Internet Search Trends on Breast Cancer Between 2017 and 2021 in Brazil: an Infoveillance Investigation

Tendências de Pesquisa na Internet sobre Câncer de Mama entre 2017 e 2021 no Brasil: uma Investigação de Infovigilância

Tendencias de Búsqueda en Internet sobre Cáncer de Mama entre 2017 y 2021 en Brasil: una Investigación de Infovigilancia

Ricardo Barbosa Lima[1]; José Cleyton de Oliveira Santos[2]; Dayane Ketlyn da Cunha Santos[3]; Pabliane Matias Lordelo Marinho[4]; Glebson Moura Silva[5]; Simone Yuriko Kameo[6]

ABSTRACT

Introduction: Internet users frequently search for information on health issues. The behavior of Internet users may be related to health contexts experienced in their communities, reflecting unmet demands. **Objective:** To analyze the interest of Internet users on breast cancer in Brazil between 2017 and 2021. **Method:** An infodemiological study was carried out using the relative search volume (ranging from 0 to 100) for the term "*câncer de mama*" (breast cancer) in Brazil, retrieved from the Google Trends tool. The annual relative search volume was compared with a non-parametric approach and significance value (*p*) of 5%. **Results:** Except between 2019 and 2020 (*p*-value>0.05), there was a significant increase in Internet searches for the term "*câncer de mama*" in Brazil (all *p*-values <0.05). The peak of interest (relative search volume=100) was found in October 2020. Internet users who searched for the term "*câncer de mama*" often searched for signs and symptoms of breast cancer. **Conclusion:** The interest of Internet users on breast cancer has increased in recent years in Brazil.
Key words: breast neoplasms/epidemiology; time factors; information seeking behavior; internet; search engine.

RESUMO

Introdução: Os usuários da Internet frequentemente buscam informações sobre questões de saúde. O comportamento desses usuários pode estar relacionado com contextos de saúde vivenciados em suas comunidades, refletindo demandas não supridas. **Objetivo:** Examinar o interesse dos usuários da Internet sobre o câncer de mama no Brasil, considerando o período entre 2017 e 2021. **Método:** Estudo infodemiológico utilizando o volume relativo de busca (variando entre 0 e 100) para o termo "câncer de mama" no Brasil, recuperado da ferramenta *Google Trends*. O volume relativo de busca entre os anos foi comparado com abordagem não paramétrica e valor de significância (*p*) de 5%. **Resultados:** Exceto entre 2019 e 2020 (*p*-valor>0,05), houve um aumento significativo nas buscas na Internet pelo termo "câncer de mama" no Brasil (todos os valores de *p*<0,05). O pico de interesse (volume relativo de busca=100) foi observado no mês de outubro de 2020. Os usuários da Internet que buscaram pelo termo "câncer de mama" frequentemente procuravam por sinais e sintomas do câncer de mama. **Conclusão:** Observou-se que o interesse desses usuários sobre câncer de mama aumentou nos últimos anos no Brasil.
Palavras-chave: neoplasias da mama/epidemiologia; fatores de tempo; comportamento de busca de informação; internet; ferramenta de busca.

RESUMEN

Introducción: Los internautas suelen buscar información sobre temas de salud. El comportamiento de los internautas puede estar relacionado con contextos de salud vividos en sus comunidades, reflejando demandas insatisfechas. **Objetivo:** Examinar el interés de los internautas sobre el cáncer de mama en Brasil, considerando el período comprendido entre 2017 y 2021. **Método:** Estudio infodemiológico utilizando el volumen relativo de búsqueda (que va de 0 a 100) para el término "*câncer de mama*" en Brasil, obtenido de la herramienta *Google Trends*. El volumen relativo de búsqueda entre los años se comparó con un enfoque no paramétrico y un valor de significación (*p*) del 5%. **Resultados:** Excepto entre 2019 y 2020 (*p*-valor>0,05), hubo un aumento significativo en las búsquedas en Internet del término "*câncer de mama*" en Brasil (todos los valores *p*<0,05). El pico de interés (volumen relativo de búsqueda=100) se observó en el mes de octubre de 2020. Los usuarios de Internet que buscaban el término "*câncer de mama*" a menudo buscaban signos y síntomas de cáncer de mama. **Conclusión:** Se observó que el interés de los internautas por el cáncer de mama ha aumentado en los últimos años en Brasil.
Palabras clave: neoplasias de la mama/epidemiología; factores de tiempo; conducta en la búsqueda de información; internet; motor de búsqueda.

