers, the program's success is obvious. A formal assessment of the PRC program conducted in 1997 by the Institute of Medicine made several recommendations for improvement.[2] Among the recommendations were better methods for tracking the centers' peer-reviewed publications and for documenting the program's contributions to public health.

Although new systems (an electronic information system and an evaluation plan named DEFINE [developing an evaluation framework: insuring national excellence]) have been established to monitor future productivity,[3,4] past accomplishments had not been systematically tracked. The effort described here is part

From the Prevention Research Centers Program (Franks, Simoes, Gray), National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention, Atlanta, Georgia, and Office of Workforce and Career Development (Singh), National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention, Atlanta, Georgia

Address correspondence and reprint requests to: Adele Franks, c/o Division of Adult and Community Health, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention, MS K-45, 4770 Buford Highway, Atlanta GA 30341. E-mail: afranks@cdc.gov.

The full text of this article is available via AJPM Online at www.ajpm_online.net.

0749-3797/06/$–see front matter
doi:10.1016/j.amepre.2005.10.025

of a larger undertaking to highlight the PRCs' accomplishments since 1994, when the program could be considered mature enough to have had an impact. A well-accepted and necessary measure of a research program's productivity is the number and type of peer-reviewed journal articles published, and a commonly used measure of scientific impact is the number of times that these articles are cited by other researchers. This paper characterizes published journal articles that the PRCs identified as being the most important to have emerged from their centers. We reviewed the number of subsequent citations of these articles in other published papers; the characteristics of the journals (circulation, impact factor, half-life) in which they were published; and the types of articles that the PRCs selected as most important. In undertaking this bibliometric analysis, we expected that some parameters would have limitations. We were nevertheless interested in exploring the feasibility of this approach to document scientific impact.

Methods

Each PRC was asked to list up to15 articles it considered the most important to have been published by its group in peer-reviewed journals between 1994 and 2004. Each was also asked to name the journals it considered to be the most influential in its field(s) of interest. Of the 28 centers, three were too new to have generated any publications and were therefore exempted. Twenty-four of the remaining 25 PRCs submitted information (96%).

More than 300 original manuscripts or publication citations were submitted. After restricting the submissions to articles published in peer-reviewed journals (i.e., excluding unpublished articles as well as published book chapters, books, and toolkits that are not tracked by citation databases), 227 articles published in 93 journals were verified through PubMed, the National Library of Medicine's biomedical publication database. Verification took place between December 2003 and March 2004. From January 2004 through April 2004, these papers were searched in the Thomson Institute of Scientific Information (ISI) Web of Science databases[5] to obtain information on ISI-derived journal characteristic measures and the subsequent citations of the papers. Specifically, two ISI databases were searched: the Science Citation Index Expanded database, which includes 5700 journals in 164 disciplines; and the Social Sciences Citation Index, which includes over 1725 journals covering 50 disciplines. For each PRC article found in these databases, information was extracted on the publishing journals, the total number of times the article was subsequently cited, and the citing journals. Of the 227 verified articles, 189 (83%) published in 76 journals were available for citation search in the Web of Science databases. Thirty-eight PRC publications were not available for citation analyses for the following reasons: three were published in 2004 in indexed journals but had not yet been cited; 7 were published in indexed journals but were not among the articles that the ISI selected for citation tracking; and 28 were published in 17 journals not indexed by the ISI. Of the latter, nine articles were published in the *Journal of Public Health Management and Practice*, three in *Ethnicity and Disease*, and two in *Nicotine and Tobacco Research*. The remaining 14 articles were published in 14 other journals (listing available on request). For the articles that lacked citation data but were published in indexed journals, journal characteristics were available and therefore included in analyses.

Bibliographic information for each article that cited the PRC article was similarly extracted. The total circulation of each publishing journal was obtained from *Ulrich's Periodicals Directory*.[6] From the ISI's *Journal Citation Report*,[7] information was gathered on each journal's ISI-computed impact factor and citing half-life. These metrics are calculated by the ISI to reflect the frequency with which the average article in each journal is cited and the article's "staying power," respectively. The impact factor is calculated by dividing the number of citations in the current year of items published in a journal in the previous two years, by the total number of substantive articles published in that journal in the same 2 years.[8] The citing half-life is calculated as the number of years before the current year that accounts for publication of 50% of the articles cited in the current year.[8] Henceforth in this paper, this measure is referred to as half-life. The number of PRCs that named a given journal as among the most influential in their fields of interest was recorded as PRC influence factor.

