



# HHS Public Access

Author manuscript

*Am J Prev Med.* Author manuscript; available in PMC 2017 February 01.

Published in final edited form as:

*Am J Prev Med.* 2015 January ; 48(1): 50–57. doi:10.1016/j.amepre.2014.08.021.

## A Bibliometric Analysis of U.S.-Based Research on the Behavioral Risk Factor Surveillance System

George M. Khalil, MPH and Carol A. Gotway Crawford, PhD

Northrop Grumman Information Systems (Khalil); and the Population Health Surveillance Branch (Crawford), Division of Population Health, CDC, Atlanta, Georgia

### Abstract

**Background**—Since Alan Pritchard defined *bibliometrics* as “the application of statistical methods to media of communication” in 1969, bibliometric analyses have become widespread. To date, however, bibliometrics has not been used to analyze publications related to the U.S. Behavioral Risk Factor Surveillance System (BRFSS).

**Purpose**—To determine the most frequently cited BRFSS-related topical areas, institutions, and journals.

**Methods**—A search of the Web of Knowledge database in 2013 identified U.S.-published studies related to BRFSS, from its start in 1984 through 2012. Search terms were *BRFSS*, *Behavioral Risk Factor Surveillance System*, or *Behavioral Risk Survey*. The resulting 1,387 articles were analyzed descriptively and produced data for VOSviewer, a computer program that plotted a relevance distance-based map and clustered keywords from text in titles and abstracts.

**Results**—Topics, journals, and publishing institutions ranged widely. Most research was clustered by content area, such as cancer screening, access to care, heart health, and quality of life. The *American Journal of Preventive Medicine* and *American Journal of Public Health* published the most BRFSS-related papers (95 and 70, respectively).

**Conclusions**—Bibliometrics can help identify the most frequently published BRFSS-related topics, publishing journals, and publishing institutions. BRFSS data are widely used, particularly by CDC and academic institutions such as the University of Washington and other universities hosting top-ranked schools of public health. Bibliometric analysis and mapping provides an innovative way of quantifying and visualizing the plethora of research conducted using BRFSS data and summarizing the contribution of this surveillance system to public health.

### Introduction

In 2014, CDC marks the 30th anniversary of the Behavioral Risk Factor Surveillance System (BRFSS), which now administers surveys in 50 U.S. states, the District of Columbia, the Commonwealth of Puerto Rico, Guam, American Samoa, Palau, and the Federated States of Micronesia.<sup>1</sup> The unique features of the BRFSS allow for estimation of risk behaviors, health conditions, and general health indicators for policymakers at state and substate

---

Address correspondence to: George M. Khalil, MPH, Northrop Grumman Corporation, 1600 Clifton Road NE, MS # 83, Atlanta GA 30333. uwm4@cdc.gov.

jurisdictions. For many of its users, the BRFSS is the sole source of this much-needed information. BRFSS collects information from more than 450,000 adult interviews each year, making it the largest continuously conducted telephone health survey system in the world.<sup>1</sup>

Since the establishment of BRFSS in 1984, many authors have used the survey's publicly available data to conduct their research. Understanding how the data are used can help ensure that the most relevant information is being collected to serve the widest range of users. Although web analytics confirm numerous BRFSS data downloads and conference attendance records indicate strong interest in BRFSS data sets, tracking actual use of BRFSS data in publications can be a challenge because of the public-use nature of the data.

Bibliometrics provides a novel, easy, and interesting way to summarize the large amount of information contained in BRFSS publications.

Bibliometrics pioneer Alan Pritchard<sup>2</sup> defined this summary technique as "the application of mathematics and statistical methods that sheds light on the processes of written communication," which is commonly used to analyze trends in publications. More recently, there have been a number of studies using bibliometrics to assess the contribution an article, topic, or researcher makes in a particular field of study. For example, researchers have used bibliometrics to track the focus of tobacco research,<sup>3</sup> track the global trends of solid waste research,<sup>4</sup> map research collaboration in the business and management field,<sup>5</sup> assess global trends in laparoscopy research,<sup>6</sup> and assess trends in remote sensing research.<sup>7</sup> This study follows the lead of previous research to gauge the contributions of the BRFSS.

A bibliometric analysis of BRFSS, which had never been done on a large health survey before this article, provides insight about BRFSS data users, the research they are conducting, and where and when these studies are being performed. The objective of this bibliometric analysis is to use visualization of similarities (VOS) mapping and other bibliometric techniques to determine the most frequently cited BRFSS-related topical areas, institutions, and journals.