[1]Universidade de São Paulo, Faculdade de Odontologia de Ribeirão Preto. Ribeirão Preto (SP), Brazil. E-mail: dentistaricardolima@gmail.com. Orcid iD: https://orcid.org/0000-0001-5274-4800

[2,3,5,6]Universidade Federal de Sergipe. Lagarto (SE), Brazil. E-mails: cleyton-121@hotmail.com; daayketlyn27@gmail.com; glebsonmoura@yahoo.com.br; simonekameo@hotmail.com. Orcid iD: https://orcid.org/0000-0001-5616-7625; Orcid iD: https://orcid.org/0000-0002-9312-4891; Orcid iD: https://orcid.org/0000-0002-4977-2787; Orcid iD: https://orcid.org/0000-0002-0035-2415

[4]Universidade de São Paulo, Escola de Enfermagem de Ribeirão Preto. Ribeirão Preto (SP), Brazil. E-mail: marinho.pabliane@gmail.com. Orcid iD: https://orcid.org/0000-0001-6190-0844

Corresponding author: Ricardo Barbosa Lima. Rua Antônio Carlos Brandão, 98, Bl. Nápoles, Apto. 207 – Centro. Petrolina (PE), Brazil. CEP 56304-110. E-mail: ricardolimadentista@gmail.com

INTRODUCTION

The worldwide spread of SARS-CoV-2 has triggered the COVID-19 pandemic with significant health impacts, including the late diagnosis and treatment of cancer. There has been an important decrease in the number of cancer diagnoses in Brazil since the beginning of the pandemic[1]. The impact of the pandemic on breast cancer is an emerging concern, as it is a common type of cancer in Brazil, predominantly affecting females[2].

In parallel, it is known that Internet users often seek information about cancer and the pandemic may have redirected the public interest towards other themes, contributing to the decline of cancer diagnoses. A decrease in interest in terms related to cancer screening, such as mammography, has already been noticed[3]. The change in the behavior of Internet users is a valuable tool to understand the health dynamics of a territory, and initiatives to find what is the public interest about health-related terms have become even more relevant, as social distancing and the reorganization of health services may have significantly influenced the health-disease process, including cancer[1,4].

Data between 2011 and 2016 have already shown potential periods to browse the Internet searching breast cancer initiatives, especially in the months of October, when *Outubro Rosa* (breast cancer awareness campaign in the month of October of every year) takes place[5]. Considering the impact of breast cancer on women's health, the *Outubro Rosa* campaign focuses on encouraging screening and early treatment, and search for health care[6]. The positive effect of this campaign on the public interest in breast cancer has already been documented with data between 2014 and 2019, demonstrating the relationship between the Internet and the search for health information[7].

However, it is possible to hypothesize that the COVID-19 pandemic may have affected the behavior of Internet users, modifying the interest in terms related to breast cancer. To the best of the current knowledge, there was no investigation on this topic with complete data between 2020 and 2021 (pandemic period) in Brazil. Therefore, the objective of this study was to investigate the interest of Internet users on the term "*câncer de mama*" in Brazil, between 2017 and 2021. The alternative hypothesis tested was: (H_1) there was a significant increase in the interest of Internet users on breast cancer between 2017 and 2021 in Brazil.

METHOD

The Google Trends[8] tool was used to retrieve the relative search volume (RSV) for the term "*câncer de*

mama" (breast cancer), following a standard procedure. RSV is a measure that reflects the public interest for a term in the Google search platform, ranging between 0 and 100 (peak of interest). When simultaneous searches occur, the RSV is normalized between periods, allowing to understand the interest in the term over time[9,10]. As this is an infodemiological study with data in the public domain and do not refer to individuals, the review by the Institutional Review Board was waived. The checklist suggested by Nuti et al.[11] was used to guide this report.

