ORIGINAL RESEARCH

Laryngoscope
Investigative Otolaryngology

Use of Google Trends to evaluate for environmental variations in search terms for benign paroxysmal positional vertigo

Josh R. Sen BA 🔗 | **Alex J. F. Tampio MD** 🔗 | **Shaelyn M. Cavanaugh BS, MPH** | **Brian D. Nicholas MD**

Department of Otolaryngology, SUNY Upstate Medical University, Syracuse, New York, USA

Correspondence
Josh R. Sen, SUNY Upstate Medical University, Department of Otolaryngology, 750 East Adams Street, Syracuse, NY 13210, USA.
Email: senj@upstate.edu

Abstract

Objective(s): Climate variables are implied in the pathogenesis of certain otologic diseases, including benign paroxysmal positional vertigo (BPPV). Using internet search data obtained through Google Trends (GT), we explored the relationship between climate patterns and symptom search frequencies for BPPV. We hypothesized that increased latitude, as a proxy for decreased sunlight exposure, would lead to increase in BPPV symptom searches.

Methods: GT searches for symptoms related to BPPV were obtained for five U.S. cities of different latitudes via the Google Trends online interface. Comparisons were made using SPSS via ANOVA analysis. Figures were made using Microsoft Excel.

Results: Searches for BPPV-related symptoms increased with increasing latitude. BPPV-related symptoms did show seasonal variations, but not in predictable manners.

Conclusion(s): GT may be a viable research tool when comparing geographical differences in searches, but may be less sensitive in detecting time dependent changes. We offer suggestions as to how big data tools may be altered for research purposes.

Level of Evidence: NA.

KEYWORDS
benign paroxysmal positional vertigo, big data, Google Trends, latitude, vitamin D

1 | INTRODUCTION

Google Trends is a publicly available big data tool that provides relative frequency of searches for specific terms a different time periods and in specific locations. Current events can spur an increase in search frequency for specific terms at specific times of year (ie, elections in the fall). Google Trends has previously been used to track influenza and dengue epidemiology in an official capacity.[1] It has also been used

to investigate interest and epidemiology in a myriad of other diseases.[2] In addition, Google Trends can be used to compare geographic and time-dependent variations. In the field of otolaryngology, prior studies have used Google Trends to evaluate environmental relationships with epistaxis, but otherwise the tool remains relatively unexplored in the field.[3] Seeking a disease that have suspected environmental relationships with symptom incidence, we settled upon Benign Paroxysmal Positional Vertigo (BPPV).

BPPV is a disorder of the inner ear that is characterized by episodic vertigo and nausea with exacerbations related to changes in head

position.[4] It is the most common type of vertigo by incidence. Risk factors for BPPV include advanced age and female sex. The pathophysiology of the disease involves displacement of otoconia within the vestibular otolith organs. Further risk factors and associations have been suggested including seasonality, temperature, osteoporosis, and vitamin D deficiency, however, these symptom-environment associations can be difficult to study when analyzing data from single institutions.[5-8] Therefore, we opted to use GT as a population-level tool to determine if these environmental factors influenced BPPV symptomatology. The risk factors listed for BPPV implicitly correlate to sun exposure, as we know that vitamin D deficiency is more common at higher latitudes.[9] Since the literature suggests that BPPV symptoms may correlate with low vitamin D levels, we hypothesized there to be an increased incidence of BPPV-related search frequencies at higher latitudes due to reduced sunlight exposure and corresponding reduced vitamin D levels. Additionally, we predicted that symptoms related to BPPV would be searched in a seasonal manner, with peaks and troughs for different terms occurring at similar times of the year.

2 | METHODS/MATERIALS

Symptoms of BPPV were searched in Google Trends for Miami, Los Angeles, Washington D.C., New York City, and Seattle. Cities were selected as they contained adequate GT data and varied in latitude at roughly equal intervals. Weekly average search frequencies were obtained from 01/01/2016 to 5/27/2018. Daily search frequencies were then obtained from 01/01/2016 to 12/31/2016. The terms searched were Dizziness, Vertigo, BPPV, Meclizine, and Nausea. Terms with insufficient Google Trends data were excluded. Analysis

over time and across geographies was performed using ANOVA analysis in SPSS. Microsoft Excel was used to make figures. This research is exempt from Institutional Review Board approval as it does not involve human or animal subjects.

