

Utility and limitations of Google searches on sensory loss as markers for new COVID-19 cases

Kim Asseo^{1#}, Fabrizio Fierro^{1#}, Yuli Slavutsky², Johannes Frasnelli³ and Masha Y Niv^{1*}

1 The Institute of Biochemistry, Food Science and Nutrition, The Robert H Smith Faculty of Agriculture, Food and Environment and The Fritz Haber Center for Molecular Dynamics, The Hebrew University of Jerusalem, Israel

2 Department of Statistics and Data Science, The Hebrew University of Jerusalem, Israel

3 Department of Anatomy, University of Québec in Trois-Rivières, Trois-Rivières, Québec, Canada

#equally contributing authors

**correspondence to masha.niv@mail.huji.ac.il*

Abstract

Evidence of smell loss in COVID-19 is growing. Researchers and analysts have suggested to use Google searches on smell loss as indicators of COVID-19 cases. However, such searches may be due to interest elicited by media coverage of the COVID-19-related smell loss, rather than attempts to understand self-symptoms. We analyzed searches related to 4 senses: smell and taste (both recently shown to be impaired in some COVID-19 patients), vision and sight (senses not currently known to be impaired in COVID-19 patients), and an additional general control (“COVID-19 symptoms”). Focusing on two countries with a large number of cases, Italy and the United States, we have compared Google Trends results per region or state to the number of new cases prevalence in that region. The analysis was performed for each of the 8 weeks ranging from March 4th till April 28th.

No correlation with vision loss or sight loss searches was identified, while taste and smell loss searches were correlated with new COVID-19 cases during a limited time window, that starts when the number of weekly new cases reached for the first time 21357 cases in Italy (11-17 March) and 47553 in the US (18-24 March).

Media effect on the specific symptoms searches was also analyzed, establishing a different impact according to the country.

Our results suggest that Google Trends for taste loss and smell loss searches captured a genuine connection between these symptoms and new COVID-19 cases prevalence in the population. However, due to variability in correlation from week to week, and overall decrease in correlation as taste and smell loss are becoming known COVID-19 symptoms, recognized now by CDC and World Health Organization, Google Trends is no longer a reliable marker for monitoring the disease spread. The “surprise rise” followed by decrease, probably attributable to knowledge saturation, should be kept in mind for future digital media analyses of potential new symptoms of COVID-19 or future pandemics.

Introduction

COVID-19 pandemic has by now hit almost all countries worldwide. Monitoring disease occurrence is a key prerequisite for combating the disease spread. Laboratory testing availability differs per country, with most of the countries unable to test the general and even the symptomatic population broadly. Furthermore, the symptoms elicited by SARS-CoV-2 are still being discovered, with the list of official symptoms being updated on a rolling basis.

Specifically, in addition to fever, cough, shortness of breath or difficulty breathing, chill, muscle pain, headache and sore throat, recent additions to [CDC symptoms](#) include “new taste and smell loss”, and World Health Organization has listed “loss of taste or smell” among “less common symptoms” of COVID-19. Smell loss, and to a lesser degree taste loss accompanying COVID-19 infection have appeared in reports of COVID-19 patients testimonies (Hopkins and Kumar, 2020), preprints of scientific papers (Bagheri et al., 2020; Menni et al., 2020; Pellegrino et al., 2020; Williams et al., 2020), peer-reviewed publications (i.e. (Bénézit et al., 2020; Eliezer et al., 2020; Moein et al., 2020; Spinato et al., 2020; Yan et al., 2020)), and are rather widely discussed by journalists.

The details of taste and smell change in relation to COVID-19 remain unclear, and scientists (including the authors of this contribution), clinicians and patients advocates have created the [GCCR consortium](#) which studies the relation between taste and smell loss and COVID-19.

Here we set out to explore the hypothesis, proposed by several groups (Brunori and Resce, 2020; Lamos et al., 2020; Walker et al., 2020) as well as discussed in the New York Times (Stephens-Davidowitz, 2020) and CNBC (Frank, 2020) that Google Trends searches on smell loss are indicative of COVID-19 cases.

