ORIGINAL ARTICLE

CrossMark

Assessing seasonal dynamics of Guillain-Barré syndrome with search engine query data

Antonino Giordano [1] · Marco Vabanesi [1] · Gloria Dalla Costa [1] · Federica Cerri [1] · Giancarlo Comi [1] ·
Vittorio Martinelli [1] · Raffaella Fazio [1]

Abstract

Background and objective In previous studies, data deriving from Google Trends showed promising correlation with disease incidence trends assessed with public health control systems. The aim of this work is to use search engine query data to investigate seasonal dynamics in Guillain-Barré syndrome (GBS) in the USA.

Methods Average Google monthly search volumes for GBS from 2008 to 2017 were analysed for the USA overall and on regional base with generalized estimating equation models. Association with monthly historical temperature variations was tested.

Results Monthly search volume for GBS displayed the greatest positive anomaly for October, clustering with September and November. Region-wide analysis confirmed this pattern and showed secondary spring (Feb/Apr) subpeaks in Pacific and Midwest. Association of GBS search volume with month-to-month temperature variations showed J-shaped relationship, with the highest peak occurring in months with greatest temperature falls, and subpeak in months with sharpest temperature rises.

Conclusions This study represents the first approach in investigating digital epidemiology of GBS and establishing possible links with traditional epidemiology. Cold season GBS peak has been observed by some traditional studies; hypothetical pathogenic relationship with infectious antecedents is supported from finding GBS peaks clustering with greatest temperature change. Further studies are needed to compare these findings to traditional public health approaches.

Keywords Guillain-Barré syndrome · Digital epidemiology · Search-engine data · Google Trends · Epidemiology · Seasonality

Introduction

Guillain-Barré syndrome (GBS) consists of a wide spectrum of acute immuno-mediated polyneuropathies characterized by acute-subacute development of muscular weakness with hyporeflexia or areflexia [1, 2].

Overall annual incidence of GBS in the world is between 1 and 2 cases for 100,000 people. In the USA, it has been estimated between 1.65 and 1.79 cases for 100,000 people [3]. The incidence of GBS seems to show a seasonal trend over the year [4]. This seasonality may depend on several environmental, climatological and demographic factors [5]. For example,

it is known that particular climate conditions favour seasonal gastrointestinal and respiratory tract infections, which are recognized as important antecedent events [6].

Nowadays, technology plays a crucial role in our society. New interest is arising in the use of digital technologies in the field of epidemiology. Google Trends (Alphabet Inc., USA) is a publicly available online tool aimed to analyse Google search keywords popularity over time and space, assessing role of factors such as country, region, related keywords and topics. This tool has a good potential in today's epidemiology, which has been explored by some recent studies. It has been shown that search peaks for seasonal influenza reflect data deriving from traditional control systems and may even precede them in detecting disease outbreaks [7]. In other studies in the field of neuro-epidemiology, Google Trends (GT) data confirmed the existence both of a latitude gradient and a seasonal trend in the incidence of multiple sclerosis, as previously observed by traditional epidemiology [8, 9].

Antonino Giordano and Marco Vabanesi contributed equally to this work.

✉ Antonino Giordano
   giordano.antonino@hsr.it

[1] Department of Neurology, San Raffaele Scientific Institute and University Hospital, Via Olgettina 48, 20132 Milan, Italy

Springer

Aims

The primary aim of this work is to use GT search results to analyse seasonal variation in GBS web search popularity in the USA, overall and in geographic subregions; we are also interested to understand if any relationship exists with the data coming from traditional epidemiology.

Material and methods

We conducted a retrospective web-based research to investigate whether a seasonal trend of GT search volumes for 'Guillain-Barré Syndrome' exists in the USA. We included searches from January 2008 to December 2017, analysing the monthly average over 10 years to smooth out any outlier peaks; we further refined the search analysing the ten most populous US states as of April 2010 Census (CA, TX, FL, NY, PA, IL, OH, GA, NC, MI) and grouping them by geographic regions (Pacific, Midwest, South West, South East, North East).

The raw index publicly provided by GT is RSV (relative search volume), with maximum monthly number of queries during the specified timeframe set to 100. GT provides also a regional interest index for the whole timeframe analysed, set to 100 for the region with maximum search frequency per unit of population. To integrate measures from different states, we calculated weighting coefficients based on US state population; then we linearly transformed data with maximum regional monthly peak set to 100.