## Methods

For this analysis, a search of the Web of Science,<sup>8</sup> a database made available through the Thomson Reuters Web of Knowledge, was conducted. The Web of Science is an online scientific indexing service that is the world's leading source of scholarly research data on citation activity. The Web of Science was used to search and analyze the work of authors citing BRFSS data as well as determine how the data were being used. BRFSS-related publications from the year of the survey's inception in 1984 through 2012, the year during which complete data were available in 2013 at the time of this analysis, were identified. The search was limited to topic key terms *BRFSS*, *Behavioral Risk Factor Surveillance System*, or *Behavioral Risk Survey*. The search was also limited to articles, meeting abstracts, and proceedings papers published in the U.S. in the English language during the 1984–2012 selected time period. During the literature search, it was determined that the first publication using BRFSS was published in 1990. Thus, this analysis focuses on the years 1990–2012.

The Web of Science service also provides a way to download expanded citations for use in other software for producing citation trend graphics. The downloaded data were used to determine the journals that published BRFSS-related research most frequently. The analysis was divided into two groups: the last decade (2003–2012) and prior (1990–2002). Journal impact factors were searched via the 2012 Journal Citation Report (JCR) for journals publishing 20 or more BRFSS-related research papers during the study period.<sup>9</sup> The 2012 JCR calculates a journal's impact factor by dividing the number of citations of that journal in the 2012 JCR divided by the number of articles published in 2010 and 2011.<sup>10</sup> An impact factor of 1.0 means that articles published in 2010 and 2011 on average were cited one time in 2012.

The downloaded Web of Science reference data were used to plot the number of BRFSS-related publications by year from 1990, when the first BRFSS-related studies were published, to 2012, the year during which complete data were available at the time of this analysis. The number of citations of BRFSS-related publications was also plotted by year. For comparison, the Web of Science was used to search for publications related to the National Health Interview Survey (NHIS), a similar but national public health surveillance system, by searching topic key terms *NHIS* or *National Health Interview Survey*. As in the BRFSS search, this search was limited to articles, meeting abstracts, and proceedings papers published in the U.S. in the English language during the 1990–2012 time period. Information extracted from the Web of Science BRFSS search was plotted against the same information obtained from the NHIS search. Trend lines were chosen by conducting simple linear regressions.

The references were also imported into VOSviewer, version 1.5.4, to plot a relevance distance-based map. VOSviewer uses a unified approach to mapping and clustering. The software maps citation terms where the overall distance between terms provides information on their relatedness based on co-occurrences in documents. The x-axis and y-axis have no special meaning because the maps may be freely rotated and flipped. It also clusters citation terms and portrays the topic by color. The density of occurrence of information is portrayed by the size of the circle. Detailed information on this software is available elsewhere.<sup>11</sup> VOSviewer was used to develop a visual representation of the most commonly cited words in BRFSS-related research. A term map was created on the basis of a text corpus from title and abstract fields using ten as the minimum number of occurrences of a term. A thesaurus was created, which allowed the software to recognize different spellings of the same word, so variations would not be shown as separate items.

## Results

Through the Web of Science search, there was a steady increase in the number of articles published using BRFSS data since the surveillance system started in 1984. From the year the first manuscript using BRFSS was published (1990) until 2012, a total of 1,278 articles, 70 meeting abstracts, and 39 proceedings papers were found. Most of the references (84.2%) were published in the last 10 years. As of 2012, publications using BRFSS were cited about 34,000 times—equating to an average of about 1,500 citations per year. The number of citations for newer articles will naturally increase because the largest number of publications

using BRFSS was in 2012 ( $n=158$ , equating to about 12.1% of publications) (Figure 1). More than 300 journals have published articles using information from the BRFSS; the *American Journal of Preventive Medicine*, *American Journal of Public Health*, *Preventive Medicine*, and *Preventing Chronic Disease* published the most manuscripts (Table 1). Most of the publications were in the fields of public, environmental, and occupational health ( $n=768$ ); general and internal medicine ( $n=268$ ); healthcare policy and services ( $n=110$ ); healthcare sciences and services ( $n=94$ ); endocrinology and metabolism ( $n=61$ ); and oncology ( $n=60$ ) (Table 2).