The data described above were imported into a Microsoft Access database that allowed for a variable number of citations for each PRC publication. Complete bibliographic information (author, article title, year published, journal, volume number, issue number, page numbers, and total number of times cited) was included for each PRC publication. For each publishing journal, data included total circulation, journal impact factor, journal half-life, and PRC influence factor. For citing journals, similar information was recorded.

Because the data were not normally distributed, descriptive statistics calculated using SAS Base Software, version 8.2 of the SAS System for Windows (SAS Institute, Cary NC, 2001) were computed based on medians for number of citations and characteristics of the publishing journals. Nonparametric Spearman correlation coefficients were calculated to assess correlation between a journal's PRC influence factor and the ISI's impact factor and between PRC influence factor and the ISI's half-life.

To categorize the types of published articles, each article was reviewed by one of this paper's authors (ALF) and classified as follows: research report (including intervention study, descriptive study based on survey data, analysis of existing data, qualitative data findings, methodologic study, or economic analysis); discussion article (including opinion piece, position statement, editorial, introduction, or methodologic discussion); dissemination article (i.e., a paper focused on dissemination of information to practitioners rather than researchers); or review article (including literature review or meta-analysis). The median number of subsequent citations for publications in each category was tabulated.

Results

The 189 PRC articles included in the ISI's database were published in 76 journals, and were subsequently cited 4628 times in 1013 journals indexed by ISI.

Journal Characteristics

From 1994 to 2004, the median circulation of the 76 publishing journals was 8000 (range 400 to 332,000), the median impact factor was 2.6 (range 0.015 to 31.74), and the median half-life was 7.2 years (range 3.5 to 11.0). The median PRC influence factor was 4.0 (range 0 to 22). None of these values showed systematic variation over time (Table 1).

Citations

The 189 publications were subsequently cited a median of five times each (range 0 to 1523). Thirty-three (approximately 17%) of the publications had not been cited at all; 30 (approximately 16%) had been cited more than 25 times each. Citation patterns suggested that peak citation frequency occurs several years after publication (data not shown). Therefore, to capture relatively mature citation histories, subsequent analysis was limited to articles published before 2001. The median number of citations of papers published prior to 2001 was 14 (range 0–1523). The publishing journals' circulation and impact factor values were most apparently related to the volume of subsequent citations (Table 2).

PRC Influence Factor

The PRCs named 99 journals as the ones most influential in their fields of interest. Of those, only 34 were named by more than one PRC. The *American Journal of Public Health* was named most frequently (by 22 centers), followed by *Journal of the American Medical Association* (by 13), *American Journal of Preventive Medicine* (by 12), *Health Education and Behavior* (by 11), *New England Journal of Medicine* (by 8), and *Journal of School Health* (by 7). (Additional information is available on request.)

Forty-six (approximately 20%) of the 227 articles were published in journals considered influential by at least half the PRCs, and 49 (22%) were published in journals not considered influential by any. Little relationship was seen between the PRC influence factor and the number of citations (Table 2).

Correlation Analysis

For the 189 articles with a known journal impact factor, the correlation of the impact factor with the PRC influence factor was small but statistically significant. For the 75 publishing journals for which both impact factor and half-life were known, the Spearman correlation coefficient for PRC influence factor with impact factor was 0.23 ($p = 0.049$). For the corresponding 917 citing journals, the Spearman correlation coefficient for PRC influence factor with impact factor was 0.09 ($p = 0.006$). No correlation was found between PRC influence factor and half-life (Spearman $r = -0.06$,

Table 1. Characteristics[a] of publishing journals for selected PRC articles,[b] by publication year

	Articles (*n*)	Median circulation (in thousands)	Minimum/maximum circulation	Median impact factor	Minimum/maximum impact factor	Median half-life	Minimum/maximum half-life	Median PRC influence factor	Minimum/maximum influence factor
1994	4	21.5	1.7/35	3.3	1.1/4.0	7.7	5.4/11.0	11.5	1/22
1995	1	332.0	332/332	16.8	16.8/16.8	6.5	6.5/6.5	13.0	13/13
1996	9	17.8	1.1/35	1.3	0.1/10.3	7.2	4.5/8.4	1.0	0/22
1997	14	5.0	1.0/332	2.6	0.4/16.8	5.8	4.5/9.2	2.0	0/13
1998	15	8.0	0.9/59.8	2.6	0.4/4.2	7.7	3.5/11.0	6.0	0/22
1999	22	3.9	0.4/59.8	1.4	0.2/3.4	7.4	4.0/11	3.0	0/12
2000	29	7.7	1.2/332	2.6	0.2/16.8	7.7	3.8/11	5.0	0/22
2001	18	15.6	1/332	3.3	0.2/31.7	7.7	5.8/11	6.0	0/22
2002	26	7.5	0.6/59.8	1.9	0.4/5.5	7.1	3.9/11	3.0	0/22
2003	18	3.2	1.2/332	2.6	0.3/16.8	6.5	4.2/11	6.0	0/13
Total	156	8.0	4/332	2.6	0.1/31.7	7.2	3.5/11	4.0	0/22