3 | RESULTS AND ANALYSIS

Analysis of searches at varying latitudes for 5 U.S. cities at varying latitude demonstrated latitude dependent differences in searches for BPPV related symptoms. Some symptoms related to BPPV increased with increasing latitude (Figure 1). These terms were "dizziness," "vertigo," and "nausea." Higher latitudes tended to demonstrate statistically significant differences when compared to Miami values as a baseline ($P < .05$) (Figure 1). "BPPV" also demonstrated increasing search frequency with increasing latitude, but this was not statistically significant (Figure 1). On the other hand, "Meclizine" showed statistically significant decreasing frequency with increasing latitude to $P < .05$ (Figure 1).

We attempted to detect seasonal changes in BPPV related searches by graphically visualizing search frequencies throughout a calendar year (Figure 2A-E). There are seasonal fluctuations with individual terms, but they are not consistent among different geographies. They also do not statistically differ from baseline January values except in 3/330 comparisons (Figure 2B,D, $P < .05$).

4 | DISCUSSION

To frame our results, it is essential to discuss the manner in which Google Trends relays search data information. Google Trends allows
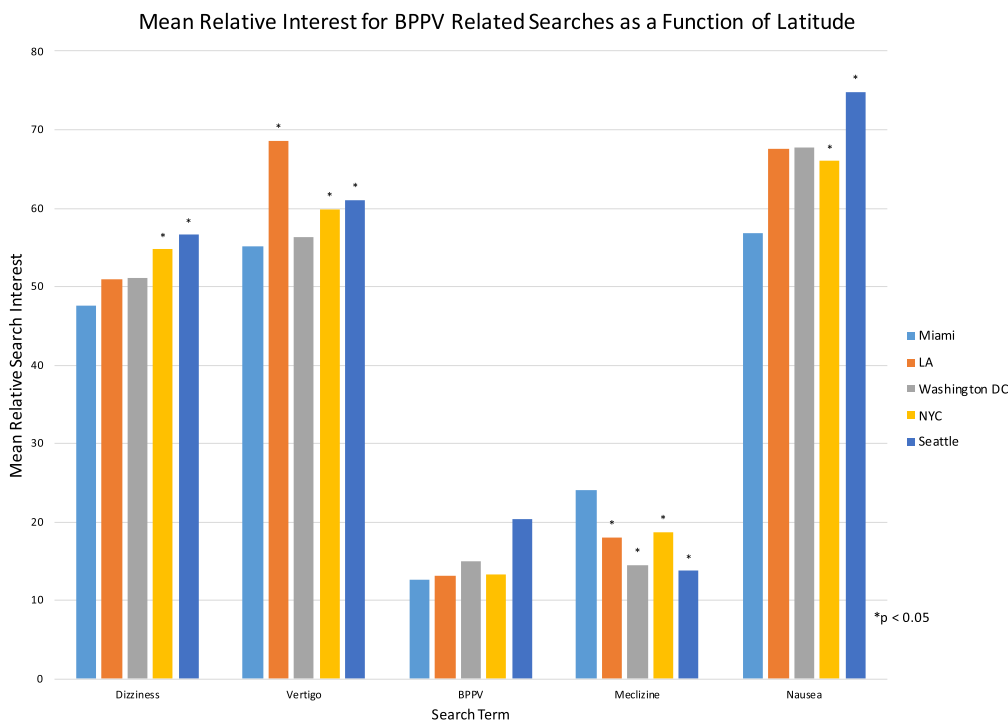
FIGURE 1 ANOVA analysis of search frequencies of BPPV related symptoms. Symptoms included were "dizziness," "vertigo," "nausea," "BPPV," and "meclizine." Cities used were Miami, LA, Washington DC, NYC, and Seattle. Miami was used as a baseline in these comparisons as it was the lowest latitude city
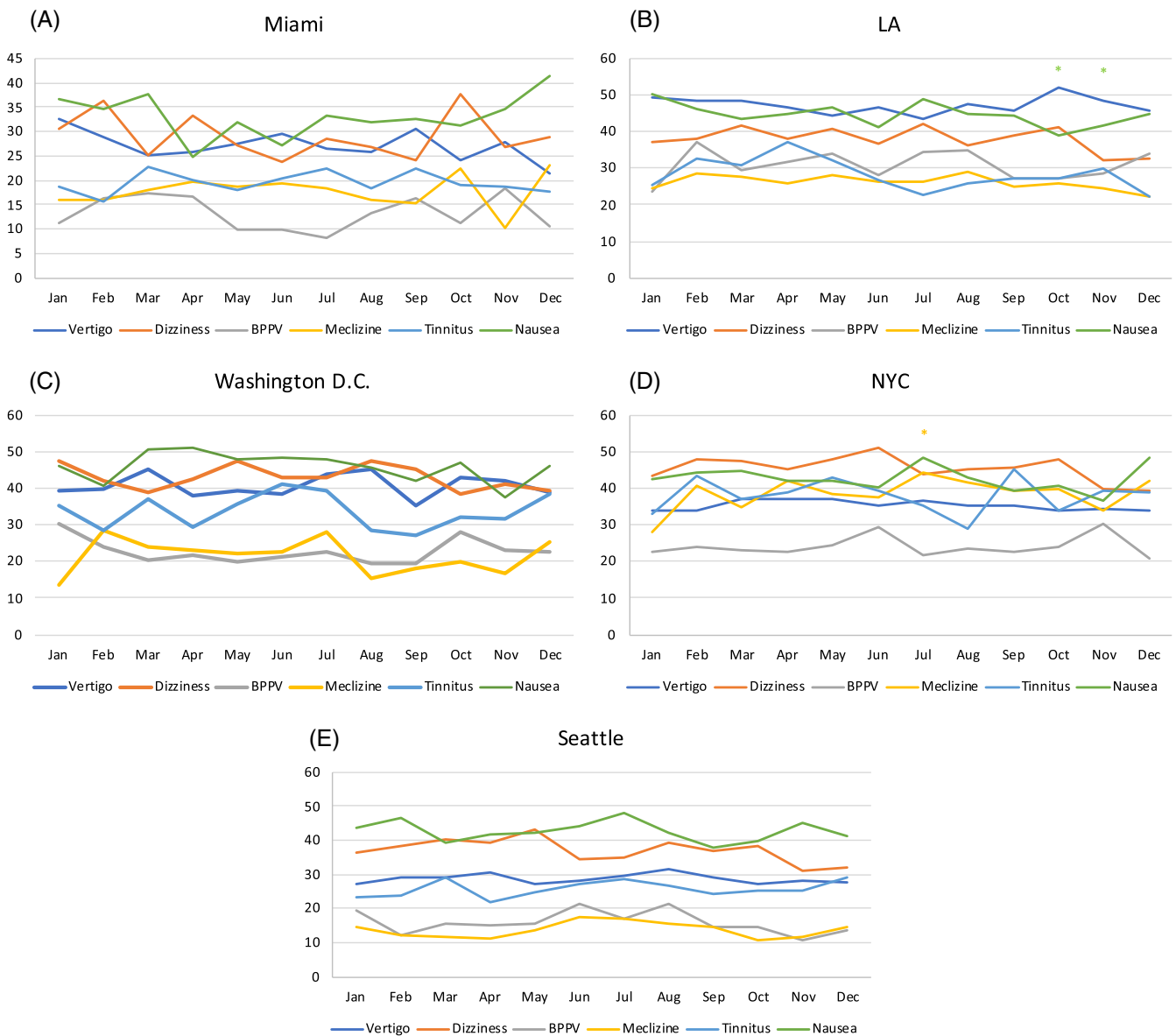
FIGURE 2 A-E, Graphical visualization of monthly averages of daily search data. ANOVA comparison of monthly averages (using January values as baseline) were statistically significant to $P < .05$ for months demonstrated by asterisk

