

Original Paper

Emerging Trends and Thematic Evolution of Breast Cancer: Knowledge Mapping and Co-Word Analysis

Ali Sadatmoosavi^{1*}, PhD; Oranus Tajedini^{2*}, PhD; Omid Esmaeili^{3*}, MD; Firouzeh Abolhasani Zadeh^{4*}, MD; Mahdiyeh Khazaneha^{5*}, PhD

¹Department of Medical Library & Information Sciences, Faculty of Management and Medical Information Sciences, Kerman University of Medical Sciences, Kerman, Iran

²Department of Knowledge and Information Science, Shahid Bahonar University of Kerman, Kerman, Iran

³School of Medicine, Shahid Mohammadi Hospital, Hormozgan University of Medical Sciences, Hormozgan, Iran

⁴Department of General Surgery, School of Medicine, Afzalipour Hospital, Kerman University of Medical Sciences, Kerman, Iran

⁵Department of Information Sciences and Medical Informatics, Kerman University of Medical Sciences, Kerman, Iran

* all authors contributed equally

Corresponding Author:

Mahdiyeh Khazaneha, PhD

Department of Information Sciences and Medical Informatics

Kerman University of Medical Sciences

Medical University Campus

Haft-Bagh Highway

Kerman, 7616913555

Iran

Phone: 98 9138444409

Email: Khazaneha.m@gmail.com

Abstract

Background: One of the requirements for scientists and researchers to enter any field of science is to have a comprehensive and accurate understanding of that discipline.

Objective: This study aims to draw a science map, provide structural analysis, explore the evolution, and determine new trends in research articles published in the field of breast cancer.

Methods: This study comprised a descriptive survey with a scientometric approach. Data were collected from MEDLINE using a search strategy based on Medical Subject Heading (MeSH) terms. This study used science mapping, which provides a visual representation and a longitudinal evolution of possible interrelations between scientific areas, documents, or authors, thus reflecting the cognitive architecture of science mapping. For this scientometric evaluation of the topic of breast cancer research, a very long period was considered for data collection. Moreover, due to the availability of numerous publications in the database, the assessment was divided into three different periods ranging from 1988 to 2020.

Results: A total of 12,577 records related to scientometric studies were extracted. The field of breast cancer research demonstrated three diagrams containing the most relevant themes for the three chronological periods evaluated. Each diagram was plotted based on the centrality and density linked to each research topic. The research output in the field was observed to revolve around 8 areas or themes: *radiation injury*, *cardiovascular disease*, *fibroadenoma*, *antineoplastic agent*, *estrogen antagonistic*, *immunohistochemistry*, *soybean*, and *epitopes*, each represented with different colors.

Conclusions: In the strategic diagrams, the themes were both well developed and important for the structuring of a research field. The first quadrant comprised motor themes of the specialty, which present strong centrality and high density (eg, corticosteroid antineoplastic age, stem cell, T-lymphocyte, protein tyrosine kinase, dietary, and phosphatidyl inositol-3-kinase). In the second quadrant of diagram, themes have well-developed internal ties but unimportant external ties, as they are of only marginal importance for the field. These themes are very specialized and peripheral (eg, DNA-binding). In the third quadrant, themes are both weakly developed and marginal. The themes in this quadrant have low density and centrality and mainly represent either emerging or declining themes (eg, ovarian neoplasm). Themes in the fourth quadrant of the strategic diagram are considered important for a research field but are not fully developed. This quadrant contains transversal and general, basic themes (eg, immunohistochemistry).

Scientometric analysis of breast cancer research can be regarded as a roadmap for future research and policymaking for this important field.

(*JMIR Cancer* 2021;7(4):e26691) doi: [10.2196/26691](https://doi.org/10.2196/26691)

KEYWORDS

scientometrics; breast cancer; co-word analysis; Scimat; science mapping

Introduction

One of the requirements for scientists and researchers to enter any field of science is to have a comprehensive and accurate understanding of that discipline [1]. Accordingly, knowledge of the concepts, history, framework, scope, components, and functions of each discipline of science, as well as analyzing and examining how these are linked in the intertwined chain of human sciences and demonstrating these links with the fields on which they are more dependent, is of key importance [2]. In general, this knowledge should facilitate the best assay to gain a comprehensive picture of the fields of activity and applications of that discipline, which should be used as a guide by those who have not yet determined their future research passageway [3].

Science mapping is the analysis of publications within a scientific field from different viewpoints; it helps visualize a general assessment of a given field [4]. By using this map, the course of changes and developments in the field can be plotted to differentiate the fields with the most and the least proximity. Science mapping is undertaken to identify points of knowledge that follow “hot topics” and current trends in a given field [5]. A science map drawn based on the scientific research outputs of a field makes it possible to study the emergence of new fields and the cessation of some saturated scientific fields [6,7]. Simply put, the purpose of a science map is to depict the results of the analysis of publications of a scientific field from different angles and to provide an overview of that field [8]. Science maps attempt to showcase the processes of growth, integration, and disintegration of different fields of science over time. Scientific domains in these maps are determined in proportion to the level of activity of scientists, and the empty spaces in the illustrative map indicate unworked or unknown domains of science. This illustration thus showcases the growth, integration, or disintegration of different scientific fields over time [9,10]. In recent times, scientometrics—as a branch of information science and a bibliometric subfield—has been used in a plethora of studies to quantitatively examine emerging research patterns in the literature [11].

One of the most widely used methods for analyzing the structure of knowledge in various fields and drawing science maps is *co-word analysis* that examines the co-occurrence of keywords in the title, abstract, or text of articles. Therefore, co-word analysis is done on a set of published articles in a specific subject area [12]. By analyzing keywords used in articles of a specific research field, we can better understand the content of the common topics in that field [13,14]. Moreover, by measuring the relative intensity of these co-occurrences, simplified representations of concept networks in a given field can be illustrated [15].

Co-word analysis can reveal the main topics of the field under study, semantic structures, and the evolution of those works over time. In a co-word analysis, it is assumed that the most frequent words have a greater impact in a field of study than words that appear less frequently. Moreover, co-word analysis allows us to reveal emerging trends and changes in paradigms to facilitate predicting the direction of future research [13]. Co-word analysis can be used as a powerful tool to enable the follow-up of structural changes and the development of the sociocognitive network. This method also helps us identify emerging topics in scientific fields and draw a clear path for future research [16]. Furthermore, these networks were mapped by running network analyses using cosine link reduction and pathfinder networking scaling techniques [17].

One of the most important topics in medical research is breast cancer. Breast cancer is the most frequent type of cancer among women, affecting approximately 2.1 million women each year. It also causes the highest number of cancer-related deaths among women [18]. A total of 268,600 new cases of invasive breast cancer were estimated to be diagnosed among women and approximately 2670 cases among men in the year 2019 [19]. In addition, an estimated 48,100 cases of ductal carcinoma in situ were estimated to be diagnosed among women. Approximately 41,760 women and 500 men were expected to die from breast cancer in 2019 [20,21].

As in other fields of science, new research studies continually emerge in the field of breast cancer, leading to advancements in the field. Many of these studies often have some similarities and overlaps. For a variety of reasons, the volume of research suddenly sees a surge in some subfields, and with such increments, thematic overlaps can occur. However, in other areas, little research may be done over months and years. Given the importance of research in the field of medicine, in general, and breast cancer, in particular, it is necessary to provide a broad picture of the status of research conducted in this field. In other words, the structure of knowledge in this field should be revealed using techniques such as co-word analysis to demonstrate how this field has developed over time, and more importantly, to better understand the emerging topics, issues, and themes that have developed in this field.

