



OPEN ACCESS

EDITED BY

V. E. Sathishkumar,
Jeonbuk National University, Republic of Korea

REVIEWED BY

Dhavalkumar Patel,
IBM Research, United States

*CORRESPONDENCE

Rowalt Alibudbud
✉ rowalt.alibudbud@dlsu.edu.ph

RECEIVED 02 April 2023

ACCEPTED 22 June 2023

PUBLISHED 04 July 2023

CITATION

Alibudbud R (2023) Wikipedia page views for health research: a review.
Front. Big Data 6:1199060.
doi: 10.3389/fdata.2023.1199060

COPYRIGHT

© 2023 Alibudbud. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Wikipedia page views for health research: a review

Rowalt Alibudbud*

Department of Sociology and Behavioral Sciences, De La Salle University, Manila, Philippines

Wikipedia is an open-source online encyclopedia and one of the most-read sources of online health information. Likewise, Wikipedia page views have also been analyzed to inform public health services and policies. The present review analyzed 29 studies utilizing Wikipedia page views for health research. Most reviewed studies were published in recent years and emanated from high-income countries. Together with Wikipedia page views, most studies also used data from other internet sources, such as Google, Twitter, YouTube, and Reddit. The reviewed studies also explored various non-communicable diseases, infectious diseases, and health interventions to describe changes in the utilization of online health information from Wikipedia, to examine the effect of public events on public interest and information usage about health-related Wikipedia pages, to estimate and predict the incidence and prevalence of diseases, to predict data from other internet data sources, to evaluate the effectiveness of health education activities, and to explore the evolution of a health topic. Given some of the limitations in replicating some of the reviewed studies, future research can specify the specific Wikipedia page or pages analyzed, the language of the Wikipedia pages examined, dates of data collection, dates explored, type of data, and whether page views were limited to Internet users and whether web crawlers and redirects to the Wikipedia page were included. Future research can also explore public interest in other commonly read health topics available in Wikipedia, develop Wikipedia-based models that can be used to predict disease incidence and improve Wikipedia-based health education activities.

KEYWORDS

Wikipedia, health research, research methodology, health informatics, internet based intervention, infodemiology, online information

1. Introduction

Wikipedia is an open-source online encyclopedia available in over 275 languages and has more than 32 million articles across various topics, including health and medicine (Heilman and West, 2015). Since its inception in 2001, it has been an influential public health platform and one of the most commonly read sources of online health information (Shafee et al., 2017). However, despite its popularity and broad content, Wikipedia faces many challenges, including its small and decreasing core editors and the academic world's skepticism (Heilman and West, 2015; Jemielniak, 2019).

Despite criticisms and skepticism from academics, a review of health-related articles on Wikipedia revealed that they commonly referenced several respected journals, such as The Lancet, The New England Journal of Medicine, Nature, Journal of the American Medical Association, and Science, were the most commonly cited sources (Heilman and West, 2015). Additionally, approximately half of its core editors are healthcare providers (Heilman and West, 2015). Furthermore, Wikipedia content often attains high rankings in Google search results (Smith, 2020; Mendes et al., 2021). Further examination of web traffic patterns has indicated that Wikipedia surpasses institutional health websites in terms of its online

presence. Notably, it garners higher web traffic compared to esteemed platforms such as the National Institutes of Health, WebMD, Mayo Clinic, National Health Service, and the World Health Organization (Heilman and West, 2015). Likewise, a scoping review reveals that 50–70% of physicians and over 90% of medical students used it as a source of health information (Shafee et al., 2017). Thus, there is high utilization of Wikipedia for health information despite criticisms. Consequently, Wikipedia page views, similar to other big data from the Internet (e.g., Google search volume), have been analyzed to inform public health services and policies (Alibudbud and Cleofas, 2022; Alibudbud, 2023b).

Eysenbach (2009) defined infodemiology as “the science of distribution and determinants of information in an electronic medium, specifically the Internet, or in a population, with the ultimate aim to inform public health and public policy.” Therefore, using Wikipedia page views, which are information from the Internet, to inform public health services and policies can be subsumed under infodemiology.

1.1. Objectives and significance

This review explored the use of Wikipedia page views for health research. Publications were summarized and described according to their year of publication, authors’ country of origin, health topic, purpose, data analysis, and other utilized big data from the Internet. By doing so, it summarizes the current state of the art and informs future research of the extent and considerations in using Wikipedia page views data.

