



ORIGINAL RESEARCH PAPER

Pulmonology

GOOGLE TRENDS AND QUERIES FOR CHRONIC RESPIRATORY DISEASES – A NEW TOOL FOR PHYSICIANS TREATING RESPIRATORY DISEASES.

KEY WORDS: Asthma, COPD, India, Big Data, Google Trends

Dr. Kartik Deshmukh	Duke FUQUA school of Management, North Carolina, USA.
Dr. Arjun Khanna *	Consultant Pulmonologist, Yashoda Superspeciality Hospital, Kaushambi, UP. *Corresponding Author
Dr. Pallavi Periwal	Consultant Pulmonologist, Aanch Hospital, Jaipur, Rajasthan.
Dr. Sushant Khurana	Medical Head, Bson Scientific, Delhi

ABSTRACT

Respiratory diseases (both acute and chronic) are among the leading of disease burden and mortality. To build further on the research, it is necessary to understand the population trends and health-seeking behavior about respiratory diseases. We aimed to evaluate real-life population trends for respiratory diseases, air pollution, and common drugs used to treat asthma/COPD. We also aimed to compare the trends of chronic respiratory diseases with other non-communicable diseases prevalent in India. One method to evaluate this behavior is analyzing internet search queries. Google indexes the internet queries with every search performed and makes this information available to the public through Google Trends. Google Trends provides Normalized search volumes for selected terms over specific time ranges and geographic regions, and we evaluated the same for queries in India. IBM-SPSS v26 and Cosinor model of R 3.6.3 were used for data analysis and assessment of seasonality respectively. The search volumes for Asthma and COPD were relatively lower than those of diabetes, hypertension, and depression. The search volumes for asthma/COPD and pneumonia showed considerable seasonality. Consistent with the seasonality in asthma and COPD, common drug combinations for treatment of asthma/COPD showed seasonal variation as well. The search volumes for Salbutamol were quite higher as compared to ICS-LABA combinations. One of the reasons for seasonal variation for respiratory diseases in India might be Air pollution, which demonstrated a spike in November every year. The peak month of searches corresponded to the winter season in India.

Conclusions: Consistent with real-life observations, Google trends for queries related to chronic respiratory diseases show seasonal variation in India. Targeted interventions, awareness programs, and SEO during the phase month (peak month) can help in increasing awareness of chronic respiratory diseases in India.

INTRODUCTION:

In the current era of “Big data” analytics in healthcare, the understanding of respiratory diseases, and the population behavior can be improved by leveraging the insights from this “Big Data”. Big data has been used in other countries in Influenza surveillance as well. ⁰ Though artificial intelligence and machine learning have been successfully assimilated in fields such as finance and engineering in India, the application of advanced analytics is not common in healthcare. Several obstacles such as lack of electronic health records, poor data quality, and lack of adequate data governance prevail in India’s healthcare sector. ⁰

Respiratory diseases (both acute and chronic) are among the leading of disease burden and mortality in India. ⁰ Studies such as POSEIDON, INSEARCH, SWORD, and AP-AIM have contributed greatly to the understanding of the characteristics of patients with respiratory disorders in Asia. ^(,) To build further on the research, it is necessary to understand the population trends and health-seeking behavior about respiratory diseases. One method to evaluate this behavior is analyzing internet search queries. It has been shown that internet search queries show seasonal and geographic variations that resemble the variations observed in the real world population. ⁰ Google, the most popular search engine, indexes the internet queries with every search performed and makes this indexed information available to the public through Google Trends. In the current study, we evaluated the Google trends data in the context of chronic respiratory diseases in India. We wanted to explore whether this data might offer actionable insights about population interests about respiratory diseases. To the best of our knowledge, no such study has been conducted in India to

date.

SUBJECTS AND METHODS:

Data Collection

Google indexes the internet queries with every search performed and makes this information available to the public through Google Trends. Google Trends provides Normalized search volumes for selected terms over specific time ranges and geographic regions. The normalized search volume is known as the “Search volume index” (SVI), which ranges from 0-100. SVI of 100 represents the peak search activity over a given period. Activity at all other periods is presented relative to that peak. Google trends allow comparison of search volumes of up to 5 terms and observe their trends over the specified time. The system eliminates repetitive search queries from the same user performed in a short period.