We focused on two countries, Italy and US, and have looked not only for searches on smell loss, but also for searches of taste loss. Media reports on these phenomena were analyzed using Media Cloud, a database collecting media articles and reporting the attention over time on a query topic.

Furthermore, we used sight loss and hearing loss as controls, since these are senses not currently known to be associated with COVID-19. Because smell and taste loss are gradually becoming recognized as COVID-19 symptoms, we also looked at searches for COVID-19 symptoms.

Results

We analyzed Google searches and the numbers of new COVID-19 cases in Italy and the US, taking into account the states (US) and the regions (Italy) that compose them. For each region or state we calculated the following parameters: Google Trends popularity index per region/state for generic (COVID-19 symptoms) and specific (taste loss, smell loss) symptoms of COVID-19, as well

as specific symptoms not known to affect patients (hearing loss, sight loss); the number of their mentions in digital media; the number of new COVID-19 cases normalized by population size per region or state inhabitants. The data were collected for 8 consecutive weeks, spanning March 4th to April 28th. Next, for each week and for each of the searches, the correlation with the normalized number of new cases was calculated for each country.

Results for weeks with a good correlation for taste loss and smell loss searches are shown in Figure 1. The volume of searches for these two keywords was high, among the other regions, in Lombardy, Emilia Romagna and Veneto, as well as New York, New Jersey and Louisiana, which are geographical sub-areas with high rates of new COVID-19 patients/inhabitants in their respective country.