We performed generalized estimating equation (GEE) analysis with search volume (for US overall and by geographic region) as dependent variable, year and month as independent variables and year as grouping variable. Post hoc Wald tests with adjustment for multiple comparisons were used to test significance of associations. Then we collected mean monthly temperatures for each state in 1981–2010 interval and tested association with RSV. These data were collected from publicly available repository from National Oceanic and Atmospheric Administration; data from the largest city of each state were used as proxy.

Privacy issues

Data coming from GT are completely anonymous, and no kind of personal information which may recall to individual web users or patients is retained by the system. Full privacy policy is available at: http://www.google.com/privacypolicy.html.

Results

Average RSV for GBS during the analysed timeframe (2008 to 2017) was 38.7 ± 8.9 arbitrary units. Analysing the USA overall with a GEE model, we verified the absence of outlier years, since no significant effect of year was observed ($p = 0.83$) on RSV. Analysing 10-year RSV average for each month, association with month of the year was highly significant for October (+ 13.74 units, 95% CI 5.60 to 21.88, $p = 0.002$), which was also an extreme outlier (> 3·IQR; Fig. 1a). The preceding and following month displayed the second highest degree of anomaly (September + 4.88 [95% CI − 0.46 to 10.22], $p = 0.17$; November + 6.02 [CI − 1.75 to 13.80], $p = 0.22$) compared to all remaining months of the year.

Analysis by geographic US regions confirmed the observation of October peak ($p < 0.03$ in all regions). This peak clustered with significant September subpeak in some regions (Pacific, Midwest); in these two regions, a secondary spring peak was observed too (April in Pacific, + 3.45 [95% CI 0.15 to 6.76], $p = 0.04$]; February in Midwest, + 5.14 [CI 0.97 to 9.31], $p = 0.016$).

For each state analysed, we correlated the mean monthly temperature variation (ΔT °C) from previous month to the monthly average RSV for GBS over years 2008–2017 (Fig. 1b). We observed a J-shaped trend, with highest RSV for records with the most negative ΔT (mostly October). Excluding the highest ΔT decile (with ΔT ≥ 5.33 °C), which formed the lesser peak of J (with 5/10 observations being April or May), we observed inverse linear correlation of RSV with ΔT. In particular, for every 1 °C month-to-month temperature reduction, mean + 1.08 (± 0.17) unit increase of RSV was observed ($p < 0.001$, linear model).

Discussion

Several traditional epidemiology studies have investigated the incidence and the seasonality of GBS all over the world. However, the results have often been contrasting. In a systematic review and meta-analysis published by Webb et al. [5], across 9386 patients from 42 studies, an increased risk of GBS in winter versus summer was observed. Shui et al. [4] reported an increased number of visits for GBS in winter and spring in the USA, compared to summer and fall. In other studies, otherwise, no clear seasonal trend was found [10, 11].

In our analysis we found a significant peak of GBS web searches in the month of October overall in the USA. This observation was also confirmed assessing seasonality in the ten most populous states, grouped in five macro-regions. We observed a trend of increased searches also in the months surrounding the October peak, mainly September, and a secondary spring peak in some regions. Month-to-month average

