

BIBLIOMETRIC ANALYSIS USING R

Tengku Muhammad Hanis Bin Tengku Mokhtar August 15, 2024



About me



Lead academic trainer at **Jom Research**

Background

- PhD (Public Health Epidemiology) from USM, 2024
- MSc (Medical Statistics) from USM, 2019
- MBBCh from Al-Azhar University, 2015

Interest:

- Medical statistics, meta-analysis, bibliometrics, scientometrics, text analysis
- Machine learning and deep learning application in medical sciences
- Application of R on health/medical data

Contact me:

- Email: tengkuhanismokhtar@gmail.com
- Website: https://tengkuhanis.netlify.app/



Jom Research promotion





Bibliometric analysis

Part 1



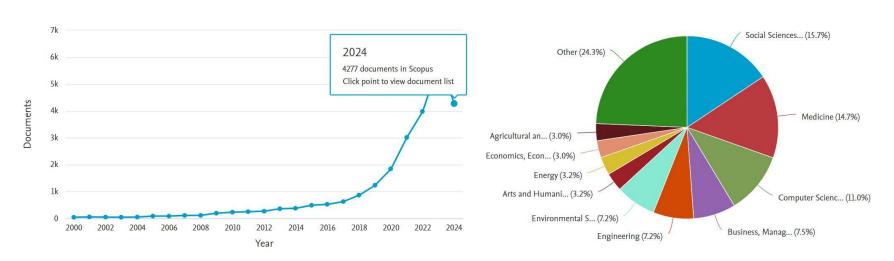
Bibliometric analysis

- Not a new analysis
- Not refers to any specific analysis, rather an umbrella term for a set of analyses
- Basically, refers to an analysis of bibliographic information
- Explorative in nature
- What we can do with bibliometric analysis:
 - Evaluate research progress
 - Quantitatively summarise research output
 - "Mapping" research contribution by author, institution, etc.
 - Explore research trends find out about over-research or under-research area
 - Etc



Bibliometric research

Bibliometric and scientometric papers on Scopus: 25, 437 (Aug 13, 2024)



Scopus data: 2000-2024



Example of bibliometric papers







Bibliometrics VS other reviews

Comparison of major review methods.

Review type	Goal	When to use	When not to use	Scope	Dataset	Analysis
Bibliometric analysis • Summarizes large quantities of bibliometric data to present the state of the intellectual structure and emerging trends of a research topic or field.		 When the scope of review is broad. When the dataset is too large for manual review. 	When the scope of review is specific. When the dataset is small and manageable enough that its content can be manually reviewed.	• Broad	• Large	Quantitative (evaluation and interpretation) Qualitative (interpretation only)
Meta-analysis	Summarizes the empirical evidence of relationship between variables while uncovering relationships not studied in existing studies.	When the focus of review is to summarize results rather than to engage with content, which may be broad or specific. When studies in the field are homogenous. When the number of homogeneous studies available is sufficiently high. When the number of homogeneous studies remaining after removing low quality studies is sufficiently high.	 When studies in the field are heterogeneous. When the number of homogenous studies is relatively low. When the number of high-quality homogeneous studies is relatively low. 	BroadSpecific	Large Small but adequate	Quantitative (evaluation and interpretation)
Systematic literature review	 Summarizes and synthesizes the findings of existing literature on a research topic or field. 	When the scope of review is specific. When the dataset is small and manageable enough that its content can be manually reviewed.	 When the scope of review is broad. When the dataset is too large for manual review. 	Specific	• Small	Qualitative (evaluation and interpretation)



- So, when to use a bibliometric analysis:
 - Data is large
 - The scope of review is broad



Data sources

Original research

BMJ Open Net survival differences of breast cancer

between stages at diagnosis and age groups in the east coast region of West Malaysia: a retrospective cohort study

Tengku Muhammad Hanis 0, Najib Majdi Yaacob, Suhaily Mohd Hairon, Sarimah Abdullah

To cite: Hanis TM, Yaacob NM, Mohd Hairon S, et al. Net survival differences of breast cancer between stages at diagnosis and age groups in the east coast region of West Malaysia: a retrospective cohort study RMI Open 2021:11:e043642 doi:10.113 bmiopen-2020-043642

 Prepublication history for this paper is available online. To view these files please visit the journal online (http://dx.doi. org/10.1136/bmiopen-2020-043642)

Received 10 August 2020 Revised 08 February 2021 Accepted 01 April 2021

Check for updates

Author(s) (or their employer(s)) 2021. Re-use nermitted under CC BY-NC No commercial re-use. See rights and permissions. Published by

¹Unit of Biostatistics and Research Methodology, School of Medical Sciences, Universit Sains Malaysia - Kampus Kesihatan, Kubang Kerian,

²Department of Community Medicine, School of Medical Sciences Universiti Sains Malaysia - Kampus Kesihata Kubang Kerian, Malaysia

Objective Estimation of the net survival of breast cancer helps in assessing breast cancer burden at a population level. Thus, this study aims to estimate the net survival of breast cancer at different cancer staging and age at diagnosis in the east coast region of West Malaysia. Setting Kelantan, Malaysia.