Objectives and specific data collection techniques for each objective were as follows:

1. Evaluate the relative search trends of five common chronic diseases in India i.e. asthma, COPD, diabetes, depression, and hypertension. This would help in getting insights about health-seeking behavior about the two most common chronic respiratory diseases in India as compared with other chronic diseases. Queries were filtered by disease, and we used the terms “asthma”, “COPD”, “diabetes”, “hypertension” and “depression” for eliciting the trends.
2. Evaluate the relative search trends and seasonality of search trends for four common respiratory diseases (both acute and chronic) in India i.e. asthma, COPD, influenza, pneumonia, and tuberculosis. This would help in comparing the health-seeking behavior of acute vs

chronic respiratory diseases in India. Queries were filtered by disease, and we used the terms “asthma”, “COPD”, “pneumonia”, and “tuberculosis”.

- Evaluate the search trends of conventional inhaled medications used for Asthma/COPD, and compare the trends of SABA with ICS-LABA combinations. Initially, we tried generating results using the generic names, but we encountered the problem of low search volumes. This was expected since it is easier for people to remember the brands rather than the generic names. Hence, we used the brand names of the brands with a maximum market share in the Indian market. ⁰ We used the brand name “Asthalin” (an age-old brand of salbutamol in India) and compared the same with brand names of common ICS-LABA in Indian markets i.e. “Seroflo”, “Foracort”, and “Maxiflo” for Salmeterol/Fluticasone, Formoterol/Budesonide, and Formoterol/Fluticasone respectively
- Lastly, we queried the Google Trends database for the term “air pollution” in India. Air pollution is an important risk factor for exacerbation of pre-existing chronic respiratory diseases, and hence it is essential to know the trends of search volumes for Air pollution ⁰

We filtered the queries from 1st January 2014 to 30th October 2019 to remove the bias introduced by the COVID-19 pandemic.

Data Analysis

Google trends data was downloaded in xls format, and the data was analysed using IBM SPSS v26. Initial analysis included calculation of mean SVI for the specified term. For the terms “asthma”, “COPD”, “depression”, “hypertension”, and “diabetes”, we divided the mean SVI of each term with that of COPD, giving an estimate of search volume of the term relative to that of COPD.

Graphs showing the trends of SVI for the term were plotted. To assess the seasonality of the trends, we subjected the data to cosinor analysis using the R version 3.6.3. Cosinor analysis consists of regression of SVI on y axis onto sine and cosine transformation of time on x axis. In this case, time refers to the month-wise data of SVI for the specified terms in specified period (Each year consisted of 12 data points for SVI corresponding to twelve months in a year). The month corresponded to the month number of the year. For ex, month 1 represented January, month 2 represented February, and so on. The resulting sinusoidal curve consists of an amplitude ([A] which refers to the magnitude of seasonal change), phase month([P] timing of the peak), and length of seasonal cycle (12 for monthly data). The low point month for any term refers to the month with the lowest volume of search queries for the particular term. The phase month refers to the peak search queries for the specified term. Cosinor model also allows for the calculation of statistical significance. The significance level was set at p<0.025 for controlling the type I error due to multiple testing in the cosinor analysis.

RESULTS:

Figure 1 shows the graphical trends of SVI of the terms “asthma”, “COPD”, “diabetes”, “hypertension” and “depression” in the selected period. Mean SVI and standard deviation for the selected five chronic diseases are mentioned in Table 1. The search frequency of diabetes was approximately 4 times (80.4±6.73 / 20.14±2.21) that of asthma and 9 times that of COPD (80.4±6.73/8.91±1.07). Similarly, the search frequency for hypertension and depression were higher than COPD and asthma. (Table 1)

