Assessing population interest in interstitial lung diseases in the USA using Google Trends

To the Editor:

The term interstitial lung diseases (ILDs) encompasses several different conditions that are characterised by inflammation and/or fibrosis of the pulmonary interstitium [1]. ILDs are associated with significant morbidity and mortality [2]. Patients receive a significant amount of health information on the internet and adoption of the internet has been steadily increasing [3]. Google is the most popular search engine with over 90% of the total online searches in the USA performed utilising it [4]. An estimated 7% of all internet searches on Google are health related, corresponding to 70 000 health-related searches occurring each minute [5]. In this study, we used Google's analysis tool, Google Trends (GT) [6], and the objective was to evaluate the public interest in ILD compared with common pulmonary diseases such as chronic obstructive pulmonry disease (COPD) and lung cancer.

GT is a publicly available tool that provides data on web queries and displays results on a normalised scale. GT determines the proportion of searches for user-specified terms among all searches performed *via* Google over a specified timeframe in a specific geographic region. The results are presented as a relative search volume (RSV). RSV is scored 1–100 representing interest relative to the peak popularity of the term in question. A score of 0 denotes that the term is below 1% of peak popularity while a score of 100 indicates peak popularity. To rule out any bias for absolute search volume measurements, RSV indirectly corrects for internet access and population size. GT excludes duplicate searches if they are attempted by the same individual in a short period of time, and for searches for the same term with apostrophes or other special characters. Most importantly, GT allows for direct comparison between RSV of different topics. The RSV for the search terms for ILD, COPD and lung cancer was directly compared relative to each other following previously used methodology [7]. The region was set to "United States" and the category was set to "health." The period was set to "Jan 1, 2011, to Dec 31, 2019."

Queries in GT can be executed as a disease/topic or search term. Diseases/topics include all keywords or terms that fall within that category. In contrast, search terms provide results that include only the keywords that are included in the term. Using "+" enables users to search for multiple keywords simultaneously. The use of quotations for the keywords limits the searches to the combination of words within the quotations.

The following search terms were selected *a priori* based on the common terms used in clinical practice:
- ILD: ILD+'interstitial lung disease'+IPF+'pulmonary fibrosis'+sarcoidosis
- COPD: COPD+"chronic obstructive pulmonary disease"+emphysema+"chronic bronchitis"
- Lung cancer: "lung cancer"

Several other search terms were evaluated with the intention to include them in the final search if the additional term increased the RSV by more than one point; however, none met this criterion. These terms included other ILD terms (*i.e.* "chronic hypersensitivity pneumonitis," "pneumoconiosis", "rheumatoid lung," "organizing pneumonia"), plural forms (*i.e.* "interstitial lung diseases") as well as related terms suggested by GT for all three diseases (*i.e.* "lung fibrosis").

Finally, the RSV for search terms was compared with the analogous RSV for the diseases/topics. Pulmonary fibrosis, COPD and lung cancer were selected as diseases/topics as they had the highest RSV compared with the matching topics identified by GT. The use of search terms yielded a higher RSV for

Shareable abstract (@ERSpublications)
Google Trends is a useful tool for evaluating the public interest in interstitial lung diseases (ILD). The interest in ILD is much lower compared to lung cancer although both conditions are associated with significant morbidity and mortality. https://bit.ly/38QzPce

ILDs and COPD and a similar RSV for lung cancer, therefore, search terms were used for all analyses instead of diseases/topics.

The prevalence of chronic respiratory diseases was obtained from the Global Burden of Disease (GBD) Database [8]. We used the Pearson correlation coefficient to examine the association between GT RSV and the disease prevalence per 100 000 people in the USA, from 2011 to 2019. Linear regression modelling was used to examine temporal trends. A p-value of $\leqslant 0.05$ was considered statistically significant. Analyses were performed with the SPSS 27.0 statistical package (IBM Corp., Armonk, NY, USA).

The prevalence and GT RSV of the common chronic respiratory diseases are shown in figure 1a. COPD has the highest prevalence (6143 per 100 000 population) with a mean RSV of 61.5. The prevalence of lung cancer is significantly lower (135 per 100 000 population) but had an RSV of 48.3. The prevalence of ILD is 200 per 100 000 population but had a mean RSV of 23.3.