Participants All breast cancer cases diagnosed in 2007 and 2011 identified from Kelantan Cancer Registry. Design This retrospective cohort study used a relative survival approach to estimate the net survival of natients with breast cancer. Thus, two data were needed; breast cancer data from Kelantan Cancer Registry and general population mortality data for Kelantan population. Primary and secondary outcome measures Net survival according to stage and age group at diagnosis at

1. 3 and 5 years following diagnosis

Results The highest net survival was observed among stage I and II breast cancer cases, while the lowest net survival was observed among stage IV breast cancer cases. In term of age at diagnosis, breast cancer cases. aged 65 and older had the best net survival compared with the other age groups

Conclusion The age at diagnosis had a minimal impact on the net survival compared with the stage at diagnosis. The finding of this study is applicable to other populations with similar breast cancer profile.

Breast cancer is the most common cancer and the leading cause of cancer-related death among women globally. In the Asia region, the incidence of breast cancer had increased at least moderately in the Eastern and Southeastern region over the last two decades.2 Several studies have suggested that the increased incidences of breast cancer in Asia are due to factors such as economic growth and adoption of a more westernised lifestyle including no breastfeeding, reduced parity and increased animal fat consumption.3 4 However, the risk of developing breast cancer in Malaysia is still lower compared with the

Strengths and limitations of this study

- > This population-based study used state cancer reqistry data to describe the net survival of breast cancer in the east coast region of West Malaysia. ► The use of relative survival annmach to estimate the
- net survival is considered as a standard practice for a population-based study. ► This study expanded an abridged life table of the general population mortality data into a complete
- life table to estimate the net survival due to unavailability of a complete life table ► The use of different method to expand the mortality data leading to a lack of standardisation in this re-
- This study lacked socioeconomic sociodemographic and clinical information which may further explain

8, respectively.5 In term of breast cancer survival, a 5-year relative survival between 2005 and 2009 in Malaysia was 67.8%, while other Asian countries such South Korea (82,7%), Indonesia (77,7%) and Thailand (71.3%) had a better survival.6 Breast cancer staging and age are important prognostic factors of breast cancer. Early-stage patients with breast cancer are expected to have better survival compared with those diagnosed at a later stage.7 Several studies considered age as an important prognostic factor of breast cancer. 8-f0 although several Malaysian studies did find it otherwise. 11-13 There is a need to provide more information on breast cancer survival, generally in the Asia Pacific region and especially in Malaysia as it will help in planning effective public health management and control of the disease.

Net survival (NS) is a hypothetical measure in which the only possible cause of death is the disease of interest. 14 15 NS aims to reflect

the Kelantan general population mortality data from the DOSM was only available in an abridged life table format. Thus, the mortality data were expanded by the researcher in this study. Other studies may use a different method to expand the mortality data leading to a lack of standardisation in this research area. Admittedly, the expanded life table remained as an approximation to an actual complete life table. However, the use of different life tables in the relative survival analysis had been shown to have a minor change in the overall estimate of the NS rate. 19 Currently, the approach used in our study may serve as an alternative for populations without a complete life table. In facts, previous cancer studies had used a similar approach to estimate an NS statistic. 58-60 Also, this study was restricted by the limited information due to the used of secondary data to further explore the finding of this study. Information such as lymph node involvement, degree of metastasis and hormone receptor status was not available in this study.

CONCLUSIONS

This study presents the NS difference of patients with breast cancer according to cancer staging and age at diagnosis in the east-coast region of West Malaysia. The best survival was observed among patients with breast cancer aged 65 and older and those diagnosed at stages I and II. The age at diagnosis appeared to had a minimal impact on the NS compared with the stage at diagnosis. The NS is multifactorial in nature, thus detailed sociodemographic and clinical information on the breast cancer profile in this region is needed to extend this finding to other neighbouring populations.

Acknowledgements The publication of this study was supported by the School of Medical Sciences, Universiti Sains Malaysia. Also, the authors would like to thank the Kelantan State Health Department for providing the data in this study.

Contributors TMH; conceptualisation, methodology, formal analysis, writingoriginal draft. NMY: conceptualisation, methodology, writing-review and editing, validation. SMH: methodology, writing-review and editing, validation. SA: validation,

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Patient and public involvement. Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research. Patient consent for publication. Not required Ethics approval This study was carried out in accordance with the Declaration

of Helsinki, Ethical approval was obtained from the Human Research Ethics Committee University Coine Malauria (USM/ ICDeM/19000420) Medical Records and Ethics Committee Ministry of Health Malaysia (NMRR-18-2675-43980(IIR)) and written approval from the Kelantan State Health Department. Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request. Data may be obtained from a third party and are not publicly available. The General population mortality data for Kelantan is available from eStatistik (https://newss.statistics. gov.my/newss-portab/ep/epLogin.seam). Additionally, the breast cancer data is available from Non-communicable Disease (NCD) unit. Kelantan State Health

Open access

Department and Medical Research and Ethics Committee Ministry of Health

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which normits others to distribute remix adant, build upon this work non-commercially and license their derivative works on different terms, provided the priningl work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial, See: http://creativecommons.org/licenses/by-nc/4.0/.