2. Methods

This review included studies utilizing Wikipedia page views for health research from PubMed, one of the world’s largest health research databases, and Scopus, one of the world’s largest research databases. Specifically, research publications in the English language utilizing Wikipedia page views for health-related topics until March 2023 were included in this review. Letters, abstracts, those primarily about non-health-related topics (e.g., conservation), and those not written in English were excluded. The keyword used to search for relevant publications were “Wikipedia” and “page” and “views”.

Figure 1 shows that 166 and 33 publications were collected from Scopus and PubMed after searching for titles, abstracts, and keywords. After collecting the articles from each database, 26 duplicate articles were removed. Then, each publication was screened for eligibility based using its abstract and title, including being written in English, the use of Wikipedia page views, and topics related to health. After excluding 142 during the eligibility screening, studies were sought and mainly assessed based on their year of publication, authors’ country of origin, health topic, purpose, data analysis, and whether they utilized other sources of internet big data. Upon further assessment, two additional publications were removed since they did not use Wikipedia page views. Thus, a total of 29 studies were included in this review.

3. Results and discussion

3.1. Publication years and countries of origin

The most productive year was 2021 (see Supplementary material), with eight publications ($n = 8$, 27.59%), followed by 2020 ($n = 5$, 17.24%) and 2022 ($n = 4$, 13.79%). The most productive country was Italy, with 12 publications (41.38%), followed by the United States ($n = 10$, 34.48%) and the United Kingdom ($n = 4$, 13.79%). Thus, most of the reviewed studies have been published in recent years. Likewise, a disparity between high-income and low- and middle-income countries was found, where most of the reviewed studies emanated from high-income countries (e.g., Italy and the United States) than low- and middle-income countries (LMICs) (e.g., Philippines and Nigeria). Therefore, future research using Wikipedia page views can be undertaken in LMICs, especially about diseases more prevalent in these countries than in high-income nations (e.g., Tuberculosis in the Philippines).

3.2. Other utilized internet data

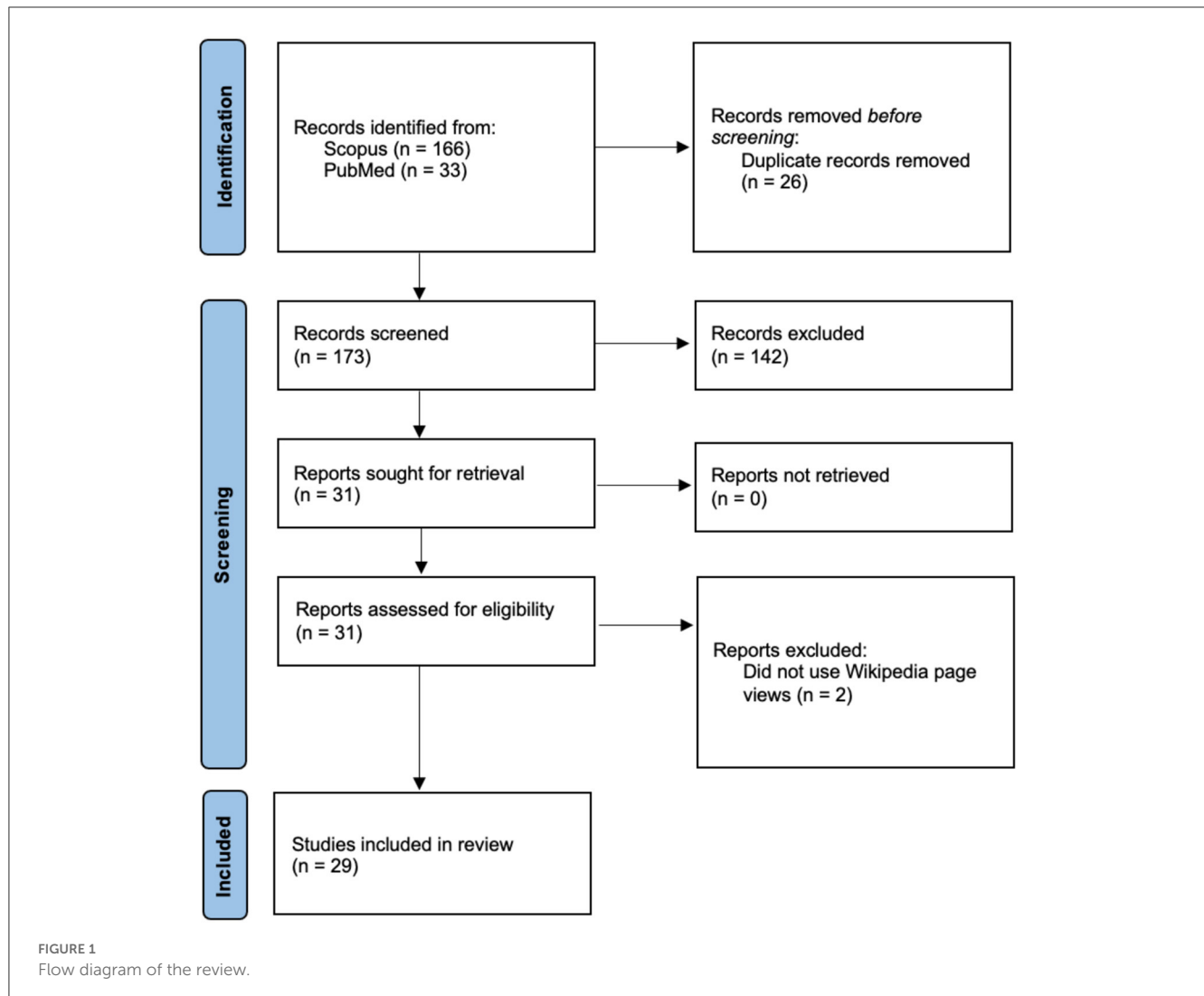
Wikipedia page views were most commonly combined with Google data (e.g., Google Trends, Google Analytics) ($n = 16$, 55.17%), followed by Twitter data (e.g., Twitter mentions) ($n = 5$, 17.24%), and Pubmed and Medline ($n = 4$, 13.79%). The reviewed studies also combined Wikipedia page views with other internet data, including online news, Youtube, Reddit, and Wikipedia Edit data (Sciascia and Radin, 2017; Gozzi et al., 2020; Szmuda et al., 2020; Wang and Zhang, 2020).