In the Google Trends platform, five search boxes were added. In each one, the same search term "*câncer de mama*" was used. Each box was configured to adjust to the location (Brazil) and period (2017, 2018, 2019, 2020 or 2021 - "*full year*"). Also, in each box, "*all query categories*" and "*web searches*" options were selected to adjust the search profile. Data were retrieved in February 2021 by a single search strategy and exported to a spreadsheet for analysis. The RSV for each year was retrieved, normalized by the Google Trends algorithm as 52 measures for each one (summarizing data from 52 weeks per year). Also, the RSV for the correlated searches made by the same users who searched for the term "*câncer de mama*" was retrieved and the most searched were presented.

As RSV is a discrete quantitative variable, a non-parametric approach was needed to compare interest across years. The Kruskal-Wallis test was used to verify if there were statistically significant differences between the years, as well as the Dwass-Steel-Critchlow-Fligner (DSCF) *post hoc* test for pairwise comparisons, both using *the jamovi* software (2.3.15, Sydney, Australia, 2020), adjusting the significance value (*p*) at 5% (α=0.05). In addition to the analytical procedure, descriptive measurements were provided for the RSV in each year: median (including 95% confidence interval by bootstrap approach; $CI_{95\%}$), quartiles (Q1: first and Q3: third), interquartile range (IQR), minimum, maximum and sum.

The trend over time (months) was estimated using the Prais-Winsten regression method to adjust the values of coefficient β_1 after $log10$ transformation. The monthly percent change (MPC; %) was estimated by the equation: $(-1+10^{\beta1})*100$. The $CI_{95\%}$ was estimated by the minimum and maximum values of β_1 applied in the equation above, both obtained by the equation: $(\beta_1) \pm$ (critical *t*-value*β_1-standard error). This analysis was performed using the PAST software (version 4.3, Oslo, Norway).

RESULTS

Table 1 summarizes the RSV for each year and in the month of October, and Figure 1 shows the RSV evolution for the term "*câncer de mama*" in Brazil between 2017 and

2021. The Kruskal-Wallis test demonstrates a statistically significant difference between the RSV results from 2017 to 2021 ($p<0.001$). In the pairwise comparisons among them, the DSCF test showed that there was no statistically significant difference in RSV between 2019 and 2020 ($p=0.111$). However, all other one-by-one comparisons were statistically significant (all $p<0.05$). In gross values (RSV), comparing 2017 and 2021, the median increase was established at 9 points ($CI_{95\%}$: 8, 9). In percentage, this variation represents 95% ($CI_{95\%}$: 75.7, 104.3).

Table 1. Relative search volume for the term *"câncer de mama"* in Brazil between 2017 and 2021

RSV	2017	2018	2019	2020	2021
Annual (median)	9[a]	11[b]	13[c]	14[c]	17[d]
Q1	7	9	11	12.8	15
Q3	10	14	15	18.5	19
IQR	3	5	4	5	4
Minimum	5	6	10	10	12
Maximum	33	40	46	100	71
Sum	549	703	830	1.036	1.107
October (median)	30	33	41	65.5	56

Captions: RSV = relative search volume; Q1 = first quartile; Q3 = third quartile; IQR = interquartile range.
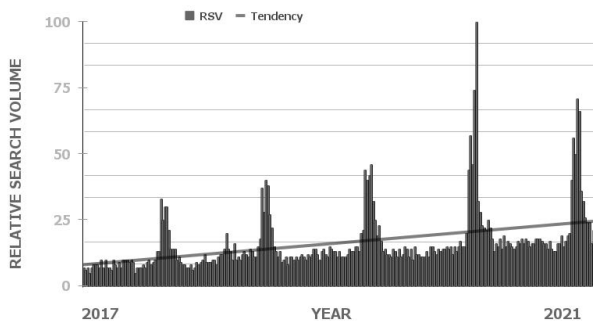(a, b, c e d) = statistically significant difference ($p<0.05$).