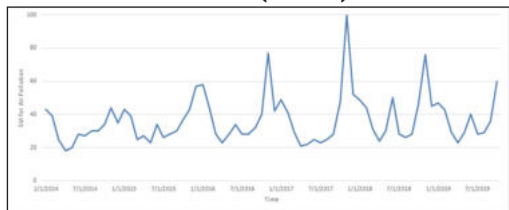


Figure 1: Trends of SVI of the terms “Asthma”, “COPD”, “Diabetes”, “Hypertension” and “Depression” from Jan 2014 to October 2019

Table 1: Mean SVI for “Asthma”, “COPD”, “Diabetes”, “Hypertension” and “Depression” from Jan 2014 to October 2019

Diseases	COPD	Asthma	Diabetes	Hypertension	Depression
Mean(±SD)	8.91±1.07	20.14±2.21	80.4±6.73	34.97±3.1	33.73±3.63
Relative search volumes vs COPD	1	2.26	9.02	3.92	3.78

The SVI of diseases “asthma”, “COPD”, “pneumonia”, and “tuberculosis” were compared and graphical trends represented in Figure 2a. Table 2 shows the mean SVI for the four diseases and points towards high search volumes of tuberculosis vs other diseases. On visual inspection, seasonal peaks were seen in the graphical trends. Seasonality in search volumes was statistically significant(p<0.005) for asthma, COPD, and pneumonia. (Table 2)

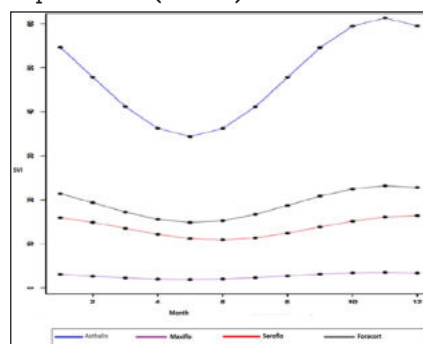


Figure 2a: Trends of SVI of the terms “Asthma”, “COPD”, “Pneumonia”, and “Tuberculosis” from Jan 2014 to October 2019

The graphical results of cosinor analysis are shown in Figure 2b.

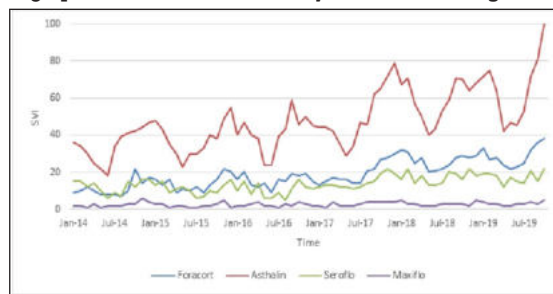


Figure 2b: Graphical results of Cosinor analysis showing seasonality in SVI(search volume index) for “Pneumonia”, “Asthma” and “COPD”

Table 2: Cosinor Analysis for seasonality and mean SVI for Tuberculosis, Asthma, COPD, and Pneumonia from Jan 2014 to October 2019

	Mean (±SD)	Amplitude	Phase Month	Low Point month	P-value
Asthma	34.57±3.7	2.77	2.7	8.7	NS(p>0.05)
COPD	15.29±1.56	3.55	11.4	5.4	P<0.005
Pneumonia	25±5.01	1.37	12.5	6.5	P<0.005
Tuberculosis	74.39±7.32	3.78	12.1	6.1	P<0.005

The trends for the brand names common drugs used for the management of asthma and COPD are shown in Figure 3a.

Table 3: Cosinor Analysis for seasonality and Mean SVI for “Foracort”, “Asthalin”, “Seroflo”, and “Maxiflo” from

Jan 2014 to October 2019

		Amplitude	Phase Month	Low Point month	P-value
Foracort (Formoterol/Budesonide)	18.94±7.78	4.15	11.1	5.1	P=0.015
Asthalin (Salbutamol)	47.49±16.46	13.45	11	5	P<0.005
Maxiflo (Formoterol/Fluticasone)	2.73±1.13	0.79	10.9	4.9	P<0.005
Seroflo (Salmeterol/Fluticasone)	13.6±4.28	2.77	11.9	5.9	P<0.005

The mean SVIs are mentioned in **Table 3**. The mean SVI of *Asthalin* was significantly higher than that of ICS-LABA combinations, indicating increased search volumes and high level of population interest for the term. Consistent with the seasonality in Asthma and COPD, the brand names for the drugs were shown to follow a seasonal pattern as well (**Table 3**). The cosinor plot for the brand names of the drugs is shown in **Figure 3b**.