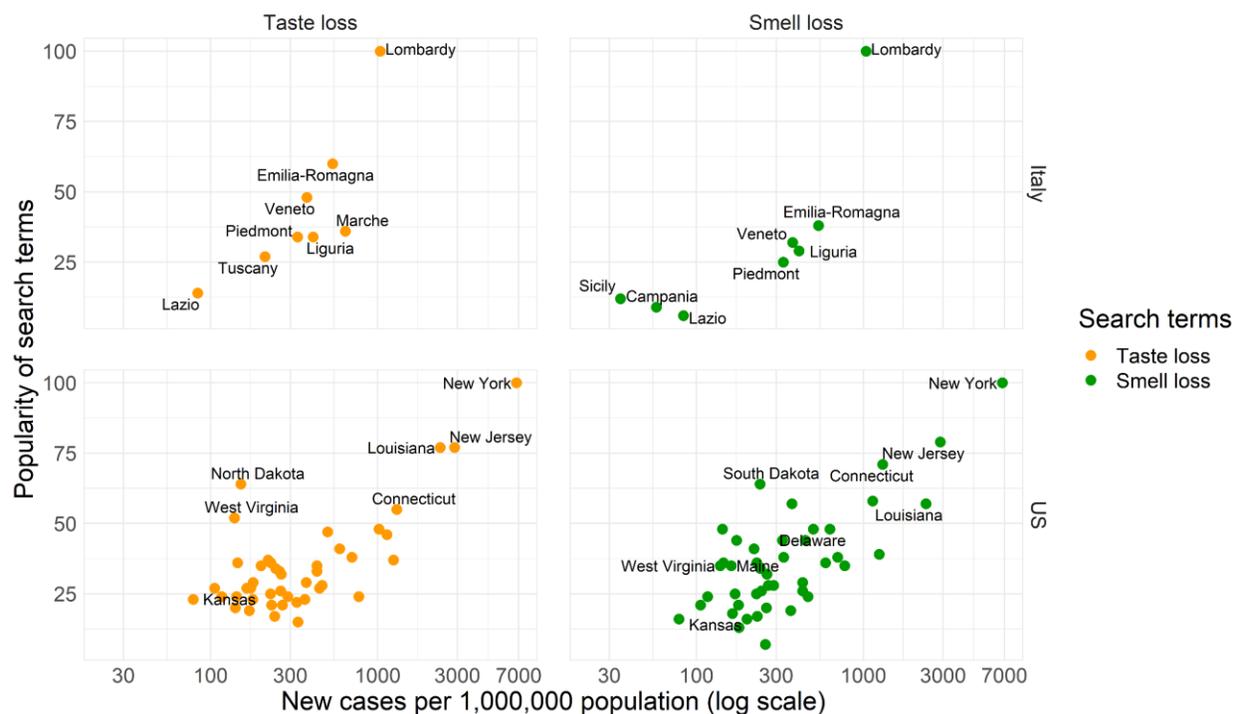


Figure 1: Data from which correlation was derived. Here shown for Italy (11-17 March) and the US (1-7 April), weeks with highest correlation for taste loss. Graphs show the taste loss (in orange) and smell loss (in green) search queries. Each point represents a different region/state. Normalized number of new cases related to the corresponding week on the x axis, popularity of the search terms on the y axis. For both graphs, not all the regions or states are shown because of the lack of popularity index for some regions/states.

Does this picture hold over time? In figure 2, we followed week by week, the correlation between the number of searches with the number of new cases, calculated for the 51 states in the US, and the 20 regions in Italy (as was exemplified for a particular week in Figure 1).

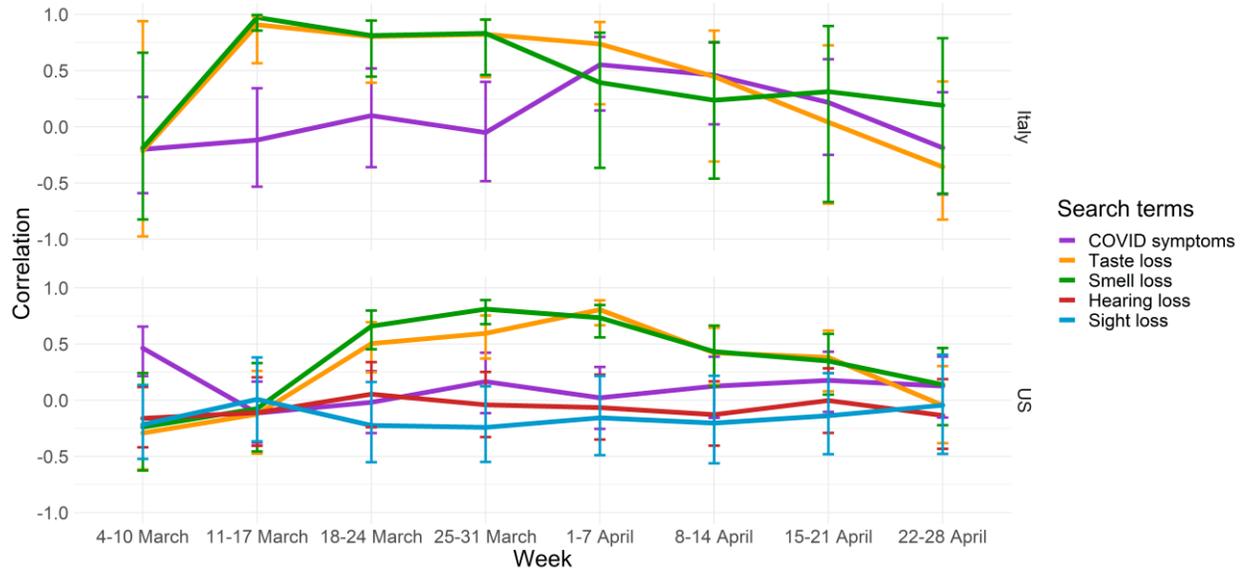


Figure 2: Correlation between number, normalized by population size, of new COVID-19 cases in each region in Italy (top graph) and in each state in the US (bottom graph) and search queries popularity index, for each week. Confidence intervals are shown.

