

Analyzing Chinese interest in osteoporosis-related diseases by using the Baidu index

Jianjun Wu¹, Yanping Lin², Jingyi Chen², Jiachun Huang², Hongxing Huang^{2*}

1. The Third School of Clinical Medicine, Guangzhou University of Chinese Medicine, Guangzhou, Guangdong Province, China.
2. The Third Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangzhou, Guangdong Province, China.

*The corresponding author: Hongxing Huang
hhx@gzucm.edu.cn

Abstract

Purpose: Understanding the public's concern about osteoporosis in China is crucial for guiding public health campaigns and educational efforts, as it is significant public health concern in country. Using the Baidu index, this study aims to examine the Chinese fascination with diseases associated with osteoporosis.

Methods: The research gathered information on the search frequency of osteoporosis-related terms on Baidu between January 2012 and December 2022. To gauge the popularity of subjects associated with osteoporosis, the Baidu index was utilized. The study analyzed the trends and patterns of search volume and identified the most popular topics related to osteoporosis.

Results: The study found an increasing trend in interest in osteoporosis-related diseases in China over the past ten years. The search volume for osteoporosis-related keywords increased steadily from 2012 to 2022. The most popular diseases related to osteoporosis were hyperosteogeny, fracture, osteoarthritis, and metabolic diseases.

Conclusions: The study offers valuable information about the level of understanding and awareness among the general population in China regarding diseases associated with osteoporosis. The growing fascination with subjects related to osteoporosis indicates that the general population is developing a greater understanding of the significance of maintaining bone health. The results of the study can provide valuable guidance for public health campaigns and educational efforts focused on enhancing the prevention and management of osteoporosis-related conditions in China.

Keywords: osteoporosis; Chinese interest; osteoporosis-related diseases; Baidu index

Osteoporosis (OP) is a condition affecting the entire skeletal system, marked by reduced bone density, deterioration of bone structure, heightened susceptibility to fractures⁽¹⁾. OP can occur at any age but is more common in postmenopausal women and the elderly⁽²⁾. According to the initial epidemiological investigation on OP in China, it has been observed that individuals aged 50 and above in China are facing a significant health concern due to osteoporosis. OP affects approximately 19.2% of individuals aged 50 and above, particularly middle-aged and elderly women who have a prevalence rate of 32.1%⁽³⁾. The prevalence of OP among individuals in their middle and old age is concerning as the aging process intensifies⁽⁴⁾.

The rise of search engines, specifically, has greatly transformed individuals' lives, enhancing the Internet's value as a resource for daily activities, education, and career opportunities. According to the 51st Statistical Report on Internet Development in China, as of December 2022, the number of Internet users in China reached approximately 1.067 billion, resulting in an Internet penetration rate of 75.6%⁽⁵⁾. It is estimated that about 81.3% of Internet users use search engines. By using this service, a total of 77.3% of people can obtain the information they require. Baidu Search, which was utilized by 90.9% of users^[6], emerged as the top choice among search engines in China.

The Baidu Index is a platform that allows the sharing of data derived from Baidu's extensive records of Internet user behavior. It enables the analysis of search trends for keywords, provides valuable insights into shifts in user demands, tracks trends in public opinion across media, and identifies digital consumer traits. The industry perspective allows for the analysis of market characteristics as well. Currently, researchers have utilized Baidu index big data to examine health data, encompassing diverse subjects such as assessing the demands for nursing services and the primary factors that affect them in urban China⁽⁶⁾, as well as evaluating the spatiotemporal trends of public panic levels in China amidst the pandemic crisis⁽⁷⁾. By examining these online search patterns, one can observe the search habits and preferences of individuals on the internet, which can provide insights into the health information sought by the public. No research has been conducted on OP using Baidu Index.

In order to understand the characteristics of popular interest in OP, this research utilizes Baidu index data platform to observe online search behavior through Internet search data. It also investigates the significance of Internet search data in tracking trends in media attention. Through this system, the public can better understand OP, prevent, and treat OP, and complement traditional OP monitoring systems.