In general, while Wikipedia is one of the most highly utilized sources of health information, internet users may also explore other health websites for their needed information (Heilman and West, 2015). To address this limitation, most reviewed studies utilized these other internet data sources to expand their coverage and understand the patterns of online information utilization.

3.3. Health topics explored using Wikipedia page views

Wikipedia page views have also been used to understand various health topics. The most common topic explored using Wikipedia page views by the reviewed studies were non-communicable diseases ($n = 11$, 37.93%), followed by communicable diseases ($n = 7$, 24.14%), factors related to health ($n = 2$, 6.90%), medications ($n = 2$, 6.90%), and a combination of the aforementioned topics ($n = 6$, 20.69%). Two (6.90%) of the reviewed studies did not indicate the specific Wikipedia pages they explored.

The specific topics explored by the reviewed studies included Dementia (Brigo et al., 2015; Alibudbud, 2023b), fencing response (Roe et al., 2023), tumors (e.g., pancreatic tumors, brain tumors, colorectal cancer) (Naik et al., 2021; Mondia et al., 2022; Gianfredi et al., 2023), substance use disorder (Alibudbud and Cleofas, 2022),



epilepsy (Brigo et al., 2015; Okumura et al., 2016), schizophrenia (Adams et al., 2020), diabetes mellitus (Potapov et al., 2021), pain (e.g., migraine, low back pain, inflammation, sciatica) (Brigo et al., 2015; Szmuda et al., 2020; Ciaffi et al., 2021; Potapov et al., 2021), cardiovascular diseases (Potapov et al., 2021), gastrointestinal conditions (Potapov et al., 2021), dermatological agents (Potapov et al., 2021), viral infections (e.g., coronavirus, COVID-19, influenza, Chikungunya) (Laurent and Vickers, 2009; Mahroum et al., 2018; Provenzano et al., 2019; Qiu et al., 2019; Gozzi et al., 2020; O'Leary and Storey, 2020; De Toni et al., 2021; Gianfredi et al., 2021; Rutovic et al., 2021; Storey and O'Leary, 2022), autoimmune conditions (e.g., Systemic Lupus Erythematosus) (Sciascia and Radin, 2017) various medications (e.g., Abacavir, Zidovudine) (Sciascia and Radin, 2017; Apollonio et al., 2018; Darrow and Borisova, 2022), different diets (e.g., vegetarian) (Nucci et al., 2021), frostbite (Laurent and Vickers, 2009), hypothermia (Laurent and Vickers, 2009), carbon monoxide poisoning (Laurent and Vickers, 2009), hyperthermia (Laurent and Vickers, 2009), sunburn (Laurent and Vickers, 2009), insect bites (Laurent and Vickers, 2009), and women's health-related topic (e.g., discrimination) (Wang and Zhang, 2020). Therefore, the

reviewed studies have been predominantly used in understanding health information utilization for various communicable and non-communicable diseases. Hence, future research can also focus on medications and health-related factors.