This study uses co-word analysis to examine articles published in the field of breast cancer and improve or continue the necessary context for correction, continuation, or promotion of the pattern of their scientific behavior by gaining an understanding of the interests and tendencies of researchers in the field over time. Accordingly, this study aims to draw a science map, provide structural analysis, explore the evolution, and find new trends in articles published in the field of breast cancer by addressing the following research questions:

1. What are the most important research areas in the field of breast cancer?
2. Under which of the 4 themes (ie, motor themes, specialized and peripheral themes, emerging or disappearing themes, and general and basic themes) are breast cancer thematic areas classified in the strategic diagram?
3. What are the most important issues in terms of frequency and intensity?
4. How have breast cancer thematic areas been developed across different periods?

Methods

Search Strategy

We used the MEDLINE database to retrieve and extract bibliographic information from breast cancer-related research articles. MEDLINE is the premier bibliographic database of the US National Library of Medicine (NLM) that contains more than 25 million references to journal articles in the field of life sciences with a concentration on biomedicine [22]. A distinctive feature of MEDLINE is that the records are indexed with NLM Medical Subject Headings (MeSH) [23,24]. With regard to subject areas, MEDLINE includes biomedicine and health care research, broadly defined to encompass those areas of life sciences, behavioral sciences, chemical sciences, and bioengineering that are needed by health professionals and others engaged in basic research and clinical care, public health, health policy development, or related educational activities. MEDLINE also covers life sciences fields vital to biomedical practitioners, researchers, and educators, including aspects of biology, environmental science, marine biology, plant and animal science, as well as biophysics and chemistry [24].

To further validate the retrieved results, the search strategy used in this study was limited to research papers published in core clinical journals. The period covered in this study included all the years covered by this database (from 1950 to March 24, 2020). In other words, this study evaluated a total sample of 12,577 research articles published across 70 years. The retrieved records were saved as full records in plain text using tab-delimited and RIS (research information systems) formats. Finally, after saving the retrieved data, the related files were integrated and saved as a single file for subsequent use.

Data Acquisition and Processing

Overview

This study has been written on the basis of co-word analysis. Bibliometric methods explore the impact of a research field, a group of researchers, or a particular paper [25]. In this study, we used science mapping, which provides a visual representation and a longitudinal evolution of the interrelations between scientific areas, documents, or authors, reflecting the cognitive architecture of science mapping [26].

We used SciMAT [22,27], which is a powerful open-source science mapping [28] software. The tool allowed us to analyze

the evolution and relevance of the literature focused on breast cancer. This tool was designed according to the science mapping analysis approach, which allows researchers to analyze a research field; detect and visualize its conceptual subdomains (particular topics or themes or general thematic areas); and develop a longitudinal framework to analyze and track the conceptual, intellectual, or social evolutions of e-government through the course of consecutive periods [29]. Different bibliometric tools are available to perform this kind of study [29], but SciMAT has some characteristics that distinguish it from other science mapping analysis tools.

SciMAT divides the analysis into four phases. A detailed explanation of these phases can be found elsewhere [7,30], although a brief description is provided below.

Detection of Research Themes

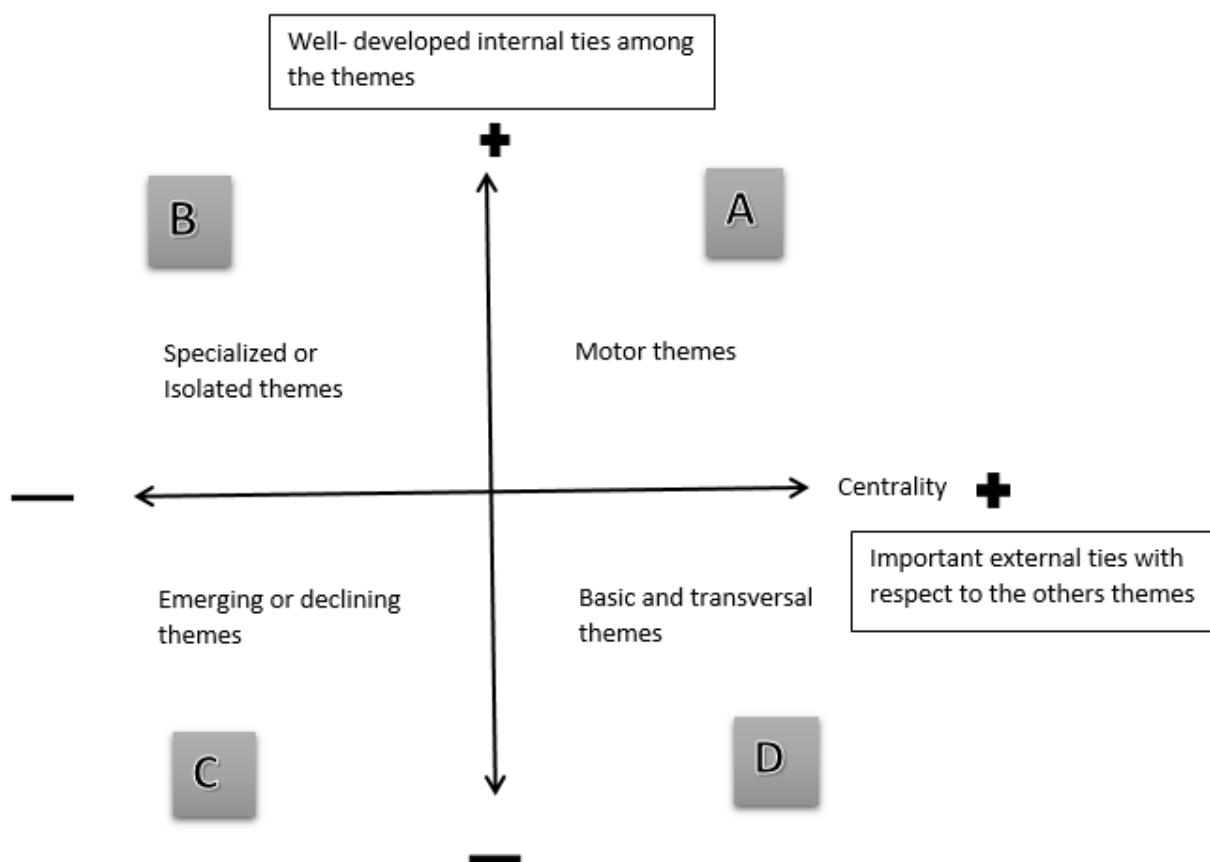
The first phase involves detection of research themes. This phase summarizes the first five steps of the workflow of science mapping analysis. In each period studied, the corresponding research themes are detected by applying a co-word analysis [31] to the raw data of from the published documents in the research field, followed by clustering of keywords to topics or themes by using the simple centers algorithm [32]. Formally, the methodological foundation of co-word analysis is based on the idea that the co-occurrence of keywords describes the content of the documents in a corpus [33]. These co-occurrences of keywords can be used to build co-word networks [34], and these networks can be associated with research themes using clustering tools. The co-occurrence frequency of two keywords is extracted from the corpus by counting the number of documents in which the two keywords appear together. Once the co-word network is built, each arc or edge will have in its weight the co-occurrence value of the linked terms. Next, the weight of each edge is transformed to normalize it (extracting the similarity relations between terms) using their keyword and co-occurrence frequencies [35].

Data Visualization

In this phase, the detected research themes are visualized using two different visualization instruments: a strategic diagram [22,36-38] and a thematic network. Each theme can be characterized by two measures [12,13]—centrality and density.

Centrality measures the degree of interaction of a network with other networks and shows the strength of external ties to other themes [39]. This value can be considered as the measure of the importance of a theme in the development of the entire research field analyzed. Density measures the internal strength of the network and reflects the strength of internal ties among all the keywords that describe the research theme. This value can be considered as a measure of the theme's development [40]. Once the centrality and density rankings are calculated, the themes can be laid out in a strategic diagram. Given both measurements, a research field can be visualized as a set of research themes, mapped in a 2D strategic diagram (Figure 1), and classified into the following four groups:

Figure 1. Strategic diagram in 2D and the 4-group classification: (A) motor themes, (B) specialized or isolated themes, (C) emerging or declining themes, (D) basic and transversal themes.



