



## Are online searches for the novel coronavirus (COVID-19) related to media or epidemiology? A cross-sectional study



Tomasz Szmuda<sup>a,1</sup>, Shan Ali<sup>b,1,\*</sup>, Tarjei Vevang Hetzger<sup>b</sup>, Philip Rosvall<sup>b</sup>, Paweł Słoniewski<sup>a</sup>

<sup>a</sup> Neurosurgery Department, Medical University of Gdansk, Dębinki 7, Gdansk, 80-952, Poland

<sup>b</sup> Scientific Circle of Neurology and Neurosurgery, Medical University of Gdansk, Dębinki 7, Gdansk, 80-952, Poland

### ARTICLE INFO

#### Article history:

Received 25 April 2020

Received in revised form 6 June 2020

Accepted 8 June 2020

#### Keywords:

Google Trends

Risk communication

COVID-19

Europe

Surveillance

### ABSTRACT

**Background:** Previous studies on the novel Coronavirus (COVID-19) have found strong correlations between online searches and the epidemiology of the disease.

**Aim:** Our aim was to determine if online searches for COVID-19 related to international media announcements or national epidemiology.

**Methods:** Searches for “coronavirus” were made on Google Trends from December 31, 2019 to April 13, 2020 for 40 European countries. The online COVID-19 searches for all countries were correlated with each other. COVID-10 epidemiology (i.e. incidence and mortality) was correlated with the national online searches. Major announcements by the World Health Organization (WHO) were taken into consideration with peaks in online searches. Correlations were made using Spearman's rank correlation coefficient.

**Results:** Overall, the online searches for COVID-19 were not correlated with the actual incidence and mortality of COVID-19. The mean Spearman correlation for incidence was 0.20 (range –0.66 to 0.76) and for mortality was 0.35 (range –0.75 to 0.85). Online searches in Europe were all strongly synchronized with each other; a mean Spearman correlation of 0.93 (range 0.62 to 0.99).

**Conclusions:** Online searches for COVID-19 in Europe are not correlated with epidemiology but strongly correlated with international WHO announcements. Our study challenges previous Google Trends studies and emphasizes the role of the WHO in raising awareness of a new disease.

© 2020 The Authors. Published by Elsevier Ltd on behalf of International Society for Infectious Diseases.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

### Introduction

On December 31, 2019, the first cases of an unknown cause of pneumonia, in Wuhan, People's Republic of China, were reported to the World Health Organization (WHO). It was later dubbed the novel coronavirus, SARS-CoV-2, which caused Coronavirus disease (COVID-19). COVID-19 rapidly spread, being declared a health emergency of national concern by WHO on January 30, 2020, and finally being declared a pandemic on March 11, 2020 (World Health Organization).

After the outbreak of COVID-19, much online research was done on the spread of the virus; some was conducted with Google Trends (among other modern analytical tools). Google trends

allows one to view the relative search frequency of various search terms worldwide. It gathers its data from the Google search engine, which is the most common search engine in the world, with a worldwide market share of over 90% (Statcounter Global Stats). Several studies have used Google Trends and/or other search engine data to track the outbreak of COVID-19 in Taiwan, China, Europe and elsewhere (Husnayain et al., 2020; Li et al., 2020; Lin et al., 2020; Mavragani, 2020; Walker et al., 2020). All of these studies have compared the online searches for COVID-19 (or related terms) with the national epidemiology of COVID-19. Significant positive correlations between the epidemiology and online search frequency were found in all these studies; the conclusion was that search engines could be used as robust tool for online surveillance. However, these studies may have been subject to bias as they did not comprehensively cover all countries on a continent and only selectively picked a handful of countries for analysis. Moreover, previous studies showed that Google Trends is influenced more by “media clamor than by true epidemiological burden (Cervellin et al., 2017).” With the heavy worldwide

\* Corresponding author at: Scientific Circle of Neurology and Neurosurgery, Medical University of Gdansk, Dębinki 7, Gdansk, 80-952, Poland.

E-mail address: [shanali@gumed.edu.pl](mailto:shanali@gumed.edu.pl) (S. Ali).

