Cancer in Indonesia: A bibliometric surveillance

Muhammad Iqhrammullah*, Randa Y. Refin, Rika I. Rasmi, Fina F. Andika, Hajjah Hajjah, Marlina Marlina and Ratna Ningsih
Faculty of Public Health, Universitas Muhammadiyah Aceh, Banda Aceh, Indonesia

*Corresponding author: m.iqhram@oia.unsyiah.ac.id

Abstract

Being archipelagic in geography and heterogenous in population contributes to the challenging comprehensive cancer surveillance in Indonesia. This present study analyzed the scientific publications on cancer epidemiology in Indonesia through bibliometric analysis aiming to complement the national survey data. The bibliometric data were retrieved from Scopus database on 9 May 2023 with pre-designed keyword combinations implying ‘cancer epidemiology in Indonesia’. The data were further presented in a table, distribution map, and visualized co-occurrence network. The visualization of co-occurring keywords was performed on VosViewer, and further analyzed qualitatively and quantitatively. We retrieved 1075 scientific records constituted of original articles (n=888, 82.6%), conference papers (n=102, 9.48%), and review articles (n=40, 3.72%) among others. The contributing affiliations are concentrated in Java Island, with Jakarta (n=603) as the province with the most publications. ‘Breast cancer’ (n=91) and ‘cervical cancer’ (n=39) emerged as the most frequently occurring keywords, and are assigned as the centers of research cluster, respectively. Keyword ‘breast cancer’ (total link strength, TLS=86) is connected with ‘colorectal cancer’, ‘metastasis’, ‘ovarian cancer’, and ‘prostate cancer’. Meanwhile, keyword ‘cervical cancer’ (TLS=35) is connected with ‘human papilloma virus’ and ‘knowledge’. The results are in line with the current epidemiological data of Indonesia, and they can provide different insights, especially on the research trend. Research and policy recommendations are presented in the last parts of this article.

Keywords: Bibliometric, cancer, epidemiology, tumor, VosViewer

Introduction

According to the International Agency for Research on Cancer (IARC), a specialized agency of the World Health Organization (WHO), the global burden of cancer continues to rise, with an estimated 19.3 million new cases and 10 million cancer-related deaths in 2020 [1]. Additionally, the latest data from the GLOBOCAN database reveals that the number of new cancer cases worldwide is projected to increase by almost 50% by 2040, reaching 28.4 million individuals [2]. It is estimated that the number of new cancer cases will rise by about 70% over the next two decades, particularly in low- and middle-income countries [3].

Cancer is a group of diseases characterized by the uncontrolled growth and spread of abnormal cells, forming tumors and/or invading nearby tissues and organs, which consequently result in damage and dysfunction [4]. These abnormally growing cells can occur in almost any part of the body and can be caused by a variety of genetic, environmental, and lifestyle factors [5]. Common types of cancer include breast cancer, lung cancer, prostate cancer, colorectal cancer, and skin cancer [6].
Tobacco use, high body mass index, low fruit and vegetable intake, and lack of physical activity are some of the major risk factors for cancer [6]. Exposure to pollution and heavy metal contamination is also a potential risk factor for cancer [7]. Studies have shown that exposure to air pollution, particularly in urban areas, is associated with an increased risk of breast cancer in postmenopausal women [8]. Additionally, high levels of heavy metals such as copper, iron, and zinc have been found in malignant breast tissues, suggesting a possible link between heavy metal exposure and breast cancer development [9].

The increasing trend of cancer incidence is also experienced in Indonesia, where more than 1.3 million cases were being treated in 2016 alone, contributing to financial burden of more than 4.8 billion USD [10]. A cancer epidemiological data (2008–2012) reported by Cipto Mangunkusumo Hospital, a national referral hospital in Indonesia, suggests that the large portion of the health burden is contributed by cervical and breast cancers [10]. Collecting epidemiologic data on cancer is one of public health urgency in Indonesia. Therefore, in the present study, we performed a cancer surveillance based on network visualization of bibliometric data from cancer-related epidemiologic study in Indonesia. Previously, bibliometric analysis has been employed to capture the raising concerns in public health sector, particularly for monkeypox [11], malaria [12], and Streptococcus pneumoniae vaccination [13]. In the case of cancer, bibliometric has also been employed for several countries, such as China [14], Middle East and North Africa Region [15], and South Africa [16]. None of the previous published studies reported the use of this simple and rapid analysis to understand the cancer epidemiology in Indonesia, suggesting the novelty of the present study.

