Eczema-related Web Search Data in Sweden: Investigating Search Patterns and the Influence of Climate [AQ1]

Linda TIZEK1,2, Hannah WECKER1,2, Simon SCHNEIDER2, Emma K. JOHANSSON1,3, Yodit GRIMAY1 and Alexander ZINK1,2
1Division of Dermatology and Venereology, Department of Medicine Solna, Karolinska Institutet, Stockholm, Sweden, 2Department of Dermatology and Allergy, Technical University of Munich, School of Medicine, Munich, Germany and 3Department of Dermatology, Karolinska University Hospital, Stockholm, Sweden

As atopic eczema is triggered by environmental factors, such as temperature, differences in disease burden between and within countries are possible. One method to study this phenomenon is to perform web-search analysis, since the internet is commonly used to retrieve health-related information. This study, investigating the Google search volume regarding eczema in Swedish counties between April 2017 and March 2021, revealed a continuous increase in [AQ2] searches and that the search volume was higher in Northern than Southern Sweden. Gotland had the most searches per 100,000 inhabitants. In general, there was a negative correlation between search volume and [AQ3] temperature ($r=-0.315, p<0.001$) and [AQ3] hours of sunshine ($r=-0.213, p<0.001$), whereas there was a positive association between search volume and [AQ4] wind ($r=0.229, p<0.001$). Search engine analysis is a rapid and cost-effective method of examining search behaviour regarding disease among the general population within a country and, thus, can enable the identification of regions with specific interests and needs.

Key words: atopic dermatitis; environment; Sweden; trigger factors; web search analysis.

Accepted Sep 6, 2022; Epub ahead of print Sep 6, 2022
Acta Derm Venereol 2022; 102: XX–XX.
DOI: 10.2340/actadv.v102.2937
Corr: Linda Tizek, Division of Dermatology and Venereology, Department of Medicine Solna, Karolinska Institutet, SE-171 77 Stockholm, Sweden. E-mail: linda.tizek@tum.de

A topic eczema (AE) is a chronic inflammatory skin disease characterized by itchy, dry, reddish, and painful skin (1, 2). These symptoms can severely impair quality of life, with AE estimated to be responsible for 123 [AQ5] disability-adjusted life-years (DALYs) on average [AQ6], making it the skin disease with the highest disease burden. Worldwide, the highest AE burden was found in Sweden, where it resulted in 327 DALYs [AQ5] in 2017 (3).

The 1-year prevalence of AE is estimated to be 22.6% in children and 17.1% in adults (4, 5). AE can be triggered by several factors, such as food allergens, viral infections, and climate conditions. For example, studies have revealed that symptoms of AE are more common when outdoor temperature, humidity, and ultraviolet (UV) exposure are low (6–8). Thus, people living in the northern parts of Europe are at a higher risk for experiencing dry and itchy skin when exposed to winter (6).

Many people affected by AE are undiagnosed or not in medical care (9, 10). One method to study the interest of a general population is search engine analysis, as the internet is one of the most important sources of health-related information (11). Correlations between Google search volume and registry data can be used to identify unmet needs, regional differences, and seasonal trends (12–15).

The aim of this study was to investigate the search volume related to eczema in the general population in Swedish counties, to determine whether there are temporal and geographical differences in frequency, and whether there is an association with weather-related parameters.

SIGNIFICANCE

Analysis of web-search behaviour is an easy-to-perform method of disease surveillance. Google data analysis can help us to understand the search behaviour of a population regarding specific topics, since relevant search terms, including their frequency, can be assessed. This study of search behaviour regarding eczema in Sweden showed that the number of searches increased continuously from 2017 to 2021, and increased enormously during the COVID-19 pandemic. People living in northern counties of Sweden searched proportionally more often compared with those living in southern counties. This could be due to a higher disease burden, which is triggered by several environmental factors, such as [AQ7] temperature and precipitation.