<sup>1</sup> Contributed equally to the manuscript.

attention that COVID-19 is receiving, it is possible that these correlations mentioned before mistakenly analyzed the increasing media clamor for COVID-19 instead of the increasing incidence of COVID-19.

Our aim was to comprehensively assess if online searches for COVID-19 are dependent on the epidemiological burden of the disease or major media announcements. To do this, we evaluated all the countries in Europe that had sufficient data. In this way, our study may confirm to deny the use of Google Trends as a tool for online surveillance in the case of COVID-19.

## Methods and materials

### Search strategy

Searches for “coronavirus” were made as “topic searches” referring to the virus instead of an exact search term. This way, various synonyms, misspellings and translations from different languages were included in the analysis. Data was collected for analysis from December 31, 2019 to April 13, 2020 under the “health” category as shown in [Figure 1](#). All data were collected on April 16, 2020.

From December 30, 2019 to April 13, 2020, Google made up 92% of the search engine market share worldwide and 93% of the market share in Europe. Thus, the online data gathered are robust ([Statcounter Global Stats](#)).

### European countries included

The 44 countries deemed to be in Europe, according to the United Nations, were taken into analysis. These countries were: Albania, Andorra, Austria, Belarus, Belgium, Bosnia and Herzegovina, Bulgaria, Croatia, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Holy See, Hungary, Iceland, Ireland, Italy, Latvia, Liechtenstein, Lithuania, Luxembourg, Malta, Moldova, Monaco, Montenegro, Netherlands, North Macedonia, Norway, Poland, Portugal, Romania, Russia, San Marino, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Ukraine and the United Kingdom. However, some of these countries that tended to be geographically smaller (Holy See, Liechtenstein, Monaco and San Marino) had insufficient data from Google Trends and were not analyzed. Thus, altogether, 40 countries were taken into the analysis.

### WHO announcements

To find WHO announcements that were considered significant to be analyzed with online health patterns we looked on the official

WHO page listing the updates. All dated announcements highlighted on the “summary” sidebar of WHO’s official updates page (on April 16, 2020) were considered significant. There were 3 of these announcements, namely: when a pneumonia of unknown cause was first reported to WHO (on 31 December, 2019), when WHO declared COVID-19 a “Public Health Emergency of International Concern” (on January 30, 2020) and when WHO announced “COVID-19” as the name for the new coronavirus disease (February 11, 2020). Next, any declarations on the epidemiological status of the virus were taken into account. This yielded one unique announcement when WHO declared COVID-19 a pandemic (on March 11). Next, any announcement(s) addressing a particular European country or Europe at-large were considered (this was done since our analysis focuses on Europe); this yielded 2 announcements: when a WHO-led team of experts travelled to Italy (on February 24, 2020) and when Europe was declared the epicenter of the pandemic (on 13 March 2020). Together this yielded 6 announcements.

### Method of analysis

First, the online COVID-19 searches for all countries were graphed and a line of best fit was drawn. Major announcements by the WHO organization were taken into consideration with peaks in online searches. Second, the online COVID-19 searches for all 40 countries were correlated with each other. This was to determine if countries all had their own unique patterns of searches or if they were the same throughout Europe. Third, the COVID-10 epidemiology (i.e. incidence and mortality) was correlated with the online searches in each specific country. This was to determine if online searches correlated more to the epidemiology of COVID-19 in a particular country or to major WHO media announcements. Epidemiological data was provided from publicly available data files from the European Centre for Disease Prevention and Control ([European Centre for Disease Prevention and Control, 2020](#)).

### Statistical analysis

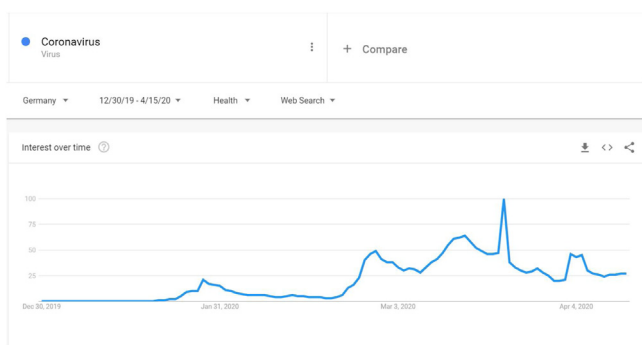
Spearman’s rank correlation coefficient was used to ascertain the strength and direction of a relationship between 2 variables.  $P < 0.05$  was deemed significant. Google Sheets (Google LLC) and Past (Hammer and Harper, Oyvind Hammer, Natural History Museum, University of Oslo) was used for statistical analysis and illustrations.