1. Motor themes that are both well developed and important for the structuring of a research field; these themes present strong centrality and high density.
2. Specialized and peripheral themes that have well-developed internal ties but unimportant external ties, as they are of only marginal importance for the field.
3. Emerging or declining themes that are both weakly developed and marginal; the themes in this quadrant have low density and low centrality.
4. Basic and transversal themes

that are important for a research field but are not developed (eg, notes in computer science) [26].

Discovery of Thematic Areas

The next phase involved temporal or longitudinal analysis. In this phase, the evolution of the research themes over a set of periods is first detected and then analyzed to identify the main general areas of evolution in the research field, their origins, and their interrelationships. This allows the discovery of the conceptual, social, or intellectual evolution of the field. SciMAT can build an evolution map [29] and an overlapping items graph (Figure 2) [41] to detect the evolution areas (see Figure 3).

Figure 2. The overlapping graph. The horizontal arrow represents the number of items shared by both periods. The upper incoming arrow represents the number of new items in the second period, and the upper outgoing arrow represents the items that are presented in the first period, but not in the second period [28,41].

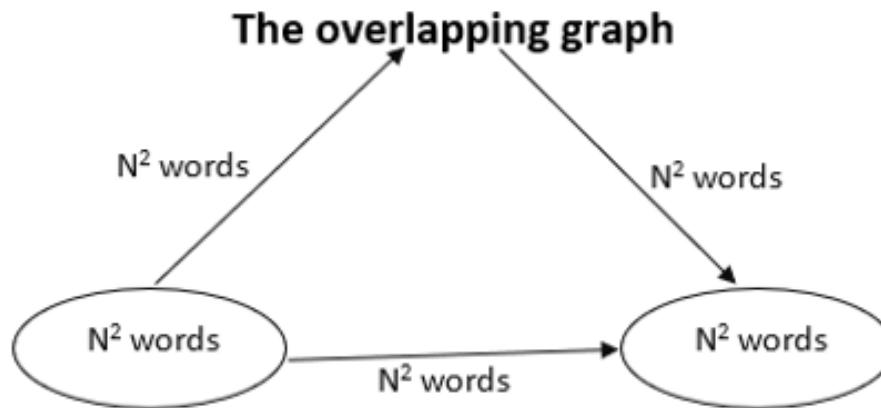
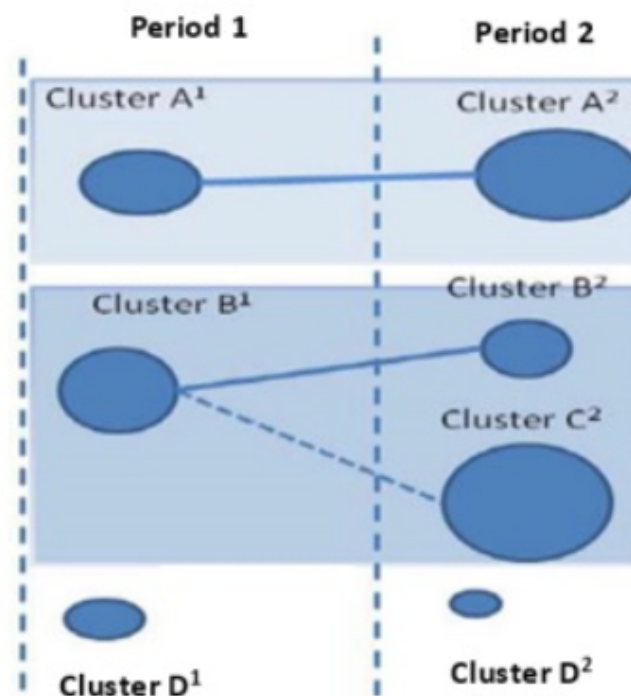


Figure 3. The evolution map: cluster D1 is discontinued, and cluster D2 is considered to be a new cluster.



For this purpose, an inclusion index is used to detect conceptual nexuses between research themes in different periods and thus identify the thematic areas in a research field. In addition, as each theme is associated with a set of documents, each thematic area can also have an associated collection of documents, obtained by combining the documents associated with its set of themes. Thus, the evolution map shows the temporal evolution of research themes of e-government, and the overlapping graph represents the number of associated keywords (Figure 2) [42].

Performance Analysis

In this phase, the relative contribution of research themes and thematic areas to the whole research field is measured (quantitatively and qualitatively) and used to establish the most prominent, most productive, and highest impact subfields. This performance analysis is developed as a complement to the analysis step of the science mapping workflow. Some

bibliometric indicators in this phase used include the number of published documents, number of citations, and the different types of h index [43,44].

Eventually, three diagrams were represented based on the three temporal visualization phases. Following the science mapping workflow, visualization techniques were used to represent a science map and the results of different analyses. In this sense, the network results from the mapping step were represented in the form a strategic map, evolution map, and overlapping graph. Finally, when the science mapping analysis was completed, experts analyzed the results and maps, using their experience and knowledge.

Ethics Approval

Since this was a metadata analysis of published work, approval from an ethics committee was not required.

Results

After retrieving a total of 12,577 records related to scientometric research, the importance of keywords (Table 1) and various journals was demonstrated (Table 2).

Overview

Table 1. Most frequently used terms in the selected articles (N=12,577).

Sr. no.	Source titles	Records, n (%)
1	Breast neoplasms	11,855 (97.9)
2	Prognosis	1919 (15.8)
3	Lymphatic metastasis	1735 (14.3)
4	Mastectomy	1520 (12.6)
5	Mammography	1435 (11.8)
6	Neoplasm staging	1420 (11.7)
7	Neoplasm recurrence local	1208 (10)
8	Neoplasm metastasis	1154 (9.5)
9	Risk factors	1115 (9.2)
10	Estrogen receptors	1113 (9.2)
11	Time factors	1014 (8.4)
12	Age factors	955 (7.9)
13	Carcinoma	889 (7.3)
14	Carcinoma ductal breast	814 (6.7)
15	Combined modality therapy	781 (6.4)
16	Axilla	779 (6.4)
17	Carcinoma intraductal noninfiltrating	753 (6.2)
18	Antineoplastic combined chemotherapy protocols	720 (5.9)
19	Lymph nodes	686 (5.7)
20	Mass screening	677 (5.6)
21	Antineoplastic agents	674 (5.6)
22	Biomarker tumor	655 (5.4)
23	Immunohistochemistry	640 (5.3)
24	Chemotherapy adjuvant	594 (4.9)
25	Neoplasm invasiveness	591 (4.9)
26	Lymph node excision	590 (4.9)
27	Tamoxifen	581 (4.8)
28	Receptor progesterone	563 (4.6)
29	Mastectomy segmental	543 (4.5)
30	Menopause	526 (4.3)

Table 2. The top 30 journals with the highest number of articles published on breast cancer (N=12,577).