[a]Circulation, impact factor, half-life, and PRC influence factor (see text for details).
[b]For 156 articles published in 1994 to 2003 in journals with all characteristics known.
PRC, prevention research center.

Table 2. Number of citations related to characteristics of journals that published PRC articles,[a] 1994–2000

Characteristic	Articles (*n*)	Citation number (median)	Citation number (minimum-maximum)
Circulation[b]			
<3000	21	6.0	(0–61)
3000–<7000	21	14.0	(0–89)
7000–<35,000	26	15.0	(1–145)
≥35,000	23	16.0	(4–1523)
ISI impact factor[c]			
<1.0	29	8.0	(0–70)
1.0–<2.6	29	15.0	(3–145)
2.6–<3.6	34	15.0	(0–114)
≥3.6	11	18.0	(6–1523)
ISI half-life[d]			
<5.6	37	10.0	(0–89)
5.6–<7.0	18	15.0	(4–1523)
7.0–<8.0	31	15.0	(0–114)
≥8.0	17	17.0	(1–145)
PRC influence factor			
0	20	13.5	(2–141)
1–4	35	10.0	(0–145)
5–13	32	11.5	(0–89)
≥13	18	16.5	(6–1523)

[a]For 105 articles published in 1994 to 2000 with known number of citations.
[b]Excludes 14 articles published in journals with unknown circulation.
[c]Excludes 2 articles published in journals with unknown impact factor.
[d]Excludes 2 articles published in journals with unknown half-life.
ISI, Institute of Scientific Information; PRC, prevention research center.

$p=0.59$ for publishing journals; Spearman $r=0.037$, $p=0.26$ for citing journals).

Repeating the correlation analysis restricted to the 34 journals considered influential by more than one PRC showed no significant correlation between PRC influence factor and the ISI impact factor (Spearman $r=0.19$, $p=0.36$) or half-life (Spearman $r=0.21$, $p=0.33$).

Types of Publications

Of the 227 articles in the database, 164 (73%) were research reports, 22 (10%) were discussion articles contributing to dialogue about public health issues, 20 (9%) were dissemination pieces, 17 (7%) were review articles, and 3 articles could not be classified by the established categories.

The number of subsequent citations varied by publication category; the highest numbers of citations of PRC articles were associated with research reports (median=6) and review articles (median=7). Dissemination pieces and discussion articles were cited less frequently (median=3 for each).

Further characterization of the PRC research papers showed that 32% reported on descriptive study findings from primary data collection (mostly survey data), 26% on original intervention studies or longitudinal descriptive studies, 18% on analyses of existing descriptive or experimental data collected by others, 12% on methodologic or statistical research, 11% on qualitative studies (primarily focus group research), and 2% on economic analyses.

Discussion

Nearly three quarters of the PRC articles selected as the most important peer-reviewed publications in the past decade were research reports, a finding that confirms the program's focus on research. The breadth of the research and the diversity of fields of expertise encompassed by the PRC program were demonstrated by the diversity of journals the PRCs put forward as the most influential. In fact, only one third of the named journals were considered influential by more than one PRC. Despite that fact, one fifth of the papers under examination were published in journals considered influential by at least half of the PRCs, and the majority of papers were published in journals considered influential by more than one PRC. Thus, the assessed publications appear to be reaching the audiences that the PRCs consider to be important consumers of their information.