for a specific term to be searched in a town, city, state, or country for a set period of time. Depending on the period of time used, the granularity of each individual data point changes. For example, if a three-month period of time is used, each individual data point will reflect 1 day of data, ultimately yielding ~90 data points for that search. If a whole year of data is searched, each data point reflects a week worth of data, so ~52 data points will be obtained. There are pros and cons to both methods. Having daily data points allows for sensitive detection of day-to-day changes, but this comes at a cost. Google Trends has an inherent minimum threshold for search data to be reported. If the true value is below that threshold, the data point reported is "0," but this does not necessarily mean that no searches were performed that day. In contrast, a week of data is more likely to meet a minimum threshold and therefore have useable data to analyze. However, the

effect of daily environmental changes on search frequencies is likely lost in week-long bins of data.

Compounded with the issues mentioned above, Google Trends provides data that is relative, not absolute. Any search that is performed yields data that ranges from 0 to 100. Each individual data point is standardized to the maximum absolute search count in a given query, a value that Google Trends does not make public. Because of this, separate search queries cannot be compared as they have not been standardized by the same maximal value. In addition, though the numbers yielded are not necessarily percentages or ratios, they are normally distributed, allowing for simple statistical analyses to be performed. Google Trends does allow for five searches to be performed simultaneously that can vary by both geographic location and time period. This sort of query will yield data that is compared to itself and

allows for granular data to be obtained for a longer period of time. It also allows for data from different locations to be compared. Even so, there are well known, and sometimes critical limitations of Google Trends and analysis of only specific types of data can be performed.

Google Trends does allow searching in many different geographic locations. As a privacy measure, many of the search frequencies do not meet the above mentioned minimum threshold, so less populated areas often yield little to no useable data. This necessitates research using Google Trends to be performed in big cities with substantial populations to allow for adequate data.

The final aspect of Google Trends that warrants discussion is the nature of the search inputs. There are two search inputs that can be made in Google Trends known as "search terms" and "topics." "Search terms" are inputs that yield data containing the entered phrase. Therefore, the data is more specific but may miss out on permutations that individuals may search when looking up their symptoms. "Topics" are a grouping of related terms, but it is unclear which terms are included and excluded in the data, so while these queries can yield broader data, they lack specificity. We chose to use search terms in our BPPV analysis (with the exception of "BPPV" for which a search term was not available). This was because we felt that variations in geography would yield large differences in searches, so we felt comfortable sacrificing the sensitivity of Topics.

Analysis of BPPV related search terms demonstrated variation in a latitude dependent manner. Prior research has demonstrated associations between BPPV and vitamin D deficiency, climate, and seasonality, and specifically, low vitamin D has been implicated in the pathogenesis of BPPV.[5-7] Using latitude as a proxy for variant vitamin D levels (related to frequency and duration of sunlight exposure), this suspected association seems to be represented in Google Trends data. We see that BPPV symptom search frequency increases with increasing latitudes, suggesting that the suspected pathophysiology may be supported by the Google Trends data. However, this trend is not evident for all the terms, namely BPPV and meclizine. There are, however, confounding variables that are difficult to control for with Google Trends data. One example is that cold/flu symptoms may be more common in colder climates, resulting in more ear related symptom searches. Another confounder is age, as BPPV tends to be more common with advancing age. Therefore, a more aged population may be expected to search more for these terms. In contrast, a younger population may be more likely to use internet search engines to research medical terms. Another confounding variable that requires attention is the racial and ethnic make-up of the chosen cities' populations. For example, our searches are limited to English, which may exclude searches from the large Spanish speaking population in Miami. Varying cultural and language differences in symptom searches are nearly impossible to control for in Google Trends, but should be kept in mind when using Google Trends to perform analyses. Overall, we find Google Trends to demonstrate the capacity to detect geographical differences in BPPV symptomatology.

