Loss of smell and taste: a new marker of COVID-19? Tracking reduced sense of smell during the coronavirus pandemic using search trends

George Cherry¹, John Rocke², Michael Chu³, Jacklyn Liu⁴, Matt Lechner⁵, Valerie J. Lund⁶ and B. Nirmal Kumar⁷

¹Independent, UK; ²Wigan and Leigh Teaching Hospitals NHS Foundation Trust, Wrigthington, UK; ³UCL Cancer Institute, University College London, London, UK; ⁴Barts Health NHS Trust, Royal London Hospital and Whippys Cross University Hospital, UK; ⁵Royal National Throat Nose and Ear Hospital, UCLH Foundation Trust, UK

ABSTRACT

Objectives: It has been demonstrated that reduction in smell and/or taste is the most predictive symptom in SARS-CoV-2/Covid-19 infection. We used Google Trends to analyze regional searches relating to loss of smell and taste across Italy, Spain, France, Brazil, and the United States of America and determined the association with reported Covid-19 cases.

Methods: In order to retrieve the data, we built a Python software program that provides access to Google Trends data via an application program interface. Daily COVID-19 case data for subregions of the five countries selected were retrieved from respective national health authorities. We sought to assess the association between raw search interest data and COVID-19 new daily cases per million for all regions individually.

Results: In total, we yielded 2188 sets of Google Trends data which included 548 time series of 4 anosmia and ageusia search concepts over the study period for 137 regions. These data indicated that differences in search interest for terms relating to anosmia and ageusia, between regions, is associated with geographical trends in new Covid-19 cases.

Conclusions: We feel that Google search trends relating to loss of smell can be utilized to identify potential Covid-19 outbreaks on a national and regional basis.

1. Introduction

The loss or reduction of the sense of smell has been widely reported as a key symptom of SARS-CoV-2/Covid-19 infection and, in a significant proportion of cases, it has been demonstrated to be the only complaint [1–3]. The positive predictive value for olfactory dysfunction (OD), or loss of smell, for Covid-19 positivity has been demonstrated to be over 60%; higher than any other associated symptoms [4]. The World Health Organization (WHO) integrated loss of smell and/or taste to their official list of associated symptoms and Public Health England have recently updated their recommendations for self-isolation to include olfactory dysfunction.

Self-isolation and social distancing are seen as key public health strategies in controlling the Covid-19 outbreak [5]. Internationally, governments are now looking to reduce restrictions on their populations to prevent economic damage. As prohibitive measures are reduced, it is likely that there will be further outbreaks of Covid-19 [6]. Moving forward, accurate tracking of Covid-19 cases and early quarantining of infected people or populations will be vital in reducing transmission and preventing secondary outbreak [6]. Effective strategies will need to be adopted by leaders and public health bodies internationally to reflect this. With a worldwide need for Polymerase Chain Reaction (PCR) tests and a significant associated economic cost, particularly in low-to-middle income countries, it is paramount to develop alternative strategies to track outbreaks aside from PCR testing and tracing [6,7].

Walker et al. have demonstrated the utility of search term analysis, using Google trends, in determining significant associations between reported Covid-19 cases on a countrywide level [8]. Using search interest data to infer population-wide behavior in developed countries has recently become possible. This is due to an increase in mobile internet usage from 17% to 78%, in the United Kingdom between 2008 and 2018 and through the adoption of unified browser Uniform Resource Locator (URL) and search bars after 2010 [9]. The two major mobile platforms use Google which drives almost all mobile search traffic through this search engine. This has reinforced a new behavior whereby people quickly and easily search for anything, wherever they are, almost exclusively with Google.

We used Google Trends to analyze regional searches relating to loss of smell and taste across Italy, Spain, France, Brazil, and the United States of America (USA) and determined the association with reported Covid-19 cases using a self-developed software programme (Python). Moving forward, regional and sub-regional tracking of outbreaks using this or a similar method could help a targeted local public health strategies and testing. This has the potential to reduce the demand on costly PCR tests and
allow areas that are not affected by new infections to avoid restrictive measures and open up their local economies and healthcare system.