MATERIALS AND METHODS

A retrospective longitudinal study using Google Ads Keyword Planner was conducted to identify the web search volume of keywords related to eczema in all Swedish counties ($n=21$). This tool is primarily used for marketing campaigns, but several studies have leveraged web search data to investigate disease-related factors among the general population (5, 12, 16–18). In the current study, the Swedish word “eksem” (eczema) was entered into the tool because it is the word mainly used by the Swedish population to refer to AE. The tool then provided a list of search terms associated with eczema (e.g. atopic eczema scalp, seborrhoeic eczema alternative treatment) including their monthly search volume (i.e. the number of searches performed) for the last 48 months. All identified keywords were qualitatively assessed to check their relevance [AQ8]. None of the keywords were excluded from analysis due to irrelevance. The total number of searches from April 2017 to March 2021 was assessed.
Data on the population statistics of the respective county (mean age, proportion of males, proportion of foreign-born inhabitants, and population density) were obtained from Statistics Sweden (19) to investigate their correlation with the number of searches. Furthermore, the influence of weather parameters, the counties were divided into northern (Dalarna, Gävleborg, Jämtland, Norrbotten, Västerbotten, and Västernorrland) and southern counties (Blekinge, Gotland, Halland, Jönköping, Kalmar, Kronoberg, Örebro, Östergötland, Skåne, Södermanland, Stockholm, Uppsala, Värmland, Västmanland, and Västra Götaland), based on a lower mean temperature in the north and higher in the south. The weather-related parameters were obtained from the Swedish Meteorological and Hydrological Institute (20), and included wind speed in m/s, temperature in °C, mean sunshine duration in h, and precipitation in mm. Since no data for entire counties were available, weather data from the city with the largest population in each county were used.

The study used publicly available data, and thus no institutional review board approval or informed consent were necessary.

Statistical analysis

The search volume of each county was calculated per 100,000 inhabitants (19). Mean values of search volume are reported with the respective standard deviation (SD). All variables were checked for normal distribution. Differences in the search volume per year were assessed using Welch’s analysis of variance (ANOVA) due to the lack of variance homogeneity. Student’s t-test was used to show differences between Northern and Southern Sweden. Wherever the variance homogeneity assumption of t-test was not fulfilled, Welch’s test was performed. To show differences between counties, the non-parametric Kruskal–Wallis test was applied, followed by pairwise comparisons with the Wilcoxon sum-rank test and Bonferroni correction as post-hoc tests with adjustment for multiple testing. Pearson’s correlation coefficient was calculated to examine the association between search volume and selected weather data. Significant results were assumed by $p < 0.05$. Statistical analysis was performed using R version 4.0.4 (21).

RESULTS

Overall, 423 eczema-related keywords were identified, with a search volume of 4,907,390 searches; i.e. 1,152,419 searches per 100,000 inhabitants were generated during the study period. In each county, the most frequently searched keyword was “eczema”, followed by “atopic eczema”.

Time course and search behaviour

In 2017 ($n=845.41 \pm 180.81$) and 2018 ($n=903.46 \pm 201.49$) the mean search volume was fewer than 1,000 searches per 100,000 inhabitants, while the average [AQ9] in 2021 was $1,687.57 \pm 252.03$ ($p < 0.001$) searches per 100,000 inhabitants. Thus, an increase of 99.6% was observed for search volume during the whole study period (Fig. 1).

Seasonal patterns were also observed in addition to the overall increasing searching volume, with search volume increasing in winter and decreasing in summer. For example, January 2020 was the month with the highest search volume in nearly all counties. The search volume then decreased until it reached its lowest value in July 2020, before increasing again at the end of the year. However, this was not the case in Gotland, where the search volume had peaks in both summer and winter (Fig. 2).

Comparison between counties

The highest mean search volume per 100,000 inhabitants was observed in Gotland ($n=1,728.32 \pm 412.87$), which was significantly higher than that in all other counties ($0.001 < p \leq 0.024$). In contrast, Skåne ($n=830.26 \pm 252.37$) had a significantly lower mean search volume per 100,000 inhabitants than all other counties except for Södermanland ($n=1,036.48 \pm 296.66$, $p=0.063$), Östergötland ($n=1,032.38 \pm 294.00$, $p=0.052$), Stockholm ($n=883.06 \pm 301.36$, $p=1.00$), and Västra Götaland ($n=848.56 \pm 290.16$, $p=1.00$, Fig. 3). Analysis revealed a positive correlation between search volume and average [AQ9] age ($r=0.396$, $p<0.001$), whereas the search volume correlated negatively with the proportion of foreign-born inhabitants ($r=-0.308$, $p=0.001$) and the population density ($r=-0.250$, $p=0.010$). No significant correlation was found with the proportion of men ($r=0.148$, $p=0.133$).

A comparison between Northern and Southern Sweden revealed that the mean search volume was higher in the north ($1,211.25 \pm 333.99$ searches/100,000 inhabitants) than in the south ($1,116.08 \pm 369.09$ searches/100,000 inhabitants, $p<0.001$).