Overall, Figure 2 illustrates that both in Italy and in the US, there are some weeks with high correlation between searches for taste and smell loss, and the normalized number of new cases. This can be due to people affected by the taste loss and/or smell loss searching for their experienced specific symptoms. But this correlation changes over time:

during the first analyzed week in Italy (4-10 March) and the first two in the US (4-10 and 11-17 March), taste loss and smell loss queries show a negative and low correlation with new COVID-19 cases. In the following weeks, the peaks with the highest correlation values are reached for Italy on 11-17 March week (0.91 for taste loss and 0.97 for smell loss) and for the US on the 25-31 March and 1-7 April weeks (0.81 for both smell loss and taste loss).

Weeks following the peaks are characterized by a decrease in correlation for the two specific symptoms keywords, where the lowest values are reached during the last analyzed week (22-28 April) (see Supplementary Figure 1 for detailed representation of data for this week).

In Italy, in the 11-17 March and 15-21 April weeks, the average number of new cases per 1,000,000 inhabitants was almost identical, 354 and 357 respectively, while the correlation decreases from 0.91 for taste loss and 0.97 for smell loss to 0.04 and 0.31, respectively (Figure 2). Similarly, from 1-7 April week to the 22-28 April week, the average number of new cases in the US remains confined in the range of 644-663, but correlation decreases from 0.81 for taste loss and 0.73 for smell loss to -0.04 and 0.14 respectively.

Hearing loss and sight loss terms, that were used as controls, show lack of correlation in the US in all of the 8 weeks under analysis, as expected (Figure 2). The corresponding searches for the Italian translation of the queries (“perdita udito” and “perdita vista”) did not produce enough results to show data relative to different regions and, consequently, display no correlation. Only a low correlation in some of the 8 weeks studied here was observed for “COVID-19 symptoms” search. This can be attributed to a general interest in the disease and its symptoms, with some contribution of searches due to people suspecting they might be sick with the disease.

Since we hypothesized that media coverage of COVID-19 related taste and smell symptoms may impact the number of searches, we next analyzed, on a daily basis and for each country as a whole, the number of new cases, as well as Google searches and the volume of media coverage of taste and smell loss.

Using the Explorer tool on the Media Cloud platform, we monitored the number of times taste and smell keywords were mentioned daily by digital news media (hence, not including radio, television or printed matter).