3.4. Purpose of using Wikipedia page views for health research

Studies utilized Wikipedia page views mainly to determine changes in the information usage of its pages (see Table 1). This curiosity toward Wikipedia page views as a metric of online health information usage may stem from its high use compared to other leading health websites, such as the World Health Organization and the National Institutes of Health (Heilman and West, 2015). The purpose and aims of the reviewed studies were categorized based on the analysis aim categorization of Nuti et al. (2014), which includes descriptive, causal reference, and surveillance. Causal inference studies aim to evaluate a hypothesized causal relationship with Wikipedia data, including statistical analysis. An

TABLE 1 Purpose and recommended methodological considerations for studies utilizing Wikipedia page views for health research.

Purpose of Wikipedia page views for health research
<ol style="list-style-type: none"> 1. To describe the changes in the utilization of online health information from Wikipedia at the country and global levels 2. To assess the impact of public events on public interest and information usage about health-related Wikipedia pages 3. To estimate and predict the incidence and prevalence of diseases 4. To predict data from other internet data sources 5. To evaluate the effectiveness of health education initiatives 6. To explore the evolution of a health topic
Recommended checklist of methodological considerations for studies utilizing Wikipedia page views for health research
Specify:
<ol style="list-style-type: none"> 1. Precise Wikipedia page or pages of study 2. Language of the Wikipedia page 3. Dates of data collection 4. Dates explored 5. Type of data (e.g., monthly or daily) 6. Whether page views were limited to Internet users 7. Whether web crawlers and redirects to the Wikipedia page were included

example of a causal inference study is [Gianfredi et al. \(2023\)](#), which used Wikipedia data to assess the impact of a celebrity's announcement of having been diagnosed with pancreatic cancer on the trend of cancer-related research on the Internet. Descriptive studies describe the temporal or geographic trends of particular Wikipedia pages. An example of a descriptive study is [Alibudbud \(2023b\)](#), which described the worldwide utilization of online information for dementia. Finally, surveillance studies evaluated the use of Wikipedia page views to forecast or monitor real-world phenomena. An example of a surveillance study is [O'Leary and Storey \(2020\)](#), which shows a model for predicting the number of people who might become infected and die from COVID-19. Additionally, this review classified several studies as experimental studies, which are studies that measure the change in page views before and after editing Wikipedia pages. An example of an experimental study is [Weiner et al. \(2019\)](#), which enhanced Wikipedia health pages using high-quality research findings and tracked the persistence of those edits and the number of page views after the enhancement to assess the reach of this initiative.

The most common aim of the reviewed studies was descriptive ($n = 13$, 44.83%), followed by causal inference ($n = 6$, 20.69%), surveillance ($n = 6$, 20.69%), and experimental ($n = 4$, 13.79%). Specifically, the present review found that data about Wikipedia page views were used to describe the changes and patterns in the utilization of online health information from Wikipedia at the country and global levels ([Laurent and Vickers, 2009](#); [Sciascia and Radin, 2017](#); [Mahroum et al., 2018](#); [Adams et al., 2020](#); [Gozzi et al., 2020](#); [Szmuda et al., 2020](#); [Ciaffi et al., 2021](#); [Nucci et al., 2021](#); [Rutovic et al., 2021](#); [Alibudbud and Cleofas, 2022](#); [Mondia et al., 2022](#); [Alibudbud, 2023b](#); [Roe et al., 2023](#)). In addition, it has also been utilized to assess the impact of public events, such as a celebrity's announcement of a disease, the death of a celebrity, media coverage of accidents and epilepsy, on public interest and information usage about different health-related Wikipedia pages ([Brigo et al., 2015](#); [Okumura et al., 2016](#); [Naik et al., 2021](#); [Gianfredi et al., 2023](#)).