## Results

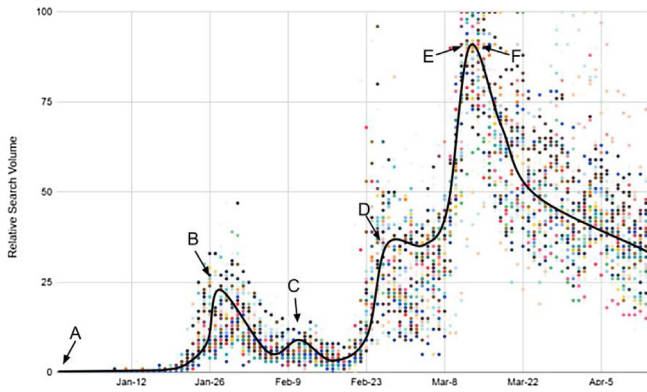
[Figure 2](#) shows that the online search trends for Coronavirus from December 31, 2019 to April 13, 2020 are similar for all 40 European countries. Most countries have a peak in searches on January 30, February 11, February 24 and March 11. On these dates the WHO made major announcements.

[Figure 3](#) shows a Spearman rank correlation analyzing the online search trends of Coronavirus in 40 European countries. The figure shows that European online searches concerning COVID-19 are synchronized with each other since all the correlations are greater than 0.60 with a mean of 0.93 (range 0.62 to 0.99). Belarus, Russia and Italy have slightly weaker correlations (around 0.06) with the rest of the European countries.

The online searches for COVID-19 were correlated with the actual incidence and mortality of COVID-19 in all 40 European countries analyzed. This is shown in [Table 1](#). The mean Spearman correlation for incidence was 0.20 (range  $-0.66$  to  $0.76$ ) and for mortality was 0.35 (range  $-0.75$  to  $0.85$ ).



**Figure 1.** A display of settings for obtaining information about online searches concerning Coronavirus. In this case the trends are shown for Germany. Information for all 40 European countries were gathered in this way.



**Figure 2.** A scatter plot of the relative online search volumes for Coronavirus of 40 European countries from December 31, 2019 to April 13, 2020. A line of best fit is drawn to show the similarities of the trends. Points marked A, B, C, D, E and F signify announcements from the World Health Organization (WHO). (A) On December 31 a pneumonia of unknown cause was first reported to WHO. (B) On January 30, WHO declared COVID-19 a “Public Health Emergency of International Concern”. (C) On February 11, WHO announced “COVID-19” as the name for the new coronavirus disease. (D) On February 24, a WHO-led team of experts travelled to Italy. (E) On March 11, WHO declared COVID-19 a pandemic. (F) on 13 March, Europe was declared the epicenter of the pandemic.

In summary, online trends are more correlated with major WHO announcements (as seen in Figures 2 and 3) than epidemiology of COVID-19 (as seen in Table 1).