Methods

Data source

The data was acquired using Baidu Index. This data sharing platform uses data from the Baidu search engine to share user behavior information. One can study keyword search trends, gain insight into shifts in user requirements, monitor trends

in public opinion in the media, and analyze the characteristics of digital consumers to understand their positioning. Additionally, market characteristics can be examined from an industry standpoint.

The study employed the subsequent data: (1) Search index: The scientific evaluation and calculation of the weighted search frequency of every keyword in the Baidu online search were conducted using keywords as the statistical subject, based on the search volume data of internet users on Baidu. (2) Information index: news connected to specific keywords that are indexed by Baidu News and published by major online media. (3) Yearly search rate on the Internet: search index / total number of netizens (from the total number of annual users in the current year of China's Internet Development Statistical Report).

The search term 'osteoporosis' was queried on the Baidu Index platform's demand map. The most popular OP-related terms were chosen, including "hyperosteogeny," "fracture," "osteoarthritis," and "metabolic diseases." excluding words that refer to diseases that are not related to osteoporosis. Collected the information index for each keyword spanning from 2018 to 2022, along with the search index for every keyword ranging from 2012 to 2022.

Statistical methods

We combined the keyword indexes searched for each year to create a single yearly search index. This index was then utilized to examine the characteristics and trends of media and public interest in OP. A one-way ANOVA was conducted for each keyword to compare the yearly, quarterly, and annual Information indexes. Additionally, we explored the correlation between the index and the year using the Pearson correlation method. The statistical difference in the correlation line angle between the groups was determined by conducting analysis of covariance after creating a scatterplot and calculating the correlation coefficient for the users' search rate each year. $P < 0.05$ (two tailed) defines statistical significance. The figures were created using Apple numbers (Apple, Inc., Cupertino, CA, USA) and GraphPad Prism version 9.5.1 (GraphPad software, Inc., Boston, Ma, USA), and all statistical analyses were performed using GraphPad Prism.

Results

Changes to search index

From 2012 to 2019, the annual search index has consistently increased for every keyword, indicating a positive correlation with the corresponding year (Pearson's correlation coefficient 0.7942; $P = 0.0186$). Despite a sharp decline in the annual search index amount for each phrase in 2020 (with an increment rate of -34.18%), the overall annual search index for each keyword has been steadily increasing over time. It has reached a significantly higher level in 2022 compared to 2019 (with

an increment rate of 47.14%).

The total annual search index for each keyword has exhibited a steady increase over the past ten years (2012-2022). In Figure 1, there was a strong negative correlation (Pearson's correlation = -0.8475, $P = 0.001$) between fracture and the year, whereas the year showed a significant positive correlation (Pearson's correlation = 0.9589, $P = 0.0001$) with osteoarthritis. There was no correlation between hyperosteogeny and metabolic diseases (hyperosteogeny: Pearson's correlation = 0.3862, $P = 0.2408$; metabolic diseases: Pearson's correlation = 0.4015, $P = 0.2210$). The hyperosteogeny search index experienced the largest growth, with an increase of 1,551,543, followed by osteoarthritis with a growth of 119,419. Hyperosteogeny had the highest growth rate (211%) among the search index, with osteoarthritis following closely behind (97%). However, the fracture search index exhibited a decrease of -42,290, accounting for a decline of 10%. Table 1 displays the progression of every search index over a span of ten years, including the correlation between each alteration and the corresponding years.

The Fisher's least significant difference method found a discrepancy in statistics between hyperosteogeny and fracture, hyperosteogeny and osteoarthritis search indexes, hyperosteogeny and metabolic diseases ($P < 0.0001$), and between fracture and metabolic diseases ($P = 0.0030$). However, fracture and osteoarthritis did not statistically differ from one another ($P = 0.0833$), osteoarthritis and metabolic diseases ($P = 0.5690$). In comparison to the other three keywords, hyperosteogeny had the greatest search index during the past ten years, followed by fracture, osteoarthritis and metabolic diseases. Table 2 illustrates the specific results.

According to Figure 2, over the last ten years, apart from the fracture search index, there has been a rise in the yearly search rate of internet users, with all linear regression slopes exceeding 0. Nevertheless, the covariance analyses indicated that there was no statistically significant difference in the slope of the linear regression between the groups. ($F = 1.598$, $P = 0.2067$).