Methods

Study design

To capture the pattern of epidemiologic research of cancer in Indonesia by means of bibliometric approach. The surveillance is based on the keyword occurrence from the published scientific report and their co-occurrence networks generated by VOSviewer.

Search strategy

The scientific literature search was carried out on a large database, Scopus (www.scopus.com) on 9 May 2023. The search was based on keywords ‘cancer’ and its synonyms combined with Indonesia (along with its provinces and large cities) and epidemiology (and related keywords such as prevalence, distribution, and risk factor). The complete keyword combination has been presented in Table 1. The search was performed on the title and abstract of the published records. No exclusion on the subject area or document type. The metadata containing information such as author(s), affiliated institutions, and keywords from author(s) and journal were extracted in a CSV file (.csv).

Network visualization

Network of keywords co-occurrence was visualized via VosViewer. The CSV (.csv) file obtained previously was read and analyzed by desktop application VosViewer version 1.6.19 (www.vosviewer.com). The visualization appeared in keywords with colors, network distance, and node size which correspond to the research cluster, co-occurrence frequency, and occurrence frequency, respectively.

Keyword analysis for type of cancer

Keyword frequency could be used to obtain the trending type of cancer being studied, where the data complement that of co-occurrence network. The keyword was selected based on its frequency. Those which do not correspond to the type of cancer were excluded. Frequencies of similar keywords were combined and presented as one keyword only. The selection was independently performed by two reviewers (M.I. and R.Y.R). Disagreement was resolved through discussion, and consultation with other authors when consensus was not reached.
Table 1. Keyword combination to search literature reporting cancer epidemiology in Indonesia

<table>
<thead>
<tr>
<th>Boolean operator</th>
<th>Keyword combination connected by “OR”</th>
</tr>
</thead>
<tbody>
<tr>
<td>First combination</td>
<td>neoplasm* OR cancer* OR tumor* OR carcinoma OR <em>blastoma AND Indonesia</em> OR Aceh* OR &quot;Bangka Belitung&quot; OR Bengkulu OR Jambi OR Lampung OR Sumatra OR Riau OR Banten OR Java* OR Jakarta OR Yogyakarta OR Kalimantan OR Borneo OR Bali* OR &quot;Nusa Tenggara&quot; OR Sulawesi OR Gorontalo OR Maluku OR Papua OR Median OR Samarinda OR Denpasar OR Makassar OR Ambon OR Jayapura AND NOT Malaysia OR &quot;Papua New Guinea&quot; AND NOT &quot;in vitro&quot; OR &quot;in vivo&quot;</td>
</tr>
</tbody>
</table>

Results

Characteristics of the retrieved records

From the literature search, a total of 1075 documents were retrieved, comprised predominantly of original articles (n=888, 82.6%), followed by conference papers (n=102, 9.48%) and review articles (n=40, 3.72%). A detailed list of document types and their respective frequencies has been presented in Table 2. The number of annual publications has been presented in Figure 1. Since 2010, publications about cancer epidemiology in Indonesia increased gradually each year until 2017. Thereafter, the increase was exponential. Yet, a declining trend was observed for 2022 publication.