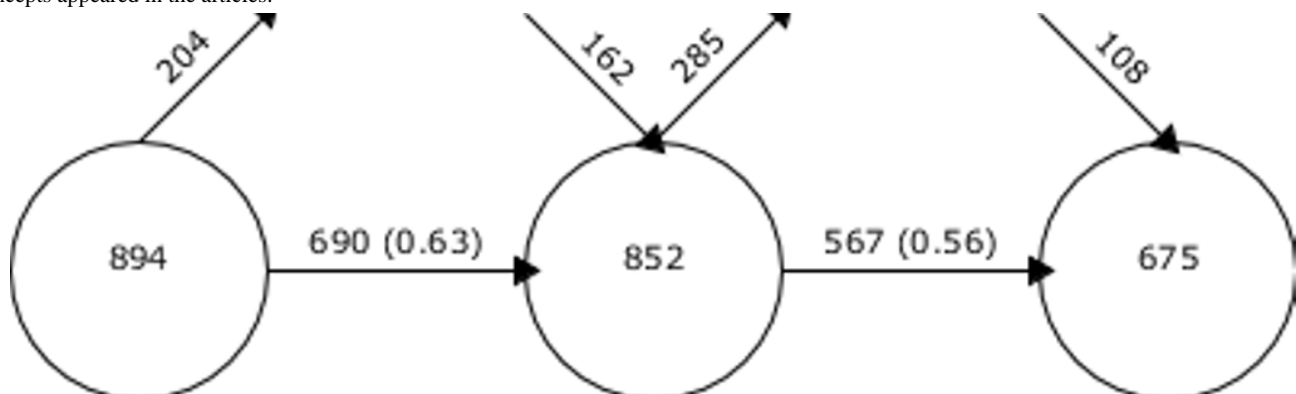
Row	Source title	Records, n (%)
1	Cancer	3321 (27.4)
2	Lancet London England	567 (4.7)
3	American Journal of Surgery	523 (4.3)
4	British Journal of Surgery	454 (3.7)
5	Radiology	453 (3.7)
6	Journal of Clinical Pathology	376 (3.2)
7	American Journal of Roentgenology	347 (2.9)
8	The New England Journal of Medicine	331 (2.7)
9	JAMA ^a	330 (2.7)
10	American Journal of Clinical Pathology	320 (2.6)
11	Annals of Surgery	295 (2.4)
12	The American Journal of Pathology	289 (2.4)
13	Medicine	256 (2.1)
14	Archives of Surgery (Chicago, Illinois) 1960	232 (1.9)
15	Archives of Pathology Laboratory Medicine	199 (1.6)
16	British Journal of Radiology	199 (1.6)
17	Endocrinology	194 (1.6)
18	Surgery	193 (1.6)
19	British Medical Journal (Clinical Research Ed.)	192 (1.6)
20	Surgery Gynecology Obstetrics	186 (1.5)
21	Journal of the American College of Surgeons	177 (1.5)
22	Journal of Clinical Endocrinology and Metabolism	153 (1.3)
23	Plastic and Reconstructive Surgery	152 (1.3)
24	British Medical Journal	145 (1.2)
25	Journal of Clinical Investigation	126 (1)
26	Southern Medical Journal	118 (1)
27	American Journal of Public Health	106 (0.9)
28	Annals of Internal Medicine	102 (0.8)
29	Surgical Clinics of North America	96 (0.8)
30	The American Journal of Clinical Nutrition	94 (0.8)

^aJAMA: Journal of the American Medical Association.

Result for the Thematic Period (1987-2020)

Figure 4 shows the number of concepts related to the thematic area of breast cancer in the three 11-year periods spanning from 1988 to March 31, 2020.

Figure 4. Thematic areas in the three evaluation periods based on centrality and density. The horizontal output arrows represent the number of concepts that served as the input for the next period; the vertical output arrows represent the number of concepts that exited a given period and were considered less important; and the vertical input arrow represents the number of concepts that received attention. In the second period, 852 new concepts appeared in the articles and 690 concepts were considered from the previous period. Of these 567 concepts entered the third evaluation period, and 675 new concepts appeared in the articles.



In the first period, the highest centrality was found for immunohistochemistry (IHC), and the highest density was related to the soybean theme. In the second period, the highest centrality was found in the antineoplastic themes, and the highest density was detected for the themes isoflavones and enzyme inhibitors. In the third period, the highest centrality was found for the antineoplastic agent theme, and the highest density was detected for the vegetable theme.

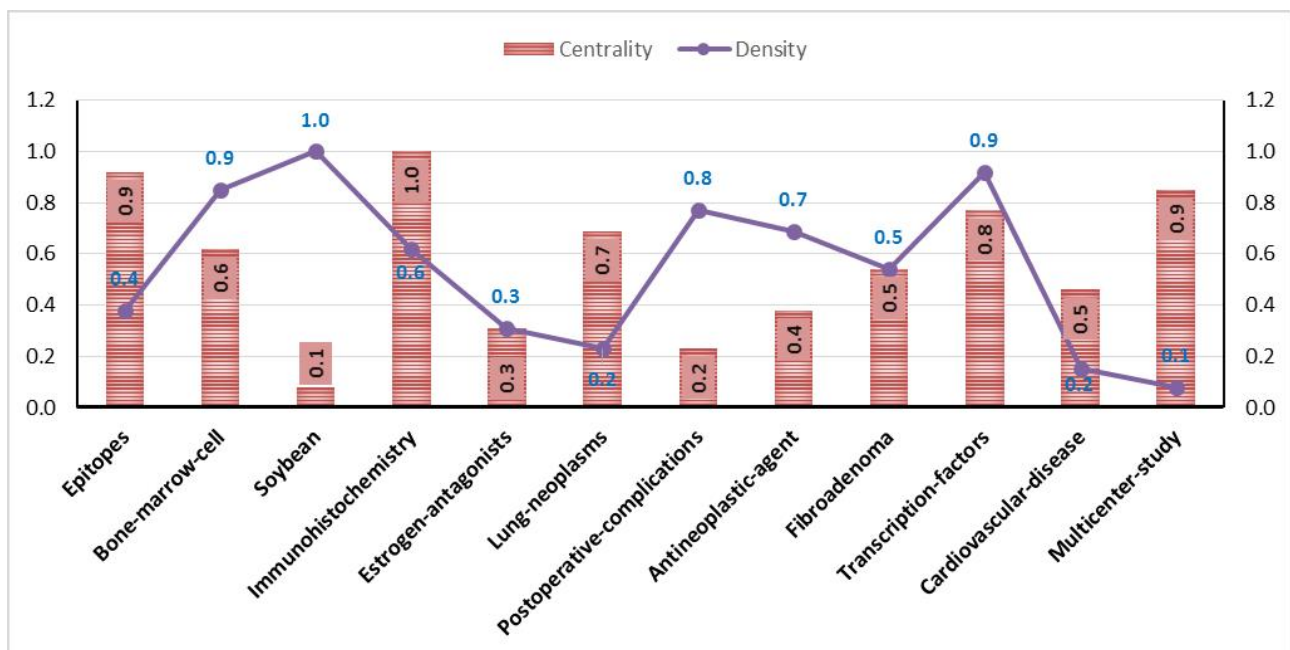
The strategic diagram of breast cancer was drawn based on the abundance of articles in the 4 thematic areas, including motor cluster, basic and transversal cluster, highly developed cluster, and emerging and declining cluster. The most important topics

were found in the motor cluster, which are displayed in 10-year periods.

First Evaluation Period (1989-1998)

In the first period, the upper-right quadrant (ie, motor cluster) comprised transcription factors, bone marrow cell, immunohistochemistry, and fibroadenoma, indicating the important role of these concepts in the field of breast cancer from 1989 to 1998 (see Figure 5). A transcription factor is a protein that controls the rate of transcription by binding to a specific DNA sequence. Immunohistochemistry is one of the best ways to detect these factors (Multimedia Appendix 1).

Figure 5. Breast cancer-related concepts identified in the first evaluation period based on density and centrality from 1988 to 1998. Date adopted from global statistics retrieved from the Web of Sciences.