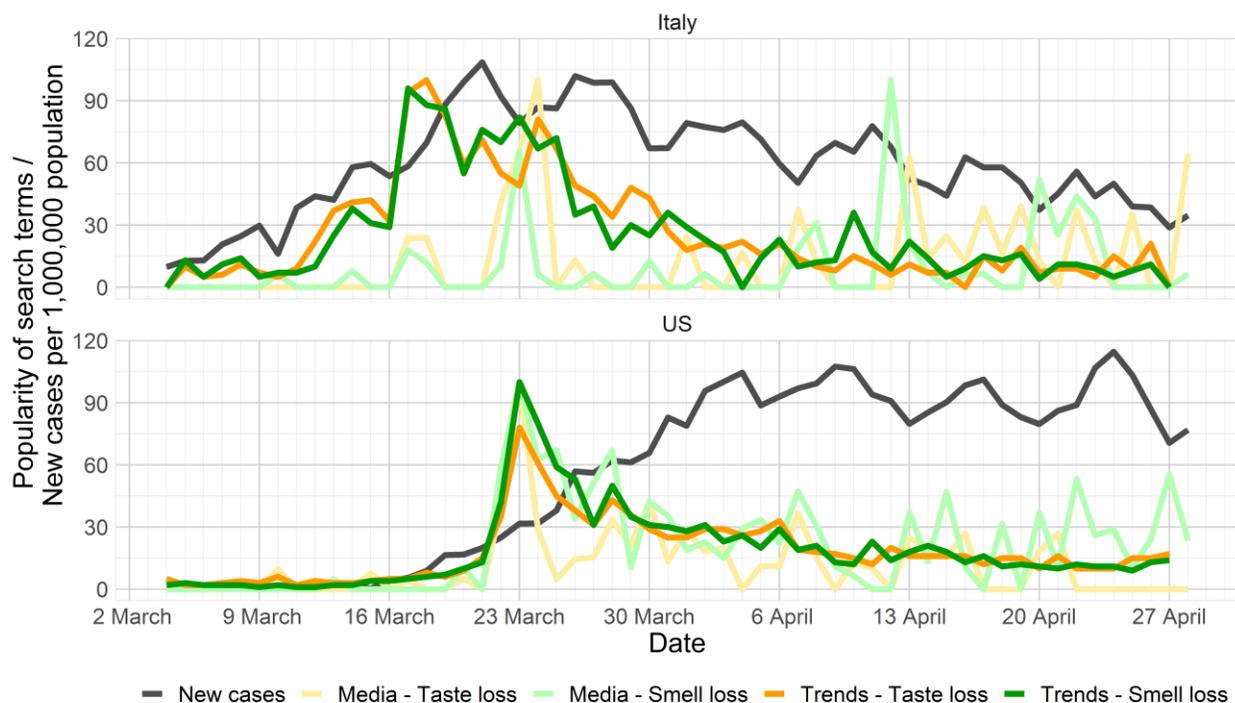


Figure 3: Comparison between Google searches volume for taste loss and smell loss queries, Media Cloud popularity of the same keywords and number of new COVID-19 cases calculated from March 4th to April 28th for Italy and the US. The Media Cloud data were normalized as for the Google search results, assigning the value of 100 to the day with the highest popularity, with the other days assigned values relative to that day. The number of new cases is relative to a population of 1,000,000 in each country.

The news about these two specific symptoms, in both Italy and the US, was reported for the first time during the 11-17 March week according to our search on Media Cloud (media coverage at the end of the 4-10 March week in both Italy and the US, actually reported the taste loss or smell loss in a different context, not related to COVID-19 symptoms) (Figure 3). The maximum media coverage in Italy is reached between the 22nd and 23rd of March, with a second higher peak for smell loss on April 12th. There is a small media peak during the 11-17 March week, which coincides with a high number of new cases. The correlation peak for Italy, reported in Figures 1 and 2, is reached during this week. An increase in popularity of searches was observed before the first media reports of March 14th (Figure 3 and Supplementary Figure 2). Indeed, the total number of cases on March 13th in Italy was 8.2 times higher than in the US (17660 for Italy vs 2147 for the US) and, consequently, the volume of searches observed until this day is also higher (Supplementary Figure 3).

In the US we observed a different situation. The Media Cloud data and the Google Trends searches are perfectly superimposed in proximity of the maximum peak. The rise of cases in the US is later than in Italy, and there is no time window in which the number of cases is surging while the taste and smell symptoms are still unknown. Therefore, in the US, it is difficult to know which part of taste and smell loss searches are driven by patients worried about their own symptoms, and which part is due to interest elicited by the media.

The correlation between the number of new cases and the volume of searches is better correlated in Italy than in the US, while the volume of searches in the US is more affected by the media trend than in Italy (Table 1). Smell and taste loss trends are more closely related to each other than to the respective media coverage of each effect, suggesting that the two symptoms are experienced together. This is of interest, as it may indirectly suggest a potential common or joint mechanism that allows SARS-CoV-2 to affect both senses together.

Table 1: Calculated correlation between the two (left and right in the correlation variables column) variables in Italy and the US, as visualized in Figure 3.

Correlation variables		Italy	US
Trends - Taste loss	Trends - Smell loss	0.93	0.97
Media - Taste loss	Trends - Taste loss	0.21	0.74
Media - Smell loss	Trends - Smell loss	0.04	0.79
New cases	Trends - Taste loss	0.61	0.27
New cases	Trends - Smell loss	0.59	0.19

Overall, the decrease in correlation between taste and smell loss searches and the number of new cases overtime is apparent both in Italy and the US (Figure 2). We believe that, with media coverage and inclusion of taste and smell loss as official symptoms of COVID-19, there are less searches for these symptoms alone, but rather a convergence with COVID-19 symptoms searches due to the broad knowledge of the symptoms.