From 2011 to 2019, the mean RSV for COPD increased from 55.2 to 78.1 ($p<0.001$, percent change 41.5%), for lung cancer it increased from 48.8 to 58.4 ($p=0.04$, percent change 19.6%), and for ILD it increased from 22.5 to 26.1 ($p=0.01$, percent change 15.9%) (figure 1b).

The prevalence per 100 000 population for COPD increased from 5703 to 6143 ($p<0.001$, percent change 7.7%), lung cancer remained stable ($p=0.25$), and ILD increased from 170 to 200 ($p<0.001$, percent change 17.6%).

The weighted correlation between prevalence and mean RSV from 2011 to 2019 for COPD was 0.93 ($p<0.001$) and for ILD was 0.72 ($p=0.03$). In 2019, the prevalence of COPD was 30.1 times higher compared with ILD, but the RSV was only 3.2 times higher. In contrast, the ILD prevalence was 1.5 times higher than for lung cancer, but the RSV was 0.5 times that for lung cancer.

ILDs are increasing in prevalence and associated with a significant burden [2]. In 2019 there were approximately 654 841 cases and 21 505 deaths from ILD in the USA. However, the public interest and awareness of ILDs are unknown. Infodemiological studies are becoming popular in medicine as more patients are obtaining their health information from the internet. GT has been used to examine public interest in numerous health conditions such as cancers and asthma [7, 9]. In this study, we examined the public interest in ILD using GT and compared it with COPD and lung cancer. Our study indicates that the online interest in ILD is relatively higher compared with COPD (compared with the prevalence). In contrast, the online interest for ILD is much lower compared with lung cancer despite patients with idiopathic pulmonary fibrosis (a major subcategory of ILD) having a dismal median survival, comparable to those with advanced lung cancer [10].

a)

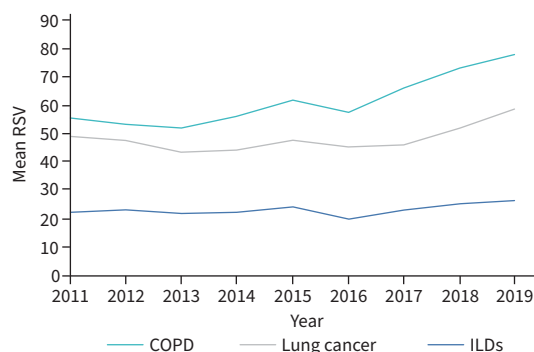| Chronic respiratory diseases	Prevalence (per 100 000 population in 2019)	Google Trends search term	Mean±SD RSV 2011–2019
COPD	6143	COPD + "chronic obstructive pulmonary disease" + emphysema + "chronic bronchitis"	61.5±9.2
Tracheal, bronchus and lung cancer	135	"Lung cancer"	48.3±4.4
ILD and pulmonary sarcoidosis	200	ILD + "interstitial lung disease" + IPF + "pulmonary fibrosis" + sarcoidosis	23.3±1.6

b)

FIGURE 1 a) Prevalence and Google relative search volume (RSV) of common chronic respiratory diseases in the USA. b) Average annual RSV for COPD, lung cancer and interstitial lung diseases in the USA from 2011 to 2019. The region was set to "United States" and the category was set to "Health." The period was set to "Jan 1, 2011 to Dec 31, 2019." Relative search volume (RSV) is scored 1–100 representing interest relative to the peak popularity of the term in question. A score of 0 denotes that the term is below 1% of peak popularity while a score of 100 indicates peak popularity.

The greater online interest in lung cancer is likely reflective of the successful efforts by organisations to promote awareness of lung cancer [11]. In addition, implementation of lung cancer screening protocols within the past decade have also likely increased the interest. Diagnosing ILD can be challenging and many providers may not be comfortable providing a definitive diagnosis to the patient [12]. Lack of a clear understanding of their diagnosis and prognosis could be one of the limiting factors in patients' ability to perform their own internet search for information. On the other hand, the lower RSV for COPD could be due to having already received sufficient knowledge from the provider as many clinicians can adequately diagnose and treat COPD. This reduces the need for patients to seek out other sources to obtain information.