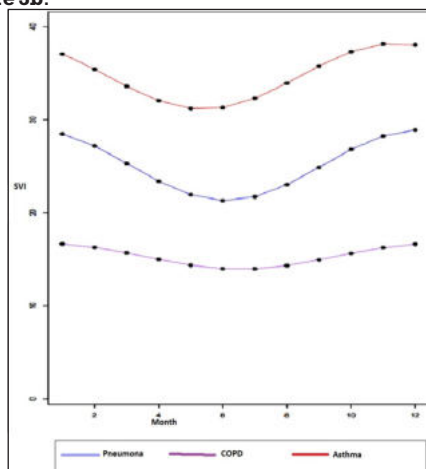


Figure 3a: Trends of SVI of the terms "Foracort", "Asthalin", "Seroflo", and "Maxiflo" from Jan 2014 to October 2019

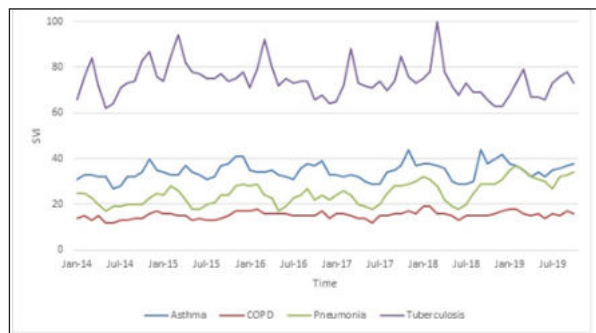


Figure 3b: Graphical results of Cosinor analysis showing seasonality in SVI (search volume index) for "Asthalin", "Maxiflo", "Foracort" and "Seroflo"

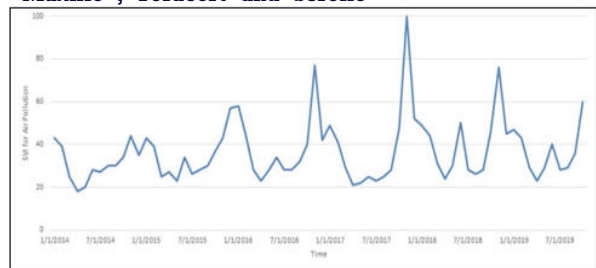


Figure 4a: Trends of SVI of the terms "Air pollution" from Jan 2014 to October 2019