Fig. 1 Google Trends relative
search volume (RSV) for
Guillain-Barré syndrome,
assessed grouping by month (**a**)
and by month-to-month mean
temperature variation (**b**)
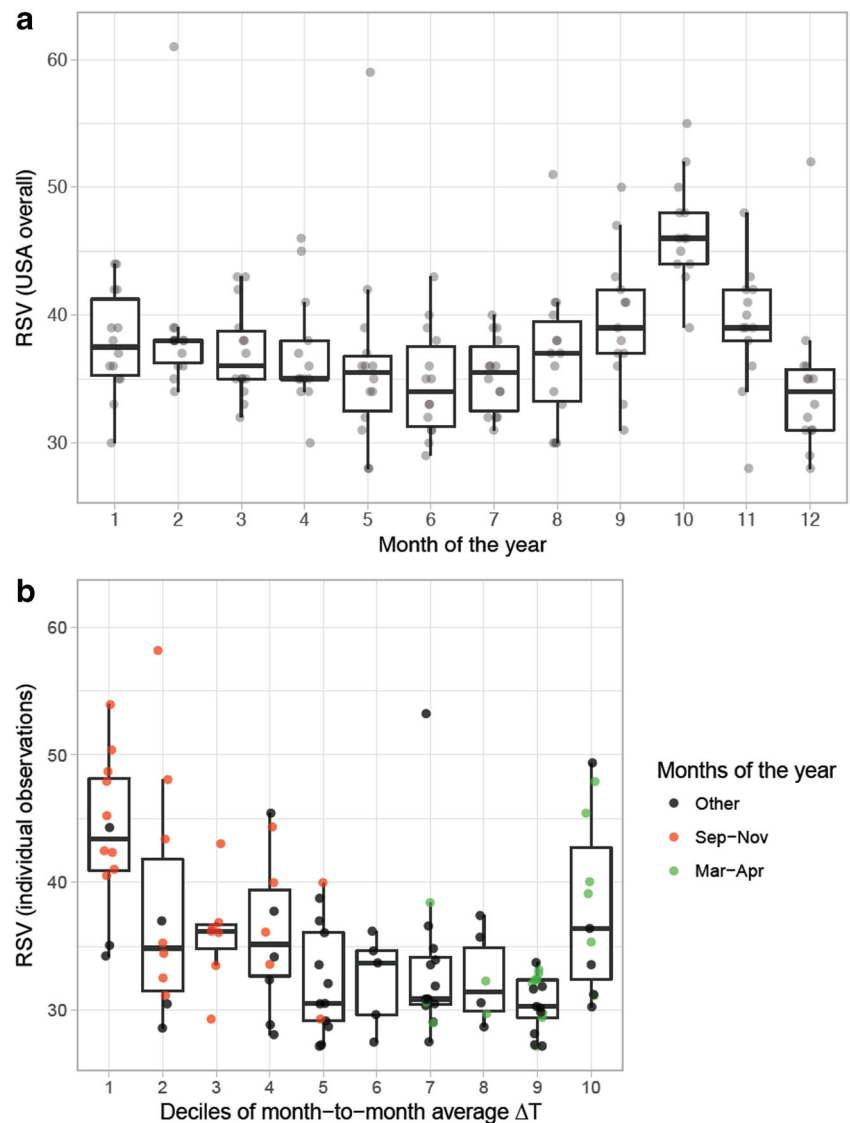
temperature variation seems to be an interesting predictor of GBS search peaks. One possible explanation for this phenomenon could be that sudden large variations of the temperatures accompany incidence peaks for seasonal infectious antecedents of GBS (e.g., upper respiratory tract infections in cold season, *Campylobacter jejuni* gastroenteritis in warm season). The interval (some weeks on average) [12] from the infectious antecedent to GBS onset and the subacute course of the disease could result in a delay from disease onset to Google search for the disease. Therefore, we hypothesize that the observed autumn peak may relate to a summer *C. jejuni*-related incidence peak [12], while the observed spring peak may be linked to a winter influenza virus-related peak [13].

Comparison of our findings to traditional epidemiology data is made difficult by the absence of analogous population-based incidence data for external validation; hence, our work constitutes only a seminal, proof-of-concept work in the field. Another limit of web search-based approach

may arise from a boost in web search popularity caused by mass-media or extensive social-media coverage of a particular topic. We minimized this bias by analysing only the 10-year-averaged monthly data. In addition, we sought proof of internal consistency by comparing data from different geographical regions in the same timeframe, thereby overcoming the effect of geographically limited hype.

One main issue regarding the previous studies of incidence and epidemiology of GBS is that they have been conducted mainly on limited size populations. Conversely, a distinctive feature of this work is that, although we used a surrogate measure of incidence, search engine data constitute a very large and geographically distributed sample.

In previous works assessing disease dynamics with web-based tools, data deriving from GT showed impressive correlation with actual incidence trends [7, 14] and reflected accurately traditional epidemiologic features of neurologic and

non-neurologic diseases [8, 9]. This data may offer timely information that, in the influenza field, was observed to possibly even precede traditional public health control systems [7, 14].

Conclusions

This study is the first search engine-based assessment to date of Guillain-Barré syndrome epidemiology. We observed repeatable seasonality patterns emerging from Google Trends searches in USA, with autumn and, at a lesser extent, spring showing the highest peaks for GBS. We correlated these peaks with the largest periods of average temperature change along the year, possibly reflecting increased susceptibility to infectious antecedents of GBS. However, at the present time, it is not possible to conclude if the data deriving from our analysis may indicate that digital epidemiology may precede traditional systems in identifying GBS outbreaks; further studies are needed to assess this topic.

