Detecting epidemiological relevance among adenoid hypertrophy, rhinosinusitis, and allergic rhinitis through Internet search

Yingchao Yang¹, Xinyi Li¹, Qiang Ma¹, Zhihui Fu¹, Kaiming Su¹*

¹Department of Otorhinolaryngology – Head & Neck Surgery, the Sixth People's Hospital affiliated to Shanghai Jiaotong University; Shanghai Key Laboratory of Sleep Disordered Breathing; Xuhui District, Shanghai, China.

Corresponding Author:

Kaiming Su, PhD
Shanghai Key Laboratory of Sleep Disordered Breathing
Yishan Road 600, 200233, Shanghai, China.
Tel: +86 18930173978
Email: 021china@sina.com

ORCID:

Yingchao Yang: 0000-0001-6493-9859
Abstract

**Purpose:** This study aimed to verify that adenoid hypertrophy (AH) and rhinosinusitis share similar epidemiologic patterns and that AH and allergic rhinitis (AR) are not related.

**Methods:** Internet search engine query data from January 2011 to December 2019 were retrieved from the Baidu index. Monthly search volume was obtained in China for the following search terms in Chinese: “adenoid hypertrophy,” “rhinosinusitis,” and “allergic rhinitis”; the data obtained were then presented as percentages. Pearson’s and Spearman’s correlation coefficients were used to detect the correlation among the search volumes of AH, rhinosinusitis, and AR. We also collected search data from the first 5 months of 2020, when segregation was implemented in China due to the coronavirus disease 2019 epidemic. Then, we compared the search data to those obtained during the same period in 2019 to detect the effects of segregation on AH and AR to varying degrees.

**Results:** Statistically significant differences were found between the search variations of AH and rhinosinusitis during 2011–2019 (R=0.643, P<0.05). However, search variations of AH and AR were negatively related (R=−0.239, P<0.05). The average monthly search volume of AH and rhinosinusitis correlated well (R=0.836, P<0.01), but no correlation was found between AH and AR. The search volume of AH and rhinosinusitis during the first 5 months in 2020 decreased, whereas that of AR increased during January–February.

**Conclusions:** AH and rhinosinusitis are epidemiologically related, whereas AH and AR are not correlated with each other.

**Keywords:** Adenoid hypertrophy, Rhinosinusitis, Allergic rhinitis, COVID-19, Baidu index, Search volume
Introduction

Adenoid hypertrophy (AH) is the most common cause of obstructive sleep apnea (OSA) in children. The pathogenesis of AH is not yet clear, but it is related to many factors, such as smoking, seasonal allergies, and respiratory infections [1]. The relationship between AH and rhinosinusitis has been confirmed in previous studies, and therapies aimed at AH are effective in the treatment of rhinosinusitis [2, 3]. However, the information on the relationship between AH and AR is not consistent. Some scholars believe that AR triggered AH [4], whereas others believe that they are irrelevant or even negatively related [5, 6].

Traditional epidemiology surveillance relies on clinical data. Currently, advanced networks provide people with access to medical knowledge, and patients can gain a basic understanding of their diseases using computers or mobile phones. Such search behaviors can be recorded and converted into search query data using an Internet search engine. Systems based on this search engine can provide real-time information about diseases and were found to have a good congruence with traditional health-care surveillance systems [7]. Studies have detected seasonal patterns of influenza epidemics, depression, restless legs symptomatology, and sleep-disordered breathing using Internet search query data [8-11]. Specific terms searched by netizens reflect their demands, and digging into Internet data could be a powerful complement to monitoring systems.

Baidu is one of the world’s largest search engines. Since Google’s withdrawal from China in 2010, Baidu became China’s main search engine. The 2018 study by Cao et al. demonstrated that the Baidu search index (BSI) of AR and its related search terms showed obvious seasonal patterns, which was consistent with epidemiological survey results. It also had a significant correlation with actual pollen concentration and outpatient visits [12]. This again supported that BSI could be an indicator of the incidence of diseases.