Jan 2014 to October 2019

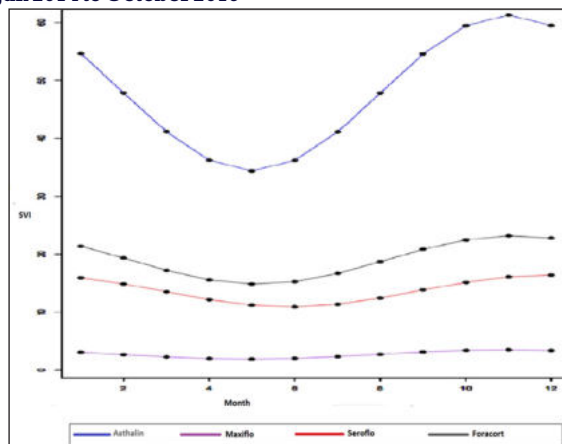
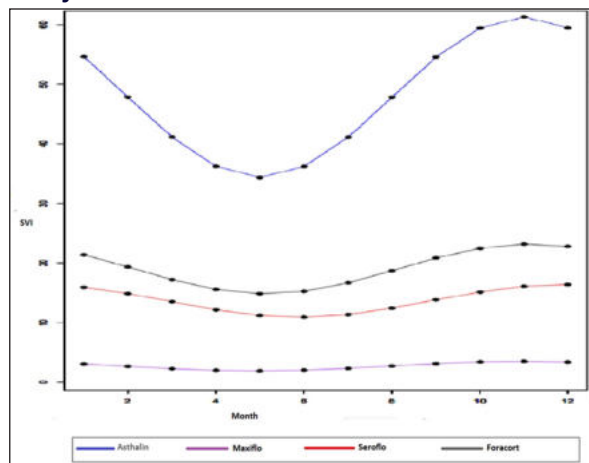
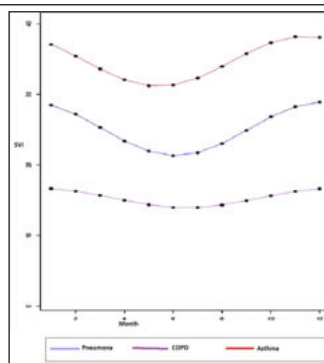
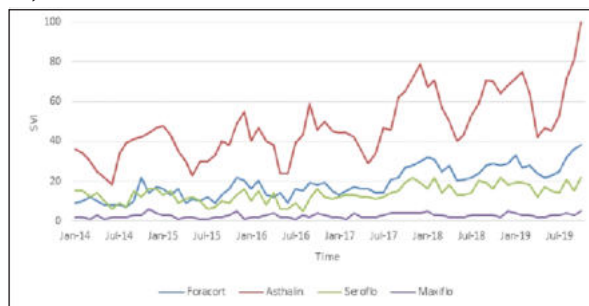


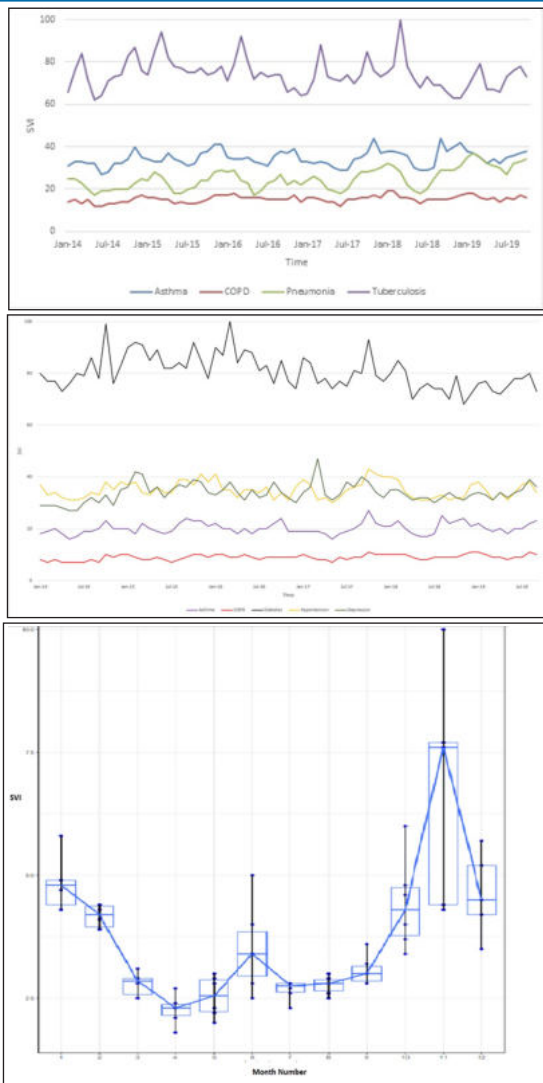
Figure 4b: Boxplot of searches for "Air pollution" in India from Jan 2014 to October 2019



The trends for Air Pollution are shown in **Figure 4a**.

Cosinor Analysis confirmed the seasonality of searches for Air Pollution (Amplitude=14.09, Phase month=10.9, Low point month=5.6, p<0.00005). The boxplot is shown in **figure 4b**, which shows a sudden significant spike in November (month 11).