Tengku Muhammad Hanis http://orcid.org/0000-0003-1466-7765

REFERENCES

- 1 Ginsburg O, Bray F, Coleman MP, et al. The global burden of women's cancers a grand challenge in global health. Lancet 2017:389:847-60
- 2 Youlden DB, Cramb SM, Vin CH, et al. Incidence and mortality of female breast cancer in the Asia-Pacific region. Cancer Biol Med 3 Yin CH. Bhoo Pathy N. Teo SH. A review of breast cancer research
- Malaysia. Med J Malaysia 2014:69:8-22. 4 Fan L, Goss PE, Strasser-Weippl K. Current status and future
- projections of breast cancer in Asia. Breast Care 2015;10:372-8. 5 Kirubakaran R. Chee Jia T. Mahamad Aris N. Awareness of breast cancer among surgical patients in a tertiary hospital in Malaysia. Asian Pac J Cancer Prev 2017;18:115–20.
- 6 Allemani C, Weir HK, Carreira H, et al. Global surveillance of cancer survival 1995-2009: analysis of individual data for 25,676,887 patients from 279 population-based registries in 67 countries (CONCORD-2), Lancet 2015;385:977-1010.
- Iqbal J, Ginsburg O, Rochon PA, et al. Differences in breast cancer stage at diagnosis and cancer-specific survival by race and ethnicity in the United States, JAMA 2015:313:165-73. 8 Wong FY, Tham WY, Nei WL, et al. Age exerts a continuous effect in
- the outcomes of Asian breast cancer patients treated with breastconserving therapy. Cancer Commun 2018;38:39-11. 9 Bharat A. Aft RL. Gao F. et al. Patient and tumor characteristics
- associated with increased mortality in young women (< or =40 years with breast cancer, J Sura Oncol 2009:100:248-51. 10 Partridge AH, Hughes ME, Warner ET, et al. Subtype-dependent
- relationship between young age at diagnosis and breast cancer eurolival . I Clin Opport 2016-34-3308-14 Ibrahim NI, Dahlui M, Aina EN, et al. Who are the breast cancer
- survivors in Malaysia? Asian Pac J Cancer Prev 2012;13:2213-8. 12 Nordin N, Yaacob NM, Abdullah NH, et al. Survival time and prognostic factors for breast cancer among/Women in north-east Peninsular Malaysia, Asian Pac J Cancer Prev 2018;19:497-502. 13 Taib NA, Akmal M, Mohamed I, Improvement in survival of breast
- cancer patients trends over two time periods in a single institution in an Asia Pacific country, Malaysia, Asian Pacific J Cancer Prev. 4 Mariotto AB, Noone A-M, Howlader N, et al. Cancer survival; an
- overview of measures, uses, and interpretation. J Natl Cancer Inst Monogr 2014;2014:145-86. 15 Lambert PC, Dickman PW, Rutherford MJ. Comparison of different
- approaches to estimating age standardized net survival. BMC Med Res Methodol 2015;15:1-13.
- 16 UKIACR. Standard operating procedure: guidelines on population based cancer survival analysis, 2016, Available: http://ukiacr.org/ publication/guidelines-population-based-cancer-survival-analysis 7 Sarfati D, Blakely T, Pearce N. Measuring cancer survival in
- populations: relative survival vs cancer-specific survival. Int J Enidemial 2010:39:598-610 18 Eloranta S, Adolfsson J, Lambert PC, et al. How can we make
- cancer survival statistics more useful for patients and clinicians an illustration using localized prostate cancer in Sweden. Cancer Causes Control 2013:24:505-15 19 Schaffar R. Rachet B, Belot A, et al. Estimation of net survival
- for cancer patients: relative survival setting more robust to some assumption violations than cause-specific setting, a sensitivity analysis on empirical data Fur J Concer 2017:72:78_83
- 20 National Cancer Intitute, Malaysian National Cancer Registry Repor 2007 - 2011. Putrajaya, 2015. http://nci.moh.gov.my

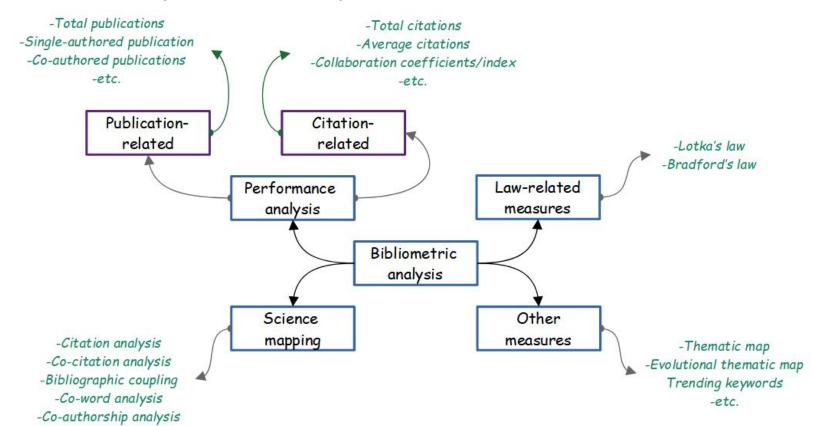


• Biliometrix compatible databases:

Source	URL	Format	Extension
Web of Science	https://www.webofknowledge.com/	 'BibTeX' 'plaintext' 'EndNote Desktop	'.bib''.txt''.ciw'
Scopus	https://www.scopus.com/	 'BibTeX' 'CSV export'	'.bib''.txt'
Dimensions	https://app.dimensions.ai/	 'Bibliometric mapping' 'Excel'	'.csv''.xlsx'
The Lens	https://lens.org/	'CSV export file'	· '.csv'
PubMed	https://pubmed.ncbi.nlm.nih.gov/	'PubMed export file'	• '.txt'
Cochrane Library	https://www.cochranelibrary.com/	'plaintext'	• '.txt'