In a secondary analysis of the different ILD terms, we noted that the term sarcoidosis was searched more frequently than pulmonary fibrosis terms (IPF+"pulmonary fibrosis") and ILD terms (ILD+"interstitial lung disease") (mean±SD RSV of 13.2±1.1 *versus* 6.8±0.8 *versus* 3.5±1.1, respectively). However, when evaluating the trends from 2011 to 2019, the searches for sarcoidosis decreased 9.1% but the search terms for pulmonary fibrosis and ILD increased (38.0% and 136%, respectively). This is encouraging and likely indicative that organisations such as the Pulmonary Fibrosis Foundation are effective in spreading awareness of the disease.

GT has provided a glimpse into the public's interest in ILD. However, there are certain limitations as it is impossible to discern the entities which are making the queries of interest in this study as they could be patients, healthcare providers or corporations. Since the data is not linked to a specific user, we cannot control for confounders that may impact search activity such as race, ethnicity, socioeconomic status and level of education. Further research needs to be conducted to evaluate the utilisation of GT and the data trove that it holds, and its utility in facilitating further research into medical diseases.

Niranjan Jeganathan [1], **Essam K. Nagori**[2] **and Matheni Sathananthan**[3]

[1]Dept of Medicine, Division of Pulmonary, Critical Care, Hyperbaric, Allergy and Sleep Medicine, Loma Linda University Health, Loma Linda, CA, USA. [2]Dept of Internal Medicine, Arrowhead Regional Medical Center, Colton, CA, USA. [3]Dept of Medicine, Loma Linda University Health, Loma Linda, CA, USA.

Corresponding author: Niranjan Jeganathan (njeganathan@llu.edu)

References

1   Cottin V, Hirani NA, Hotchkin DL, *et al.* Presentation, diagnosis and clinical course of the spectrum of progressive-fibrosing interstitial lung diseases. *Eur Respir Rev* 2018; 27: 180076.

2   Jeganathan N, Sathananthan M. The prevalence and burden of interstitial lung diseases in the USA. *ERJ Open Res* 2022; 8: 00630-2021.

3   Fox S, Duggan M. Health online 2013. Pew Research Center, 2013. Date last accessed: 21 January 2022. Date last updated: 15 January 2013. www.pewresearch.org/internet/2013/01/15/health-online-2013

4   mozilla. More than 90% of the world uses Google Search. Date last accessed: 21 January 2022. Date last updated: April 2018. https://internethealthreport.org/2018/90-of-the-world-uses-google-search

5   Murphy M. Dr. Google will see you now: Search giant wants to cash in on your medical queries. Date last accessed: 21 January 2022. Date last updated: 10 March 2019. www.telegraph.co.uk/technology/2019/03/10/google-sifting-one-billion-health-questions-day

6   Google Trends. Date last accessed: 15 March 2022. https://trends.google.com/trends

7   Mavragani A, Ochoa G. Google Trends in infodemiology and infoveillance: methodology framework. *JMIR Public Health Surveill* 2019; 5: e13439.

8   Institute for Health Metrics and Evaluation, Global Health Data Exchange. GBD Results tool. Date last accessed: 18 January 2022. http://ghdx.healthdata.org/gbd-results-tool

9   Schootman M, Toor A, Cavazos-Rehg P, *et al.* The utility of Google Trends data to examine interest in cancer screening. *BMJ Open* 2015; 5: e006678.

10   Ballester B, Milara J, Cortijo J. Idiopathic pulmonary fibrosis and lung cancer: mechanisms and molecular targets. *Int J Mol Sci* 2019; 20: 593.

11   Saab MM, FitzGerald S, Noonan B, *et al.* Promoting lung cancer awareness, help-seeking and early detection: a systematic review of interventions. *Health Promot Int* 2021; 36: 1656–1671.

12   Cosgrove GP, Bianchi P, Danese S, *et al.* Barriers to timely diagnosis of interstitial lung disease in the real world: the INTENSITY survey. *BMC Pulm Med* 2018; 18: 9.