Changes to information index

From 2018 to 2022, the information index exhibited a comparable pattern to that of the search index. Table 3 displays the fluctuation in the information index of each keyword over a period of five years. When compared to osteoarthritis and metabolic illnesses ($P = 0.0458$), the information index of fracture showed significant differences from those conditions ($P = 0.0343$) (See Supplemental Table 1).

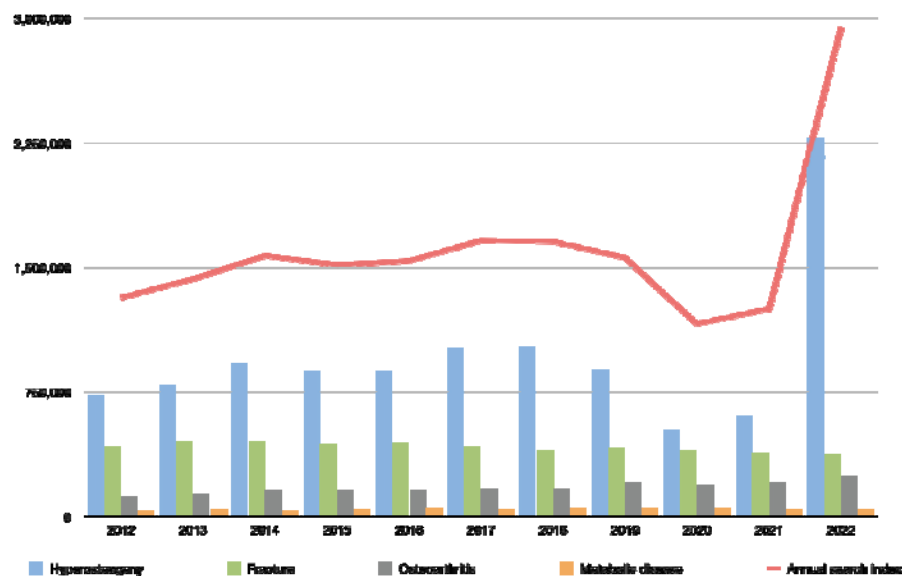


Figure 1. Annual changes to search index.

| Keywords | Hyperosteoegeny | Fracture | Osteoarthritis | Metabolic diseases | Annual search index |
|--------------------------------|-----------------|----------|----------------|--------------------|---------------------|
| 2012 | 733,285 | 422,670 | 123,005 | 39,785 | 1,318,745 |
| 2013 | 798,616 | 453,544 | 139,412 | 41,132 | 1,432,704 |
| 2014 | 923,832 | 452,452 | 155,792 | 39,312 | 1,571,388 |
| 2015 | 882,700 | 436,800 | 155,064 | 40,768 | 1,515,332 |
| 2016 | 880,015 | 448,585 | 161,252 | 50,005 | 1,539,857 |
| 2017 | 1,020,656 | 425,880 | 171,080 | 46,592 | 1,664,208 |
| 2018 | 1,026,844 | 402,948 | 170,716 | 56,784 | 1,657,292 |
| 2019 | 884,884 | 414,960 | 205,660 | 53,508 | 1,559,012 |
| 2020 | 524,505 | 395,660 | 192,355 | 49,275 | 1,161,795 |
| 2021 | 612,612 | 386,204 | 210,756 | 42,952 | 1,252,524 |
| 2022 | 2,284,828 | 380,380 | 242,424 | 41,860 | 2,949,492 |
| Increment | 1,551,543 | -42,290 | 119,419 | 2,075 | 1,630,747 |
| Increment rate (%) | 211 | -10 | 97 | 5 | 123 |
| Correlation coefficient | 0.3862 | -0.8475 | 0.9589 | 0.4015 | 0.4063 |
| P | 0.2408 | <0.001 | <0.0001 | 0.2210 | 0.2149 |

Table 1. Analyses of the search index's fundamental status and correlation for the period from 2012 to 2022.

| Keywords | | P | 95% CI | |
|----------------|-------------------|---------|--------|---------|
| Hyperosteogeny | Fracture | <0.0001 | 273739 | 808569 |
| | Osteoarthritis | <0.0001 | 518517 | 1053348 |
| | Metabolic disease | <0.0001 | 648112 | 1182943 |
| Fracture | Osteoarthritis | 0.0833 | -22637 | 512194 |
| | Metabolic disease | 0.0030 | 106958 | 641789 |
| Osteoarthritis | Metabolic disease | 0.5690 | 137821 | 397010 |

Table 2. Comparisons from search index keywords.