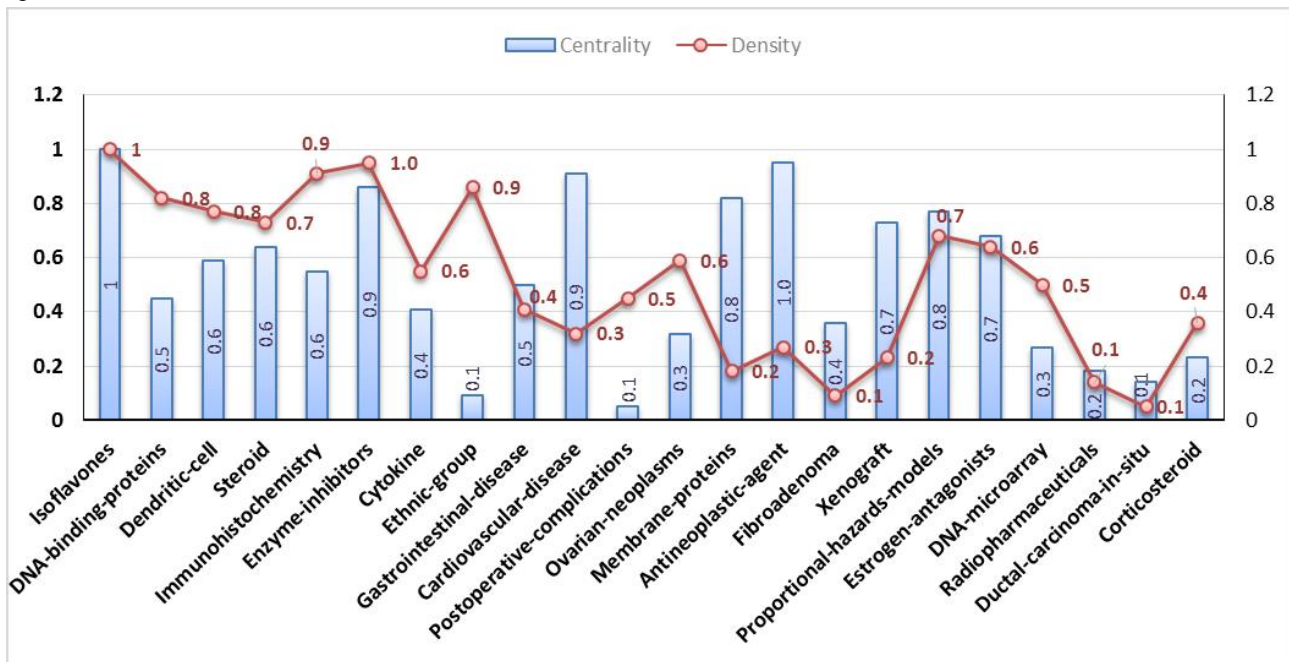


Second Evaluation Period (1999-2009)

The concepts related to the motor theme included isoflavones, enzyme inhibitors, immunohistochemistry, estrogen, proportional hazard model, and steroid. Soy isoflavones are

enzyme inhibitors similar to lipoxigenase. Moreover, there is a close relation between suppression of dendritic cell maturation and functions by isoflavones (phytoestrogen; see Figure 6). Therefore, soy isoflavones can bind to estrogen receptors and act as an estrogen antagonist (Multimedia Appendix 2).

Figure 6. Breast cancer–related concepts identified in the second evaluation period based on density and centrality from 1999 to 2009. Date adopted from global statistics retrieved from the Web of Sciences.

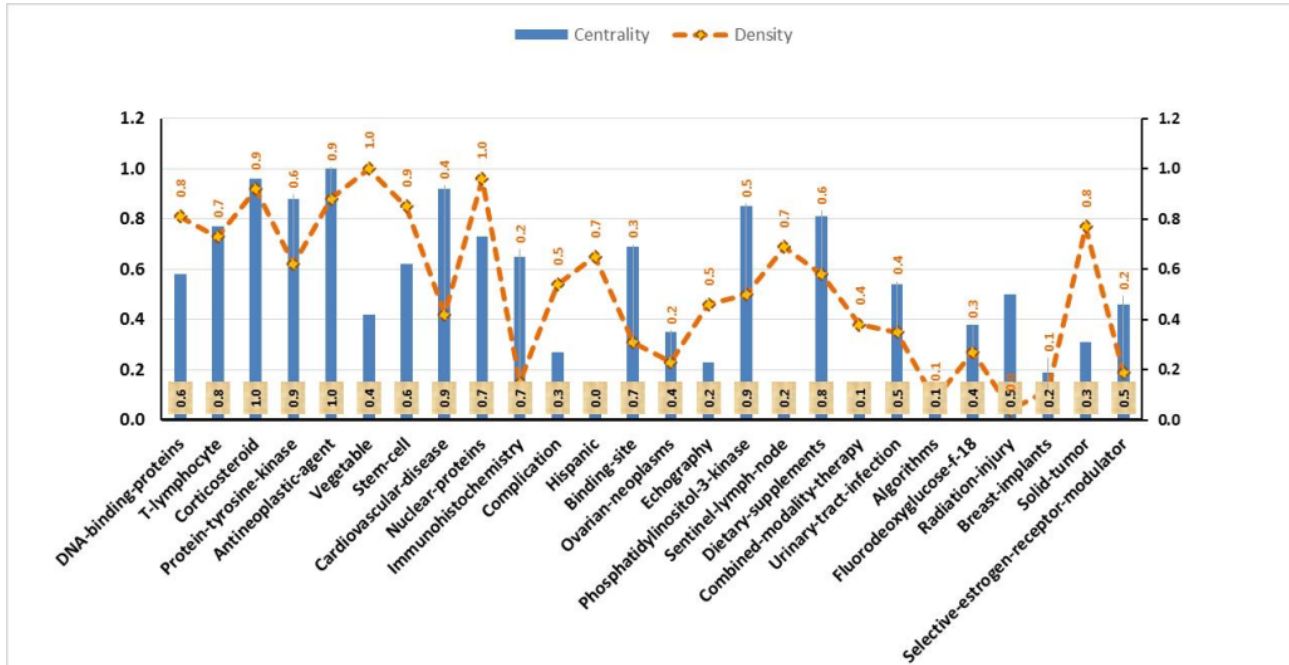


Third Evaluation Period (2010-2020)

The concepts of the motor theme in the third evaluation period included revealing corticosteroid antineoplastic age, stem cell,

T-lymphocyte, protein tyrosine kinase, dietary, and phosphatidylinositol-3-kinase, indicating the importance of these topics in this period (Figure 7).

Figure 7. Breast cancer–related concepts identified in the third evaluation period based on density and centrality from 2010 to 2020. Date adopted from global statistics retrieved from the Web of Sciences.



Steroids are important biodynamic agents and can be used as a particular agent for receptor-mediated diseases just like for breast cancer. Furthermore, infiltrative T-lymphocytes are related to invasive breast cancer.

Protein tyrosine phosphatases have a crucial role in the regulation of stem cell renewal and differentiation. Some studies have shown relations between DNA-binding protein oxidation

and dietary supplements that contain plant extracts and vitamins (Multimedia Appendix 3).