Meanwhile, in December 2019, some hospitals in Wuhan City, Hubei Province, discovered cases of
unexplained pneumonia in persons with a history of exposure to the South China Seafood Market, and the cases were subsequently confirmed to be infected with severe acute respiratory syndrome coronavirus 2, a strain of coronavirus that causes coronavirus disease 2019 (COVID-19). The number of newly discovered cases increased exponentially, and this rapidly spreading novel virus has become the most widespread global pandemic that humans have encountered in the past century. In response to COVID-19, the Chinese government sealed Wuhan on January 23, 2020, to April 8, 2020, and urged the whole nation to isolate at home or in the hospital. Thus, a unique model that provides the opportunity to investigate the effect of segregation on diminishing the morbidities of AH and rhinosinusitis was created.

Therefore, this study aimed to use the Internet search data provided by the Baidu index to verify that AH and rhinosinusitis shared similar seasonal patterns that were irrelevant to AR as well as to analyze the different degrees of effects of segregation on reducing the morbidities of the three diseases during the COVID-19 epidemic in China.

Methods

Data collection

Internet search engine query data on the terms “adenoid hypertrophy,” “rhinosinusitis,” and “allergic rhinitis” in Chinese from January 2011 to December 2019 were retrieved from the Baidu index. We also added the data of the first 5 months of 2020 to the timeframe of our study so that we could observe the effect of aggregation by comparing search volumes during the same period of 2019 and 2020. Since 2020 is a leap year, we excluded the data of the extra day in February to ensure that the amount of data is consistent with those of 2019.

The daily search volume in our study was defined as the average number of people searching for the term per day, that is, the average daily search volume for the term. It was obtained from the Baidu index.
and then added to the monthly search volume. Following data normalization against the total search volume, results are displayed on a scale from 0% to 100%, with individual values over time calculated by dividing each monthly volume by the yearly search volume; thus, the data were presented as percentages.

**Data analysis** The main variables examined were the monthly search volume of AH, rhinosinusitis, and AR in Chinese from January 2011 to May 2020 in percentage form. Statistical analysis was performed using SPSS Statistics version 26.0 (IBM Corp., Armonk, NY, USA). Continuous variables were tested for normality and then analyzed by non-parametric Kruskal-Wallis test, as needed. To analyze the correlation between quantitative variables, Pearson’s and Spearman’s correlation coefficients were used.

**Results**

A total of 339 monthly values, in percentage form were obtained: AH (n=113), rhinosinusitis (n=113), and AR (n=113).

The tendency charts of the three search terms from January 2011 to December 2019, before the outbreak of the COVID-19 epidemic, are presented in Fig. 1. The figure shows that two trendlines, representing AH and rhinosinusitis searches, coincided well. Furthermore, the seasonal variations of the two trendlines are statistically relevant (R=0.643, P<0.05). It also showed opposite fluctuations of AH and AR searches. The searches for AR significantly increased in August to September, whereas searches for AH reached the minimum during this period. Statistical analysis showed that searches for AH and AR were negatively related (R=−0.239, P<0.05).

The average monthly search volumes for the three terms during the 9-year timeframe are presented in Fig. 2. The highest and lowest average monthly search volumes for AH were in December and September, respectively.
The volume search for AH had one small and one big valley, which were in February and September, respectively. The searches for rhinosinusitis showed a similar trend, which was statistically correlated with AH (R=0.836, P<0.01), but the trend was more gradual from August to October. The curve of the search for AR had two peaks, April–May and August–September. As searches for AR increased from June, the searches for AH decreased. Such contrast appeared again when searches for AR decreased from September and searches for AH increased at the same time.

The first 5-month search volumes of the three terms in 2019 and 2020 are presented in Fig. 3. The search volumes for all three terms decreased during the first 5 months of 2020 compared with those in 2019. This phenomenon was abnormal because the total search volume for AH and rhinosinusitis increased annually since 2011. Although the growth of the search for AR was not as significant as those for AH and rhinosinusitis, it increased steadily. The search volumes for AH and rhinosinusitis in the first 5 months of 2020 decreased by 41.21% (P<0.001) and 54.31% (P<0.001), respectively, than the search volumes during the same period in 2019. However, the rate of decline for AR was only 28.08% (P<0.001), which was significantly lower than those of AH and rhinosinusitis. The searches for AR from January to February in 2020 were even higher than those in 2019.