The plot for the graph that displays the total number of publications per year (Figure 2A) shows that the BRFSS (solid black line) and NHIS (dotted gray line) have also had a steady increase in publications over time, and that around the year 2008, the total number of publications using the BRFSS identified in this analysis exceeded those using NHIS. The plot of citations comparing the BRFSS and NHIS shows that there has been a considerable rate of increase in the citations for both the BRFSS and NHIS over time (Figure 2B). Overall, the NHIS has more citations than BRFSS during the study period.

Bibliometric maps are used to determine the terms that are often used in the same publications. One such application of how VOSviewer has been used for a cluster analysis is an analysis of terms related to *Nature* and *Science* editorials.<sup>12</sup> According to the bibliometric map (Figure 3), most of the research during the 1990–2012 time period was conducted in content areas such as cancer screening (right corner in dark blue); state or regional studies (top right in green); cardiovascular health (bottom center in pink); obesity (near center in light blue); quality of life (left in red); and alcohol consumption (top left in yellow); but research topics ranged widely from behavioral sciences to determinants of health. The authors of articles using the BRFSS were mostly affiliated with CDC (502 publications) and academic centers such as the University of Washington (81 publications) and Harvard University (43 publications) (Table 3).

## Discussion

BRFSS data are being used increasingly to publish on emerging public health research needs, as the data provide public-use, state-based, and selected local data on a wide range of topics related to behavioral health outcomes and risk factors. However, given the public-use nature of BRFSS data, and the easy access from CDC's BRFSS website, it had been difficult to quantify the impact of this data to the field of public health. Through a search of the Web of Science,<sup>8</sup> a steady increase was found in the number of articles published using BRFSS data since the surveillance began in 1984. In fact, a majority of the research reports have been published during the last 10 years. The last year of data included in this study (2012) saw the greatest number of publications using the BRFSS. By looking at citations and publications, it is evident through the analysis of the literature that BRFSS data greatly contribute to the advancement of public health.

As the frequency by which BRFSS data are being used became evident, how BRFSS data are being used became of specific interest. The co-word cluster mapping using VOSviewer visualization of similarities techniques allowed this to be investigated by mining the title and

abstracts of research using BRFSS. Content areas publishing the most articles using BRFSS were gleaned. The findings suggest that BRFSS data are being used for a wide range of topics including regional analysis, cardiovascular health, quality of life, and alcohol consumption.

This analysis helped BRFSS determine which organizations published the most research using BRFSS data and which journals published them. The *American Journal of Preventive Medicine*, which had an impact factor of 3.94, published the most BRFSS-related research (Table 1). According to this analysis, CDC published the most research using BRFSS data (36.2%); however, academic institutions, particularly those hosting top-ranked schools of public health, also used BRFSS data to publish research (Table 3).

Every effort was made to be as accurate as possible in this project; however, bibliometric analyses do have some associated limitations. First, the Web of Science sometimes counts the same word as different if it has been spelled differently or if it appears with varying abbreviations from journal to journal. To account for this, names of institutions and keywords were manually reviewed, and the built-in VOSviewer thesaurus option was used to re-categorize words when this potential inaccuracy occurred. Finally, because bibliometrics assesses the quantity of scientific publications—not quality—an article's content was not considered for its strength, value, or individual contribution to the literature. This research did find, however, that most of the articles identified for this study were published in reputable peer-reviewed journals.

The utility of bibliometric analyses goes beyond information and library science. This study may encourage other agencies to evaluate how their data are being used. Bibliometrics of citations is useful for an entity such as CDC BRFSS in identifying, diagraming, and comparing the most frequently published BRFSS-related authors, publishing journals, topics, and content. Bibliometrics allows the analysis and visualization of the type of research that has been done and by which journals so that BRFSS stakeholders can direct their efforts if needed.

## Acknowledgments

Although the authors do not believe the results of this analysis were in any way influenced by conflicting interests, they would like to report that they are supported by CDC's Behavioral Risk Factor Surveillance System.

George M. Khalil is a Northrop Grumman contractor and is financially supported by CDC's Behavioral Risk Factor Surveillance System.