When the number of publications is mapped according to the author’s location or institution, an uneven distribution can be observed (Figure 2). A province had publications as many as 603 papers, while others had as low as 1 or even 0 paper. The distribution is concentrated in Java Island, particularly in Jakarta (n=603), followed by Yogyakarta (n=310), Jawa Timur (n=277), Jawa Barat (n=172), and Jawa Tengah (n=157). The top contributors from Jakarta included Dr. Cipto Mangunkusumo Hospital, Dharmais Cancer Hospital, and Universitas Indonesia. Sumatera Utara (n=103), Sulawesi Selatan (n=94), and Bali (n=78) emerged as the top three provinces outside Java Island that publish epidemiologic studies of cancer. Even though Sulawesi Selatan among the top list, most provinces in the eastern part of Indonesia had relatively lower number of publications that is less than 10 papers (Figure 2).

![Figure 1. Annual number of publications reporting about cancer epidemiology in Indonesia.](http://doi.org/10.52225/narra.v1i2.86)

Figure 2. Distribution map of cancer epidemiology research in Indonesia.

Table 2. Types of documents recorded in the CSV (.csv) file

<table>
<thead>
<tr>
<th>Document type</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original Article</td>
<td>888 (82.61)</td>
</tr>
<tr>
<td>Conference Paper</td>
<td>102 (9.49)</td>
</tr>
<tr>
<td>Review</td>
<td>40 (3.72)</td>
</tr>
<tr>
<td>Conference Review</td>
<td>15 (1.39)</td>
</tr>
<tr>
<td>Book Chapter</td>
<td>14 (1.3)</td>
</tr>
<tr>
<td>Erratum</td>
<td>4 (0.37)</td>
</tr>
<tr>
<td>Short Survey</td>
<td>3 (0.28)</td>
</tr>
<tr>
<td>Note</td>
<td>3 (0.28)</td>
</tr>
<tr>
<td>Editorial</td>
<td>2 (0.19)</td>
</tr>
<tr>
<td>Data Paper</td>
<td>2 (0.19)</td>
</tr>
<tr>
<td>Letter</td>
<td>1 (0.09)</td>
</tr>
<tr>
<td>Book</td>
<td>1 (0.09)</td>
</tr>
</tbody>
</table>

Most occurring keywords – type of cancer

The numbers of occurrences of keywords that correspond to the type of cancer has been presented in Figure 3. Breast cancer had the highest occurrence with n=91, followed by cervical cancer (n=39). With the occurrence of n=27, nasopharyngeal cancer was found in the third position, which is slightly higher than colorectal cancer (n=25). Other cancers in this list are lung cancer, oral cancer, hepatocellular carcinoma, prostate cancer, brain tumor, ovarian cancer, and head and neck cancer.

Keyword occurrence and co-occurrence

The network visualizations of the keywords co-occurrences have been presented in Figure 4. Restriction of minimum threshold of 5 occurrence was applied high quality visualization. Of 2538 keywords, 71 met the threshold. The most used keyword in cancer epidemiology research in Indonesia was ‘Indonesia’ (n=118) and ‘breast cancer’ (n=91). These keywords were also observed to have n>50 occurrences. Other than the two keywords, ‘cancer’ (n=40), ‘cervical cancer’ (n=39) and epidemiology (n=27), were also emerged as the top keywords.

In the co-occurrence network, breast cancer appears as the most connected keywords and is the center of the ‘red’ cluster. This is further corroborated by the overlay visualization, where the keyword has intense red color. Keyword ‘breast cancer’ is connected with other keywords such as ‘colorectal cancer’, ‘metastasis’, ‘ovarian cancer’, and ‘prostate cancer’. Keyword ‘Indonesia’ is connected to two clusters. The first cluster (light blue) is populated by ‘hepatocellular carcinoma’, ‘hepatitis B virus’, and ‘liver disease’. Another cluster (cyan) that is connected with ‘Indonesia’ is comprised of ‘nasopharyngeal carcinoma, and ‘oral carcinoma’. In orange cluster, the center is ‘cervical cancer’ connected closely to ‘human papilloma virus’ and ‘knowledge’. Cervical cancer is
also a predominant keyword as confirmed by the overlay visualization. In another major cluster (yellow), keywords ‘gastric cancer’ and ‘Helicobacter pylori’ were found.