Wikipedia page views have also been used to compare and correlate with established epidemiological data and the burden of diseases with moderate to strong correlations (e.g., data from Istituto Superiore di Sanit) ([Provenzano et al., 2019, 2021](#); [Qiu et al., 2019](#); [Gianfredi et al., 2021](#)). In addition, it has been used in developing models that can be used to estimate and predict the incidence and prevalence of diseases such as influenza and coronavirus ([O'Leary and Storey, 2020](#); [De Toni et al., 2021](#)). Likewise, it has been utilized to predict data from other internet data sources, such as the sentiment of tweets ([Storey and O'Leary, 2022](#)). Thus, the reviewed studies support that models using Wikipedia page views, similar to other sources of internet big data (e.g., Google Trends) ([Alibudbud, 2023a](#)), can be developed to forecast outbreaks of various health conditions.

Wikipedia has also been used to evaluate the effectiveness of institutional and school-based health education initiatives (e.g., Cochrane Russia Initiative) ([Adams et al., 2020](#); [Potapov et al., 2021](#)). For example, the studies of [Apollonio et al. \(2018\)](#) and [Weiner et al. \(2019\)](#) showed that educational activities could be supplemented by having students edit Wikipedia pages and using their page views as activity indicators. Interestingly, the study by [Wang and Zhang \(2020\)](#) also used Wikipedia page views to explore the evolution of a particular health topic, Women's health.

Generally, the reviewed studies also showed that Wikipedia use for health-related information has changed over the years, which can persist in the future ([Mahroum et al., 2018](#); [Alibudbud and Cleofas, 2022](#); [Darrow and Borisova, 2022](#); [Alibudbud, 2023b](#)). For instance, [Alibudbud \(2023b\)](#) predicts a decreasing utilization of Wikipedia for online dementia information, while [Mahroum et al. \(2018\)](#), [Alibudbud and Cleofas \(2022\)](#), and [Darrow and Borisova \(2022\)](#) showed an increasing trend of public utilization of online information from Wikipedia for substance use disorder, drugs, and chikungunya, respectively. Therefore, the reviewed studies show that previous notions of widespread use of Wikipedia for health information may vary depending on the health topic itself. The review also supports that future research can explore other health topics and areas to fully understand the utilization of Wikipedia for health information.

3.5. Data analysis of Wikipedia page views for health research

The reviewed studies were also categorized according to their data analysis using the data analysis categorization by [Mavragani et al. \(2018\)](#) of Google Trends data. This categorization includes visualization, seasonality, correlations, forecasting, modeling, and statistical tools. Studies considered under the visualization category include those with any form of visualization (e.g., figures and screenshots). Studies categorized under seasonality included those that explored the seasonality of their respective topic. Studies that have examined correlations are included in the correlations category. These correlations may be between Wikipedia data and other web-based sources (e.g., Google Trends). Forecasting studies include those that predicted future Wikipedia page views (e.g., ARIMA). Modeling studies employed some form of modeling using Wikipedia data (e.g., Structural Equation Modeling). For

this review, the other statistical tools category includes studies, which utilized statistical tools aside from the ones in the previous categories (*t*-test and Wilcoxon sign rank test). The most common data analysis used by the reviewed studies was visualization ($n = 20$, 68.97%), followed by correlations ($n = 8$, 27.59%), modeling ($n = 7$, 24.14%), seasonality ($n = 4$, 13.79%), and forecasting ($n = 3$, 10.34%). About a quarter utilized other statistical tools ($n = 7$, 24.14%). Thus, similar to other utilized big data from the Internet used in health studies, such as Google Trends, future studies may further explore forecasting the use of Wikipedia for health information (Mavragani et al., 2018).

3.6. Recommended methodological considerations for future studies

Some of the reviewed studies may also be difficult to replicate due to some limitations in methodological information. These limitations in methodological information have also been observed in studies that use other big data on the Internet, such as Google Trends (Alibudbud, 2023a). For instance, some of the reviewed research, especially those studying a large amount of Wikipedia pages, did not mention or supplement their publication with the specific Wikipedia pages under study. Therefore, the details needed may not be enough to replicate their studies. In this regard, common methodological considerations that may enable replicability among the reviewed studies can be adapted in future studies using Wikipedia page views (Laurent and Vickers, 2009; Sciascia and Radin, 2017; Mahroum et al., 2018; Adams et al., 2020; Gozzi et al., 2020; Szmuda et al., 2020; Ciaffi et al., 2021; Nucci et al., 2021; Rutovic et al., 2021; Alibudbud and Cleofas, 2022; Mondia et al., 2022; Alibudbud, 2023b; Roe et al., 2023). As shown in Table 1, these methodological considerations can include specifying the precise Wikipedia page of study, the language of the Wikipedia page, the dates of data collection, the dates explored, the type of data (e.g., monthly or daily), and whether page views were limited to Internet users or web-crawlers and redirects to the Wikipedia page were included in the analysis.