Discussion

Principal Findings

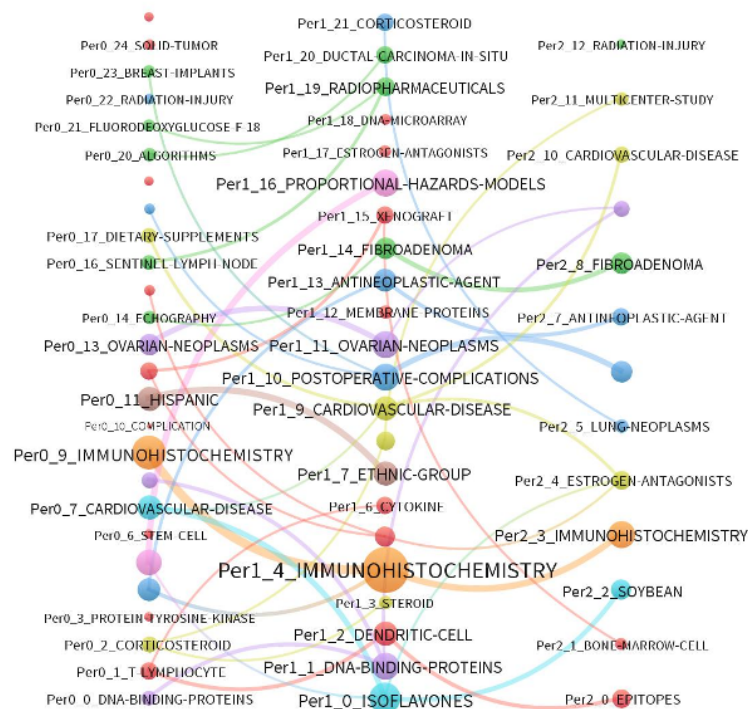
After retrieving a total of 12,577 records related to scientometric studies, we were able to demonstrate important keywords,

including *breast neoplasms, prognosis, lymphatic, metastasis, mastectomy, mammography, and neoplasm staging*. Indeed, examples of 30 journals with the highest number of articles published on breast cancer were *Cancer, Lancet (London, England), The American Journal of Surgery, British Journal of Surgery, Radiology, Journal of Clinical Pathology, American Journal of Roentgenology, The New England Journal of Medicine, JAMA (Journal of the American Medical Association), and The American Journal of Clinical Pathology*. Moreover, across the three evaluation periods, the themes with the highest density and centrality were observed in the first period, with the highest centrality was related to immunohistochemistry and the highest density related to the soybean theme. In the second evaluation period, the highest centrality was associated with the

antineoplastic themes, and the highest density was observed in the isoflavones and enzyme inhibitor themes. In the third evaluation period, the highest centrality was related to the antineoplastic agent theme, and the highest density was observed in the vegetable theme.

As in this scientometric evaluation of breast cancer topic, a very long period was considered for data collection. Moreover, due to the multiplicity of the publications, this assessment was divided into 3 decades from 1989 to 2020. The results of this study confirmed the progression of studies in recent decades and different concentrations of assessments completed in different years. Moreover, communications between these themes were also shown (Figure 8).

Figure 8. Thematic trends in the field of breast cancer from 1988 to 2020. Date adopted from global statistics retrieved from the Web of Sciences.



Emerging Themes from the Three Evaluation Periods

Using SciMAT [22] and Vos-viewer [45], the research output in the field was observed to revolve around 8 areas. As shown in Multimedia Appendix 3, the themes in the rightmost column included *radiation injury, cardiovascular disease, fibroadenoma, antineoplastic agent, estrogen antagonistic, immunohistochemistry, soybean, and epitopes*, as indicated with different colors. Thematic links are demonstrated by a solid line. The size of these nodes is proportionate to the number of documents under each theme. In addition, the color of the nodes indicates different areas.

As seen in Figure 4, the analyzed research output is categorized by solid cohesion. Most of the identified topics have been gathered via thematic nodes. They arise from a topic appearing in the previous period and show a continuous evolution with almost no jumps or gaps.

Regarding the starting period, thematic areas started in the first period (Figure 5, Figure 7, right panel). Thus, they can be considered as the primary subjects in breast cancer. Furthermore, in the second period, a new thematic area emerged: *ethnic group, proportional hazards, corticosteroids, postoperative, ovarian neoplasm, ethnic group, and cytokine*. Indeed, the emerged thematic areas play an essential role in the development of the field. Regarding the theme composition, the thematic areas of *immunohistochemistry* are mainly composed of motor themes across all three periods. Furthermore, in the third period, *ethnicity* evolved to *Hispanic*. In addition, topics such as stem cell, solid tumor, breast implant, echography, and protein tyrosine kinase emerged in this field with some of them evolving from the second period.

The relationship between IHC and cancer biology, which is now better known, has influenced axillary lymph node dissection (ALND). Tumor biological factors are different in each tumor

if tumor tend to metastasize to visceral or lymph nodes depends on tumor biological features. With advanced knowledge and understanding of tumor biology, systemic therapy and targeted therapy policies have changed. At present, the decision to initiate and prescribe chemotherapy (ie, systemic therapy) is influenced by the tumor stage and tumor biological factors, as well as the patient's lymphatic status. For example, in some cases, a tumor is diagnosed by screening mammography in the early stages, and there is no lymphatic involvement. Decisions to continue adjuvant treatments depends on the biological factors of the tumor. Biological factors play a key role on the decision to start neoadjuvant therapy. For example, triple-negative, and Her2neu-positive tumors have a dermatological response to neoadjuvant therapy.

In patients with a positive sentinel lymph node biopsy (SLNB), tumor biological factors such as ER/PR/Her2neu are prognostic factors, and it is of interest to know whether ALND changes the patient's cervix. According to the AMAROS trial [46], axillary radiotherapy was comparable to axillary dissection for local axillary control and even had fewer side effects. In patients with T1 and T2 masses, who had a positive SLNB and received axillary radiotherapy, overall survival and disease-free survival were similar to those who underwent axillary dissection. Therefore, SLNB is currently recommended for many patients with breast cancer. Currently, based on the IBCSG23-01 study, the National Comprehensive Cancer Network guidelines recommend only radiotherapy for patients with a positive SLNB (micrometastasis), without axillary dissection. Thus, SLNB has currently replaced ALND in many cases.

The false-negative rate is low in cases where a dual agent is used and at least more than two SLNs are found in patients with clinical lymph nodes (N1). Lymph node biopsy can be performed for patients undergoing neoadjuvant therapy. However, dual-agent therapy is preferably used when finding at least two lymph nodes in patients with preneoadjuvant clinical lymph node N1.

According to the AJCC (American Joint Committee on Cancer) staging, biomarkers such as ER/PR/Her2neu are recommended to be effective.

Pathological Analysis

Currently, the basis of breast cancer treatment is complete knowledge of its progression and biological factors (ER/PR/Her2neu). These factors affect the stage of the disease and also indicate the likelihood of tumor recurrence. They can also assist in response to selected treatments.

Conclusion

Eventually, scientometric analysis can showcase the current state of the science. Similarly, co-word analysis determines the frequency of words and thus indicates the most important research topics of a field. Using these methods, the characteristics and challenges of research fields and scientific disciplines can be determined. In addition, scientometric analysis of breast cancer research can be regarded as a roadmap for future research and policymaking in this important field of study.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Strategic diagram for the second evaluation period (1999 to 2009).

[\[PNG File , 102 KB-Multimedia Appendix 1\]](#)

Multimedia Appendix 2

Strategic diagram for the second evaluation period (1999 to 2009).

[\[PNG File , 180 KB-Multimedia Appendix 2\]](#)

Multimedia Appendix 3

Strategic diagram for the third evaluation period (2010 to 2020).