![Figure 3. Most occurring keywords – type of cancer.](image)

### Discussion

#### Bibliometric surveillance

Our data suggests that researchers focused on investigating the epidemiology of breast cancer, followed by cervical cancer. The two cancers were prevalent among Indonesian patients based on the five-year data (2008–2012) from Cipto Mangunkusumo Hospital, Jakarta, Indonesia [10]. Breast cancer itself occupies 19.2% of total cancer incidence in Indonesia, making it the most prevalent among other cancers [17]. Data from GLOBOCAN 2018 suggests that Indonesia has a high prevalence of cervical cancer compared to other Southeast Asian countries [18]. In the present study, we also found keyword ‘nasopharyngeal carcinoma’ and ‘colorectal cancer’ being frequently occurred. In a single center study, nasopharyngeal carcinoma is found as the most prevalent among other head and neck cancers [19]. Colorectal cancer, especially that in the colon, is known as one in the top 10 list of the most prevalent cancer in Indonesia [20]. Taken altogether, the bibliometric trend is in agreement with the epidemiology data, suggesting that the research is likely to be carried out in response to the incidence trend of cancer in Indonesia.

We also found that the cancer epidemiology studies were mostly reported by authors with affiliations located in Jakarta and its surrounding Provinces (Java Island). This also indicates that more studies had been carried out in the foregoing locations, as compared to the rest of the provinces. Unfortunately, the publication distribution is not in line with the distribution of cancer prevalence in Indonesia. For example, cervical cancer is more prevalent in Maluku Utara and Papua Barat [21], located in the eastern part of Indonesia. Reasons for low research activities and high cancer prevalence in the stated provinces could be associated with the absence of research facilities, healthcare facilities, and education institutions that supposedly are capable of carrying health promotion programs to reduce cancer risk factor as well as the screening and treatment [22].

From the visualization of keywords co-occurrence in the present study, we found some interesting connections: (1) ‘breast cancer’ with ‘colorectal cancer’, ‘metastasis’, ‘ovarian cancer’, and ‘prostate cancer’; (2) ‘cervical cancer’ with ‘human papilloma virus’; and (3) keywords ‘gastric cancer’ and ‘Helicobacter pylori’.
Figure 4. Co-occurrence network of keywords of cancer epidemiology in Indonesia (A) and its overlay display (B). In network visualization, connected keywords with the same color represent a research hotspot. In overlay visualization, the color indicates the occurrence frequency of the keywords, and the positions determine their closeness. The visualizations were generated on VosViewer (minimum occurrence = 5).
Breast cancer and ovarian cancer might have strong connection because the research subjects are women [23, 24]. Another possibility is due to metastasis, where the keyword is also present in the cluster. There is around 3–30% of propensity that breast cancer can metastasize to ovary [25], but it is not as high as other metastatic sites such as lung, bone, brain, and liver [26]. There are also cases where breast cancer occur as the results of metastasis of gastric and ovarian cancer [27]. Metastatic ovarian tumor has been suggested to be predominantly originated from non-gynecologic organs (mostly from gastrointestinal tract) as compared to gynecologic organs (viz. uterine body, uterine cervix, and fallopian tube) accounting for 40.6% and 59.4% of the total cases (n=64), respectively [28].

Oncogenic infection of human papilloma virus (HPV) is major cause of cervical cancer, responsible to approximately 70% of the total cases worldwide [29]. Hence, no wonder that the two keywords are strongly connected. As for keywords ‘gastric cancer’ and ‘Helicobacter pylori’, the two are closely related topics considering H. pylori can infect stomach and cause cancer [30]. Molecular pathway of H. pylori in causing gastric cancer involves the activation of cell signaling pathways that regulate cell proliferation concomitant to the exposure of bacterial virulence factors (such as cytotoxin- associated gene A, vacuolating cytotoxin A, and outer membrane proteins) [30]. A study from Japan that was participated by over 10,000 gastric cancer patients found that the development of cancer from H. pylori infection is also attributed to genetic factors [31]. Taken altogether, cancer-causing infections remain as the concerning issue in Indonesia.