Association with weather-related parameters

The search volume decreased monotonically with increasing temperature and sunshine hours. In contrast, an increase in search volume with increasing wind speed was found. For precipitation, a slight linear decrease
in searches with increasing precipitation, whereby the northern trend line was constant and the southern was decreasing, was noted (Fig. 4). Thus, a significant negative association between search volume per 100,000 inhabitants and temperature ($r=-0.315$, $p<0.001$), and sunshine hours ($r=-0.213$, $p<0.001$) was observed, whereas wind speed was positively associated with the search volume per 100,000 inhabitants ($r=0.229$, $p<0.001$). No significant correlation was found with precipitation ($r=-0.055$, $p=0.131$).

**DISCUSSION**

In the general population of Swedish counties, it was found that the mean search volume related to eczema was higher in Northern than in Southern Sweden. However,
Gotland, in Southern Sweden, had the highest search volume per 100,000 inhabitants of all counties. The average search volume per year increased by nearly 100% from 2017 to 2021 and was usually higher in winter than in summer. Thus, a negative association between temperature, and sunshine hours was detected.

In Sweden the search volume per 100,000 inhabitants was greater than twice that in Germany (17). One explanation for the higher relative search volume in Sweden may be the higher disease burden: in 2019, a DALY rate of AE of 131.45 per 100,000 people was reported for Sweden and 115.44 for Germany (22). These results indicate that web search data may present another method to investigate and compare the burden of disease in different geographical areas, with the added advantage of being easy to collect and inexpensive.

In general, the DALY rate of AE remained relatively constant in Sweden between 2017 and 2019 (22). The search volume, however, increased continuously during the study period, reflecting the increasing importance of the internet as a source of health information (23). This was especially the case during the COVID-19 pandemic (24). The number of personal consultations decreased considerably, while the usage of online solutions, such as teledermatology and web searches, increased significantly (25, 26). Performed by physicians, teledermatology was shown to be comparable to conventional in-person consultations and to could reduce the need for personal appointments (27). One study, however, showed that many websites reached via search engine do not provide evidence-based information (28). Since Iglesias-Puzas et al. (28) showed that websites that they categorized as providing imprecise information have a higher mean number of interactions than websites providing precise information, it is important that the public is aware of websites with trustworthy information, particularly when seeking information on managing a disease.

Studies have shown that there is a positive correlation between Google search volume and the prevalence of some diseases (15, 29). Because, to date, there is no register-based study representing the prevalence of AE in the respective Swedish counties, no correlation analysis between search volume and disease prevalence could be performed. Although the prevalence of AE is generally higher in children than in adults (4, 5, 13), a positive association was found between average age and the number of searches in this study. A possible explanation for this finding could be that parents of affected children commonly used the internet, because children are not able to conduct online searches on their own. People living in rural areas have been shown to use search engines less frequently than those living in urban areas (23), which was also noted in the current study, as the search volume correlated negatively with population density. Furthermore, a negative association with proportion of foreign-born inhabitants was found. Since only the Swedish term was used to determine eczema-related search volume in this study, this may explain why the search volume was lower in counties with a higher proportion of foreign-born inhabitants. Considering the Swedish counties, the search volume was surprisingly high in Gotland.

Fig. 4. Scatterplot of the eczema-related search volume and (a) temperature, (b) wind, (c) precipitation, and (d) sunshine in North and South Sweden with linear trend lines.
in comparison with that in all other southern counties. This could be due to the high number of people from other Swedish regions visiting Gotland in the summer, leading to a temporary increase in the number of people conducting online searches who may not be registered as residents. Thus, the calculation of the search volume per 100,000 inhabitants may be distorted. Supporting this, Gotland had a different seasonal pattern in search trends, as there were peaks in search volume in summer and winter. In addition to genetic factors, clinical experience has shown that environmental factors are responsible for differences in AE burden, since the prevalence can vary widely between and within countries (30, 31). Several studies have examined the relationship between weather parameters and flare-ups of AE. Although one study has indicated that the burden of AE is higher in summer due to increased sweating (31), the majority of studies have suggested that the risk of AE flares increases with low outdoor temperature (6–8). February and October were found to be the months with the highest incidence of AE (6). The current study found a negative correlation between search volume and [AQ13] temperature, which also indicates that the need for information and thereby the disease burden were higher in winter. Comparable results were presented by a German study using search engine data for AE (17). Another environmental factor that was reported to influence AE symptoms is sun exposure. According to the literature, sun exposure has a protective effect on AE (6). We observed that the search volume decreased with increasing sunshine duration. Regarding precipitation, studies performed in conventional medical settings have demonstrated that the prevalence of AE symptoms correlated positively with [AQ14] precipitation (6, 7). In contrast, this study showed that search volume negatively correlated with precipitation, which was also demonstrated by the results of the aforementioned German study on internet search volume (17) [AQ15]. One explanation for these different findings may be that the data used in this study are only available on a monthly basis and thus may not accurately reflect the influence of precipitation on AE as accurately as the influence of other environmental parameters. Nevertheless, using publicly available observational data is a viable approach to studying the burden of a disease in a general population.