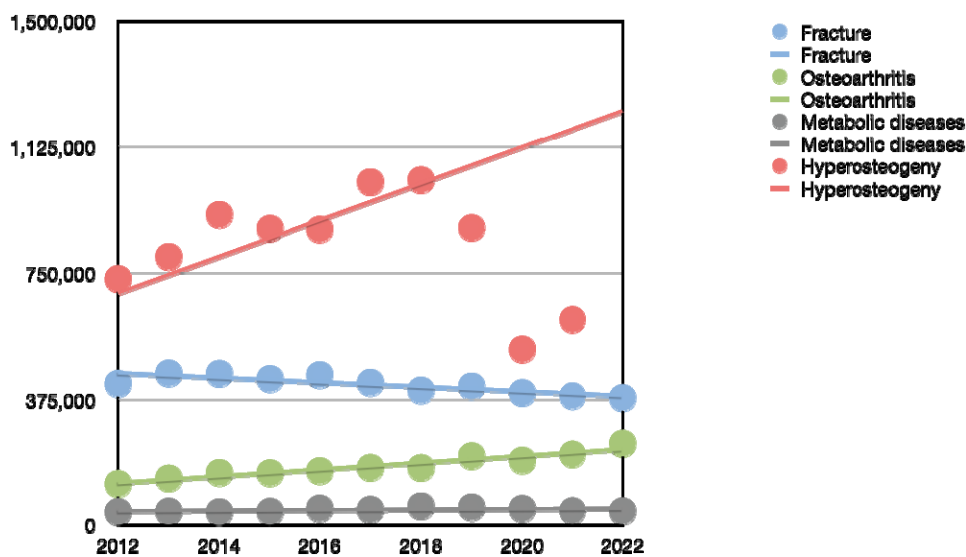


Figure 2. Annual netizen search rate scatter plot.

| Keywords | 2018 | 2019 | 2020 | 2021 | 2022 |
|--------------------|---------|---------|---------|--------|---------|
| Hyperosteogeny | 17,391 | 76,879 | 70,876 | 41,419 | 43,966 |
| Fracture | 261,952 | 208,384 | 113,447 | 93,617 | 493,838 |
| Osteoarthritis | 7803 | 18,361 | 18,540 | 11,291 | 33,260 |
| Metabolic diseases | 0 | 1 | 0 | 0 | 60 |

Table 3. Basic information index situation from 2018 to 2022.

Discussion

During the past decade, the study discovered an increase in public awareness regarding OP; however, the level of attention towards specific diseases varied. Hyperosteogeny and fracture garnered considerably more attention when compared to osteoarthritis and metabolic diseases. Furthermore, the search index for hyperosteogeny achieved a superior ranking compared to the remaining three keywords,

surpassing the second place by more than twice the magnitude. This may be connected to the recent rise in the occurrence of hyperosteogeny. Hyperosteogeny and osteoporosis have a strong connection, as the depletion of calcium in bones can result in osteoporosis, while the irregular accumulation of calcium in bones can lead to hyperosteogeny⁽⁸⁻¹⁰⁾.

OP is a condition that tends to be hidden, with many individuals not experiencing any symptoms in the beginning stages⁽¹¹⁾. However, when discomfort arises, patients often take the initiative to gather relevant information through personal research^(12, 13). The fact that Baidu is China's most widely used search engine means that its search results can fairly reflect the needs of its users. The number of confirmed COVID-19 cases in China can be monitored using the illness prediction tool developed collaboratively by Baidu and the Center for Systems Science and Engineering at Johns Hopkins University⁽¹⁴⁾. In addition to conventional detection methods, it can also be utilized to predict the outbreak pattern of diseases⁽¹⁵⁾. The utilization of the Baidu index in OP-related research has not been fully exploited in China. Our research, which is the initial investigation into the behavior and interest of Chinese internet users in OP, validates the potential of utilizing internet search data to depict the real situation of OP patients in China.