**Research recommendation**

Herein, we observed several major research hotspots indicated by clusters in the keywords co-occurrence network visualization. Researchers in Indonesia have explored breast cancer and cervical cancer extensively. However, the extrinsic and intrinsic factors associated with the pathomechanism of the two cancers, especially their metastasis, are interesting to study. Effects of resection on the metastasis should be studied, and surgical techniques should be improved to reduce its incidence. There are some topics that are still understudied such as oral and nasopharyngeal cancers that are connected with ‘Indonesia’ as a non-specific keyword. In the future, researchers should explore the morbidity and mortality rates, characteristics, and risk factors of these diseases. Hepatocellular carcinoma has its own cluster, but the number is too small to be considered as the cluster center, necessitating more research on this topic. Given that the cancer-causing infections (such as HPV, hepatitis B virus, H. pylori) as one of the concerning issues, research should focus on their surveillance (including genomic surveillance), control, and prevention. Knowledge that determines the preventive efficacy for cancer-causing infections should be further studied and further developed as health promotion tools.

**Policy recommendation**

Health promotions should focus on preventing cancer-causing infections. With growing number of attributable publications, government and related organizations should be able to develop health promotion tools such as web-based or even social media-based educational contents. For example, a systematic review has highlighted the important role of social media in shaping awareness, attitude, and knowledge on HPV vaccination [32]. For regions with the least research activities, hence having small body of evidence, government should consider the affirmative allocation of research funding. Further, empowering the researchers’ capacity and building research infrastructures in the foregoing provinces (especially in Maluku Utara and Papua Barat) are of importance.

Self-assessment for possible cancer cases should be promoted, and examinations in clinical settings should be encouraged. Digitalization of this self-assessment can be considered as people have become more aware of the technology. Self-assessment mobile applications have been developed and evaluated for lung cancer, skin cancer, breast cancer, and so on [33, 34]. Further, digitalization could suppress the financial and human resources required to realize health promotion programs. Last but not least, the government is also responsible for providing accessible cancer screening. Long and difficult screening process for breast cancer and cervical cancer (such as mammography or Pap smears) have been notified by researchers [35].
**Strengths and limitations**

Limitations of using bibliometric study include the present of ‘noise’ from irrelevant research. Though keyword combinations have been used, there is still probability for irrelevant research to be included in the extracted data; however, the number was not predominant. Secondly, the study could only provide the trend, and not necessarily the accurate data which limit the interpretation and application of the findings. In reading the keyword co-occurrence analysis, unintended keywords are likely to be included contributing to the biased result, but this problem has been resolved by discussion among authors in the interpretation. Regardless, bibliometric could be considered as a powerful tool to observe epidemiology studies since it can be performed both nationally and internationally in a short amount of time. The findings could be used to advice future research of cancer epidemiology in Indonesia.

**Conclusion**

Data from the bibliometric study can be used to depict the concerns of researchers on specific cancer which is an important epidemiological insight. Further, the study could inform government regarding the trend of cancer epidemiology research in Indonesia, in which the data could be used for allocations of funding or other type of incentives. The findings from the present study reveal that breast cancer is the most epidemiologically studied cancer and followed by cervical cancer. These are in line with the national survey data, suggesting that the research project has been in accordance with the national agenda. Unfortunately, there are several provinces that are not active in studying cancer epidemiology. Therefore, in the future, the government can incentivize the researchers from the aforementioned provinces to increase their research and publication activities.

**Ethics approval**

Not required

**Acknowledgments**

Authors appreciate the supervision from Dr. Maidar of the Universitas Muhammadiyah Aceh.

**Competing interests**

All authors declare that they have no conflicts of interest.

**Funding**

This study received no external funding.

**Underlying data**

All data underlying the results are available from the corresponding author upon reasonable request.

**How to cite**


**References**


32. Ortiz RR, Smith A, Coyne-Beasley T. A systematic literature review to examine the potential for social media to impact HPV vaccine uptake and awareness, knowledge, and attitudes about HPV and HPV vaccination. Hum Vaccin Immunother 2019;15(7–8):1465–1475.