This study has some limitations. The analysis only considered data from people who have access to the internet, who use Google, and who perform online searches in Swedish. Thus, the search behaviour of people who live in Sweden but do not have Swedish as their first language is not adequately covered. Furthermore, the tool does not provide any demographic information about the search engine users (e.g. age, sex, education). Thus, it is not possible to determine which searches were conducted by affected individuals, their relatives, or healthcare workers. As different groups are likely to have different motivations for their searches (e.g. parents having a child with AD vs parents who are just worried [AQ16]), demographic information would have provided further insight into various search patterns. Data from other social media, such as Instagram, which is often used by younger people, was not collected, and this may have led to skewed data. Another limitation is that the search volume is provided and estimated by Google without any possibility for verification by users. Furthermore, the calculated search volume per 100,000 is based on the number of inhabitants according to population registers and not on the actual number of people staying in the respective region. This can lead to a falsely high search volume in areas with increased tourism.

This study demonstrated that leveraging web search data is an easy-to-use approach to identify the needs of a general population, revealing regions and periods of time with particularly high interest in eczema. Since these cost-effective data are available in real-time compared with data from conventional medical settings, they can provide valuable insight into recent trends in disease activity as well as factors of importance for flare-ups. [AQ17] Considering the negative association with several weather-related parameters, it may be useful to provide more eczema health information online during the colder months. In addition, the study indicated that there is an increasing usage of search engines to find information on AE. Therefore, in order to better reach their target audience, public awareness campaigns should strongly consider the increasing relevance of the internet.

ACKNOWLEDGEMENTS

The study was funded by the Division of Dermatology and Allergy, Department of Medicine Solna, Karolinska Institutet, Stockholm, Sweden and the Department of Dermatology and Allergy, Techni- cal University of Munich, Munich, Germany.

Conflicts of interest: EKJ has received speaker honoraria from and/or been a consultant for AbbVie, ACO, Galderma, Galenica, LEO Pharma, Novartis, and Sanofi-Genzyme. AZ received an unrestricted research grant from Leo Pharma and has been an advisor and/or received speaker’s honoraria of the following companies related to eczema: Eli Lilly, Leo Pharma, Pfizer, and Sanofi-Aventis. All other authors have no conflicts of interest to declare.

REFERENCES

Author queries

Dear Author,
Some questions have arisen during the preparation of your manuscript for typesetting. These are marked in the text by [AQ#]. Please consider the points below and make any corrections required.

**AQ1:** A better term would be “weather-related factors”, not “climate” (as this is the global pattern of weather). Please change title wording.

**AQ2:** “number of searches”? 

**AQ3:** Sense? Insert “increasing”? 

**AQ4:** Sense? Change to “increasing wind speed”? 

**AQ5:** Units? Is this DALYs per 100,000 population? (Clearly it can’t be “per person”). 

**AQ6:** Delete the term “average”, change to “a mean of …DALYS”? Use mean, median, or mode, not “average”.

**AQ7:** “increased”? 

**AQ8:** In order that your study can be replicated, please explain how you performed this qualitative check. i.e. Method.

**AQ9:** “mean”? Use mean, median, or mode, not “average”.

**AQ10:** To be clear, do you mean “Thus, there was a negative correlation between search volume and the following factors: increasing temperature and increasing sunshine hours.”

**AQ11:** Do you mean “inaccurate”? 

**AQ12:** Do you mean “accurate”? 

**AQ13:** Sense? Insert “increasing”? 

**AQ14:** Sense? Insert “higher levels of ….”? 

**AQ15:** Comments from the reviewer: I believe “there was no significant correlation between search volume and precipitation”, would state that instead.

**AQ16:** Insert “about AD” or “about their child”? 

**AQ17:** Given the large number of limitations discussed above, it would be helpful if you added a sentence here explaining this in your Conclusion, such as “However, data collected from analysis of search engine use should be analysed with caution, based on the limitations discussed above.”

Many thanks.