Discussion

In this study, we show the presence of a correlation in Italy and the US between Google searches for specific and new symptoms of COVID-19 (taste loss and smell loss) and the number of people affected by the SARS-CoV-2 virus. On average, regions with a high percentage of patients among the total number of inhabitants tend to search for specific symptoms more often than in regions with low rates when the number of new cases reach for the first time a relatively high volume (21357 for Italy on 11-17 March and 47553 for the US 18-24 March). Nevertheless, relying solely on Google trends for taste loss and smell loss searches is not a reliable strategy, even though suggested by Goldman and Sachs (Frank, 2020) to monitor the spread of the SARS-CoV-2. Indeed, this correlation varies among the weeks analyzed here, even when the number of new cases in a country is high, and decreases in the recent weeks, potentially because the specific symptoms are becoming well established by now.

Media had a strong impact on the volume of searches for taste loss and smell loss in the US, and today taste and smell loss in COVID-19 patients is rather widely known. On the other hand, in Italy the media effect on searches was slightly milder due to the advanced state of the pandemic when the news started to appear on digital media. The Italian data likely suggests genuine interest based on self-symptoms even before they became broadly established.

Interestingly, we also found a strong correlation between searches for taste loss and searches for smell loss. This is in line with recent findings that the degree of Covid-19-related anosmia (loss of smell) and ageusia (loss of taste) correlate closely in affected individuals (Yan et al., 2020). Hence, both senses seem to be affected simultaneously in COVID-19 patients, a finding that may provide clues for mechanisms of action of the virus.

Strategies for geographic monitoring of COVID-19 hotspots are being developed (Giordano et al., 2020), self-reporting apps becoming increasingly common (Mayor, 2020; Rossman et al., 2020), and epidemiological tools, such as sewage monitoring (Medema et al., 2020) are being introduced. We believe that social media monitoring of taste and smell loss searches is unlikely to remain a useful marker for COVID-19 cases, as these symptoms become widely known, and will likely give way to other tools. More generally, as additional new symptoms may be discovered and future pandemics may be expected, usage of digital media searches on pandemic symptoms for disease should take into account the potential “surprise rise and knowledge saturation” curve.

Methods

Correlation data

The data on new COVID-19 cases per each of the 20 Italian regions were obtained from the Italian Ministry of Health website (<http://www.salute.gov.it/portale/nuovocoronavirus/homeNuovoCoronavirus.jsp?lingua=english>), and from the Johns Hopkins Coronavirus Resource Center for the 51 US states (<https://coronavirus.jhu.edu/us-map>).

The data was normalized per region or state population. The inhabitants of every Italian region were retrieved from the last data available on the Istituto Nazionale di Statistica (National Institute for Statistics) website (<http://dati.istat.it/Index.aspx?lang=en&SubSessionId=1d073136-f11a-4329-a1ed-3a3920a1ec32>). For the USA, the population of every US state was calculated from (Bialek et al., 2020).

The search words described in Table 2 were used as input in Google Trends (<https://www.google.com/trends>) and the searches data for each region or state were collected. Additional terms and different combinations of keywords were tested, resulting in no evident differences in the calculated correlations from the results obtained with the Table 2 terms, identified as the most popular. Google Trends provides the normalized number of searches according to the population living in the country's sub-area and assigns a popularity index to the keyword searches that spans from 0 to 100. 100 is assigned to the region in which the keyword reaches the maximum volume of searches for the dates and countries selected, with no relation to the other keywords searched in that comparison.

Table 2: Search terms used for Italian and English searches on Google Trends. The first column is the reference keyword used to summarize the corresponding searches in the main text. The “+” character allows to sum up data for multiple searches with a logical OR.