**Discussion**

The present study investigated the epidemiological relevance among AH, rhinosinusitis, and AR using an Internet database. Our results showed that the search volume of AH and rhinosinusitis varied in a statistically consistent manner, and this consistency was also present during the period of segregation in China in response to COVID-19. Meanwhile, the search volumes for AH and AR did not show statistical consistency; hence, further research is need to determine their relationship.
The use of Internet data for analyzing infectious disease epidemic has become an important topic in the prevention and control of infectious diseases. Online surveillance tools based on Google Trends have been explored in recent years. As Baidu is the largest search tool in China, its search queries could be a good representation of the needs of the population, particularly in regions with a high Internet penetration rate. The BSI has been shown to significantly correlate with the incidence of human immunodeficiency virus (HIV)/acquired immunodeficiency syndrome (AIDS) [7]. Previous studies [11, 7-10, 13] based on search engine query data focused on the regularities of specific diseases and on building analysis models to predict their outbreaks. Our study analyzed three different diseases, AH, rhinosinusitis, and AR, and investigated the interrelationships among them.

AH continuously affects the occurrence and development of rhinosinusitis in several ways, including posterior nasal obstruction, hindering nasal drainage, and constantly releasing pathogenic bacteria [14, 15]. Rhinosinusitis, in turn, can lead to AH through inflammatory stimulation. A study by Ameli et al. showed that large adenoids may be associated with the absence of allergy [16], whereas some scholars believed that AR triggers AH [4]. Our study provides a new perspective on the relationships among AH, rhinosinusitis, and AR based on Internet search query data. In our study, AH and rhinosinusitis shared similar patterns, which were unrelated to AR. This contrast may indicate that AH and AR are not related. Further studies are needed to explain the interrelationships between AR and AH.

After the outbreak of COVID-19 in China, schools were shut down and classes were held online until May 2020. This could explain the decline in morbidity of AH and rhinosinusitis to some extent. Furthermore, after public activities resumed, a reexamination of some children with AH revealed that their adenoid volume was reduced. Dave et al. found that, approximately 3 weeks after the enactment of a statewide shelter-in-place order, the average number of cumulative cases decreased by approximately 53.5% [17], again demonstrating the effect of segregation
in reducing pathogen transmission. Considering the decline in morbidity during winter and summer holidays, we assumed that avoiding the aggregation of people can reduce the morbidity of AH and rhinosinusitis, but it was not very effective for AR. The morbidities of AH and AR were not related.

Another important measure was wearing a facemask. Studies have been conducted to confirm the effect of facemask. Brienen et al. concluded that extensive use of facemask could help in delaying an influenza pandemic and in reducing the reproduction number [18]. Stutt et al. reported the results of two mathematical models and concluded that the use of facemask combined with physical distancing or periods of lockdown may be a way of managing the COVID-19 pandemic and reopening economic activity [19]. Hence, the use of facemask was strictly mandated in China during the COVID-19 pandemic. Since facemask could block a variety of pathogen particles, we believe that it contributed to decreasing morbidities of AH and rhinosinusitis.

Our study has several limitations. First, the vast geographic variations and large population of China made it unrealistic to develop surveillance networks for AH, rhinosinusitis, and AR around the country. There were only supervision systems for serious infectious diseases such as HIV/AIDS and newly eruptive COVID-19 epidemic. Thus, our study lacked clinical data. Second, the search index of each term is easily influenced by continuous changes in the search behavior of individuals. Third, the increase in search behaviors may be a result of social propaganda, events, and media reports. Finally, search behaviors depend on Internet access, which is uneven throughout China, and less than half of the population in rural areas do not have access. Thus, information on epidemic in such areas could not be included in the statistics. Moreover, the users of Baidu are mainly young, which leads to age bias. However, with the popularization of cell phones and computers, regional and age biases would gradually decrease. According to the 45th China Statistical Report on Internet Development (available from China Internet Network Information Center), China had 904 million netizens, and the Internet penetration rate reached 64.5%
Meanwhile, according to the China Statistical Report on Search Engine in 2019, 81.3% of Chinese netizens used search engines to obtain information. It was estimated that 70.5% of netizens used the search engine to seek specialized information in the field of medicine and/or law [21]. Surveillance systems based on BSI could compensate for the lack of clinical data to some extent, thus providing a novel option for monitoring morbidities of diseases and for investigating the relationships among diseases.