3.7. Limitations of the present review

Although this review provided information on several uses of Wikipedia page views, its findings should be interpreted in light of its limitations. This review explored two of the world's largest research databases. Thus, future reviews can examine other databases that may contain studies about Wikipedia page views and health topics. Second, this review utilizes a limited number of keywords. Different keywords, such as "WikiTrends" and "Wiki", can also be explored in future studies. Third, this review solely considered publications that included mentions of Wikipedia in their titles, abstracts, and keywords. As a result, studies that focused on Wikipedia but only mentioned it in their maintext, such as the article by Rustagi and Patel (2020), were not considered in the review. Therefore, the limited search scope may have overlooked other studies approaching the topic

from different angles. Fourth, this review explored limited study characteristics. Future studies can explore other important study characteristics, such as the statistical analyses used in examining Wikipedia page views.

3.8. Conclusion

Wikipedia a widely read source of online health information. This review analyzed 29 studies utilizing Wikipedia page views for health research. Most of the reviewed studies have been published in recent years. Most reviewed studies also emanated from high-income countries. Alongside Wikipedia page views, these studies commonly incorporated data from Google, Twitter, YouTube, Reddit, and online news sources. The reviewed studies also predominantly explored non-communicable diseases and communicable diseases. Additionally, the utilization of Wikipedia page views in health research encompassed various purposes, including describing changes in online health information utilization, examining the impact of public events on public interest and information usage, estimating disease incidence and prevalence, predicting data from other internet sources, evaluating the effectiveness of health education initiatives, and exploring the evolution of health topics.

To address the limitations in replicating some of the reviewed studies, future studies can specify several methodological aspects, including the specific Wikipedia page(s) analyzed, the language of the Wikipedia pages examined, data collection dates, dates explored, type of data, and the inclusion web crawlers and redirects to the Wikipedia page(s). Because the pattern of Wikipedia usage varies depending on the health topic and the presence of public events, future research can look into other commonly read health topics. Future research can also develop models using Wikipedia page views that can be used to predict disease outbreaks and forecast the utilization of online health information. In addition, health education activities can be developed and explored using Wikipedia page views.

Author contributions

RA had substantial contributions to the design, drafting, revision, acquisition, interpretation, and final approval of the data and work.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of