Summary of analysis





Type of analysis

- Performance analysis account for contribution in research area
 - Publication-related metrics (proxy of productivity):
 - Total publications
 - Single-authored publications
 - Co-authored publications
 - etc
 - Citation-related metrics (proxy of influence)
 - Total citations
 - Average citations
 - Collaboration index/collaboration coefficient
 - etc



- 2. Science mapping assess relationship in research area (author, article, etc)
 - Citation analysis most influential publications
 - Co-citation analysis
 - Relationship among cited publication (paper 1 and paper 2 cited in paper 3, paper 1 & 2 = co-cited paper)
 - foundation/basic theme (look in the past)
 - Bibliographic coupling
 - Relationship among publications with same references (at least one reference)
 - Latest development



- Co-word analysis
 - Certain words (keywords, tittle, etc) shared by publications
 - May indicate future trend or support co-citation analysis and bibliographic coupling
- Co-authorship analysis
 - Interaction among researchers
 - Also can be applied to institutions, countries

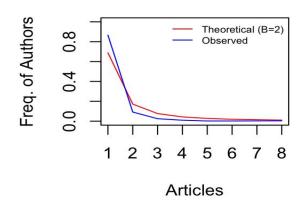


Bibliometric technique taxonomy	Unit of analysis used	Kind of relation
Bibliographic Coupling	AuthorDocumentJournal	 Common references in authors' oeuvres Common references in documents Common references in journals' oeuvres
Co-citation	AuthorReferenceJournal	Co-cited authorsCo-cited documentsCo-cited journals
Co-author	 Author Country from affiliation Institution from affiliation 	 Co-occurrence of authors in the author list of a document Co-occurrence of countries in the address list of a document Co-occurrence of institutions in the address list of a document
Co-word	 Keyword, or term extracted from title, abstract or document's body 	Co-occurrence of terms in a document

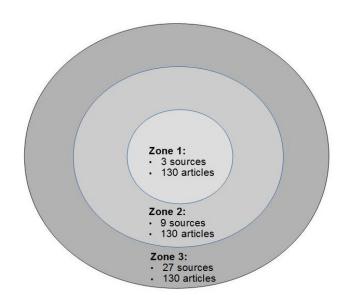


3. Theory/law-related metrics

Scientific Productivity



Lotka's law



Bradford's law



4. Other measures

- Thematic map
- Evolutional thematic map
- Trending keywords
- etc

	Density
Highly developed and isolated themes	Motor themes
<u> </u>	Centrality
Emerging or declining themes	Basic and transversal themes

a The strategic diagram.

Thematic map



Extension of bibliometrics

- 1. Basic textual analysis:
 - TF-IDF (Term Frequency-Inverse Document Frequency)
 - Wordcloud
- 2. Clustering (general)
 - k-means
 - Hierarchical clustering
 - etc
- 3. Topic models (more specific to corpus analysis)
 - Latent dirichlet allocation (LDA)
 - Latent semantic analysis (LSA)
 - Correlated topic model (CTM)
- 4. Network analysis
- Use R (<u>bibliometrix</u>) and Python (<u>pyBibX</u>)



Some idea on a bibliometric paper

- General bibliometrics
 - Bibliometric assessment of breast cancer research
- Specific bibliometrics limited to topics (need to review), number of papers, regions
 - Breast cancer and machine learning
 - Top 100 papers in machine learning
- Bibliometric comparison
 - Between countries, regions, etc



pen Access

BMJ Open Bibliometric analysis of the top-cited gastroenterology and hepatology articles

Samy A Azer, 1 Sarah Azer2

To cite: Azer SA. Azer S. Bibliometric analysis of the top-cited gastroenterology and hepatology articles. BMJ Open 2016:6:e009889 doi:10.1136/bmjopen-2015-

 Prepublication history and additional material is available. To view please visit the journal (http://dx.doi.org/ 10.1136/bmjopen-2015-

Received 1 September 2015 cented 18 December 2015

ARSTRACT

Objective: To identify the top-cited articles in gastroenterology and hepatology, and analyse their

Methods: Two searches were conducted in the Science Citation Index Expanded database; a search of 69 journals under the category 'Gastroenterology and Hepatology' (list A) and a keyword search of all journals (list B). The search results were analysed and the inter-rater coefficient of agreement between evaluators was measured using Cohen k.

Results: The number of citations varied from 1049 to 2959 in list A and from 1929 to 5500 in list B. In both lists, the majority of articles were research papers. No significant correlations were found between the number of citations and the number of years since publication (R2=0.00992, p=0.473 and R2=0.00202,

- Strengths and limitations of this study . Two searches were conducted in the Science Citation Index Expanded database.
- The search was based on journals with high impact factor and only those in the English
- Analysis explored a range of parameters in the

measurement to assess the work of researchers and impact of research,3 and to rank researchers on the basis of differences in citation indices.3 4 Recently, Nicholson and Ioannidis explored whether there is a lip between highly cited research and

Women's Health

RESEARCH ARTICLE

Open Access

Contribution of Arab countries to breast cancer research: comparison with non-Arab Middle Eastern countries

Waleed M Sweileh1*, Sa'ed H Zyoud2, Samah W Al-Jabi2 and Ansam F Sawalha1

Abstract

Background: Breast cancer is one of the most common types of cancers affecting women worldwide. The main objective of this study was to assess and compare research activity in breast cancer in Arab countries with non-Arab Middle Eastern countries.