[\[PNG File , 166 KB-Multimedia Appendix 3\]](#)

References

1. Birkle C, Pendlebury D, Schnell J, Adams J. Web of Science as a data source for research on scientific and scholarly activity. *Quantitative Science Studies* 2020 Feb;1(1):363-376 [FREE Full text] [doi: [10.1162/qss_a_00018](https://doi.org/10.1162/qss_a_00018)]
2. Al Sultan A, Henson JH, Lickteig D. Assessing preservice elementary teachers' conceptual understanding of scientific literacy. *Teaching and Teacher Education* 2021 Jun;102:103327 [FREE Full text] [doi: [10.1016/j.tate.2021.103327](https://doi.org/10.1016/j.tate.2021.103327)]
3. Hallinger P, Kovačević J. A Bibliometric Review of Research on Educational Administration: Science Mapping the Literature, 1960 to 2018. *Review of Educational Research* 2019 Feb 15;89(3):335-369 [FREE Full text] [doi: [10.3102/0034654319830380](https://doi.org/10.3102/0034654319830380)]
4. Tibaná-Herrera G, Fernández-Bajón M, de MF. Mapping a Research Field: Analyzing the Research Fronts in an Emerging Discipline. *Scientometrics*. 2018 Jul 18. URL: <https://doi.org/10.5772/intechopen.76731> [accessed 2018-07-18]

5. Bales M, Wright D, Oxley P, Wheeler T. Bibliometric visualization and analysis software: State of the art, workflows, and best practices. 2020. URL: <http://creativecommons.org/licenses/by-nc-sa/4.0/> [accessed 2020-06-24]
6. Liao L, He X, Zhang H, Chua T. Attributed social network embedding. *IEEE Transactions on Knowledge and Data Engineering*;30(12). 2018 Mar 27. URL: <https://doi.org/10.1109/TKDE.2018.2819980> [accessed 2020-03-27]
7. Khazaneha M, Osareh F, Shafiee K. Trend Linking of Multiple System Atrophy: A Scientometric Study. *EMIDDT 2021 May 04*;21(4):700-710 [FREE Full text] [doi: [10.2174/1871530320666200607194810](https://doi.org/10.2174/1871530320666200607194810)]
8. El Emam K, Arbuckle L, Jonker E, Anderson K. Two h-index benchmarks for evaluating the publication performance of medical informatics researchers. *J Med Internet Res* 2012 Oct 18;14(5):e144 [FREE Full text] [doi: [10.2196/jmir.2177](https://doi.org/10.2196/jmir.2177)]
9. Masculo F, op den Buijs J, Simons M, Harma A. Natural language processing of medical alert service notes reveals reasons for emergency admissions. *iproc* 2019 Oct 2;5(1):e15225 [FREE Full text] [doi: [10.2196/15225](https://doi.org/10.2196/15225)]
10. Börner K, Scharnhorst A. Visual conceptualizations and models of science. *Journal of Informetrics* 2009 Jul;3(3):161-172 [FREE Full text] [doi: [10.1016/j.joi.2009.03.008](https://doi.org/10.1016/j.joi.2009.03.008)]
11. Ghanbari MK, Behzadifar M, Doshmangir L, Martini M, Bakhtiari A, Alikhani M, et al. Mapping research trends of universal health coverage from 1990 to 2019: bibliometric analysis. *JMIR Public Health Surveill* 2021 Jan 11;7(1):e24569 [FREE Full text] [doi: [10.2196/24569](https://doi.org/10.2196/24569)] [Medline: [33427687](https://pubmed.ncbi.nlm.nih.gov/33427687/)]
12. Sedighi M. Application of word co-occurrence analysis method in mapping of the scientific fields (case study: the field of Informetrics). *Library Review*; 65(1-2) 2016 Feb 1:52-64 [FREE Full text] [doi: [10.1108/lr-07-2015-0075](https://doi.org/10.1108/lr-07-2015-0075)]
13. Leung X, Sun J, Bai B. Bibliometrics of social media research: A co-citation and co-word analysis. *International Journal of Hospitality Management* Sep 1 2017:66 [FREE Full text] [doi: [10.1016/j.ijhm.2017.06.012](https://doi.org/10.1016/j.ijhm.2017.06.012)]
14. Janssens F, Leta J, Glänzel W, De Moor B. Towards mapping library and information science. *Information Processing & Management* 2006 Dec 1;42(6):1614-1642 [FREE Full text] [doi: [10.1016/j.ipm.2006.03.025](https://doi.org/10.1016/j.ipm.2006.03.025)]
15. Lyu X, Hu J, Dong W, Xu X. Intellectual structure and evolutionary trends of precision medicine research: cword analysis. *JMIR Med Inform* 2020 Mar 04;8(2):e11287 [FREE Full text] [doi: [10.2196/11287](https://doi.org/10.2196/11287)] [Medline: [32014844](https://pubmed.ncbi.nlm.nih.gov/32014844/)]
16. Randhawa K, Wilden R, Hohberger J. A bibliometric review of open innovation: setting a research agenda. *J Prod Innov Manag* 2016 Mar 02;33(6):750-772 [FREE Full text] [doi: [10.1111/jpim.12312](https://doi.org/10.1111/jpim.12312)]
17. Waqas A, Teoh SH, Lapão LV, Messina LA, Correia JC. Harnessing telemedicine for the provision of health care: bibliometric and scientometric analysis. *J Med Internet Res* 2020 Oct 02;22(10):e18835 [FREE Full text] [doi: [10.2196/18835](https://doi.org/10.2196/18835)] [Medline: [33006571](https://pubmed.ncbi.nlm.nih.gov/33006571/)]
18. Rebner M, Pai V. Breast cancer screening recommendations: African American women are at a disadvantage. *Journal of Breast Imaging* 2020 Sep;2(5). URL: <https://doi.org/10.1093/jbi/wbaa067> [accessed 2020-06-23]
19. Iacoviello L, Bonaccio M, de Gaetano G, Donati MB. Epidemiology of breast cancer, a paradigm of the "common soil" hypothesis. *Semin Cancer Biol* 2021 Jul;72:4-10 [FREE Full text] [doi: [10.1016/j.semcancer.2020.02.010](https://doi.org/10.1016/j.semcancer.2020.02.010)] [Medline: [32087245](https://pubmed.ncbi.nlm.nih.gov/32087245/)]
20. Society A. American Cancer Society. URL: <https://www.cancer.org/research/cancer-facts-statistics/breast-cancer-facts-figures.html> [accessed 2020-07-22]
21. Flores-Luevano S, Shokar NK, Dwivedi AK, Shokar GS, Defeu SN. Breast cancer fear among Mexican American women in the United States. *Breast Cancer (Auckl)* 2020 Aug 26;14:1178223420952745 [FREE Full text] [doi: [10.1177/1178223420952745](https://doi.org/10.1177/1178223420952745)] [Medline: [32922022](https://pubmed.ncbi.nlm.nih.gov/32922022/)]
22. Cobo M, López-Herrera A, Herrera-Viedma E, Herrera F. SciMAT: A new science mapping analysis software tool. *J Am Soc Inf Sci Tec* 2012 Jul 02;63(8):1609-1630 [FREE Full text] [doi: [10.1002/asi.22688](https://doi.org/10.1002/asi.22688)]
23. Lowe HJ. Understanding and using the medical subject headings (MeSH) vocabulary to perform literature searches. *JAMA* 1994 Apr 13;271(14):1103 [FREE Full text] [doi: [10.1001/jama.1994.03510380059038](https://doi.org/10.1001/jama.1994.03510380059038)]
24. MEDLINE: Overview. National Library of Medicine. URL: https://www.nlm.nih.gov/medline/medline_overview.html [accessed 2020-06-22]
25. Zupic I, Čater T. Bibliometric methods in management and organization. *Organizational Research Methods* 2014 Dec 22;18(3):429-472 [FREE Full text] [doi: [10.1177/1094428114562629](https://doi.org/10.1177/1094428114562629)]
26. Santana M, Morales-Sánchez R, Pasamar S. Mapping the link between corporate social responsibility (CSR) and human resource management (HRM): how is this relationship measured? *Sustainability* 2020 Feb 24;12(4):1678 [FREE Full text] [doi: [10.3390/su12041678](https://doi.org/10.3390/su12041678)]
27. SciMAT – Science Mapping Analysis Tool. URL: <https://sci2s.ugr.es/scimat/> [accessed 2021-09-30]
28. Gu D, Yang X, Deng S, Liang C, Wang X, Wu J, et al. Tracking knowledge evolution in cloud health care research: knowledge map and common word analysis. *J Med Internet Res* 2020 Feb 25;22(2):e15142 [FREE Full text] [doi: [10.2196/15142](https://doi.org/10.2196/15142)] [Medline: [32130115](https://pubmed.ncbi.nlm.nih.gov/32130115/)]
29. Cobo M, López-Herrera A, Herrera-Viedma E, Herrera F. Science mapping software tools: review, analysis, and cooperative study among tools. *J Am Soc Inf Sci* 2011 May 02;62(7):1382-1402 [FREE Full text] [doi: [10.1002/asi.21525](https://doi.org/10.1002/asi.21525)]
30. Moral-Muñoz JA, Cobo M, Peis E, Arroyo-Morales M, Herrera-Viedma E. Complement Ther Med 2014 Apr;22(2):409-418 [FREE Full text] [doi: [10.1016/j.ctim.2014.02.003](https://doi.org/10.1016/j.ctim.2014.02.003)] [Medline: [24731913](https://pubmed.ncbi.nlm.nih.gov/24731913/)]