Publications appeared in journals that varied widely in their circulation, impact factor, half-life, and PRC influence factor. These characteristics did not systematically change across the decade. The correlation between the ISI's impact factor and the PRC influence factor was minimal, suggesting that the ISI's metric and the PRCs' judgment are not closely aligned, although this is not the first time a lack of correspondence between a group of scientists' judgments and the ISI's impact factor has been noted.[9] The ISI journal impact factor measure has many critics who caution against its use in evaluating the impact of any researcher, research group, institution, or country. For example, it has been pointed out that the magnitude of citation counts to a given journal does not apply equally to individual papers published in it, and that the method of calculating the impact factor skews the measure in favor of journals with short publication lag time and those containing active discussion forums for timely issues.[10]

That the idiosyncrasies of the impact factor make it subject to misinterpretation and misuse has been duly acknowledged by its creator.[11,12] One of the limitations of the impact factor is that it should not be used to compare publications in different disciplines since citation frequencies vary greatly across disciplines. This limitation raises an interesting dilemma for efforts such as ours to characterize published papers in a field as diverse as public health prevention research with its disparate fields of inquiry (e.g., tobacco control, physical activity, sexual behavior) and types of journals (e.g., general medicine, public health, pediatrics, geriatrics).

If one uses the ISI journal category of "Public, Environmental, and Occupational Health" for comparison, the journal characteristics associated with the articles assessed in this paper appear quite favorable. The journals that published the assessed PRC papers had both a greater median impact factor (median 2.6) than the 89 journals in the "Public, Environmental, and Occupational Health" category in the Science Edition (median 1.3) and the 56 journals in the Social Science Edition (median 1.0); and they had a longer median half-life (7.2 years compared with 6.3 and 6.2, respectively).[13] However, reviewing the journals included in this ISI category shows a decided lack of uniformity in the disciplines included. Thus, there is good reason to agree with others who have invoked the need for careful delineation of journals important to public health improvement.[14]

Another potential approach is to compare with characteristics of journals publishing papers used in the development of evidence-based guidelines. One analysis[15] of citations used in the promulgation of guidelines for clinical preventive services found that the most highly cited journals (56 journals that had each been cited more than five times in the body of evidence) had a median impact factor of 2.8, quite similar to our reported median of 2.6 for journals in which the PRC papers were published. These authors asserted that journals focusing on preventive services tend to have lower impact factors than journals in other disciplines. They proposed that assessment of researcher and journal influence in the field of preventive medicine might more appropriately focus on the frequency of citation by evidence-based practice guidelines than on ISI rankings.[15]

A surprising finding of the current study was that 17 of the journals publishing the assessed PRC papers were not among the ISI's source journals at the time of this study. In fact, seven of these nonindexed journals were deemed highly influential by at least one PRC and accounted for 18 of the articles under study. The absence of data for these papers and the fact that seven additional PRC articles in indexed journals were not selected by the ISI for citation tracking limits the utility of our study and raises questions about the ISI's selection criteria. State public health or medical journals constituted one category relevant to the PRCs but not indexed by the ISI. In addition, new journals may not have a track record of reliable publication considered sufficient by the ISI editors. Furthermore, topical areas not traditionally considered within the realm of public health, science, or social science—such as social marketing—may not be well covered by the ISI databases.

Even if the impact factor were discarded as a pertinent measure because of its inherent biases, information on counts of citations of published papers would remain of interest. Even then, because it is the ISI's source journals that provide the basis for citation tracking, the failure to include all journals and all articles relevant to the PRC program presents a substantial unforeseen problem. Furthermore, the absence of relevant benchmarks for citation frequency in the field of prevention research is an important obstacle to using even a more accurate citation measure to assess scientific impact.

Nonetheless, any publicly funded research enterprise must be accountable for its productivity; publishing and being cited in the professional literature is a well-accepted, even mandatory, aspect of demonstrating accomplishment. The limitations imposed by the absence of some important journals and papers from the ISI databases, and the lack of relevant benchmarks, are therefore of great concern.

This exploration of a bibliometric approach to assessing the scientific impact of publications generated by the PRC program demonstrates its feasibility; it could be expanded to include all PRC publications for a more robust characterization. However, our analysis makes apparent the shortcomings of such an approach. First, any such effort is incomplete if all important journals are not included in the citation tracking databases. Second, no benchmark exists for bibliometric measures in the field of public health, and benchmarks from other fields should not be used, because citation patterns vary extensively among disciplines. Addressing these shortcomings will require knowledge about public health practitioners' preferences among professional journals, and which they consider most influential in guiding their practice. Once important journals are identified, these could be combined with other journals important to public health researchers and decision makers and grouped into a meaningful public health category whose citations could be tracked as a group. Eventually, benchmarks could be established for papers in this collection of journals and used for comparison as a measure of the scientific impact of individual papers, researchers, or research groups.