Methods: Publications about "breast cancer" as a research topic were retrieved using the ISI Web of Science database. Analysis was confined to original research and review articles. Research productivity was assessed by assessing number of publications and time trend of these publications, names of journals, citation analysis, top 10 active institutions as well as country contribution to breast cancer research. The quantity and quality of publications from Arab countries in addition to 3 other Middle East countries (Turkey, Iran and Israel) were assessed and compared using the h-index tool,

Results: A total of 1658 original research and review articles about "breast cancer" were published from Arab countries al research productivity from Arah countries in the field of "breast cancer" was pedicible but showed a size!

Original research

BMJ Global Health A bibliometric analysis of COVID-19 research in Africa

Fatuma Hassan Guleid, 1 Robinson Oyando, 2 Evelyn Kabia, 2 Audrey Mumbi, 2 Samuel Akech.3 Edwine Barasa 0 2,4

To cite: Guleid FH. Ovando R. Kabia E, et al. A bibliometric analysis of COVID-19 research

in Africa. BM / Global Health 2021;6:e005690. doi:10.1136/ bmigh-2021-005690

Handling editor Seye Abimbola Additional supplemental view, please visit the journal line (http://dx.doi.org/10. bmigh-2021-005690).

Background The COVID-19 pandemic has led to an

antioxidants

unprecedented global research effort to build a body of knowledge that can inform mitigation strategies. We carried out a bibliometric analysis to describe the COVID-19 research output in Africa in terms of setting, study design, research themes and author affiliation. Methods We searched for articles published between 1 December 2019 and 3 January 2021 from various material is published online only. databases including PubMed, African Journals Online, medRxiv, Collabovid, the WHO global research database

and Google. All article types and study design were

Key questions

What is already known?

► Africa's contribution to global health research is low (1.3%) considering the high burden of infectious disease on the continent.

What are the new findings?

- Africa is contributing to the generation of COVID-19 knowledge by publishing primary and secondary research articles and editorial and commentary-type ► African authors have made significant contribu



RESEARCH ARTICLE

University-Industry Collaboration in China and the USA: A Bibliometric Comparison

Ping Zhou¹*, Robert Tijssen²€, Loet Leydesdorff³€

1 Department of Information Resources Management, School of Public Affairs. Zheiiang University. Hangzhou, Zheijang, China, 2 Center for Science and Technology Studies (CWTS), Leiden University, Leiden, The Netherlands, 3 Amsterdam School of Communications Research (ASCoR), University of Amsterdam, Amsterdam, The Netherlands

These authors contributed equally to this work.

* pingzhou@zju.edu.cn



Abstract

To cite: Zhai X, Cui J, Shao J,

spinal ultrasound: a systematic

Prepublication history and

aper are available online. To

w these files, please visit

surnal online (http://dx.do

136/hmionen-2016-

additional material for this

omjopen-2016-015317

Researc

BMJ Open Global research trends in spinal ultrasound: a systematic bibliometric analysis

Xiao Zhai, 1 Jin Cui, 1,2 Jie Shao, 1 Qijin Wang, 3 Xiao Chen, 1 Xianzhao Wei, 1 Xiaoyi Zhou, 1 Ziqiang Chen, 1 Yushu Bai, 1 Ming Li1

Taylor McElroy 1,2,3 and Antiño R. Allen 1,2,3,*

Stress and Chemobrain: 1990-2019

Division of Radiation Health, University of Arkansas for Medical Sciences, Little Rock, AR 72205, USA; TMMCELROY@uams.edu

A Bibliometric Review of Publications on Oxidative

- Department of Pharmaceutical Sciences, University of Arkansas for Medical Sciences, Little Rock, AR 72205, USA
- Department of Neurobiology & Developmental Sciences, University of Arkansas for Medical Sciences, Little Rock, AR 72205, USA
- Correspondence: aallen@uams.edu; Tel.: +1-501-686-7553

Received: 17 April 2020; Accepted: 15 May 2020; Published: 18 May 2020



MDPI

Abstract: Oxidative stress is considered one of the possible mechanisms behind chemobrain or the cognitive dysfunction persistent after chemotherapy treatment. Breast cancer patients have reported chemobrain symptoms since the 1990s. In this present bibliometric review, we employed

et al. Global research trends in Background In recent years, there has been increased interest in the use of ultrasound technology in the bibliometric analysis, BMJ Open evaluation of spinal and paraspinal regions. 2017:7:e015317. doi:10.1136/ Objective This study aimed to investigate trends in spinal

ultrasound research from 1994 to 2015 and compare the contributions of such research from different countries and

Study design Bibliometric analysis Setting Publications related to spinal ultrasound from 1994 to 2015 were retrieved from the Web of Science

Strengths and limitations of this study

- This bibliometric study included data on publication number, citation frequency, relative research interest, mean number of citations per paper, and
- The fitting mathematical model was used to analyse the time trend of the publications by using GraphPad
- VOSviewer, a Java programme, was used to visualise and analyse hotspots, masterpiece clusters, and







Bibliometric analysis

Part 2



Steps to do a bibliometric paper (quite similar to systematic review)

- 1. Develop research questions
- 2. Define:
 - Scope limit by country/year/etc
 - Objectives more specific than RQ
 - Selection criteria related to RQ and objectives
- 3. Specify bibliometric techniques (based on objectives)
- 4. Select databases
- Develop search terms always check with the databases whether the terms valid or not
- 6. Test run your search terms on the databases
- 7. Search and extract bibliographic data from databases
- 8. Review downloaded data/abstracts if needed (brief review)