Similarly, vitamin D deficiency can be precipitated by seasonal variations in sunlight exposure. Seasonality and temperature have also been implicated in the pathogenesis of BPPV.[6] Saeed et al. demonstrated that BPPV cases often peak in seasons with colder temperatures.[5] While Google Trends data seems to demonstrate fluctuations over the course of the year, we did not see the expected peaks demonstrated in prior research. Troughs and peaks, if present at all, occurred at different times of the year for different terms, not at similar times as might be expected. This data suggests that Google trends may be less sensitive in detecting seasonal or time dependent changes in BPPV symptoms. It may also suggest that seasonal climactic variations may not affect BPPV related symptoms. The lack of a demonstrable temporal association among search terms is further reflected in our findings for MD.

Overall, Google Trends seems to demonstrate the sensitivity to detect geographical differences in search term frequency. This is useful in the field of otolaryngology, as climate and environment are often implicated in otologic, rhinologic, and allergic pathologies. Seasonal fluctuations demonstrated in prior literature were not as clearly reflected in online data, suggesting that Google Trends may not be a sensitive enough tool to detect time-dependent variations in data. It may be that Google Trends lacks the sensitivity to detect minute changes, especially for relatively uncommon symptoms. The alternative conclusion is that climate and geographical variations may not influence otologic symptom onset, and that prior findings of this nature may have been statistically but not clinically significant. Ultimately, we do not believe that our findings negate previously developed hypotheses. There may, however, be small modifications that would allow Google Trends data to benefit clinical research.

First, absolute search frequencies, as opposed to relative, would be obtained through the data tool—this would allow for stronger analyses and results. In addition, allowing for frequency data over longer periods of time to be obtained would permit comparisons over multiple years. Other researchers should exercise caution depending on the types of variables being used. With ever evolving methods of capturing and storing data, use of viable big data tools may prove beneficial in health care research. As such, we encourage further exploration of this and other such big data tools.

5 | CONCLUSIONS

Google Trends demonstrated the capacity to detect differences in BPPV related symptoms across US latitudes, but failed to detect a seasonal variation. Further research with Google Trends is necessary to elucidate its true potential as a research tool.

ORCID
Josh R. Sen https://orcid.org/0000-0002-7211-4280
Alex J. F. Tampio https://orcid.org/0000-0001-9612-892X

BIBLIOGRAPHY

1. Google Flu Trends. https://www.google.org/flutrends/about. Accessed April 7, 2019.
2. Nuti SV, Wayda B, Ranasinghe I, et al. The use of google trends in health care research: a systematic review. *PLoS One*. 2014;9(10):e109583.
3. Unsal AA, Dubal PM, Pfaff JA, Friedel ME, Eloy JA, Kountakis SE. Doctor Google: correlating internet search trends for epistaxis with metropolitan climates. *Am J Otolaryngol*. 2019;40(3):358-363.
4. Kim JS, Zee DS. Clinical practice. Benign paroxysmal positional vertigo. *N Engl J Med*. 2014;370(12):1138-1147.
5. Saeed BMN, Omari AF. Climatic variations and benign paroxysmal positional vertigo. *J Otol*. 2016;11(1):33-37.
6. Buki B, Ecker M, Junger H, Lundberg YW. Vitamin D deficiency and benign paroxysmal positioning vertigo. *Med Hypotheses*. 2013;80(2):201-204.
7. Holick MF. Vitamin D deficiency. *N Engl J Med*. 2007;357(3):266-281.
8. Byun H, Chung JH, Lee SH, Park CW, Kim EM, Kim I. Increased risk of benign paroxysmal positional vertigo in osteoporosis: a nationwide population-based cohort study. *Sci Rep*. 2019;9(1):3469.
9. Wacker M, Holick MF. Sunlight and vitamin D: a global perspective for health. *Dermatoendocrinology*. 2013;5(1):51-108.