DISCUSSION:

To the best of our knowledge, our study is the first study to use Google Trends for queries on chronic respiratory diseases in India. Studies using Google Trends have been published for diseases such as influenza, OSA, allergic rhinitis, cardiac disease, and restless leg syndrome in other countries.⁽¹⁾

An estimated 55 million and 37 million people are living with COPD and asthma in India.⁽²⁾ However, in the current study, the SVI of asthma and COPD was consistently below that of diabetes, hypertension, and depression. Lack of awareness is a common problem with obstructive airway diseases, and the problem is even worse with COPD. Physicians and patients both are likely to be unaware of COPD, leading to sub-optimal management and increased burden.⁽³⁾ These results indicate a low interest of the population in seeking information about asthma and COPD as compared to hypertension, diabetes, and depression, despite a high burden. It was a surprising observation that the SVI of asthma and COPD was below that of depression, which is often considered a neglected and stigmatized disease in the Indian population.

When we analyzed the trends of interest for respiratory disease i.e. asthma, COPD, pneumonia, and tuberculosis, we observed a high population interest in tuberculosis vs other selected diseases. This might be due to the pre-existing awareness of the disease and extensive Indian government campaigns about tuberculosis since 2015. In the trends for tuberculosis, we could notice a spike in March 2015, which can be attributed to the launch of the “TB Harega Desh Jeetega”

campaign in March 2015. COPD remained the least searched disease among the four diseases; the mean search volumes for asthma were 2.26 times those of COPD, indicating an increased interest and awareness of asthma as compared to COPD in the population. Cosinor analysis proved the statistically significant seasonal variation in asthma, pneumonia, and COPD search queries. This relates well with the clinical data, which shows worsening of asthma and COPD in winters, and increased hospitalizations for pneumonia at low temperatures.^(4,5) In our study, the phase month (or the peak month) for search queries for asthma, COPD, and pneumonia lay between 11.4 to 12.5, pointing to November and December. These two months are often characterized by winters in most of the Indian states.

As seen with asthma and COPD, the search volumes for “Asthalin”, “Seroflo”, “Foracort”, and “Maxiflo” also showed seasonal variations. The peak month for searches lay between 10.9 to 11.9 corresponding to the end of October and End of November. There is a significant overlap between the peak months for queries on asthma/COPD and queries for the selected four medications. One explanation might be the increased diagnosis and increased prescription of the selected medications due to the worsening of underlying asthma or COPD. This might lead to increased patient/caregiver interest in the drugs for the management of asthma and COPD.

On visual inspection of the trends, the queries for Asthalin exceed those of ICS-LABAs.

In the current study, the query for “Asthalin” was not specific to Salbutamol inhaler but also included syrup and tablets. Nevertheless, Asthalin continues to be more popular than ICS-LABAs and is evident in the prescriptions of many asthmatic patients even today. The search volume for Asthalin was more than 3 times than that of the brand for Formoterol/Budesonide. Regardless of the formulation, this finding is of grave concern since the oral formulation of Salbutamol (Asthalin tablet and Asthalin Syrup) contributes to an increased asthma burden in India, and (Asthalin Inhaler) inhaled Salbutamol is no longer recommended by the current GINA guidelines.⁽⁶⁾ Among the ICS-LABAs, Formoterol/Budesonide(Foracort) has higher search volumes vs Salmeterol/Fluticasone and Formoterol/Fluticasone. This is quite consistent with clinical practice, with most physicians preferring Formoterol/ Budesonide because of substantial data and quick onset of action.

The trends for air pollution might attributed to the smog and increased air pollution in the during winter every year. Increased air pollution might lead to exacerbation of underlying Asthma/COPD, and thus lead to increased interest of population in winter. Along with the temperature changes and seasonal alteration of circulating microbes, air pollution is an important factor in India that may account for increased burden, leading to increased interest and increased prescription of drugs for asthma and COPD. This might explain the proximity of phase month for Air pollution to the phase month for asthma and COPD.