Authors' contributions A.G. and M.V. equally contributed to the conception and design of the study; acquisition, analysis and interpretation of data; and drafting of the article. G.D.C. contributed to the conception and design of the study. F.C., G.C., V.M. and R.F. contributed to the analysis and interpretation of data. All authors contributed to the critical revision of the article.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Van den Berg B, Walgaard C, Drenthen J, Fokke C, Jacobs BC, van Doorn PA (2014) Guillain-Barré syndrome: pathogenesis, diagnosis, treatment and prognosis. Nat Rev Neurol 10(8):469–482. https://doi.org/10.1038/nrneurol.2014.121

2. Wijdicks EF, Klein CJ (2017) Guillain-Barré syndrome. Mayo Clin Proc 92(3):467–479. https://doi.org/10.1016/j.mayocp.2016.12.002

3. Alshekhlee A, Hussain Z, Sultan B, Katirji B (2008) Guillain-Barré syndrome: incidence and mortality rates in US hospitals. Neurology 70(18):1608–1613. https://doi.org/10.1212/01.wnl.0000310983.38724.d4

4. Shui IM, Rett MD, Weintraub E, Marcy M, Amato AA, Sheikh SI, Ho D, Lee GM, Yih WK, Vaccine Safety Datalink Research Team (2012) Guillain-Barré syndrome incidence in a large United States cohort (2000-2009). Neuroepidemiology 39(2):109–115. https://doi.org/10.1159/000339248

5. Webb AJ, Brain SA, Wood R, Rinaldi S, Turner MR (2015) Seasonal variation in Guillain-Barré syndrome: a systematic review, meta-analysis and Oxfordshire cohort study. J Neurol Neurosurg Psychiatry 86(11):1196–1201. https://doi.org/10.1136/jnnp-2014-309056

6. Kwong JC, Vasa PP, Campitelli MA, Hawken S, Wilson K, Rosella LC, Stukel TA, Crowcroft NS, McGeer AJ, Zinman L, Deeks SL (2013) Risk of Guillain-Barré syndrome after seasonal influenza vaccination and influenza health-care encounters: a self-controlled study. Lancet Infect Dis 13(9):769–776. https://doi.org/10.1016/S1473-3099(13)70104-X

7. 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

8. Dalla Costa G, Giordano A, Romeo M, Sangalli F, Comi G, Martinelli V (2018) Digital epidemiology confirms a latitude gradient of MS in France. Mult Scler Relat Disord 20:129–131. https://doi.org/10.1016/j.msard.2018.01.009

9. Moccia M, Palladino R, Falco A, Saccà F, Lanzillo R, Brescia Morra V (2016) Google Trends: new evidence for seasonality of multiple sclerosis. J Neurol Neurosurg Psychiatry 87(9):1028–1029. https://doi.org/10.1136/jnnp-2016-313260

10. Sipilä JOT, Soilu-Hänninen M, Ruuskanen JO, Rautava P, Kytö V (2017) Epidemiology of Guillain-Barré syndrome in Finland 2004-2014. J Peripher Nerv Syst 22(4):440–445. https://doi.org/10.1111/jns.12239

11. Sivadon-Tardy V, Orlikowski D, Rozenberg F, Caudie C, Sharshar T, Lebon P, Annane D, Raphaël JC, Porcher R, Gaillard JL (2006) Guillain-Barré syndrome, greater Paris area. Emerg Infect Dis 12(6):990–993. https://doi.org/10.3201/eid1206.051369

12. Rees J, Soudain S, Gregson NA, Hughes RAC (1995) *Campylobacter jejuni* infection and Guillain-Barré syndrome. N Engl J Med 333:1374–1379. https://doi.org/10.1056/NEJM199511233332102

13. Lehmann HC, Hartung HP, Kieseier BC, Hughes RA (2010) Guillain-Barré syndrome after exposure to influenza virus. Lancet Infect Dis 10(9):643–651. https://doi.org/10.1016/S1473-3099(10)70140-7

14. Olson DR, Konty KJ, Paladini M, Viboud C, Simonsen L (2013) Reassessing Google Flu Trends data for detection of seasonal and pandemic influenza: a comparative epidemiological study at three geographic scales. PLoS Comput Biol 9(10):e1003256. Published online 2013 Oct 17. https://doi.org/10.1371/journal.pcbi.1003256