2. Methods

2.1. Data selection

We chose the five countries worst affected by COVID-19 according to Johns Hopkins global COVID-19 data; USA, UK, Spain, Italy, and France [10]. However, Google Trends sub-divides the UK into the home nations only (England, Scotland, Wales, and Northern Ireland) making further regional analysis impossible. Therefore, Brazil was used in place of the UK.

2.2. Google trends data

Google Trends is an open access platform providing data on the amount of search requests performed using Google. For a chosen search term, date range and geographical region Google Trends data contains the relative changes in volume of search requests over time and between regions. Of particular usefulness, the geographical boundaries use ISO-3166 standard national subregions which typically match COVID-19 reporting from the parent countries.

Google Trends data is normalized over the time period and geography in question against all other searches that took place. A value of 100 indicates the highest popularity for that term and 0 the lowest popularity. Similarly, for geographical comparisons, a score of 100 for a region indicates the highest relative popularity and 0 the lowest – for that country or region within a country [11].

We defined four concepts that individuals may perform a Google search for when experiencing anosmia or ageusia:

- Loss of sense of smell
- Sense of smell
- Loss of sense of taste
- Sense of taste

These four concepts were manually translated using the Google Trends Related Queries tool in order to find the appropriate semantic translation and grammar – as opposed to a literal translation. Table 1 shows the results of this.

National subregions were defined by their ISO-3166 codes – with the exception of France as Google Trends uses the pre-2016 regional boundaries [12]. The time period of interest for Google Trends data started when each country began recording sub-regional COVID-19 data and ended on 17 May 2020. The USA

was an exception as the previous method exceeded Google Trends 90-day window for daily data aggregations, therefore the date any US state passed 10 cases of COVID-19 was taken as the start date. Search interest data were aggregated on a daily basis within the period of interest.

In order to retrieve the data, we built a Python software program, which includes the pytrends open source library that provides access to Google Trends data via an application program interface (API) [13]. Using the pytrends request methods we collected search interest data over time for each translated term per subregion of each country. A second set of search interest data was collected for individual days within the time period per search term and country which provided relative values for search interest between subregions.

2.3. Regional COVID-19 cases and population data collection

Daily COVID-19 case data for subregions of the five countries selected were retrieved from respective national health authorities [14–18]. Regional population data were supplied with COVID-19 data from Brazil and the USA. Census population data were retrieved separately for Italy, Spain, and France [19–21].

French COVID-19 and population data were reorganized from department level grouping into the pre-2016 regional structure in order to match the search interest geographies used by Google Trends.

2.4. Analysis

Further data processing was performed using our self-developed software program. Unadjusted search interest data could be used for within region analyses over time. To conduct between region analyses, search interest over time per region was combined with national search interest data providing relative weightings between regions producing a figure for daily weighted search interest.

COVID-19 data were normalized against the population of each region producing a value for new cases per million. Normalizing COVID-19 case data were required to adjust for population differences in order to make between region comparisons between regions within a country.

Raw search interest data, weighted search interest data, and COVID-19 new daily cases per million data were smoothed using a 7-day moving-mean to reduce the noise introduced by recording errors and low data volumes.

We then sought to assess the association between smoothed raw search interest data and smoothed COVID-19 new daily cases per million for all regions individually. Following this we looked broadly at the number of regions with significant associations and the strength of any association.

A second analysis pooled the smoothed, weighted search interest data for all regions within a country and the smoothed COVID-19 new daily cases per million. This second analysis was an attempt to understand whether search interest data could provide insights into geographical as well as temporal changes in anosmia and ageusia secondary to Covid-19.
3. Results

In total 137 regions were included in the study; Italy (20), USA (51), Spain (17), France (22) and Brazil (27). Running until the 17 May 2020, the following start dates and number of days are shown for each country:

- Italy: 24 February 2020–83 days
- USA: 21 February 2020–86 days
- Spain: 20 February 2020–87 days
- France: 24 February 2020–83 days
- Brazil: 26 February 2020–81 days

In total this yielded 2188 sets of Google Trends data which included 548 time series of the 4 anosmia and ageusia search concepts over the study period for all 137 regions, aggregated per day. The remaining 1640 sets of Google Trends data consisted of the regional comparison data per country and search concept on each date within the study period.