The results from this study can be used for identifying the right time for awareness programs for respiratory diseases in India. Based on the results from this study, it seems logical to increase the frequency of awareness programs in winters in India. Since the volume of queries for asthma, COPD and pneumonia increase in winters, Search Engine Optimization(SEO) can be used to direct the internet traffic to appropriate and reliable health education content.

The results of this study bear a policy level implication for chronic respiratory diseases in India. Despite a high prevalence, chronic respiratory diseases often don't get the attention they deserve and are often poorly managed, which leads to increased morbidity and mortality. Appropriate policies to generate awareness about the diseases and their

management are the need of the time. There are several limitations in the current study. First, Google Trends does not capture the demographic details, and hence it is not possible to do an additional analysis depending on demographic parameters. Second, it cannot be proved that the trends in Google Trends are representative of actual disease activity. However, a study has shown the correlation between Google queries and epidemiologic data in 2005-06 for the flu season, and this was the basis for assumption in the current study as well. Third, we used the English terms such as "asthma", "COPD", "hypertension", "diabetes", and "depression" for non-English speaking Indian population. The rural population might not be aware of such complex terms, and hence the current study might not reflect the status across all places in India. Fourth, confounding factors such as Christmas holidays and winter vacations might lead to increased search queries because of the availability of free time. Fifth, the use of brand names isn't ideal because all patients are not prescribed the same brand. However, the selection of brands that prescribed by a majority of physicians might have alleviated this problem to an extent.

To conclude, Google Trends can provide important insights in chronic respiratory diseases in India. The population interest for asthma and COPD is much lesser than that for diabetes, depression and hypertension. The search queries for asthma, COPD and pneumonia follow a seasonal trend, with increased search volumes in winter. The brands of commonly used drugs for management of Asthma and COPD demonstrate seasonal variation as well. Salbutamol still seems to be commonly used vs ICS-LABAs in India. Air pollution, an important risk factor for respiratory diseases, peaks in winter which coincides with the peak of chronic respiratory diseases. Targeted interventions, awareness programs and SEO during the phase month (peak month) can help in increasing awareness of chronic respiratory diseases in India.

Institutional Review Board approval: The study doesn't study involve Human Subjects and/or Human Subjects Data, and hence doesn't qualify for Institutional review board approval/ Ethics committee approval. This study is based on aggregate data, freely available in public domain.

REFERENCES

1. Yang S, Santillana M, Kou SC. Accurate estimation of influenza epidemics using Google search data via ARGO. *Proc Natl Acad Sci U S A*. 2015 24;112(47):14473-8.
2. Kruse CS, Goswamy R, Raval Y, Marawi S. Challenges and Opportunities of Big Data in Health Care: A Systematic Review. *JMIR Med Inform*. 2016 21;4(4):e38.
3. India State-Level Disease Burden Initiative Collaborators. Nations within a nation: variations in epidemiological transition across the states of India, 1990-2016 in the Global Burden of Disease Study. *Lancet*. 2017 2;390 (10111): 2437-2460.
4. Salvi S, Apte K, Madas S, Barne M, Chhowala S, Sethi T, Aggarwal K, Agrawal A, Gogtay J. Symptoms and medical conditions in 204 912 patients visiting primary health-care practitioners in India: a 1-day point prevalence study (the POSEIDON study). *Lancet Glob Health*. 2015 ;3(12):e776-84.
5. Singh V, Sharma BB. Respiratory disease burden in India: Indian chest society SWORD survey. *Lung India*. 2018;35(6):459-460.
6. Jindal SK, Aggarwal AN, Gupta D, Agarwal R, Kumar R, Kaur T, Chaudhry K, Shah B. Indian study on epidemiology of asthma, respiratory symptoms and chronic bronchitis in adults (INSEARCH). *Int J Tuberc Lung Dis*. 2012; 16 (9): 1270-7.
7. Salvi SS, Apte KK, Dhar R, Shetty P, Faruqi RA, Thompson PJ, Guleria R. Asthma Insights and Management in India: Lessons Learnt from the Asia Pacific - Asthma Insights and Management (AP-AIM) Study. *J Assoc Physicians India*. 2015;63(9):36-43.
8. Kumar N, Pandey A, Garg N, Sampene E, Lavie CJ. Seasonal and Geographic Patterns in Seeking
9. Cardiovascular Health Information: An Analysis of the Online Search Trends. *Mayo Clin Proc*. 2018 ;93(9):1185-1190.
11. "Indian Pharma Market registers 11.9% growth in September 2019", October 12 2019. *Pharmabiz*, [http:// pharmabiz.com/ Article Details. asp? aid= 118 717&sid=1](http://pharmabiz.com/ArticleDetails.aspx?aid=118717&sid=1)
12. Jiang XQ, Mei XD, Feng D. Air pollution and chronic airway diseases: what should people know and do? *J Thorac Dis*. 2016;8(1):E31-40.
13. Bousquet J, Agache I, Anto JM, Bergmann KC, Bachert C, Annesi-Maesano I, Bousquet PJ, D'Amato G, Demoly P, De Vries G, Eller E, Fokkens WJ, Fonseca J, Haahtela T, Hellings PW, Just J, Keil T, Klimek L, Kuna P, Lodrup Carlsen KC, Mõsges R, Murray R, Nekam K, Onorato G, Papadopoulos NG, Samolinski B, Schmid-Grendelmeier P, Thibaudon M, Tomazic P, Triggiani M, Valiulis A, Valovirta E, Van Eerd M, Wickman M, Zuberbier T, Sheikh A. Google Trends