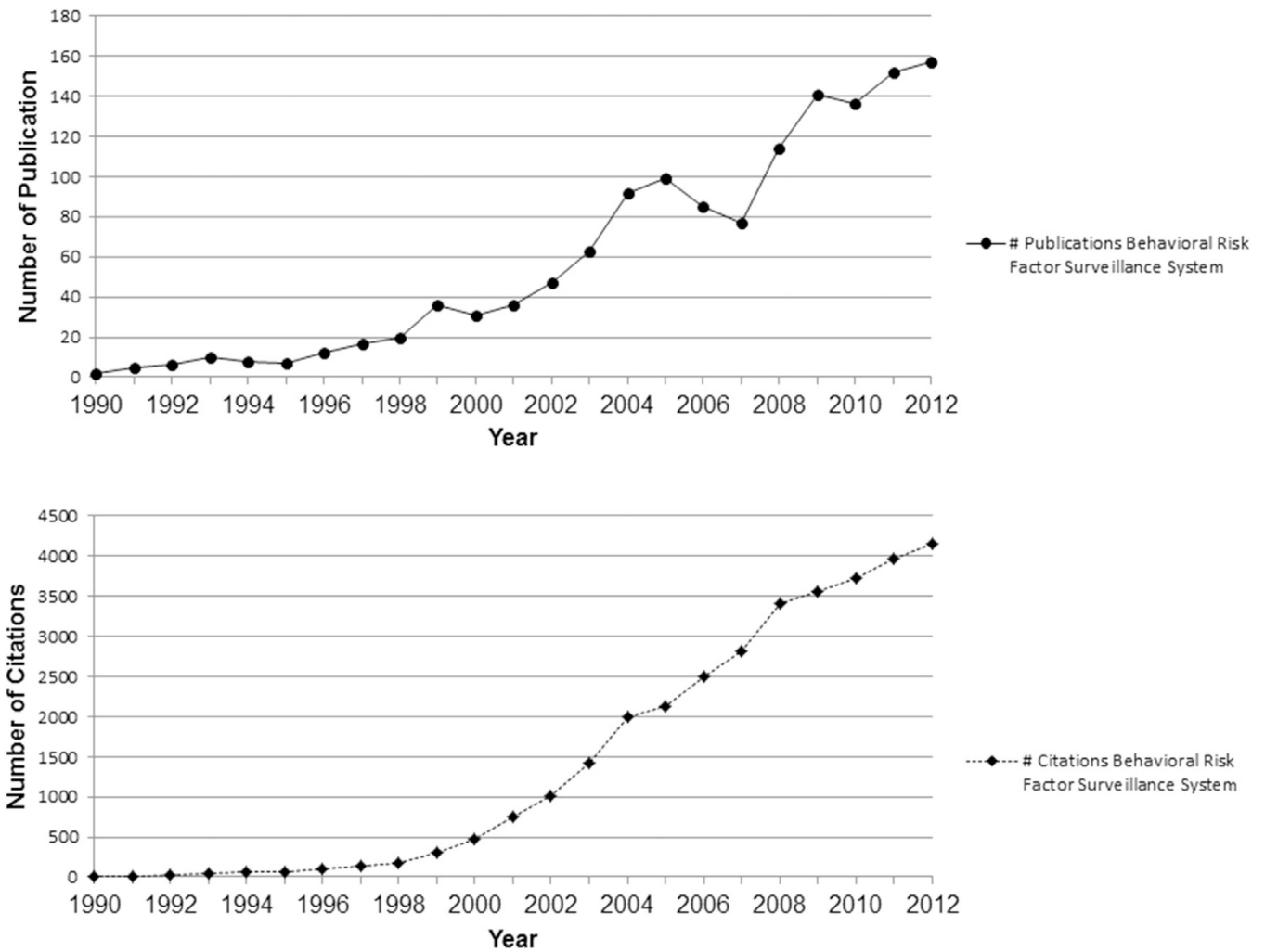
Carol A. Gotway Crawford is an employee of CDC's Behavioral Risk Factor Surveillance System.

The findings in this report are those of the authors and do not necessarily represent the official position of CDC.