their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Adams, C. E., Montgomery, A. A., Aburrow, T., Bloomfield, S., Briley, P. M., Carew, E., et al. (2020). Adding evidence of the effects of treatments into relevant Wikipedia pages: a randomised trial. *BMJ Open* 10, e033655. doi: 10.1136/bmjopen-2019-033655
- Alibudbud, R. (2023a). Google trends for health research: its advantages, application, methodological considerations, and limitations in Psychiatric and Mental Health Infodemiology. *Front. Big Data* 6, 1132764. doi: 10.3389/fdata.2023.1132764
- Alibudbud, R. (2023b). The worldwide utilization of online information about dementia from 2004 to 2022: an infodemiological study of Google and Wikipedia. *Issues Ment. Health Nurs.* 44, 209–217. doi: 10.1080/01612840.2023.2186697
- Alibudbud, R., and Cleofas, J. V. (2022). Global utilization of online information for substance use disorder: an infodemiological study of Google and Wikipedia from 2004 to 2022. *J. Nurs. Scholars.* 55, 665–680. doi: 10.1111/jnu.12844
- Apollonio, D. E., Broyde, K., Azzam, A., De Guia, M., Heilman, J., and Brock, T. (2018). Pharmacy students can improve access to quality medicines information by editing Wikipedia articles. *BMC Med. Educ.* 18, 265. doi: 10.1186/s12909-018-1375-z
- Brigo, F., Igwe, S. C., Nardone, R., Lochner, P., Tezzon, F., and Otte, W. M. (2015). Wikipedia and neurological disorders. *J. Clin. Neurosci.* 22, 1170–1172. doi: 10.1016/j.jocn.2015.02.006
- Ciaffi, J., Meliconi, R., Landini, M. P., Mancarella, L., Brusi, V., Faldini, C., et al. (2021). Seasonality of back pain in Italy: an infodemiology study. *Int. J. Environ. Res. Public Health* 18, 1325. doi: 10.3390/ijerph18031325
- Darrow, J. J., and Borisova, E. (2022). Communication of drug efficacy information via a popular online platform. *J. Am. Board Family Med.* 35, 833–835. doi: 10.3122/jabfm.2022.04.210539
- De Toni, G., Consonni, C., and Montresor, A. (2021). A general method for estimating the prevalence of influenza-like-symptoms with Wikipedia data. *PLoS ONE* 16, e0256858. doi: 10.1371/journal.pone.0256858
- Eysenbach, G. (2009). Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. *J. Med. Internet Res.* 11, e11. doi: 10.2196/jmir.1157
- Gianfredi, V., Nucci, D., Nardi, M., Santangelo, O. E., and Provenzano, S. (2023). Using Google trends and Wikipedia to investigate the global public's interest in the pancreatic cancer diagnosis of a celebrity. *Int. J. Environ. Res. Public Health* 20, 2106. doi: 10.3390/ijerph20032106
- Gianfredi, V., Santangelo, O. E., and Provenzano, S. (2021). Correlation between flu and Wikipedia's pages visualization. *Acta Biomed.* 92, e2021056. doi: 10.23750/abm.v92i1.9790
- Gozzi, N., Tizzani, M., Starnini, M., Ciulla, F., Paolotti, D., Panisson, A., et al. (2020). Collective response to media coverage of the COVID-19 pandemic on Reddit and Wikipedia: mixed-methods analysis. *J. Med. Internet Res.* 22, e21597. doi: 10.2196/21597
- Heilman, J. M., and West, A. G. (2015). Wikipedia and medicine: quantifying readership, editors, and the significance of natural language. *J. Med. Internet Res.* 17, e62. doi: 10.2196/jmir.4069
- Jemielniak, D. (2019). Wikipedia: why is the common knowledge resource still neglected by academics? *GigaScience* 8, giz139. doi: 10.1093/gigascience/giz139
- Laurent, M. R., and Vickers, T. J. (2009). Seeking health information online: does Wikipedia matter? *J. Am. Med. Assoc.* 301, 471–479. doi: 10.1093/jama.M3059
- Mahroum, N., Adawi, M., Sharif, K., Waknin, R., Mahagna, H., Bisharat, B., et al. (2018). Public reaction to Chikungunya outbreaks in Italy—insights from an extensive novel data streams-based structural equation modeling analysis. *PLoS ONE* 13, e0197337. doi: 10.1371/journal.pone.0197337
- Mavragani, A., Ochoa, G., and Tsararakis, K. P. (2018). Assessing the methods, tools, and statistical approaches in google trends research: systematic review. *J. Med. Internet Res.* 20, e270. doi: 10.2196/jmir.9366
- Mendes, T. B., Dawson, J., Evenstein Sigalov, S., Kleiman, N., Hird, K., Terenius, O., et al. (2021). Wikipedia in health professional schools: from an opponent to an Ally. *Med. Sci. Educator* 31, 2209–2216. doi: 10.1007/s40670-021-01408-6
- Mondia, M. W. L., Espiritu, A. I., and Jamora, R. D. G. (2022). Brain tumor infodemiology: worldwide online health-seeking behavior using Google trends and Wikipedia pageviews. *Front. Oncol.* 12, 855534. doi: 10.3389/fonc.2022.855534