31. Li F, Li M, Guan P, Ma S, Cui L. Mapping publication trends and identifying hot spots of research on Internet health information seeking behavior: a quantitative and co-word biclustering analysis. *J Med Internet Res* 2015 Mar 25;17(3):e81 [FREE Full text] [doi: [10.2196/jmir.3326](https://doi.org/10.2196/jmir.3326)] [Medline: [25830358](https://pubmed.ncbi.nlm.nih.gov/25830358/)]
32. Martínez MA, Cobo M, Herrera M, Herrera-Viedma E. Analyzing the scientific evolution of social work using science mapping. *Research on Social Work Practice* 2014 Feb 11;25(2):257-277 [FREE Full text] [doi: [10.1177/1049731514522101](https://doi.org/10.1177/1049731514522101)]
33. De LHA, Muñoz-Leiva F, Bakucz M. Past themes and future trends in medical tourism research: a co-word analysis. *Tourism Management*;65 2018 Apr 1:A [FREE Full text] [doi: [10.1016/j.tourman.2017.10.001](https://doi.org/10.1016/j.tourman.2017.10.001)]
34. Hellsten I, Leydesdorff L. Automated analysis of actor–topic networks on twitter: new approaches to the analysis of socio - semantic networks. *J Assoc Inf Sci Technol* 2019 Mar 18;71(1):3-15 [FREE Full text] [doi: [10.1002/asi.24207](https://doi.org/10.1002/asi.24207)]
35. Bordoloi M, Biswas S. Graph based sentiment analysis using keyword rank based polarity assignment. *Multimed Tools Appl* 2020 Jul 24;79(47-48):36033-36062 [FREE Full text] [doi: [10.1007/s11042-020-09289-4](https://doi.org/10.1007/s11042-020-09289-4)]
36. Lopez-Robles J, Otegi-Olaso J, Gamboa-Rosales N, Gamboa-Rosales H, Robles-Berumen H, Gamboa-Rosales A. Visualizing and mapping the project management research areas within the International Journal of Project Management: a bibliometric analysis from 1983 to 2018. In *Research and education in project management-REPM*, Bilbao (Spain), 21-22 February. 2019. URL: <http://eprints.rclis.org/39670/> [accessed 2020-08-13]
37. Gray C, Malins J. *Visualizing Research: A Guide to the Research Process in Art and Design* (1st ed.). London: Routledge; 2016:A.
38. Herrera-Viedma E, Martinez M, Herrera M. Bibliometric tools for discovering information in database. 2016 Jun 14 Presented at: International conference on industrial, engineering and other applications of applied intelligent systems (pp.). Springer, Cham; 2016 Aug 2; Cham p. 193-203 URL: https://doi.org/10.1007/978-3-319-42007-3_17 [doi: [10.1007/978-3-319-42007-3_17](https://doi.org/10.1007/978-3-319-42007-3_17)]
39. Ji Y, Tao W, Rim H. Mapping corporate social responsibility research in communication: A network and bibliometric analysis. *Public Relations Review* 2020 Dec 1;46(5) 2020 Dec 05:A [FREE Full text] [doi: [10.1016/j.pubrev.2020.101963](https://doi.org/10.1016/j.pubrev.2020.101963)]
40. Sardi A, Sorano E, Cantino V, Garengo P. Big data and performance measurement research: Trends, evolution and future opportunities. In: *Measuring Business Excellence*. England: Emerald Publishing Limited; May 20, 2020.
41. Nikzad M. Factors affecting continuance in the scientific activity of Iranian authors. Article in Persian. *International Journal of Information Science and Management* 2017;15(2):43 [FREE Full text]
42. Alcaide–Muñoz L, Rodríguez–Bolívar MP, Cobo M, Herrera–Viedma E. Analysing the scientific evolution of e-Government using a science mapping approach. *Government Information Quarterly* 2017 Sep;34(3):545-555 [FREE Full text] [doi: [10.1016/j.giq.2017.05.002](https://doi.org/10.1016/j.giq.2017.05.002)]
43. Phadermrod B, Crowder R, Wills G. Importance-performance analysis based SWOT analysis. *Int J Inf Manage* 2019 Feb;44:194-203 [FREE Full text] [doi: [10.1016/j.ijinfomgt.2016.03.009](https://doi.org/10.1016/j.ijinfomgt.2016.03.009)]
44. Costas R, Franssen T. Reflections around 'the cautionary use' of the h-index: response to Teixeira da Silva and Dobránszki. *Scientometrics* 2018 Mar 9;115(2):1125-1130 [FREE Full text] [doi: [10.1007/s11192-018-2683-0](https://doi.org/10.1007/s11192-018-2683-0)] [Medline: [29628537](https://pubmed.ncbi.nlm.nih.gov/29628537/)]
45. van Eck NJ, Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics* 2010 Aug 31;84(2):523-538 [FREE Full text] [doi: [10.1007/s11192-009-0146-3](https://doi.org/10.1007/s11192-009-0146-3)] [Medline: [20585380](https://pubmed.ncbi.nlm.nih.gov/20585380/)]
46. Donker M, van Tienhoven G, Straver P, Meijnen P, van de Velde CJH, Mansel RE, et al. Radiotherapy or surgery of the axilla after a positive sentinel node in breast cancer (EORTC 10981-22023 AMAROS): a randomised, multicentre, open-label, phase 3 non-inferiority trial. *Lancet Oncol* 2014 Nov;15(12):1303-1310 [FREE Full text] [doi: [10.1016/S1470-2045\(14\)70460-7](https://doi.org/10.1016/S1470-2045(14)70460-7)] [Medline: [25439688](https://pubmed.ncbi.nlm.nih.gov/25439688/)]

Abbreviations

- AJCC:** American Joint Committee on Cancer
- ALND:** axillary lymph node dissection
- IHC:** immunohistochemistry
- JAMA:** Journal of American Medical Association
- MeSH:** Medical Subject Headings
- NLM:** National Library of Medicine
- SLNB:** sentinel lymph node biopsy
- RIS:** research information system

Edited by D Vollmer Dahlke; submitted 22.12.20; peer-reviewed by G Kolostoumpis, G Balazs; comments to author 18.02.21; revised version received 08.04.21; accepted 30.05.21; published 28.10.21

Please cite as:

Sadatmoosavi A, Tajedini O, Esmaeili O, Abolhasani Zadeh F, Khazaneha M

Emerging Trends and Thematic Evolution of Breast Cancer: Knowledge Mapping and Co-Word Analysis

JMIR Cancer 2021;7(4):e26691

URL: <https://cancer.jmir.org/2021/4/e26691>

doi: [10.2196/26691](https://doi.org/10.2196/26691)

PMID:

©Ali Sadatmoosavi, Oranus Tajedini, Omid Esmaeili, Firouzeh Abolhasani Zadeh, Mahdiyeh Khazaneha. Originally published in JMIR Cancer (<https://cancer.jmir.org>), 28.10.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on <https://cancer.jmir.org/>, as well as this copyright and license information must be included.