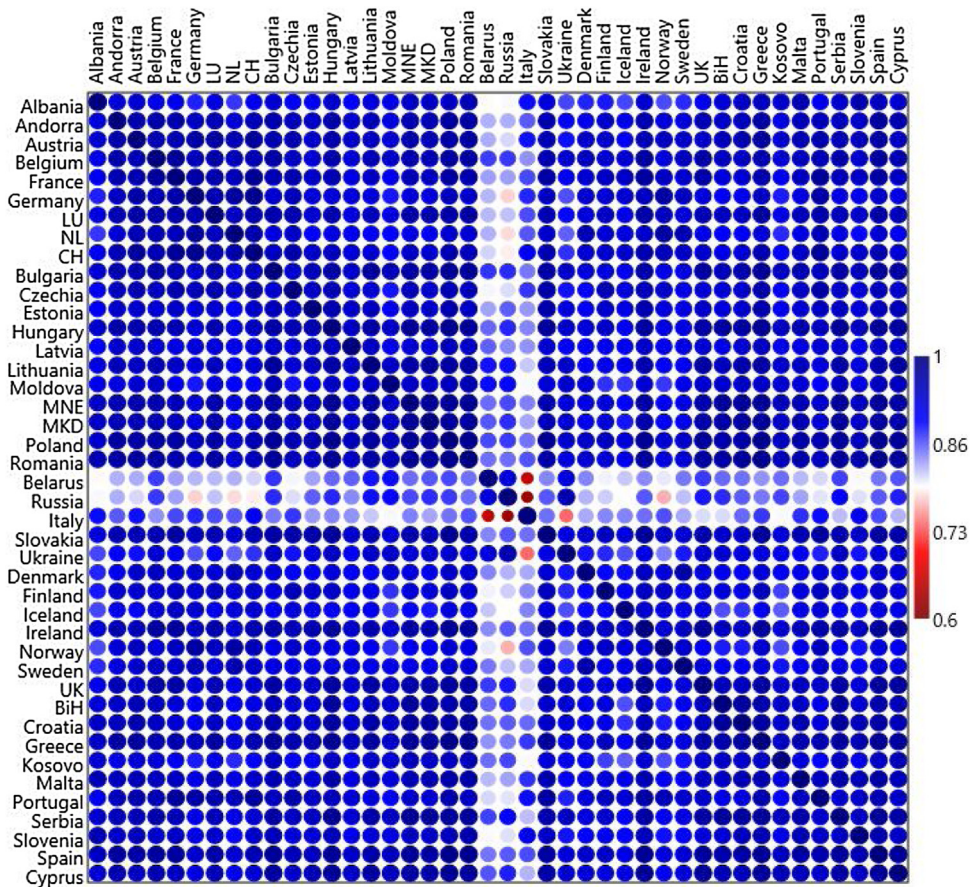
**Discussion**

*Key findings*

We found that online searches for COVID-19 in Europe are weakly correlated with the epidemiology (i.e. incidence and mortality) but strongly correlated with WHO announcements. In other words, the WHO greatly determines the search trends across Europe. This finding emphasizes the role of the WHO in raising awareness of a new disease. While several Google Trends studies have already been done on COVID-19, our findings are novel as they highlight that search trends seem to follow international major media announcements rather than national epidemiology. This challenges the current scientific evidence concerning COVID-19 (Husnayain et al., 2020; Li et al., 2020; Lin et al., 2020; Mavragani, 2020; Walker et al., 2020). Moreover, this study analyzed 40 European countries using a comprehensive selection criterion; to date, this is the largest sample pool ever taken into analysis for an online COVID-19 study.

*Context*

We understand that Google trends may be used for real-time surveillance to some extent as many studies have been done in the past suggesting this (Carneiro and Mylonakis, 2009; Seifter et al., 2010; Pervaiz et al., 2012; Alicino et al., 2015; Al-garadi et al., 2016; Hassid et al., 2017; Husnayain et al., 2019; Santangelo et al., 2019).



**Figure 3.** Spearman's rank correlation coefficient values are represented on a spectrum of blue or red dots; blue values are closer to 1.0 (a perfect positive correlation) and red values are closer to 0.6 (a slightly weaker positive correlation). Note that all countries indicate a positive correlation in search trends for Coronavirus (since they are all above 0.6) from December 31, 2019 to April 13, 2020 in selected European countries. Abbreviations: LU, Luxembourg; UK, United Kingdom; NL, Netherlands; CH, Switzerland; BiH, Bosnia and Herzegovina; MNE, Montenegro; MKD, North Macedonia.

**Table 1**

Spearman's rank correlation coefficient values for correlations between the online searches for coronavirus and the mortality and the incidence of coronavirus in 40 European countries from December 31, 2019 to April 13, 2020. Abbreviations: UK, United Kingdom; BiH, Bosnia and Herzegovina; MKD, North Macedonia.