## References

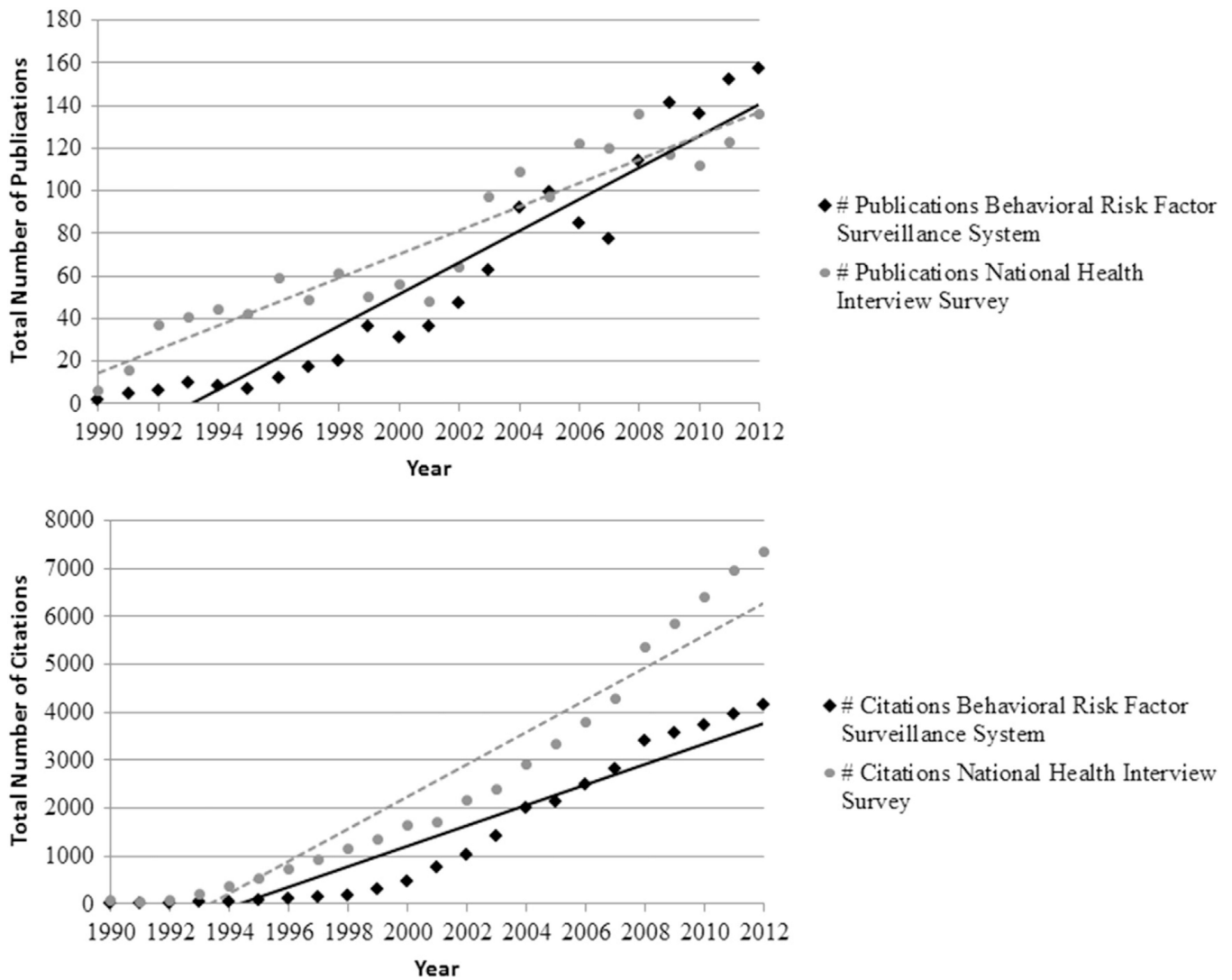
1. CDC. Behavioral Risk Factor Surveillance System. About the Behavioral Risk Factor Surveillance System (BRFSS). 2013. [www.cdc.gov/brfss/about/about\\_brfss.htm](http://www.cdc.gov/brfss/about/about_brfss.htm)
2. Pritchard A. Statistical bibliography or bibliometrics? *J Doc.* 1969; 25(4):348–349.
3. Cohen JE, Chaiton MO, Planinac LC. Taking stock a bibliometric analysis of the focus of tobacco research from the 1980s to the 2000s. *Am J Prev Med.* 2010; 39(4):352–356. [PubMed: 20837286]