Figure 1. Relative search volume for the term *"câncer de mama"* in Brazil between 2017 and 2021

It is possible to conclude that the peak of interest (RSV equal to 100) was registered in October 2020. When qualitatively comparing the annual RSV and in the weeks of the referenced month, higher values were found in this month in relation to the other periods of the year. Hence, with the exception of the interval between 2019 and 2020, it is possible to detect the increased interest on the Internet for the term *"câncer de mama"* in Brazil, especially in the months of October. The stratified analysis between the years was confirmed by the tendency analysis in the Prais-Winsten regression method. The MPC was established

at 0.37% ($CI_{95\%}$: 0.04, 0.70, $r^2=0.333$, *P*-value<0.001), indicating an increasing and significant tendency over time (2017-2021).

Evaluating the Internet searches made by the same users at Google platform for the term *"câncer de mama"*, the interest was constant for signs and symptoms for all the years. The RSV for this topic between 2017 and 2021 was 54 (*"sintomas de câncer de mama"* or *"câncer de mama sintomas"*), 70 (*"sintomas"*, *"sintomas de câncer de mama"* or *"câncer de mama sintomas"*), 86 (*"sintomas de câncer de mama"* or *"câncer de mama sintomas"*), 73 (*"sintomas de câncer de mama"* or *"câncer de mama sintomas"*), and 79 (*"sintomas de câncer de mama"* or *"câncer de mama sintomas"*) respectively. This suggests that users searching for this information often want to be informed about signs and symptoms.

DISCUSSION

H_1 was fully accepted. The outcome observed supports the increasing interest on breast cancer issues in Brazil by Internet users (especially for signs and symptoms), even in the COVID-19 pandemic scenario, when the peak of interest occurred. Furthermore, it supports the possible impact of the *Outubro Rosa* campaign, considering that during this period an increase in searches on the term *"câncer de mama"* was noticed.

The literature has already extensively discussed the increased use of the Internet during the COVID-19 pandemic and the impact of shared information on health issues. The interest of Internet users on breast cancer raises the necessity of monitoring the content available on the web, because it is a valuable tool for sharing high-impact content and promoting health education on breast cancer issues[12,13]. Previous investigations have already demonstrated the Internet as a tool to support women with breast cancer. Nowadays, considering the facilities to access this tool, strategies to offer information on adverse events related to the cancer treatment, as well as emotional support, may be relevant[14,15].

Moreover, also considering the increased interest of Internet users, this opportunity is even more promising, expanding access to health information. However, it must be taken into account that a level of digital literacy is necessary to search and understand the health information available on the Internet. The language, the relevance of the content and its veracity, as well as the credibility of those who provide the content, are aspects of the problem that involves the search for health information on the Internet[16].

A limitation of this study is that there are scarce information about Internet users who searched for

breast cancer. In addition, individuals who do not have Internet access or the ability to search for information did not influence this outcome, as well as alternative terms (e.g. regional expressions or misspellings) were not considered. Future investigations should evaluate the content available on the Internet about breast cancer in Portuguese, quantitatively and qualitatively and monitor the tendency over the years to understand changes in user interests.

CONCLUSION

The interest of Internet users on breast cancer has increased in recent years in Brazil, including the COVID-19 pandemic. It may reflect the public interest, reinforcing the necessity to understand the relationship between the illness and Internet searches for health issues over time.

CONTRIBUTIONS

Ricardo Barbosa Lima (conception/design, planning, data collection, analysis, interpretation, wording and review); José Cleyton de Oliveira Santos and Dayane Ketlyn da Cunha Santos (conception/design, analysis, interpretation, wording and review); Pabliane Matias Lordelo Marinho, Glebson Moura Silva and Simone Yuriko Kameo (conception/design, analysis, interpretation and review). All the authors approved the version to be published.

DECLARATION OF CONFLICT OF INTERESTS

There is no conflict of interests to declare.

FUNDING SOURCES

None.

