This work was supported by a grant from the US Department of Veterans Affairs/Office of Research and Development as part of funding for VASeqCURE, which in turn received funding from the American Rescue Plan Act funds, with additional support from Central Texas Veterans Health Care System, Temple, Texas, USA.

About the Author

Dr. Choi is a research scientist at the Central Texas Veterans Health Care System, Temple, Texas, USA. His primary research interests focus on infectious disease and whole-genome sequencing.

References

1. Centers for Disease Control and Prevention. Epidemiology, testing, and management of extensively drug-resistant shigellosis [cited 2023 March 22]. https://emergency.cdc.gov/coca/calls/2023/callinfo_022823.asp
2. Ud-Din AI, Wahid SU, Latif HA, Shahnaij M, Akter M, Azmi IJ, et al. Changing trends in the prevalence of *Shigella* species: emergence of multi-drug resistant *Shigella sonnei* biotype g in Bangladesh. PLoS One. 2013;8:e82601. https://doi.org/10.1371/journal.pone.0082601
3. Lee I, Ouk Kim Y, Park SC, Chun J. OrthoANI: An improved algorithm and software for calculating average nucleotide identity. Int J Syst Evol Microbiol. 2016;66:1100–3. https://doi.org/10.1099/ijsem.0.000760
4. Larsen MV, Cosentino S, Lukjancenko O, Saputra D, Rasmussen S, Hasman H, et al. Benchmarking of methods for genomic taxonomy. J Clin Microbiol. 2014;52:1529–39. https://doi.org/10.1128/JCM.02981-13
5. van den Beld MJC, Reubsaet FAG, Pijnacker R, Harpal A, Kuiling S, Heerkens EM, et al. A multifactorial approach for surveillance of *Shigella* spp. and entero-invasive *Escherichia coli* is important for detecting (inter)national clusters. Front Microbiol. 2020;11:564103. https://doi.org/10.3389/fmicb.2020.564103
6. Larsen MV, Cosentino S, Rasmussen S, Friis C, Hasman H, Marvig RL, et al. Multilocus sequence typing of total-genome-sequenced bacteria. J Clin Microbiol. 2012;50:1355–61. https://doi.org/10.1128/JCM.06094-11
7. Clausen PTLC, Aarestrup FM, Lund O. Rapid and precise alignment of raw reads against redundant databases with KMA. BMC Bioinformatics. 2018;19:307. https://doi.org/10.1186/s12859-018-2336-6
8. Bortolaia V, Kaas RS, Ruppe E, Roberts MC, Schwarz S, Cattoir V, et al. ResFinder 4.0 for predictions of phenotypes from genotypes. J Antimicrob Chemother. 2020;75:3491–500. https://doi.org/10.1093/jac/dkaa345
9. Malberg Tetzschner AM, Johnson JR, Johnston BD, Lund O, Scheutz F. In silico genotyping of *Escherichia coli* isolates for extraintestinal virulence genes by use of whole-genome sequencing data. J Clin Microbiol. 2020;58:e01269-20. https://doi.org/10.1128/JCM.01269-20
10. Thorley K, Charles H, Greig DR, Prochazka M, Mason LCE, Baker KS, et.al. Emergence of extensively drug-resistant and multidrug-resistant *Shigella flexneri* serotype 2a associated with sexual transmission among gay, bisexual, and other men who have sex with men, in England: a descriptive epidemiological study. Lancet Infect Dis. 2023:S1473–3099:00807–6. https://doi.org/10.1016/S1473-3099(22)00807-6

Address for correspondence: Chetan Jinadatha, Central Texas Veterans Health Care System, 1901 S Veterans Dr, Temple, TX 76504, USA; email: Chetan.Jinadatha@va.gov

Longitudinal Association of COVID-19 Hospitalization and Death with Online Search for Loss of Smell or Taste