4. Yang L, Chen Z, Liu T, Gong Z, Yu Y, Wang J. Global trends of solid waste research from 1997 to 2011 by using bibliometric analysis. *Scientometrics*. 2013; 96(1):133–146.
5. Kumar S, Jan JM. Mapping research collaborations in the business and management field in Malaysia, 1980–2010. *Scientometrics*. 2013; 97(3):491–517.
6. Cao Y, Zhou S, Wang G. A bibliometric analysis of global laparoscopy research trends during 1997–2011. *Scientometrics*. 2013; 96(3):717–730.
7. Zhuang Y, Liu X, Nguyen T, He Q, Hong S. Global remote sensing research trends during 1991–2010: a bibliometric analysis. *Scientometrics*. 2013; 96(1):203–219.
8. Thomson Reuters. Web of Science. 2013 Web of Knowledge. [apps.webofknowledge.com/UA\\_GeneralSearch\\_input.do?highlighted\\_tab=UA&product=UA&search\\_mode=GeneralSearch&SID=2DWTB9LxDaWISWr7AFe&last\\_prod=WOS&highlighted\\_tab=UA&cacheurl=no](https://apps.webofknowledge.com/UA_GeneralSearch_input.do?highlighted_tab=UA&product=UA&search_mode=GeneralSearch&SID=2DWTB9LxDaWISWr7AFe&last_prod=WOS&highlighted_tab=UA&cacheurl=no).
9. Thomson Reuters. Journal citation reports. ISI Web of Knowledge. 2012 [admin-apps.webofknowledge.com/JCR/JCR?=home](https://apps.webofknowledge.com/JCR/JCR?=home).
10. Thomson Reuters. Journal citation reports: impact factor. *J Impact Factor*. 2012 [admin-apps.webofknowledge.com/JCR/help/h\\_impfact.htm](https://apps.webofknowledge.com/JCR/help/h_impfact.htm).
11. Van Eck NJ, Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*. 2010; 84(2):523–538. [PubMed: 20585380]
12. Waaijer CJF, van Bochove CA, van Eck NJ. On the map: Nature and Science editorials. *Scientometrics*. 2011; 86(1):99–112. [PubMed: 21212822]



**Figure 1.**  
**(A)** Behavioral Risk Factor Surveillance System–related publications by year, 1990–2012.  
**(B)** Behavioral Risk Factor Surveillance System–related citations by year, 1990–2012.  
*Source:* Thomson Reuters. Web of Science. [apps.webofknowledge.com/](http://apps.webofknowledge.com/).  
*Note:* Figure 1A is a graph of the number of publications using BRFSS by year. The figure shows an increasing number of publications over time, with fewer than 20 in the early 1990s to nearly 160 in 2012. Figure 1B is a graph of the number of citations using BRFSS by year. The figure shows an increasing number of citations over time, with fewer than 100 citations in the early 1990s to more than 4,000 citations in 2012. BRFSS, Behavioral Risk Factor Surveillance System.





**Figure 2.** (A) Behavioral Risk Factor Surveillance System– and National Health Interview Survey–related publications by year, 1990–2012. (B) Behavioral Risk Factor Surveillance System– and National Health Interview Survey–related citations by year, 1990–2012.

*Source:* Thomson Reuters. Web of Science. [apps.webofknowledge.com/](http://apps.webofknowledge.com/).

*Note:* Figure 2A compares BRFSS and NHIS publications by plotting the number of publications by year from 1990 to 2012. It shows that there has been a considerable increase in the number of publications for both the BRFSS and the NHIS over time. The BRFSS went from an average of 18.23 publications per year between 1990 and 2002 to an average of 111.6 publications per year from 2003 to 2012. The NHIS went from an average of 44.1 publications per year between 1990 and 2002 to an average of 116.9 publications per year from 2003 to 2012. From 1990 to 2007, the NHIS had more publications than BRFSS, but starting in 2008, the number of publications using BRFSS exceeded those using NHIS. Figure 2B compares BRFSS and NHIS citations by plotting the number of citations by year from 1990 to 2012. It shows that there has been a considerable increase in the number of