- 9. Apply selection criteria if needed
- 10. Screen the papers if needed
- 11. Run bibliometric analysis and report findings



Develop a plan for bibliometric paper

- The purpose of the plan:
 - To guide you during the analysis and writing of the bibliometric paper
 - To record everything
- The plan should contains:
 - Proposed title
 - Objectives and bibliometric techniques (in a table)
 - Search terms
 - Databases
 - Screening strategy
 - Targeted journal
 - Key/main papers
- Example of the plan



Choosing a journal

- Some journal quite difficult to publish a bibliometric paper
 - Editor not familiar with a bibliometric paper
 - No reviewer with an expertise to review a bibliometric paper
- Solution:
 - Use <u>JANE website</u>
 - Make sure the journal has published a bibliometric paper previously



Result presentation

- Not all results of the analysis should be presented
- Make sure the result is aligned with the objectives
- Since we have a lot of results, try to put related result under a heading

while the number of journals increased as the number of zones increased. Bradford's law of scattering was estimated through the bibliometrix package.

Research Trends

Research trends were evaluated using a thematic map and trending keywords analysis. Both analyses were implemented through the bibliometrix package. Thematic map or strategic theme was proposed by Cobo et al. in 2011. The author's keywords which were extracted from the metadata of the database were used to construct the thematic map. Only the top 100 keywords with a minimum appearance of eight were included in the analysis to identify the thematic clusters.

Trending keywords analysis used the author's keywords retrieved from the Scopus database. Keywords with a minimum frequency of two from the publications between 2010 and 2020 were included in the analysis. Only the

top five keywords for each year were displayed in the plot.

RESULTS

DESCRIPTIVE	BIBLIOMETRIC
RESULT	

After removing three publications, there were 340 publications, ranging from 1982 to 2021 included in this study. The publications consisted of 312 research articles, 21 review papers and 7 conference papers. Additionally, Chan (1982) authored the earliest publication related to breast cancer research in Malaysia.

Main Characteristics of the Included Studies

The average citation per document was 15.8, and the range of citations was between 0 and 244. There were 8 single-authored papers, while the remaining 332 publications were multi-authored papers. The annual

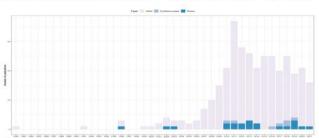


Figure 2: Frequency of publications related to breast cancer research in Malaysia according to the year of publication.

percentage of the growth rate of publications was 7.4%. Figure 2 shows a distribution of publications according to the year of publication. About 21.8% of publications were funded and the remaining 78.2% were not funded. Additionally, the number of authors per publication ranged between 1 and 26 and about 57.9% of the publications in this study had 3 to 6 authors.

Distribution of Authors

There were 1,222 authors involved in publications of breast cancer research in Malaysia. Yip CH was the top author with 69 publications related to breast cancer research in Malaysia ranging from 1996 until 2019, followed by Taib NA and Taib NAM with 56 and 21 publications, respectively. Figure 3 presents the top 10 most productive authors over time, taking into account the total citation per year on top of the number of articles published.

Distribution of Journals

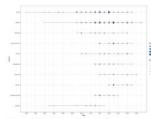


Figure 3: Top 10 most productive authors over time in the publication of breast cancer research in Malaysia.

Table 1: Core journals in breast cancer research in Malaysia.

Journals	Rank	No. of papers (%)	
Asian Pacific Journal of Cancer Prevention	1	67 (19.7)	
PLOS One	2	12 (3.5)	
World Journal of Surgery	2	12 (3.5)	
Medical Journal of Malaysia	4	9 (2.6)	
Singapore Medical Journal	4	9 (2.6)	
Malaysian Journal of Medical Sciences	6	7 (2.1)	

There were 146 journals related to breast cancer research in Malaysia. The six core journals published 34.1% of papers out of total publications in this study. Table 1 showed the list of all core journals involved in related publications, while Table 2 presents a zone of journals according to Bradford's law of scattering.

Institutional collaborations

About 360 institutions were involved in breast cancer research in Malaysia and was summarised in the institutional collaboration plot in Figure 4. However, only the top 20 institutions were displayed in the plot. Additionally, node degree reflected the number of connections each institution had, which is reflected by

Table 2: Zones of journals related to breast cancer research in Malaysia.

Zones	No of journals	No of papers
Zone 1 (core journals)	6	116
Zone 2	40	112
Zone 3	112	112

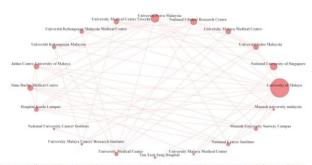


Figure 4: Collaboration among the top 20 institutions related to breast cancer research in Malaysia

the size of nodes in Figure 4. The most active institution in the research area was the University of Malaya with a node degree of 345, followed by the National University of Singapore (148), Universiti Sains Malaysia (132), Universiti Malaya Medical Centre (116) and National Clinical Research Centre RESEARCH TRENDS

Med & Health Dec 2022:17(2): 158-168

Thematic clusters

Figure 5 presented 11 clusters of breast

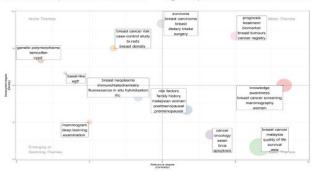


Figure 5: Thematic map of breast cancer research in Malaysia. cypd=cytochrome P450 2D6 or CYP2D6, egfr=epidermal growth factor receptor, bi-rads= breast imaging reporting and data system, ihc= immunohistochemistry, brca=breast cancer gene.

cancer research in Malaysia. The upperright quadrant represented a motor theme, and the lower-right quadrant represented basic themes. Additionally, the upper-left quadrant reflected niche themes, while the lower-left quadrant represented emerging or declining themes. The most frequent keywords were used to represent each cluster. There was one cluster in the motor themes, three clusters in the basic themes, three clusters in the niche themes and two clusters in emerging themes. Additionally, there were two in-between clusters, one cluster between the motor and basic themes and another one between motor and niche themes.