3.1. Within region analyses

The search interest data with 7-day moving-mean, although continuous, would still have been bound at an upper and lower limit. In addition, Kolmogorov-Smirnov testing demonstrated the data was nonparametric. Therefore, Spearman’s rank correlation test was used in all regions and for all 4 translated search concepts. Daily search interest data (7-day moving-mean) were tested against daily new Covid-19 cases per million (7-day moving-mean) over time within each region.

For seven sets of results Google Trends returned an insufficient data response. This affected Molise, Basilicata, and Valle d’Aosta in Italy and North Dakota and Wyoming in the USA. Therefore, in all other combinations of regions and search concepts 541 Spearman’s rank correlation tests were performed – Table 2 shows a summary of the results.

A strong ($r_s > 0.65$) or moderate ($r_s 0.3$ to 0.65) positive correlation between new daily Covid-19 cases and search interest for ‘loss of sense of smell’ was observed in 110 regions (82%), for ‘loss of sense of taste’ 113 regions (84%), ‘sense of smell’ 107 regions (78%) and for ‘sense of taste’ 83 regions (61%). Of this total of 413 strong to moderate test results, 381 had a significance value of less than 0.001. Our data shows 121 regions (88%) as having a moderate or strong correlation between either ‘loss of sense of smell’ or ‘loss of sense of taste’ and new daily Covid-19 cases.

Weakly positive or negative correlations ($r_s 0.3$ to $-0.3$) were observed in 10 regions for ‘loss of smell’ (7.5%), 7 regions for ‘loss of sense of taste’ (5.1%), 12 regions for ‘sense of smell’ (8.8%) and 10 regions for ‘sense of taste’ (7.4%).

Moderate negative correlations ($r_s -0.3$ to $-0.65$) were not observed in any regions for ‘loss of sense of smell’, 2 regions for ‘loss of sense of taste’ (1.5%), 3 regions for ‘sense of smell’ (2.2%) and 5 regions for ‘sense of taste’ (3.7%). No strongly negative correlations were observed.

Brazil had the highest number of strong or moderate correlations at 97 out of 108 (90%), then France with 78 out of 88 (89%), Spain had 56 out of 68 (82%), Italy 52 out of 75 (69%) and then the USA with 130 out of 202 (64%). The USA had a tendency toward more regions with moderate than strong correlations, whereas for regions from other countries the reverse is true. Italy had the highest proportion of weak correlations with 8 (11%) and the USA the highest amount of moderate negative correlations (3.5%).

Our data indicate that in general for any individual region there is a temporal relationship between Google search volume for terms relating to anosmia and ageusia and new cases of Covid-19. This phenomenon appears strongest for search terms specifically mentioning loss of either smell or taste.

3.2. Between region analyses

Weighted search interest scores were produced by applying daily regional weighting data to the original time series data. Kolmogorov-Smirnov testing indicated the weighted search

| Search concept        | Country | Strong positive | Moderate positive | Weak positive | Weak negative | Non-significant |
|-----------------------|---------|-----------------|-------------------|--------------|--------------|----------------|
| Loss of sense of smell| Brazil  | 15              | 7                 | 2            | 0            | 0              |
|                       | Spain   | 8               | 8                 | 1            | 0            | 0              |
|                       | France  | 12              | 9                 | 0            | 0            | 0              |
|                       | Italy   | 8               | 5                 | 2            | 0            | 0              |
|                       | USA     | 6               | 32                | 4            | 1            | 0              |
| Loss of sense of taste| Brazil  | 17              | 6                 | 0            | 0            | 0              |
|                       | Spain   | 9               | 7                 | 1            | 0            | 0              |
|                       | France  | 17              | 5                 | 0            | 0            | 0              |
|                       | Italy   | 9               | 7                 | 1            | 0            | 0              |
|                       | USA     | 7               | 29                | 5            | 0            | 0              |
| Sense of smell        | Brazil  | 16              | 9                 | 1            | 0            | 0              |
|                       | Spain   | 11              | 5                 | 0            | 0            | 0              |
|                       | France  | 16              | 2                 | 2            | 0            | 0              |
|                       | Italy   | 11              | 5                 | 2            | 0            | 0              |
|                       | USA     | 5               | 27                | 5            | 2            | 0              |
| Sense of taste        | Brazil  | 19              | 8                 | 0            | 0            | 0              |
|                       | Spain   | 3               | 5                 | 1            | 0            | 0              |
|                       | France  | 6               | 11                | 2            | 0            | 0              |
|                       | Italy   | 3               | 4                 | 1            | 2            | 2              |
|                       | USA     | 3               | 21                | 3            | 0            | 12             |