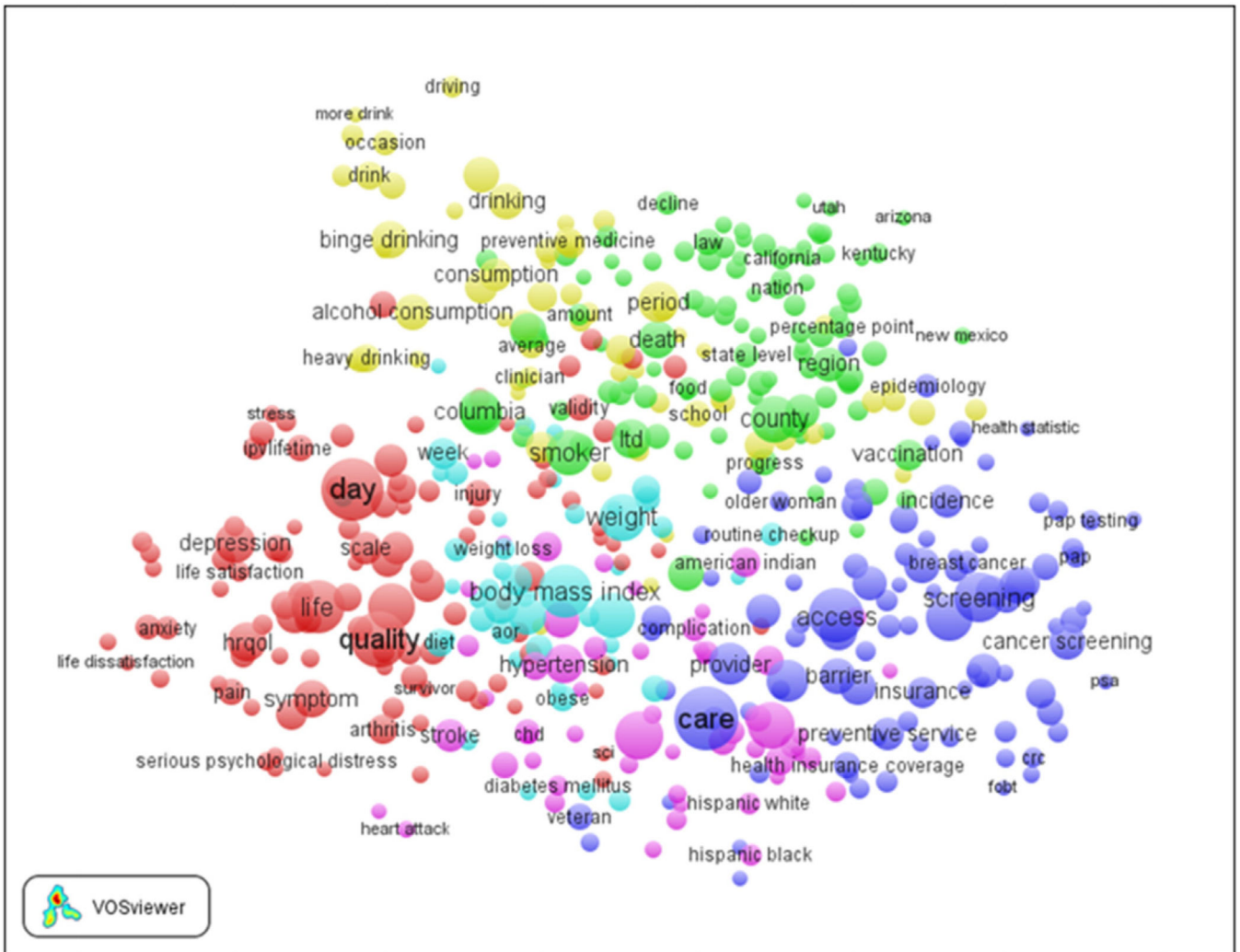
citations for both the BRFSS and the NHIS over time. The BRFSS went from an average of 245.3 citations per year between 1990 and 2002 to an average of 2,966.6 citations per year from 2003 to 2012. The NHIS went from an average of 846.8 citations per year between 1990 and 2002 to an average of 4,863.5 citations per year from 2003 to 2012. The NHIS has had more citations than the BRFSS during the study period. BRFSS, Behavioral Risk Factor Surveillance System; NHIS, National Health Interview Survey.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



**Figure 3.**

Mapping of keywords based on a co-occurrence network of Behavioral Risk Factor Surveillance System–related research, 1990–2012.

*Source:* Leiden University. VOSviewer. [www.vosviewer.com/](http://www.vosviewer.com/).

*Note:* VOSviewer uses a unified approach to mapping and clustering. The software maps information where the overall distance between terms, provides information on the relatedness of terms based on co-occurrences in documents. It clusters information and portrays the topic by color. At [bit.ly/N6c405](http://bit.ly/N6c405) this density map can be opened with VOSviewer and examined in detail.

Figure 3 shows that based on the colors of this map, most research was conducted in the following content areas: cancer screening (right corner in dark blue); state or regional studies (top right in green); cardiovascular health (bottom center in pink); obesity (near center in light blue); quality of life (left in red); and alcohol consumption (top left in yellow).