Regarding the information index, the findings revealed that fracture received more attention from the media than other keywords. This demonstrates how, during the last five years, the public now has access to more fracture-related information via the media. Overall, the OP was receiving more media coverage. The explanation may be connected to the Internet's quick development, the population's aging trend, increased hygiene awareness, the emergence of "Internet Plus Healthcare," and the rise of OP-related media.

Despite the fact that the population affected by OP incurred significant healthcare costs in China⁽¹⁶⁾, there has consistently been a dearth of information regarding the outbreak of OP in the country. It is extremely challenging to grasp the true prevalence of OP and understand the requirements or characteristics of OP patients. In recent years, there has been research on web-based digital disease monitoring methods due to the rising popularity of Internet usage and the public's increasing dependence on Internet search engines for obtaining health information. Search engines can disclose sensitive information about diseases before disclosing a diagnosis as a query tool, helping to improve disease control. Big data from the Internet has many potential uses in the medical industry and might be used to enhance and expand on present clinical and observational data⁽¹⁷⁻¹⁹⁾. Given the quick growth of Internet access and engines of search, network analysis of data might prove considered as a supplement to today's more traditional illness monitoring techniques.

Limitations

There are a few other restrictions on this study. First of all, we only pay attention to OP attention from users of the Baidu search engine; We neglected to take into

account OP attention from users of social media or users of other search engines. This may only represent the concern of the public. Second, the Baidu index could be influenced by sampling bias. Despite the significant increase in China's Internet penetration rate, there remains a noticeable preference towards socioeconomically and educationally privileged groups among netizens. Thirdly, the study's China-specific focus, the data had to include a small number of foreigners. Furthermore, assessing the efficiency and reliability of the Baidu Index proves challenging due to the undisclosed details of the method's particular criteria, despite the implementation of a weighted filtering approach for processing all Baidu Index data. Fourth, due to privacy protection law, Information index data was not made public before 2018. To ensure data variation, future studies should examine data obtained from different engines of search and media. To enhance the objectivity of the research findings, it is imperative to meticulously analyze the data, thereby mitigating the influence of confounding variables.

Conclusions

From 2012 to 2022, there was a rising trend in the rate of online searches for OP. However, amidst the COVID-19 pandemic, both the search rate and the information index exhibited a decline. By utilizing the Baidu index, individuals can monitor the online behaviors and areas of interest of Chinese Internet users. Improving our understanding of disease prevalence, patient education, and the availability of online resources could be beneficial. Tracking people's requests for information about OP using data on searching on the Internet trends can be valuable. Furthermore, apart from serving as a scientific basis for the prevention and management of OP in China, it can also be employed as a research method to gain further insights into the characteristics and preferences of the patients.

Between 2012 and 2022, the online search rate for OP showed an increasing trend; however, during COVID-19, it decreased, as well as the information index showed a comparable pattern. Using the Baidu index, one may track Chinese Internet users' online activities and OP interest. This could enhance our knowledge of illness incidence rates, education of patients, and the accessibility of resources on the Internet. Tracking people's requests for information about OP using data on searching on the Internet trends can be valuable. In addition to providing a scientific foundation for the prevention and treatment of OP in China, it can be utilized as a research technique to find out more about the traits and interests of the patients.

Abbreviations

OP osteoporosis

Declarations

Ethics approval and consent to participate

The Baidu Index belong to public databases. Users can download relevant data for free for research and publish relevant articles. Our study is based on open-source data, so there are no ethical issues and other conflicts of interest.

Consent for publication

Not applicable

Availability of data and materials

The Baidu Index Data Platform provides the information used in this study. (<https://index.baidu.com/v2/main/index.html#/help>).

Competing interests

The authors declare no competing interests.

Funding

This study did not receive any funding.

Authors' contributions

All authors have contributed to the study of concepts and designs. Manuscript writing and data analysis by J. - J.W., Y. - P.L. performed the word revision and formatting adjustments. The remaining authors carried out the paper revision and grammar editing work. Y. - P.L., J. -Y.C., and J. - C.H. provided technical and thesis writing guidance support throughout the process. Corresponding author H. - X.H. reviewed the article and is responsible for communicating with the editor. All authors read and approved the final manuscript.