**Table 2.** Summary of Spearman’s rank correlation test outcomes for search interest in terms relating to anosmia and ageusia and new daily Covid-19 cases per million (both data as 7-day moving-mean) the table shows counts of regions within each country and result group. $r_s$: Spearman’s rank correlation coefficient. * Significance level of p < 0.05.
interest data pooled by country was not normally distributed. Spearman’s rank correlation analysis was used to assess the relationship between weighted search interest (7-day moving mean) and daily new Covid-19 cases per million (7-day moving mean) over time and regions within each country. The results are shown in Table 3.

All four search concepts demonstrated a significant (p < 0.05), positive correlation with new daily Covid-19 cases. However, ‘loss of sense of smell’ and ‘loss of sense of taste’ had a higher range for $r_s$ when compared to ‘sense of smell’ and ‘sense of taste’. This finding conforms to our previous observation that interest in terms including loss of either smell or taste more closely follow population changes in Covid-19 cases.

The most obvious differences between countries are that Italy has a slightly lower range for the test statistic ($r_s = 0.141$ to 0.510) and Brazil a higher range ($r_s = 0.601$ to 0.637). These differences appear reduced with search interest for ‘loss of sense of smell’ and ‘loss of sense of taste’.

The data suggests for any given point in time within a single country that regions with higher Covid-19 cases will also have higher search interest for terms relating to anosmia and ageusia.

Within region analysis reveals that when considering a single region, higher numbers of Covid-19 cases typically correspond with higher search interest for terms relating to anosmia and ageusia. Between region analysis demonstrates an association between the relative increase in Covid-19 cases and relative increase in search interest for anosmia and ageusia that shows consistency across all regions within a country. Taken together this data is evidence of a widespread association between Covid-19 cases and search interest for anosmia and ageusia that is observed in regional geographies inside the studied countries.

### 4. Discussion

As often stated, a global pandemic is a series of smaller national epidemics which themselves consist of a series of smaller outbreaks. Walker, et al.’s findings demonstrate an association between search interest in loss of smell and new Covid-19 cases at a national level and we have now found evidence of this at a regional level reflecting smaller disease outbreaks [8]. When analyzing whether the relative size of search interest changes reflected numbers of Covid-19 cases between regions we were also able to demonstrate positive correlations.

Observed regional and national variations in the strength of the association we describe may be attributed to factors we could not account for. Such factors would include differences in internet usage behaviors, the level public awareness of Covid-19 symptoms and internet availability.

As we now know the significance of the reduction of sense of smell in Covid-19 positivity, finding ways of tracking the reporting of this symptom is important. Mobile application-based reporting tools, such as the one produced by Menni et al., could be one method of tracking cases based on symptomatology [22]. However, allowing access to geo-locations on a population wide level is controversial due to impacts on the right to privacy and as such may prove unpopular. Tracking search terms does not identify an individual and does not interfere with this human right.

Tracking symptomatology through Google is a newly developed technique and has transferability to other areas of medicine and future pandemics depending on the disease phenotype. Due to the recent increase in widespread usage of mobile internet and the use of Google on the biggest platforms as the primary search engine, this data has become more reliable when compared to previous years. Comparison with previous pandemics, due to this behavioral change in internet interaction, was not undertaken in this study for that reason.

Covid-19 has affected numerous nations with significantly disparate Gross Domestic Products (GDGs) and public health budgets. Sudan, for example, spends £25 per capita on health compared with £2989 in the UK [23,24]. Countries, which are not able to invest in PCR testing or future antibody assays, will look for more cost-effective means to track their outbreaks and reduce transmission. The use of mobile internet, in general, is high in these low-to-middle income countries and; 75% of the population of Sudan have access [25].