The most challenging shortcoming of any bibliometric approach to assessing a prevention research program is that it is inherently insufficient for measuring the actual impact on public health practice or population health. Having an influence on public health practitioners' adoption of evidence-based strategies to improve public health is arguably the most important impact the PRCs, or any prevention research program, can have. Certainly any measure based on citation counts can fail to capture changes in public health practice that may have a substantial impact on populations. To measure public health impact will require an understanding of the intermediary steps from findings or recommendations to influencing public health practice (i.e., through dissemination research) and the development of methods to measure these intermediate steps. In addition, measures of the population impact of public health programs need to be devel-

What This Study Adds . . .

Citations of published papers and characteristics of journals that they were published in are frequently used in academia to measure scientific impact.

This article reports an application of such bibliometric methods to assess publications from prevention research centers.

Although feasible, this approach was found to be of limited utility because there were no suitable benchmarks for comparison, influential journals were not all included in the relevant databases, and public health impact could not be addressed.

oped. A multifaceted approach to assessing impact has been suggested,[16] including review by an expert panel with post-publication peer review as a component. An approach that should be considered to assess the productivity and value of prevention research programs would involve the constitution of objective panels of public health researchers, practitioners, and decision makers. Such panels would likely consider publication, subsequent citation, or other bibliometric measures reported in this paper, as important, although they would likely include additional measures appropriate to specific fields of inquiry, such as use in development of evidence-based guidelines, influence on policy decisions, and the number of people affected by a program supported by research findings. Efforts to develop measures of public health impact, and efforts to improve bibliometric measures of scientific impact, are both important to the continuing evolution of prevention research and public health practice.

We wish to acknowledge the painstaking efforts of Reba Norman, Gail Cruse, and Brenda Mazzocchi of the Technical Information and Editorial Services Branch, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention, in gathering the data for this project. We are also indebted to Jeanne Casner for invaluable technical assistance and to Gabrielle Fowler for graciously filling multiple roles as needs arose.

No financial conflict of interest was reported by the authors of this paper.

References

1. Public Law 98-551. Health Promotion and Disease Prevention Amendments of 1984. (October 30, 1984).
2. Institute of Medicine. Linking research and public health practice: a review of CDC's program of centers for research and demonstration of health promotion and disease prevention. Washington DC: National Academy Press, 1997.
3. Anderson LA, Brownson R, McLinden D, Gwaltney MK. Developing an evaluation framework: insuring national excellence for the CDC Prevention Research Centers (Project DEFINE). Paper presented at American Public Health Association, 130th annual meeting and exposition, Philadelphia PA, 2002.
4. COSMOS Corporation. A national evaluation plan for CDC's prevention research centers program: final report. Bethesda MD: COSMOS Corporation, 2003.
5. Thomson Institute for Scientific Information, Philadelphia PA. Available at: www.scientific.thomson.com/products/wos
6. Available at: www.ulrichsweb.com/ulrichsweb/. Accessed Spring 2004.
7. Thomson Scientific. Information about Journal Citation Reports. Available at: www.isinet.com/media/presentrep/tspdf/jcrqrc.pdf. Accessed Spring 2004.
8. Thomson Scientific. Glossary of Thomson scientific terminology. Available at: www.thomsonisi.com/glossary/. Accessed Spring 2004.
9. Foster WR. Letter to the editor. Lancet 1995;346:1301.
10. Seglen PO. Why the impact factor of journals should not be used for evaluating research. BMJ 1997;314:498–502.
11. Garfield E. Which medical journals have the greatest impact? Ann Intern Med 1986;105:313–20.
12. Garfield E. Journal impact factor: a brief review. CMAJ 1999;161:979–80.
13. Thomson Scientific. ISI Journal Citation reports (commercial database accessed August 10, 2004). Available at: www.isinet.com/media/presentrep/tspdf/jcrquc.pdf. Accessed August 10, 2004..
14. Porta M, Fernandez E, Murillo J, Alguacil J, Copete JL. The bibliographic "impact factor": the total number of citations and related bibliometric indicators—the need to focus on journals of public health and preventive medicine. Soz Praventivmed 2004;49:15–8.
15. Nakayama T, Fukuhara S, Tsutani K, Yamazaki S. Comparison between impact factors and citation in evidence-based practice guidelines. JAMA 2003;290:755–6.
16. Liu JLY. Research quality can be assessed by using combination of approaches. BMJ 2003;326:931.