REFERENCES

1. Marques NP, Silveira DMM, Marques NCT, et al. Cancer diagnosis in Brazil in the COVID-19 era. Semin Oncol. 2021;48(2):156-9. doi: https://doi.org/10.1053/j.seminoncol.2020.12.002

2. Ferreira MC, Vale DB, Barros MBA. Incidence and mortality from breast and cervical cancer in a Brazilian town. Rev Saude Publica. 2021;55:67. doi: https://doi.org/10.11606/s1518-8787.2021055003085

3. Snyder A, Jang S, Nazari IS, et al. Google search volume trends for cancer screening terms during the COVID-19 pandemic. J Med Screen. 2021;28(2):210-2. doi: https://doi.org/10.1177/0969141321999426

4. Berning P, Maybody M, Deipolyi AR, et al. Internet search trends relevant to interventional oncology: a Google Trends study (2004-2020). J Vasc Interv Radiol. 2021;32(10):1445-1448.e1. doi: https://doi.org/10.1016/j.jvir.2021.07.005

5. Vasconcellos-Silva PR, Carvalho DBF, Trajano V, et al. Using Google Trends data to study public interest in breast cancer screening in Brazil: why not a pink February? JMIR Public Health Surveill. 2017;3(2):e17. doi: https://doi.org/10.2196/publichealth.7015

6. Sá MFS. Pink October and breast cancer in Brazil. Rev Bras Ginecol Obstet. 2021;43(10):725-7. doi: https://doi.org/10.1055/s-0041-1739451

7. Quintanilha LF, Souza LN, Sanches D, et al. The impact of cancer campaigns in Brazil: a Google Trends analysis. Ecancermedicalscience. 2019;13:963. doi: https://doi.org/10.3332/ecancer.2019.963

8. Google Trends [Internet]. Mountain View, CA: Google; [1999]. [acesso 2022 mar 25]. Disponível em: https://trends.google.com.br/trends

9. Mavragani A, Ochoa G. Google Trends in infodemiology and infoveillance: methodology framework. JMIR Public Health Surveill. 2019;5(2):e13439. doi: https://doi.org/10.2196/13439

10. Mavragani A, Ochoa G, Tsagarakis KP. Assessing the methods, tools, and statistical approaches in Google Trends research: systematic review. J Med Internet Res. 2018;20(11):e270. doi: https://doi.org/10.2196/jmir.9366

11. Nuti SV, Wayda B, Ranasinghe I, et al. The use of Google Trends in health care research: a systematic review. PLoS One. 2014;9(10):e109583. doi: https://doi.org/10.1371/journal.pone.0109583

12. Venegas-Vera AV, Colbert GB, Lerma EV. Positive and negative impact of social media in the COVID-19 era. Rev Cardiovasc Med. 2020;21(4):561-4. doi: https://doi.org/10.31083/j.rcm.2020.04.195

13. Gottlieb M, Dyer S. Information and disinformation: social media in the COVID-19 crisis. Acad Emerg Med. 2020;27(7):640-1. doi: https://doi.org/10.1111/acem.14036

14. Kemp E, Koczwara B, Turner J, et al. Internet use and preferences among women living with advanced breast cancer. Breast J. 2019;25(2):290-5. doi: https://doi.org/10.1111/tbj.13203

15. Kemp E, Koczwara B, Butow P, et al. Online information and support needs of women with advanced breast cancer: a qualitative analysis. Support Care Cancer. 2018;26(10):3489-96. doi: https://doi.org/10.1007/s00520-018-4206-1

16. Lee K, Hoti K, Hughes JD, et al. Dr Google and the consumer: a qualitative study exploring the navigational needs and online health information-seeking behaviors of consumers with chronic health conditions. J Med

Internet Res. 2014;16(12):e262. doi: https://doi.org/10.2196/jmir.3706

Associate-Editor: Fernando Lopes Tavares de Lima. Orcid iD: https://orcid.org/0000-0002-8618-7608
Scientific-Editor: Anke Bergmann. Orcid iD: https://orcid.org/0000-0002-1972-8777