Country	Spearman's Rank correlation coefficient	
	Searches and mortality	Searches and incidence
Albania	-0.49	-0.57
Andorra	-0.11	-0.61
Austria	0.46	0.77
Belarus	0.38	0.68
Belgium	0.61	0.81
BiH	-0.54	-0.53
Bulgaria	-0.19	-0.14
Croatia	0.34	0.74
Czechia	0.26	0.69
Denmark	0.52	0.82
Estonia	0.24	0.72
Finland	0.25	0.63
France	0.73	0.83
Germany	0.51	0.80
Greece	0.58	0.79
Hungary	-0.3	-0.44
Iceland	0.1	0.73
Ireland	0.5	0.77
Italy	0.67	0.73
Latvia	-0.1	-0.16
Lithuania	0.39	0.71
Luxembourg	0.4	0.69
Malta	-0.21	-0.55
Moldova	0.28	-0.08
Montenegro	-0.33	-0.30
Netherlands	0.55	0.72
MKD	0.71	0.41
Norway	0.36	0.74
Poland	-0.54	-0.45
Portugal	-0.66	-0.53
Romania	0.45	0.78
Russia	0.59	0.79
Serbia	-0.65	-0.75
Slovakia	-0.17	-0.3
Slovenia	-0.45	0.22
Spain	0.76	0.85
Sweden	0.53	0.78
Switzerland	0.59	0.76
Ukraine	-0.32	-0.25
UK	0.74	0.84

However, Google Trends is most effective in tracking a new disease outbreak in the absence of related media coverage. In our study, we showed how online search spikes for COVID-19 were likely in reaction to WHO media reports. Past studies have shown that media influences online search trends more than epidemiology (Cervellin et al., 2017). We encourage future scientists to analyze online searches with epidemiology, but also to take into consideration media coverage in their analysis to gain a more wholesome picture of the situation.

A 5 country analysis (of Italy, Spain, France, Germany, and the United Kingdom) found that online searches of COVID-19 strongly correlated with the epidemiology (Mavragani, 2020). Our study confirms this relatively strong positive correlation (as seen in Table 1), however, most other European countries had a much lower correlation as the mean correlation among the 40 countries were 0.33 for incidence and 0.17 for mortality. Therefore, we believe this to be a major selection bias as negative results were not reported. Our study has a more robust method as we indiscriminately chose all the countries in Europe that had sufficient data.

Several studies have been conducted on infectious disease outbreaks. Lin et al. found that searches for “wash hands” strongly correlated (Pearson's correlation coefficient of  $-0.70$ ) with a lower spreading speed of COVID-19 (Lin et al., 2020). A study found that Middle East respiratory syndrome and COVID-19 spread correlated

with Google searches (Shin et al., 2016; Husnayain et al., 2020; Li et al., 2020). Walker et al. found a strong correlation between online searches related to the loss of smell (a symptom of COVID-19) with COVID-19 epidemiology (Walker et al., 2020). However, all these studies were prone to selection bias as their sample pool of analyzed countries was small, as explained before. Moreover, we suggest that online search trends simply overlapped with epidemiology since disease incidence increases at the same time when major media announcements are made. In other words, Google trends is better at analyzing media clamor than actual disease incidence.

In our study, we found that the online search patterns for Europe were all very similar. However, Belarus, Russia, and Italy had slightly different online search patterns (as seen in Figure 3). Notably, Google has a market share of 93% in Europe, however, it only has a market share of 55% in Russia and 78% in Belarus (Statcounter Global Stats). Thus, Russia's and Belarus's poor correlation with the rest of the European countries analyzed may have resulted from this difference of the search engine used. Italy was the first European country with a major outbreak of COVID-19; this may explain why Italy had a slightly different online search trend than the rest of western Europe.

#### Limitations of analysis

Testing for COVID-19 varies among each European country; thus, the reported incidence among each country may be far lower than the actual incidence of COVID-19. Therefore, for this study we also took into account the daily mortality rate, which may be more consistent with the actual number of fatalities.

The demographics (i.e. age, gender, location or education level) of online searches and COVID-19 cases were not taken into analysis. This information is not currently available; however, future studies may analyze this.

We focused our media analysis only on WHO announcements since the WHO is responsible for international public health, because it has established a prominent authority on the international stage, and because national policy in a greatly influenced by WHO suggestions. However, other national and international announcements (that may have influenced public online search patterns) were not analyzed. This is a limitation of the study. While it is beyond the scope of this paper, we encourage future researchers to develop a computer bot (that crawls the web and automatically records relevant information) to go through hundreds of announcements (and in various languages) from local governments, news outlets, public figures, prominent social media accounts, etc. to conduct a more complete analysis.