Breast cancer research related to prognosis, biomarker and cancer registry were the most developed and central research area in Malaysia. However, research areas related to family history, risk factors, quality of life, and survival of breast cancer patients reflected basic themes of breast cancer research in Malaysia. Research clusters in this theme had been studied for a long time and were once a central research area in breast cancer research in Malaysia. As shown in Figure 5, there was one in-between cluster that had become less central in breast cancer research in Malaysia.

The three clusters in the niche themes may indicate a highly developed and isolated research area. However, the clusters in this theme may migrate into a motor theme as one in-between cluster was observed between the niche and motor themes. Thus, the research clusters in niche themes may become a central research area of breast cancer

in Malaysia in future. Additionally, there were two emerging clusters related to breast cancer research in Malaysia, Researches related to deep learning and mammogram rose in popularity, while research related to immunohistochemistry (IHC) and fluorescence in situ hybridisation (FISH) had also showed an emerging trend.

Trending Keywords

Table 3 presents the top author keywords for publications in breast cancer research in Malaysia. Breast cancer research in Malaysia which related to awareness, quality of life, patient's survival and risk factors was starting to increase an established research domain. These keywords coincided with research clusters in basic themes in Figure 5.

Figure 6 illustrates trending keywords in breast cancer research in Malaysia for the last 10 years. The most prevalent keyword was 'immunohistochemistry' and 'breast neoplasm spanned' from

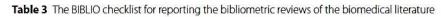
Table 3: Top author keywords for breast cancer

Author keywords	Frequency	
Breast cancer	179	
Malaysia	56	
Knowledge	15	
Quality of life	12	
Survival	12	
Asia	11	
Awareness	10	
Chemotherapy	9	
Polymorphism	9	
Risk factors	8	



Tips to write a bibliometric paper

- Make sure your research questions are critical enough
- Know what you can and can't do with the bibliometric analysis
 - Limitation of the software
 - Limitation of your knowledge
- Bibliometric analysis is quite an explorative in nature thus, be creative!
- Read past related papers know what else have not been addressed
- Find main paper reference that is close to what you want to do and try to adopt its approach/methods
- Follow <u>BIBLIO quideline</u> for reporting a bibliometric review/paper



Section/topic	Item no	Checklist item	Reported on page no
Title			9
Identification	1	Identify the report as a bibliometric review in the title	
Issues/topics	2	Indicate the key issues/topics under investigation and coverage of time period	
Abstract			
Structured summary	3	Structured summary including (as applicable): background, methods, results (key findings), and conclusions	
Introduction/background			
Justification/rationale/explanation	4	Present review of existing knowledge and epidemiological information	
Objectives	5	Statement of the objective(s) or question(s)	
Methods			
Search engines (data sources)	6	Describe all information sources (such as electronic databases, contact with study authors, trial registers, or other gray literature sources)	
Search strategy	7	Keywords and systematization criteria (date of search, language, type of document) for the search	
Time period	8	The period that the review covers and the justification	
Eligibility criteria	9	Describe all inclusion and exclusion criteria, languages, study design, type of publication, and time period	
Data refinement (data selection procedure) 10		Remove the irrelevant articles; inspection to eliminate duplicate and unrelated articles (after evaluation of the title, abstract, and content)	
Quality assessment (optional)	11	Assessment of papers by three authors and the use of assessing checklists	
Data synthesis	12	Describe the methods used for summarizing, handling, synthesis, tabulations, or schematic displays. Describe how the data were analyzed	
Results			
Descriptive findings (statistics)	13	 Provide details of the search and selection process in a flow diagram Number of citations retrieved (number of publication, year of publication, type of documents, country of publication, articles with the highest impact, most impactful authors, most impactful articles, authors with the highest production, top journals, top institutions,) 	
Schematic map and trend	14	Summarize and/or present the schematic maps and trends using an appropriate software to present citations, journals, authors, top journals, time trends, emerging literature, and any relevant indicators (as applicable) [64–68]	78



Table 3 (continued)