- Naik, H., Johnson, M. D. D., and Johnson, M. R. (2021). Internet interest in colon cancer following the death of Chadwick Boseman: infoveillance study. *J. Med. Internet Res.* 23, e27052. doi: 10.2196/27052
- Nucci, D., Santangelo, O. E., Nardi, M., Provenzano, S., and Gianfredi, V. (2021). Wikipedia, Google trends and diet: assessment of temporal trends in the internet users' searches in Italy before and during COVID-19 pandemic. *Nutrients* 13, 3683. doi: 10.3390/nu13113683
- Nuti, S. V., Wayda, B., Ranasinghe, I., Wang, S., Dreyer, R. P., Chen, S. I., et al. (2014). The use of Google trends in health care research: a systematic review. *PLoS ONE* 9, e109583. doi: 10.1371/journal.pone.0109583
- Okumura, A., Abe, S., Kurahashi, H., Takasu, M., Ikeno, M., Nakazawa, M., et al. (2016). Worsening of attitudes toward epilepsy following less influential media coverage of epilepsy-related car accidents: an infodemiological approach. *Epilepsy Behav.* 64, 206–211. doi: 10.1016/j.yebeh.2016.09.026
- O'Leary, D. E., and Storey, V. C. (2020). A Google–Wikipedia–Twitter model as a leading indicator of the numbers of coronavirus deaths. *Intelligent Syst. Account. Finance Manage.* 27, 151–158. doi: 10.1002/isaf.1482
- Potapov, A. S., Alexandrova, E. G., Yudina, E. V., and Ziganshina, L. E. (2021). Improving the Russian-language Wikipedia articles on medicines using new knowledge Cochrane. *Kazan Med. J.* 102, 459–473. doi: 10.17816/KMJ2021-459
- Provenzano, S., Gianfredi, V., and Santangelo, O. E. (2021). Insight the data: Wikipedia's researches and real cases of arboviruses in Italy. *Public Health* 192, 21–29. doi: 10.1016/j.puhe.2020.12.010
- Provenzano, S., Santangelo, O. E., Giordano, D., Alagna, E., Piazza, D., Genovese, D., et al. (2019). Predicting disease outbreaks: evaluating measles infection with Wikipedia trends. *Recenti Prog. Med.* 110, 292–296. doi: 10.1701/3182.31610
- Qiu, R., Hadzikadic, M., Yu, S., and Yao, L. (2019). Estimating disease burden using Internet data. *Health Inform. J.* 25, 1863–1877. doi: 10.1177/1460458218810743
- Roe, K. L., Giordano, K. R., Ezzell, G. A., and Lifshitz, J. (2023). Public awareness of the fencing response as an indicator of traumatic brain injury: quantitative study of Twitter and Wikipedia data. *JMIR Format. Res.* 7, e39061. doi: 10.2196/39061
- Rustagi, S., and Patel, D. (2020). “DiNer-on building multilingual disease-news profiler,” in *Transactions on Large-Scale Data-and Knowledge-Centered Systems XLIII* (Berlin, Heidelberg: Springer), 114–137. doi: 10.1007/978-3-662-62199-8_5
- Rutovic, S., Fumagalli, A. I., Lutsenko, I., and Corea, F. (2021). Public interest in neurological diseases on Wikipedia during coronavirus disease (COVID-19) pandemic. *Neurol. Int.* 13, 59–63. doi: 10.3390/neurolint13010006
- Sciascia, S., and Radin, M. (2017). What can Google and Wikipedia can tell us about a disease? Big Data trends analysis in Systemic Lupus Erythematosus. *Int. J. Med. Inform.* 107, 65–69. doi: 10.1016/j.ijmedinf.2017.09.002
- Shafee, T., Masukume, G., Kipersztok, L., Das, D., Häggström, M., and Heilman, J. (2017). Evolution of Wikipedia's medical content: past, present and future. *J. Epidemiol. Commun. Health* 71, 1122–1129. doi: 10.1136/jech-2016-208601
- Smith, D. A. (2020). Situating Wikipedia as a health information resource in various contexts: a scoping review. *PLoS ONE* 15, e0228786. doi: 10.1371/journal.pone.0228786
- Storey, V. C., and O'Leary, D. E. (2022). Text analysis of evolving emotions and sentiments in COVID-19 Twitter communication. *Cognit. Comput.* 2022, 1–24. doi: 10.1007/s12559-022-10025-3
- Szmuda, T., Ali, S., Czyz, M., and Sloniewski, P. (2020). Sciatica: internet search trends. *Eur. J. Transl. Clin. Med.* 3, 49–52. doi: 10.31373/ejtc/119130
- Wang, Y., and Zhang, J. (2020). Investigation of women's health on Wikipedia—a temporal analysis of women's health topic. *Informatics* 7, 22. doi: 10.3390/informatics7030022
- Weiner, S. S., Horbacewicz, J., Rasberry, L., and Bensing-Brody, Y. (2019). Improving the quality of consumer health information on Wikipedia: case series. *J. Med. Internet Res.* 21, e12450. doi: 10.2196/12450

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fdata.2023.1199060/full#supplementary-material>