	Italian keywords	English keywords
Taste loss	perdita gusto + perdita del gusto	taste loss + loss of taste
Smell loss	perdita olfatto + perdita dell'olfatto	smell loss + loss of smell
Sight loss	perdita vista + perdita della vista	vision loss + sight loss
Hearing loss	perdita udito + perdita dell'udito	hearing loss + loss of hearing
COVID symptoms	sintomi coronavirus + sintomi covid19 + sintomi covid + coronavirus sintomi + covid sintomi	COVID symptoms + coronavirus symptoms + COVID-19 symptoms + COVID19 symptoms + Coronavirus symptoms

The correlation between the number of new COVID-19 cases per 1,000,000 population and the Google Trends value is simply Pearson correlation.

The collected data were used to build the graphs in Figures 1 and 2 using the software RStudio (R Core Team, 2013).

To make sure the correlation wasn't dominated by an outlier in the data, a test was done on the US by removing New York - the state with the highest number of new COVID-19 cases per population, and recalculating the correlation from the new data to find the change in correlation insignificant.

The correlations shown in Table 1 were calculated using the Pearson correlation.

Media impact data

First, we calculated the popularity index of the taste loss and smell loss queries on a daily basis from March 4th to April 28th. As for the regions/states Google Trends searches, the data is automatically normalized assigning the day with the highest volume of searches the value of 100, and other days are assigned values relative to that day.

Next, we used the Media Cloud webserver (<http://www.mediacloud.org>) to obtain an estimate for the number of times a certain keyword appeared on digital news on a daily basis in the period between March 4th to April 28th. By using the same keywords as defined in Table 2 for taste loss and smell loss, we obtained the normalized number of appearances in digital news. The collection of media used to search our keywords are the "Italy - National" and the "U.S. Top Sources 2018" available on the Media Cloud website.

In order to superimpose the Media Cloud results to the Google Trends one, as done in Figure 3, Media Cloud data were further normalized in the manner Google Trends are normalized: a value of 100 was assigned to the day with the highest media coverage peak of a particular search query, and other days assigned values relative to that day. The number of new cases shown in the same figure were normalized per 1,000,000 population.

Figure 3 was generated using the software Rstudio (R Core Team, 2013).

Acknowledgement – We thank the Center for Interdisciplinary Data Science Research, The Hebrew University, for support, Maria Veldhuizen for discussions, and Noam Lahav and Eitan Margulis for help in the initial stages of this project. MYN is funded by ISF grant #1129/19 and is a member of COST actions Mu.Ta.Lig (CA15135) and ERNEST (CA18133). MYN, KA, FF and JF are members the Global Consortium of Chemosensory Research, the GCCR.