#### Future directions

Future studies may choose to focus on analyzing Asian countries. This analysis would require the use of additional search engines (i.e. Baidu, Yandex RU and Sogou), since they are more popular than Google in parts of Asia. Additionally, the online analysis may be expanded not only to various search engines but also to public health behavior on various other social media platforms (i.e. Facebook, Twitter, Instagram).

#### Public health implications

Each country in Europe acted independently of each other in regards to quarantine, closing borders with other countries and closing public areas (e.g. schools, movie theaters, shops), canceling mass events and testing for COVID-19 among the population. However, the online search trends among all the 40 European

countries analyzed were strong. This shows that although the WHO only plays an advisory role, it still exerts major influence on the public perception of a disease across Europe.

Google trends has classically been used in other studies to track the spread of diseases. However, our study showed that it may be a better tool for tracking the dissemination of information. Global and national institutions may use Google Trends to see if their announcements created a spike in public searches online. By tracking the public response to an announcement, institutions may evaluate the level of public unrest. This method may also help identify mis-information circulating throughout the Internet. During the COVID-19 pandemic, online telecommunications are emphasized and utilized more often to limit the spread of disease (Szmuda et al., 2020b). This allows online surveillance, such as with Google Trends, to be even more powerful during this time as populations rely more heavily on online sources of information. To ensure patient safety and help reduce the spread of mis-information, we encourage physicians to share understandable, high-quality and reliable health information online using trusted sources (Basch et al., 2020; Szmuda et al., 2020a, 2020e, 2020d, 2020c).

## Conclusion

In our paper, we challenged recent literature which found that Google Trends was able to surveil the epidemiology of COVID-19. Though a comprehensive methodology, our study shows that these findings in previous publications are most likely incidental. Instead, Google trends, in the case of COVID-19, may be better suited to surveil the public response to major news announcements such as those from the WHO.

## Conflict of interest

None reported.

## Funding Source

None.

## Ethical approval

None required as all data was publicly available.

## Funding statement

None.

## References

- Al-garadi MA, Khan MS, Varathan KD, Mujtaba G, Al-Kabsi AM. Using online social networks to track a pandemic: A systematic review. *J Biomed Informatics* 2016;62:1–11 Academic Press Inc..
- Alicino C, Bragazzi NL, Faccio V, Amicizia D, Panatto D, Gasparini R, et al. Assessing Ebola-related web search behaviour: Insights and implications from an analytical study of Google Trends-based query volumes. *Infect Dis Poverty* 2015;4(1).
- Basch CH, Hillyer GC, Meleo-Erwin ZC, Jaime C, Mohlman J, Basch CE. Preventive behaviors conveyed on YouTube to mitigate transmission of COVID-19: cross-sectional study. *JMIR Public Heal Surveill* 2020;6(2):e18807.