**Table 1**

Journals most frequently publishing Behavioral Risk Factor Surveillance System–related research and their impact factors, 1990–2002 and 2003–2012

Journal	1990–2002	2003–2012	Total <i>n</i> (%)	2012 impact factor
<i>American Journal of Preventive Medicine</i>	26	69	95 (6.8)	3.94
<i>American Journal of Public Health</i>	29	41	70 (5.0)	3.93
<i>Preventive Medicine</i>	9	45	54 (3.9)	3.50
<i>Preventing Chronic Disease</i>	0	53	53 (3.8)	1.59
<i>Public Health Reports</i>	12	26	38 (2.7)	1.42
<i>Ethnicity &amp; Disease</i>	0	28	28 (2.0)	1.12
<i>Journal of Women's Health</i>	2	25	27 (1.9)	1.42
<i>American Journal of Epidemiology</i>	4	22	26 (1.9)	4.78
<i>Medicine and Science in Sports and Exercise</i>	4	18	22 (1.6)	4.48
<i>American Journal of Health Promotion</i>	0	20	20 (1.4)	2.37

Sources: Thomson Reuters. Web of Science. [apps.webofknowledge.com/](https://apps.webofknowledge.com/) and Thomson Reuters. Journal Citation Reports. ISI Web of Knowledge. 2012. [admin-apps.webofknowledge.com/JCR/JCR?=&home](https://admin-apps.webofknowledge.com/JCR/JCR?=&home).

Note: This table lists the journals most frequently publishing BRFSS-related research. The *American Journal of Preventive Medicine* (impact factor=3.94); *The American Journal of Public Health* (impact factor=3.93); *Preventive Medicine* (impact factor=3.50); and *Preventing Chronic Disease* (impact factor=1.59) published the most BRFSS-related manuscripts.

BRFSS, Behavioral Risk Factor Surveillance System.

**Table 2**

Top 20 Web of Science categories for publications using Behavioral Risk Factor Surveillance System, 1990–2002 and 2003–2012

Web of Science categories	No. of publications 1990–2002	No. of publications 2003–2012	Total no. of publications
Public, environmental and occupational health	122	646	768
Medicine, general & internal	65	203	268
Health policy and services	15	95	110
Healthcare sciences and services	12	82	94
Endocrinology and metabolism	7	54	61
Oncology	10	50	60
Obstetrics and gynecology	9	39	48
Women's studies	4	42	46
Psychiatry	0	41	41
Nutrition and dietetics	5	35	40
Rehabilitation	5	29	34
Sport sciences	7	27	34
Clinical neurology	2	29	31
Substance abuse	4	27	31
Economics	2	22	24
Gerontology	3	21	24
Respiratory system	1	22	23
Social sciences interdisciplinary	2	18	20
Psychology, multidisciplinary	0	18	18
Rheumatology	2	16	18

Source: Thomson Reuters. Web of Science. [apps.webofknowledge.com/](https://apps.webofknowledge.com/).

Note: Categories are not exclusive; one publication may have multiple categories. This table lists the fields of public that BRFSS data were used for. Most publications were in the fields of environmental and occupational health ( $n=767$ ); general and internal medicine ( $n=268$ ); healthcare policy and services ( $n=110$ ); healthcare sciences and services ( $n=94$ ); endocrinology and metabolism ( $n=61$ ); and oncology ( $n=60$ ).

BRFSS, Behavioral Risk Factor Surveillance System.

**Table 3**

Top 20 organizations publishing Behavioral Risk Factor Surveillance System research, 1990–2002 and 2003–2012

Organizations	1990–2002	2003–2012	No. of publications	% of total publications
CDC	98	404	502	36.2
University of Washington	4	77	81	5.8
Harvard University	5	38	43	3.1
University of North Carolina	3	36	39	2.8
University of Illinois	4	29	33	2.4
Emory University	6	26	32	2.3
Columbia University	6	24	30	2.2
Brown University	3	26	29	2.1
University of Michigan	2	25	27	1.9
Boston University	2	24	26	1.9
University of South Carolina	2	24	26	1.9
West Virginia University	4	22	26	1.9
St. Louis University	14	11	25	1.8
University of Florida	1	23	24	1.7
Medical University of South Carolina	2	21	23	1.7
Massachusetts Department of Public Health	7	14	21	1.5
University of Rochester	0	19	19	1.4
Ohio State University	0	18	18	1.4
Substance Abuse and Mental Health Services Administration	1	17	18	1.4
University of Pennsylvania	1	17	18	1.4

Source: Thomson Reuters. Web of Science. [apps.who.int/iris/handle/10665/204866](https://apps.who.int/iris/handle/10665/204866).

Note: This table lists the institutions most frequently publishing BRFSS research. The top five organizations include CDC (36.2% of publications); University of Washington (5.8% of publications); Harvard University (3.1% of publications); University of North Carolina (2.8% of publications); and University of Illinois (2.4% of publications).

BRFSS, Behavioral Risk Factor Surveillance System.