Acknowledgements

We appreciated the Baidu Index Platform for providing the data for the manuscript.

References

1. Compston JE, McClung MR, Leslie WD. Osteoporosis. *Lancet*. 2019;393(10169):364-76.
2. Lane NE. Epidemiology, etiology, and diagnosis of osteoporosis. *Am J Obstet Gynecol*. 2006;194(2 Suppl):S3-11.
3. National Health Commission of the People's Republic of China. Epidemiological Survey of Osteoporosis in China and Release of the Results of the "Healthy Skeletons" Special Action https://ncncd.chinacdc.cn/zxd/201810/t20181020_195265.htm (2018) [

4. Wang L, Yu W, Yin X, Cui L, Tang S, Jiang N, et al. Prevalence of Osteoporosis and Fracture in China: The China Osteoporosis Prevalence Study. *JAMA Netw Open*. 2021;4(8):e2121106.
5. China Internet Network Information Center(CNNIC). The 51th China Statistical Report on Internet Development. <https://www3.cnnic.cn/n4/2023/0303/c88-10757.html> [
6. Bian J, He M, Luo J, Gao Q, Zhang D. Spatio-temporal Analysis of Demand for Elderly Care Services and Its Key Influencing Factors in Chinese Cities: A Web Search Engine Approach. *Inquiry*. 2023;60:469580231159751.
7. Chen Y, Liu Y, Yan Y. Revealing the spatiotemporal characteristics of the general public's panic levels during the pandemic crisis in China. *Trans GIS*. 2022.
8. Armas LA, Recker RR. Pathophysiology of osteoporosis: new mechanistic insights. *Endocrinol Metab Clin North Am*. 2012;41(3):475-86.
9. Torres L, Klingberg E, Nurkkala M, Carlsten H, Forsblad-d'Elia H. Hepatocyte growth factor is a potential biomarker for osteoproliferation and osteoporosis in ankylosing spondylitis. *Osteoporos Int*. 2019;30(2):441-9.
10. Deminger A, Klingberg E, Lorentzon M, Geijer M, Göthlin J, Hedberg M, et al. Which measuring site in ankylosing spondylitis is best to detect bone loss and what predicts the decline: results from a 5-year prospective study. *Arthritis Res Ther*. 2017;19(1):273.
11. Compston JE, McClung MR, Leslie WD. Osteoporosis. *The Lancet*. 2019;393(10169):364-76.
12. Hernlund E, Svedbom A, Ivergård M, Compston J, Cooper C, Stenmark J, et al. Osteoporosis in the European Union: medical management, epidemiology and economic

burden. A report prepared in collaboration with the International Osteoporosis Foundation (IOF) and the European Federation of Pharmaceutical Industry Associations (EFPIA). Arch Osteoporos. 2013;8(1):136.

13. Freedman KB, Kaplan FS, Bilker WB, Strom BL, Lowe RA. Treatment of osteoporosis: are physicians missing an opportunity? J Bone Joint Surg Am. 2000;82(8):1063-70.

14. Tu B, Wei L, Jia Y, Qian J. Using Baidu search values to monitor and predict the confirmed cases of COVID-19 in China: - evidence from Baidu index. BMC Infect Dis. 2021;21(1):98.

15. Liu K, Wang T, Yang Z, Huang X, Milinovich GJ, Lu Y, et al. Using Baidu Search Index to Predict Dengue Outbreak in China. Sci Rep. 2016;6:38040.

16. Si L, Winzenberg TM, Jiang Q, Chen M, Palmer AJ. Projection of osteoporosis-related fractures and costs in China: 2010-2050. Osteoporos Int. 2015;26(7):1929-37.

17. Saunte DML, Jemec GBE. Hidradenitis Suppurativa: Advances in Diagnosis and Treatment. Jama. 2017;318(20):2019-32.

18. Dai S, Han L. Influenza surveillance with Baidu index and attention-based long short-term memory model. PLoS One. 2023;18(1):e0280834.

19. Mendonça MD, Caetano A, Viana-Baptista M. "Dr Google" will see you now - time trends in online searches on headache. Cephalalgia. 2018;38(2):407-8.