In conclusion, AH and rhinosinusitis shared similar seasonal patterns, which were irrelevant to AR. The decrease in public aggregation was effective in reducing the morbidities of AH and rhinosinusitis but not those of AR.

**Declarations**

**Funding:** This study is supported by the National Natural Science Foundation (81570910;81974142), Shanghai Pujiang Program (18PJD036), Shanghai Municipal Commission of Science and Technology (18DZ2260200).

**Conflicts of interest/Competing interests:** The authors declare that they have no conflict of interest.

**Ethics approval:** Ethics committee approval was not required for this study because the study was based on Internet search data.

**Consent to participate:** The study was based on Internet search data, and no personal information was revealed, thereby maintaining confidentiality.

**Consent for publication:** Not applicable

**Availability of data and material:** The Baidu index is publicly available at [https://index.baidu.com](https://index.baidu.com).

**Code availability:** SPSS Statistics version 26.0 is available online [https://www.ibm.com/support/pages/downloading-ibm-spss-statistics-26](https://www.ibm.com/support/pages/downloading-ibm-spss-statistics-26).
Authors’ contributions: All authors contributed to the study conception and design. Data collection and analysis were performed by Yingchao Yang. The first draft of the manuscript was written by Yingchao Yang and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

References

1. Ersu R, Arman AR, Save D, Karadag B, Karakoc F, Berkem M, Dagli E (2004) Prevalence of snoring and symptoms of sleep-disordered breathing in primary school children in istanbul. Chest 126 (1):19-24. doi:10.1378/chest.126.1.19
2. AM B, AM S, G F, C R, C P (2019) Adenoidal Disease and Chronic Rhinosinusitis in Children-Is there a Link? Journal of clinical medicine 8 (10). doi:10.3390/jcm8101528
3. Tuncer U, Aydogan B, Soylu L, Simsek M, Akcali C, Kucukcan A (2004) Chronic rhinosinusitis and adenoid hypertrophy in children. Am J Otolaryngol 25 (1):5-10. doi:10.1016/j.amjoto.2003.10.003
4. Modrzynski M, Zawisza E (2007) The influence of birch pollination on the adenoid size in children with intermittent allergic rhinitis. Int J Pediatr Otorhinolaryngol 71 (7):1017-1023. doi:10.1016/j.ipjol.2007.02.018
5. Eren E, Arslanoğlu S, Erdem SB, Nacaroğlu T, Karkiner C, Can D, Önal K (2015) Chicken or the egg: the dilemma of allergic rhinitis versus adenoid hypertrophy. Rhinology 53 (2):154-159. doi:10.4193/Rhin14.013
6. Ameli F, Brocchetti F, Tosca MA, Schiavetti I, Ciprandi G (2014) Tonsil volume and allergic rhinitis in children. Allergy Rhinol (Providence) 5 (3):137-142. doi:10.2500/ar.2014.5.0095
7. He G, Chen Y, Chen B, Wang H, Shen L, Liu L, Suolang D, Zhang B, Ju G, Zhang L, Du S, Jiang X, Pan Y, Min Z (2018) Using the Baidu Search Index to Predict the Incidence of HIV/AIDS in China. Sci Rep 8 (1):9038. doi:10.1038/s41598-018-27413-1
8. Ingram DG, Matthews CK, Plante DT (2015) Seasonal trends in sleep-disordered breathing: evidence from Internet search engine query data. Sleep Breath 19 (1):79-84. doi:10.1007/s11325-014-0965-1
9. Ingram DG, Plante DT (2013) Seasonal trends in restless legs symptomatology: evidence from Internet search query data. Sleep Med 14 (12):1364-1368. doi:10.1016/j.sleep.2013.06.016
10. Yang A, Huang N, Peng C, Tsai S (2010) Do seasons have an influence on the incidence of depression? The use of an internet search engine query data as a proxy of human affect. PloS one 5 (10):e13728. doi:10.1371/journal.pone.0013728
11. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L (2009) Detecting influenza epidemics using search engine query data. Nature 457 (7232):1012-1014. doi:10.1038/nature07634
12. Cao YJ, Xu R, Tao XY, Shi JB, Chen FH (2019) [Investigation of temporal and spatial characteristics of allergic rhinitis epidemics by Baidu Index in China]. Zhonghua Er Bi Yan Hou Tou Jing Wai Ke Za Zhi 54 (12):888-893. doi:10.3760/cma.j.issn.1673-0860.2019.12.002
13. Li K, Liu M, Feng Y, Ning C, Ou W, Sun J, Wei W, Liang H, Shao Y (2019) Using Baidu Search Engine to Monitor AIDS Epidemics Inform for Targeted intervention of HIV/AIDS in China. Sci Rep 9 (1):320. doi:10.1038/s41598-018-35685-w
14. Lee D, Rosenfeld RM (1997) Adenoid bacteriology and sinonasal symptoms in children. Otolaryngol Head Neck Surg 116 (3):301-307. doi:10.1016/s0194-59989770264-x
15. Belcher R, Virgin F (2019) The Role of the Adenoids in Pediatric Chronic Rhinosinusitis. Med Sci (Basel) 7 (2). doi:10.3390/medsci7020035
16. Ameli F, Brocchetti F, Tosca M, Signori A, Ciprandi G (2013) Adenoidal hypertrophy and allergic rhinitis: is there an inverse relationship? American journal of rhinology & allergy 27 (1):e5-10. doi:10.2500/ajra.2013.27.3854
17. Dave D, Friedson A, Matsuzawa K, Sabia J (2020) When Do Shelter-in-Place Orders Fight COVID-19 Best? Policy Heterogeneity Across States and Adoption Time. Economic inquiry. doi:10.1111/ecin.12944
18. Brienen NC, Timen A, Wallinga J, van Steenbergen JE, Teunis PF (2010) The effect of mask use on the spread of influenza during a pandemic. Risk Anal 30 (8):1210-1218. doi:10.1111/j.1539-6924.2010.01428.x
19. Stutt R, Retkute R, Bradley M, Gilligan C, Colvin J (2020) A modelling framework to assess the likely effectiveness of facemasks in combination with 'lock-down' in managing the COVID-19 pandemic. Proceedings Mathematical, physical, and engineering sciences 476 (2238):20200376. doi:10.1098/rspa.2020.0376
20. China Internet Network Information Center (CNNIC). The 45th China statistical report on internet development; 2019. Available from: http://www.cnnic.net.cn/hlwfzyj/. Accessed April 28, 2020.
21. China Internet Network Information Center (CNNIC). China statistical report on search engine; 2019. Available from: http://www.cnnic.net.cn/hlwfzyj/. Accessed Oct 25, 2019.

Figure legends

**Fig. 1** Monthly search volume (in percentage) for (a) AH and rhinosinusitis and (b) AH and AR during 2011–2019.

The search variations for AH and rhinosinusitis were statistically relevant (R=0.643, P<0.05) and fluctuated in similar patterns. The search variations for AH and AR were negatively related (R=−0.239, P<0.05) and fluctuated oppositely to a certain extent. AH, adenoid hypertrophy; AR, allergic rhinitis

**Fig. 2** Average monthly search volume for AH, rhinosinusitis, and AR. The search variations of AH and rhinosinusitis coincided well (R=0.836, P<0.01). They both declined from December to February and from May to September and then increased from February to May and from September to December. The search variations of AH and AR were not correlated (P>0.05). Although searches for AR increased from June to August, searches for AH declined. Although searches for AR decreased from September to December, searches for AH increased. AH, adenoid hypertrophy; AR, allergic rhinitis
Fig. 3 Comparison of the search volume for (a) AH, (b) rhinosinusitis, and (c) AR during the first 5 months of 2019 and 2020. The search volume of AH and rhinosinusitis in the first 5 months of 2020 decreased by 41.21% and 54.31%, respectively, compared with the search volume during the same period in 2019. The total search volume for AR decreased by only 28.08% and even increased in January and February.