Section/topic	Item no	Checklist item	Reported on page no
Tabulation and summarizing the findings	15	General recommendation: Studies under consideration could be summarized and organized by different subtitles and different scenarios. Regardless, results need to be presented in separate tables covering each subtitle. The followings are some options that could help to summarize the findings Option 1: - Start the presentation with a historical view [when and who first published on the topic] - Report on review papers. The result should be listed in a separate table. Also, specify the review type (scoping review, narrative review, systematic review, and meta-analysis) - Summarize the findings according to the study designs and main study types Option 2: - Start the presentation with a historical view [when and who first published on the topic] - Report on review papers. The result should be listed in a separate table. Also, indicate the review type (scoping review, narrative review, systematic review, and meta-analysis) should be specified - Summarize the findings according to outcome measures or populations. For example, see [63] Option 3: - Start the presentation with a historical view [when and who first published on the topic] - Report on review papers. The result should be listed in a separate table. Also, specify the review papers. The result should be listed in a separate table. Also, specify the review papers. The result should be listed in a separate table. Also, specify the review papers. The result should be listed in a separate table. Also, specify the review papers. The result should be listed in a separate table. Also, and meta-analysis) - Summarize the findings according to concept [28] Option 4: - Start the presentation with a historical view [when and who first published on the topic] - Report on review papers. The result should be listed in a separate table, and also specify the review type (scoping review, narrative review, systematic review, and meta-analysis) - Summarize the findings according to different subtitles relevant to the main topic [26]	
Synthesis of findings	16	Synthesize the findings as much as possible, find the gap, and propose a model, hypothesis, etc. (if applicable)	
Discussion		NA 10 15 12 15 15 15	
Summary of evidence	17	Summarize the main findings. The findings should be presented in more "general" or "accessible" terms	
Interpretation	18	Include interpretation consistent with results. Explanations for observed outcomes, similarities, and differences reported would be essential	
Strengths and limitations	19	Discuss the strengths and limitations	
Conclusion(s)	20	Provide a general interpretation of the results with respect to the review questions and objectives, as well as potential implications	









Bibliometric analysis

Hands-on in R and biblioshiny



Field tags/column names

Field Tag	Description				
AU	Authors				
TI	Document Title				
so	Publication Name (or Source)				
ال	ISO Source Abbreviation				
DT	Document Type				
DE	Authors' Keywords				
ID	Keywords associated by SCOPUS or ISI database				
AB	Abstract				
C1	Author Address				
RP	Reprint Address				
CR	Cited References				
TC	Times Cited				
PY	Year				
sc	Subject Category				
UT	Unique Article Identifier				
DB	Bibliographic Database				

More field tags - https://www.bibliometrix.36df



Biliometrix compatible databases

Source	URL	Format	Extension
Web of Science	https://www.webofknowledge.com/	 'BibTeX' 'plaintext' 'EndNote Desktop	'.bib''.txt''.ciw'
Scopus	https://www.scopus.com/	 'BibTeX' 'CSV export'	'.bib''.txt'
Dimensions	https://app.dimensions.ai/	 'Bibliometric mapping' 'Excel'	'.csv''.xlsx'
The Lens	https://lens.org/	'CSV export file'	· '.csv'
PubMed	https://pubmed.ncbi.nlm.nih.gov/	'PubMed export file'	。 '.txt'
Cochrane Library	https://www.cochranelibrary.com/	'plaintext'	• '.txt'



Hands-on in WOS

- Use advanced search: TITLE("male breast cancer")
- Download either in plaintext or BibTeX format



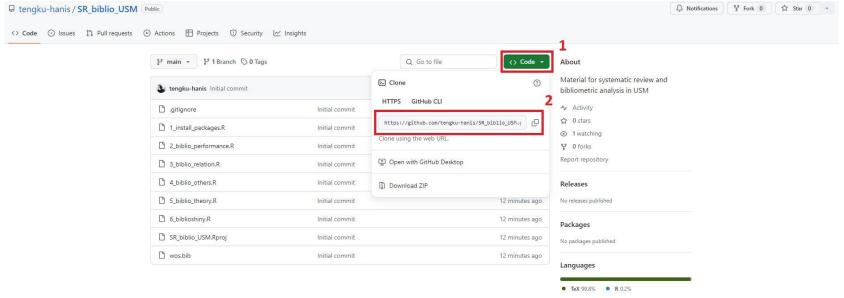
Hands-on in R - setting Posit Cloud

- Can either use:
 - RStudio IDE in your PC/laptop
 - Posit Cloud <u>sign up for a free account</u>



Hands-on in R - bibliometric analysis

- Go to https://github.com/tengku-hanis/SR_biblio_USM
- 2. Click $\langle \rangle$ Code \rightarrow local \rightarrow Clone \rightarrow HTTPS \rightarrow copy the link





- Log in to Posit Cloud
- Click New Project → New Project from Git Repository → paste the url



Suggested readings/references

- Bibliometrix website -<u>https://www.bibliometrix.org/vignettes/Introduction_to_bibliometrix.html</u>
- Aria, M., & Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. Journal of Informetrics, 11(4), 959–975. https://doi.org/10.1016/j.joi.2017.08.007
- Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021).
 How to conduct a bibliometric analysis: An overview and guidelines.
 Journal of Business Research, 133, 285–296.
 https://doi.org/10.1016/j.jbusres.2021.04.070
- Grant, M.J. and Booth, A. (2009), A typology of reviews: an analysis of 14 review types and associated methodologies. Health Information & Libraries Journal, 26: 91-108. https://doi.org/10.1111/j.1471-1842.2009.00848.x



- Hood, W.W., Wilson, C.S. The Literature of Bibliometrics, Scientometrics, and Informetrics. Scientometrics 52, 291 (2001).
 https://doi.org/10.1023/A:1017919924342
- Montazeri, A., Mohammadi, S., M.Hesari, P. et al. Preliminary guideline for reporting bibliometric reviews of the biomedical literature (BIBLIO): a minimum requirements. Syst Rev 12, 239 (2023).

https://doi.org/10.1186/s13643-023-02410-2