- Carneiro HA, Mylonakis E. Google Trends: a web-based tool for real-time surveillance of disease outbreaks. *Clin Infect Dis* 2009;49(10):1557–64.
- Cervellin G, Comelli I, Lippi G. Is Google Trends a reliable tool for digital epidemiology? Insights from different clinical settings. *J Epidemiol Glob Health* [Internet] 2017;7(3):185–9 [cited 2019 May 18]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28756828>.
- European Centre for Disease Prevention and Control. Download today's data on the geographic distribution of COVID-19 cases worldwide [Internet]. European Centre for Disease Prevention and Control; 2020 [cited 2020 Mar 26]. Available from: <https://www.ecdc.europa.eu/en/publications-data/download-todays-data-geographic-distribution-covid-19-cases-worldwide>.
- Hassid BG, Day LW, Awad MA, Sewell JL, Osterberg EC, Breyer BN. Using search engine query data to explore the epidemiology of common gastrointestinal symptoms. *Dig Dis Sci* [Internet] 2017;62(3):588–92 [cited 2020 Apr 17]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27878646>.
- Husnayain A, Fuad A, Lazuardi L. Correlation between Google Trends on dengue fever and national surveillance report in Indonesia. *Glob Health Action* 2019;12(1).
- Husnayain A, Fuad A, Su EC-Y. Applications of Google search trends for risk communication in infectious disease management: A case study of COVID-19 outbreak in Taiwan. *Int J Infect Dis* 2020;.
- Li C, Chen LJ, Chen X, Zhang M, Pang CP, Chen H. Retrospective analysis of the possibility of predicting the COVID-19 outbreak from Internet searches and social media data, China, 2020. *Int J Infect Dis* 2020;12:25(10):2000199.
- Lin YH, Liu CH, Chiu YC. Google searches for the keywords of "wash hands" predict the speed of national spread of COVID-19 outbreak among 21 countries. *Brain Behav Immun* 2020;.
- Mavragani A. Tracking COVID-19 in Europe: an infodemiology study. *JMIR public Heal Surveill* [Internet] 2020; [cited 2020 Apr 17]; Available from: <http://www.ncbi.nlm.nih.gov/pubmed/32250957>.
- Pervaiz F, Pervaiz M, Rehman NA, Saif U. FluBreaks: Early epidemic detection from Google flu trends. *J Med Internet Res* 2012;14(5).
- Santangelo OE, Provenzano S, Piazza D, Giordano D, Calamusa G, Firenze A. Digital epidemiology: assessment of measles infection through Google Trends mechanism in Italy. *Ann di Ig.* 2019;31(4):385–91.
- Seifter A, Schwarzwald A, Geis K, Aucott J. The utility of "Google Trends" for epidemiological research: Lyme disease as an example. *Geospat Health* 2010;4(2):135–7.
- Shin S-Y, Seo D-W, An J, Kwak H, Kim S-H, Gwack J, et al. High correlation of Middle East respiratory syndrome spread with Google search and Twitter trends in Korea. *Sci Rep* [Internet] 2016; [cited 2020 Mar 26];6(1):32920. Available from: <http://www.nature.com/articles/srep32920>.
- Statcounter Global Stats. Search Engine Market Share Worldwide [Internet]. [cited 2020 Mar 2]. Available from: <https://gs.statcounter.com/search-engine-market-share>.
- Szmuda T, Ali S, Słoniewski P. Letter to the Editor. The future of patient safety in neurological surgery. *J Neurosurg* [Internet] 2020a; [cited 2020 Apr 17];1-2. Available from: <https://thejns.org/view/journals/j-neurosurg/aop/article-10.3171-2019.11.JNS193005/article-10.3171-2019.11.JNS193005.xml>.
- Szmuda T, Ali S, Słoniewski P, Group NsW. Telemedicine in neurosurgery during the novel coronavirus (COVID-19) pandemic. *Neurol Neurochir Pol* [Internet] 2020b; [cited 2020 Apr 23]; Available from: <http://www.ncbi.nlm.nih.gov/pubmed/32319670>.
- Szmuda T, Özdemir C, Ali S, Singh A, Syed MT, Słoniewski P. Readability of online patient education material for the novel coronavirus disease (COVID-19): a cross-sectional health literacy study. *Public Health* [Internet] 2020c; [cited 2020 May 31]; Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0033350620302031>.
- Szmuda T, Özdemir C, Fedorow K, Ali S, Słoniewski P. YouTube as a source of information for narcolepsy: A content-quality and optimization analysis. *J Sleep Res* [Internet] 2020d:e13053 [cited 2020 Apr 23];. Available from: <https://pubmed.ncbi.nlm.nih.gov/32315117/>.
- Szmuda T, Rosvall P, Hetzger TV, Ali S, Słoniewski P. YouTube as a source of patient information for hydrocephalus: a content-quality and optimization analysis. *World Neurosurg* [Internet] 2020e; [cited 2020 Mar 7]; Available from: <https://www.ncbi.nlm.nih.gov/pubmed/32147561>.
- Walker A, Hopkins C, Surda P. The use of google trends to investigate the loss of smell related searches during COVID-19 outbreak. *Int Forum Allergy Rhinol* [Internet] 2020; Apr 11 [cited 2020 Apr 17]; Available from: <http://www.ncbi.nlm.nih.gov/pubmed/32279437>.
- World Health Organization. Coronavirus (COVID-19) events as they happen [Internet]. [cited 2020 Mar 17]. Available from: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/events-as-they-happen>.