References

- Bagheri, S.H.R., Asghari, A.M., Farhadi, M., Shamshiri, A.R., Kabir, A., Kamrava, S.K., Jalessi, M., Mohebbi, A., Alizadeh, R., and Honarmand, A.A. (2020). Coincidence of COVID-19 epidemic and olfactory dysfunction outbreak. medRxiv.
- Bénézit, F., Le Turnier, P., Declerck, C., Paillé, C., Revest, M., Dubée, V., Tattevin, P., Arvieux, C., Baldeyrou, M., and Chapplain, J.-M. (2020). Utility of hyposmia and hypogeusia for the diagnosis of COVID-19. *The Lancet Infectious Diseases*.
- Bialek, S., Bowen, V., Chow, N., Curns, A., Gierke, R., Hall, A., Hughes, M., Pilishvili, T., Ritchey, M., and Roguski, K. (2020). Geographic differences in COVID-19 cases, deaths, and incidence—United States, February 12–April 7, 2020.
- Brunori, P., and Resce, G. (2020). Searching for the peak Google Trends and the Covid-19 outbreak in Italy (Dipartimento di Economia e Finanza-Università degli Studi di Bari" Aldo Moro").
- Eliezer, M., Hautefort, C., Hamel, A.-L., Verillaud, B., Herman, P., Houdart, E., and Eloit, C. (2020). Sudden and complete olfactory loss function as a possible symptom of covid-19. *JAMA Otolaryngology–Head & Neck Surgery*.
- Frank, T. (2020). Goldman says fewer 'loss of smell' Google queries suggest better COVID outlook. Consumer News and Business Channel. <https://www.cnn.com/2020/04/15/goldman-says-fewer-loss-of-smell-google-searches-is-a-positive-sign-for-pandemic.html>
- Giordano, G., Blanchini, F., Bruno, R., Colaneri, P., Di Filippo, A., Di Matteo, A., and Colaneri, M. (2020). Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy. *Nature Medicine*, 1-6.
- Hopkins, C., and Kumar, N. (2020). Loss of sense of smell as marker of COVID-19 infection. *ENT UK*.
- Lamos, V., Moura, S., Yom-Tov, E., Cox, I.J., McKendry, R., and Edelstein, M. (2020). Tracking COVID-19 using online search. arXiv preprint arXiv:200308086.
- Mayor, S. (2020). Covid-19: Researchers launch app to track spread of symptoms in the UK (British Medical Journal Publishing Group).
- Medema, G., Heijnen, L., Elsinga, G., Italiaander, R., and Brouwer, A. (2020). Presence of SARS-Coronavirus-2 in sewage. medRxiv.
- Menni, C., Valdes, A., Freydin, M.B., Ganesh, S., Moustafa, J.E.-S., Visconti, A., Hysi, P., Bowyer, R.C., Mangino, M., and Falchi, M. (2020). Loss of smell and taste in combination with other symptoms is a strong predictor of COVID-19 infection. medRxiv.
- Moein, S.T., Hashemian, S.M., Mansourafshar, B., Khorram-Tousi, A., Tabarsi, P., and Doty, R.L. (2020). Smell dysfunction: a biomarker for COVID-19. Paper presented at: Int Forum Allergy Rh (Wiley Online Library).
- Pellegrino, R., Cooper, K., Di Pizio, A., Bhutani, S., and Parma, V. (2020). Corona Viruses and the Chemical Senses: Past, Present, and Future.
- R Core Team, R.C.T. (2013). R: A language and environment for statistical computing.
- Rossmann, H., Keshet, A., Shilo, S., Gavrieli, A., Bauman, T., Cohen, O., Shelly, E., Balicer, R., Geiger, B., and Dor, Y. (2020). A framework for identifying regional outbreak and spread of COVID-19 from one-minute population-wide surveys. *Nature Medicine*, 1-4.
- Spinato, G., Fabbris, C., Polesel, J., Cazzador, D., Borsetto, D., Hopkins, C., and Boscolo-Rizzo, P. (2020). Alterations in Smell or Taste in Mildly Symptomatic Outpatients With SARS-CoV-2 Infection. *JAMA*.

Stephens-Davidowitz, S. (2020). Google Searches Can Help Us Find Emerging Covid-19 Outbreaks. The New York Times. <https://www.nytimes.com/2020/04/05/opinion/coronavirus-google-searches.html>

Walker, A., Hopkins, C., and Surda, P. (2020). The use of google trends to investigate the loss of smell related searches during COVID-19 outbreak. Paper presented at: Int Forum Allergy Rh (Wiley Online Library).

Williams, F.M., Freydin, M., Mangino, M., Couvreur, S., Visconti, A., Bowyer, R.C., Le Roy, C.I., Falchi, M., Sudre, C., and Davies, R. (2020). Self-reported symptoms of covid-19 including symptoms most predictive of SARS-CoV-2 infection, are heritable. medRxiv.

Yan, C.H., Faraji, F., Prajapati, D.P., Boone, C.E., and DeConde, A.S. (2020). Association of chemosensory dysfunction and Covid-19 in patients presenting with influenza-like symptoms. Paper presented at: Int Forum Allergy Rh (Wiley Online Library).