Excess mortality is now being adopted as the most accurate measure of Covid-19-related mortality due to differences in testing between nations and the relatively low sensitivity of the standard PCR test [26,27]. Loss of smell search terms have the potential to identify active cases on an international basis as long as access to the internet and subsequent search term data is freely available. This is something that is more easily accessible and available in less developed countries with healthcare infrastructure that is not well resourced. We do not yet know if similar trends will continue in countries who have already been impacted by a large number of Covid-19 cases but this can be assessed prospectively going forwards and modeling can be subsequently adapted.

### 5. Conclusion

We have demonstrated that there is clear association between Google Trends search terms relating to loss of smell and taste and Covid-19 cases both on a regional, national, and international basis. We feel that Google search trends relating to loss of smell can
be utilized to identify potential Covid-19 outbreaks on a regional basis within countries. This could help the implementation of targeted public health measures in these areas; which will be of particular benefit in low-to-middle income countries where testing is not widely available. If regional outbreaks are identified, spread throughout a country could be prevented whilst allowing other areas to continue with reduced restrictions therefore diminishing negative impact on economic growth and non-Covid-19-related healthcare activity which is of significant concern in itself.

Github repository
https://github.com/GeorgeCherry/anosmia_search_interest_covid-19

Funding
This paper was not funded.

Declaration of interest
The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

Reviewer disclosures
Peer reviewers on this manuscript have no relevant financial or other relationships to disclose.

Author contributions
GC contributed to study design, developed the software which retrieved the data, analyzed the data and drafted the Methods and Results section of the manuscript. JR contributed to the study design and drafted the Introduction, Discussion, and Conclusion sections of the manuscript. MC, JL, ML, VL, and BNK revised the manuscript and approved the final version for publication.

References
Papers of special note have been highlighted as either of interest (•) or of considerable interest (••) to readers.

1. Gane SB, Kelly C, Hopkins C. Isolated sudden onset anosmia in COVID-19 infection. A novel syndrome? Rhinology Journal. 2020 April 02;58(3):299–301.
2. Mao L, Jia H, Wang M, et al. Neurologic manifestations of hospitalized patients with coronavirus disease 2019 in Wuhan, China. JAMA Neurol. 2020 April 10;77(6):683.
3. Lechien JR, Cabaraux P, Chiesa-Estomba CM, et al. Objective olfactory evaluation of self-reported loss of smell in a case series of 86 COVID-19 patients. Head Neck. 2020 May;21:1583–1590.
4. Rocke J, Hopkins C, Philpott C, et al. Is loss of sense of smell as a diagnostic marker in COVID-19? A systematic review and meta-analysis. Authorthe 2020 [cited 2020 Apr 20]. DOI:10.22541/au.158757120.03114031
   •• This systematic review, currently under peer-review, identified the positive predictive value for olfactory dysfunction and a positive Covid-19 PCR swab. The studies included in the review agree that olfactory dysfunction is a key symptom in Covid-19 patients and is more prevalent than cough and fever in this cohort.
5. Lewnard JA, Lo NC. Scientific and ethical basis for social-distancing interventions against COVID-19. Lancet Infect Dis. 2020 March;23:631–633.
6. Anderson RM, Heesterbeek H, Klinkenberg D, et al. How will country-based mitigation measures influence the course of the COVID-19 epidemic? Lancet. 2020 March 21;395(10228):931–934.