- terms reporting rhinitis and related topics differ in European countries. *Allergy*. 2017;72(8):1261-1266.
14. Nuti SV, Wayda B, Ranasinghe I, Wang S, Dreyer RP, Chen SI, Murugiah K. The use of google trends in health care research: a systematic review. *PLoS One*. 2014;9(10):e109583.
15. India State-Level Disease Burden Initiative CRD Collaborators. The burden of chronic respiratory diseases and their heterogeneity across the states of India: the Global Burden of Disease Study 1990-2016. *Lancet Glob Health*. 2018;6(12):e1363-e1374.
16. Kaur I, Aggarwal B, Gogtay J. Understanding perception of chronic obstructive pulmonary disease among general practitioners, physicians, and pulmonologists in India: Results from a face-to-face survey. *Perspect Clin Res*. 2016;7(2):100-5.
17. Thakrar R, Alaparthy GK, Kumar SK, Vaishali K, Zulfequeer CP, Aanad R. Awareness in patients with COPD about the disease and pulmonary rehabilitation: A survey. *Lung India*. 2014;31(2):134-8.
18. Lin HC, Lin CC, Chen CS, Lin HC. Seasonality of pneumonia admissions and its association with climate: an eight-year nationwide population-based study. *Chronobiol Int*. 2009;26(8):1647-59.
19. Donaldson GC, Wedzicha JA. The causes and consequences of seasonal variation in COPD exacerbations. *Int J Chron Obstruct Pulmon Dis*. 2014 6;9:1101-10.
20. Wisniewski JA, McLaughlin AP, Stenger PJ, Patrie J, Brown MA, El-Dahr JM, Platts-Mills TA, Byrd NJ, Heymann PW. A comparison of seasonal trends in asthma exacerbations among children from geographic regions with different climates. *Allergy Asthma Proc*. 2016;37(6):475-481.
21. GINA 2020 Main report. [https://ginasthma.org/wp-content/ uploads/ 2020/ 06/GINA-2020-report_20_06_04-1-wms.pdf](https://ginasthma.org/wp-content/uploads/2020/06/GINA-2020-report_20_06_04-1-wms.pdf)
22. Raja Dhar, Sushmeeta Chhowala, Meena Lopez, Rashmi Hegde, Sapna Madas, Sundeep Salvi, Jaideep Gogtay. Does The High Use Of Oral Antiasthmatics Contribute To The Elevated Burden Of Asthma In India? *European Respiratory Journal* 2019 54: Oa3562;
23. Eysenbach G. Infodemiology: tracking flu-related searches on the web for syndromic surveillance. *AMIA Annu Symp Proc.*:2006:244-8.