https://github.com/GeorgeCherry/anosmia_search_interest_covid-19

availability,

%2F22541%3Au.158757120.03114031

7. Corral JE, Hoogenboom SA, Kroner PT, et al. COVID-19 polymerase chain reaction testing before endoscopy: An economic analysis. Gastrointest Endosc. 2020 April 8. DOI:10.1016/j.gie.2020.04.049.
8. 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. 2020 April 11;10:839–847.
   •• Walker et al were the first group to investigate Google trends on a national basis. They found positive correlations in search terms for loss of smell and Covid-19 cases.
9. Ofcom. Communications market report. 2018 [cited 2020 Apr 20]. Available from https://www.ofcom.org.uk/__data/assets/pdf_file/0022/117256-CMR-2018-narrative-report.pdf
10. Johns Hopkins University & Medicine. Coronavirus resource centre. [Internet]. Maryland: Johns Hopkins University; [cited 2020 April 20]. Available from https://coronavirus.jhu.edu/data/new-cases.
11. Google. FAQ about google trends data - trends help. [Internet]. California: Google; [cited 2020 May 11]. Available from https://support.google.com/trends/answer/4365533.
   • This link explains how Google Trends data is collected and how the popularity of a search term is scored on a regional basis.
12. International Organization for Standardisation. ISO online browsing platform (OBP). [Internet]. Geneva: International Organization for Standardization; [cited 2020 May 21]. Available from https://www.iso.org/obp/ui.
13. The Python Package Index. Pytrends - PyPi. [Internet]. Minneapolis: General Mills; [cited 2020 May 25]. Available from https://pypi.org/project/pytrends.
14. Presidencia del Consiglio dei Ministri - Departamento della Protezione Civile. Dati COVID-19 Italia. [Internet]. Rome: Dipartimento della Protezione Civile; [cited 2020 May 18]. Available from https://github.com/pcm-dpc/COVID-19/tree/master/dati-regioni.
15. Github. Time series summary. [Internet]. Maryland: Johns Hopkins - Centre for Systems Science and Engineering; [cited 2020 May 18]. Available from https://github.com/CSSEGISandData/COVID-19/tree/master/cesse_covid_19_data/cesse_covid_19_time_series.
16. Centro Nacional de Epidemiologia. COVID-19. [Internet]. Madrid: Centro Nacional de Epidemiologia; [cited 2020 May 19]. Available from https://cnecovid.isciii.es/covid19/#documentaci%C3%B3n-y-datos.
17. Republique Francaise. Données des urgences hospitalières et de SOS médics relatives à l’épidémie de COVID-19. [Internet]. Paris: Santé publique France; [cited 2020 May 19]. Available from https://www.data.gouv.fr/fr/datasets/donnees-des-urgences-hospitalieres-et-de-sos-medecins-relatives-a-lepidemie-de-covid-19/#.
18. Ministerio da Saúde. Coronavirus Brasil. [Internet]. Brasilia: Brasil pelo Ministério da Saúde; [cited 2020 May 19]. Available from https://covid.saude.gov.br/.
19. Instituto Nazionale di Statistica. Istat - Population and households. [Internet]. Rome: Istituto Nazionale di Statistica; [cited 2020 May 9]. Available from https://www.istat.it/en/population-and-households-data-and-indicators.
20. Instituto Nacional de Estadística. Resultados semestrales. [Internet]. Madrid: Centro Nacional de Epidemiología; [cited2020 May 11]. Available from https://www.ine.es/dynt3/inebase/es/index.htm?padre=19548&capsel=1900.
21. Institut national de la statistique et des études économiques. Populations légales. Paris: L’Institut national de la statistique et des études économiques; 2017. [Internet] cited 2020 May 16. Available from https://www.insee.fr/fr/statistiques/426530?sommaire=4265511
22. Menni C, Valdés AM, Freidin MB, et al. Loss of smell and taste in combination with other symptoms is a strong predictor of COVID-19 infection. Nature Medicine. 2020 May 11.
23. Barroy H, Musango L, Hsu J, et al. Public financing for health in Africa: from Abuja to the SDGs. 2016, cited 2020 Apr 20. Available from: www.data.gouv.fr/fr/datasets/donnees-des-urgences-hospitalieres-et-de-sos-medecins-relatives-a-lepidemie-de-covid-19/#.
24. Cooper J. How does UK healthcare spending compare with other countries? Office for National Statistics. 2019. Available at https://www.ons.gov.uk/

25. The World Bank. Mobile cellular subscriptions (per 100 people) - Egypt, Arab rep, Ethiopia, South Sudan, Sudan; 2020 [cited 2020 Apr 20]. Available at https://data.worldbank.org/indicator/IT.CEL.SETS.P2?locations=EG-ET-SS-SD

26. Wang W, Xu Y, Gao R, et al. Detection of SARS-CoV-2 in different types of clinical specimens. JAMA. 2020 March;11:2020.
   • Wang et al demonstrate the relatively low sensitivity of the current Covid-19 PCR test which reinforces the need for other strategies in tracking outbreaks other than testing alone.

27. Leon DA, Shkolnikov VM, Smeeth L, et al. COVID-19: a need for real-time monitoring of weekly excess deaths. Lancet. 2020;395 (10234):e81–6736.