Derek Toomre, Sasikiran Kandula, Jeffrey Shaman

Author affiliations: Yale University School of Medicine, New Haven, Connecticut, USA (D. Toomre); Columbia University Mailman School of Public Health, New York, New York, USA (S. Kandula, J. Shaman); Columbia University School of Climate, New York (J. Shaman)

DOI: https://doi.org/10.3201/eid2908.230071

Surveillance of COVID-19 is challenging but critical for mitigating disease, particularly if predictive of future disease burden. We report a robust multiyear lead-lag association between internet search activity for loss of smell or taste and COVID-19–associated hospitalization and deaths. These search data could help predict COVID-19 surges.

A challenge throughout the COVID-19 pandemic has been forecasting surges in hospitalizations and deaths so that health officials can plan and mitigate accordingly. However, effective COVID-19 surveillance and forecasting has been complicated by numerous factors: reported new cases variably underestimate true incidence; wastewater surveillance of SARS-CoV-2 is limited; variants have different virulence levels (*1*); and the risk for severe outcomes depends on previous immunizations, infections, and duration of the immune response, which is increasingly heterogeneous and variant-dependent. Ideally, independent proxies could help surveil the risk for increases in levels of severe COVID-19 disease; however, such proxies should be predictive and include a sufficient lead-lag relationship to enable public health mitigation. We investigated a possible lead-lag relationship between Google searches for "loss of smell" and "loss of taste" and COVID-19 hospitalizations and deaths.

Online search activity has previously been shown to have some predictive power for other diseases (*2*). Multiple symptoms are associated with COVID-19, but "new loss of smell or taste" is highly specific (odds ratio ≈10) (*3*). Loss of taste is confounded because flavor occurs partly through retronasal olfaction, and most persons do not differentiate between changes in taste versus flavor. In psychophysical smell and taste tests of persons with acute COVID-19, 72% had an olfactory defect and 19% had a gustatory defect (*4*). Early studies in the pandemic noted a correlation between Google Trends searches for loss of smell and taste and COVID-19 cases (*5*,*6*). This correlation occurred even before anosmia was publicly recognized as a COVID-19 symptom (*6*), underscoring the possibility that olfactory and gustatory symptoms are useful indicators for COVID-19 surveillance.

SARS-CoV-2–induced olfactory dysfunction has been studied at the cellular level and in human trials (*7*). Nasal sustentacular epithelial cells adjacent to olfactory neurons have high angiotensin-converting enzyme 2 receptor levels and are a key site of virus replication. SARS-CoV-2 enters cells either by fusing at the cell surface or in endosomes (*7*). Those 2 pathways vary between cell and tissue types; respiratory and olfactory epithelial cells use endosomal and cell surface pathways, and cell surface pathways require activation by cell surface proteases (e.g., TMPRSS2) (*7*). Mutations associated with Omicron caused it to be TMPRSS2-resistant (*8*) and display enhanced replication in the upper respiratory tract, consistent with less severe lung disease, lower mortality rates (*9*), and less frequent self-reported olfactory dysfunction (*10*). A hypothetical correlate is that olfactory dysfunction might be a proxy for general risk for infection of lung cells at the population level. Given this potential link, we examined whether Internet searches for "loss of smell" and "loss of taste" correlate with waves of COVID-19 deaths with a lead-lag relationship, and if so, whether that correlation is maintained across different waves of COVID-19 variants.

To robustly test for a potential association, we analyzed Google Trends searches for "loss of smell" and "loss of taste" across 5 different English-speaking countries and 3 different years (2020, 2021, and 2022) and examined the correlation to reported COVID-19 hospitalizations and deaths (Figure). We retrieved weekly query frequencies for "loss of smell" (or "anosmia") and "loss of taste" (or "ageusia") from the Google Extended Trends API for Australia, Canada, South Africa, the United Kingdom, and the United States. Using public sources, we computed weekly COVID-19–associated mortality and hospitalization

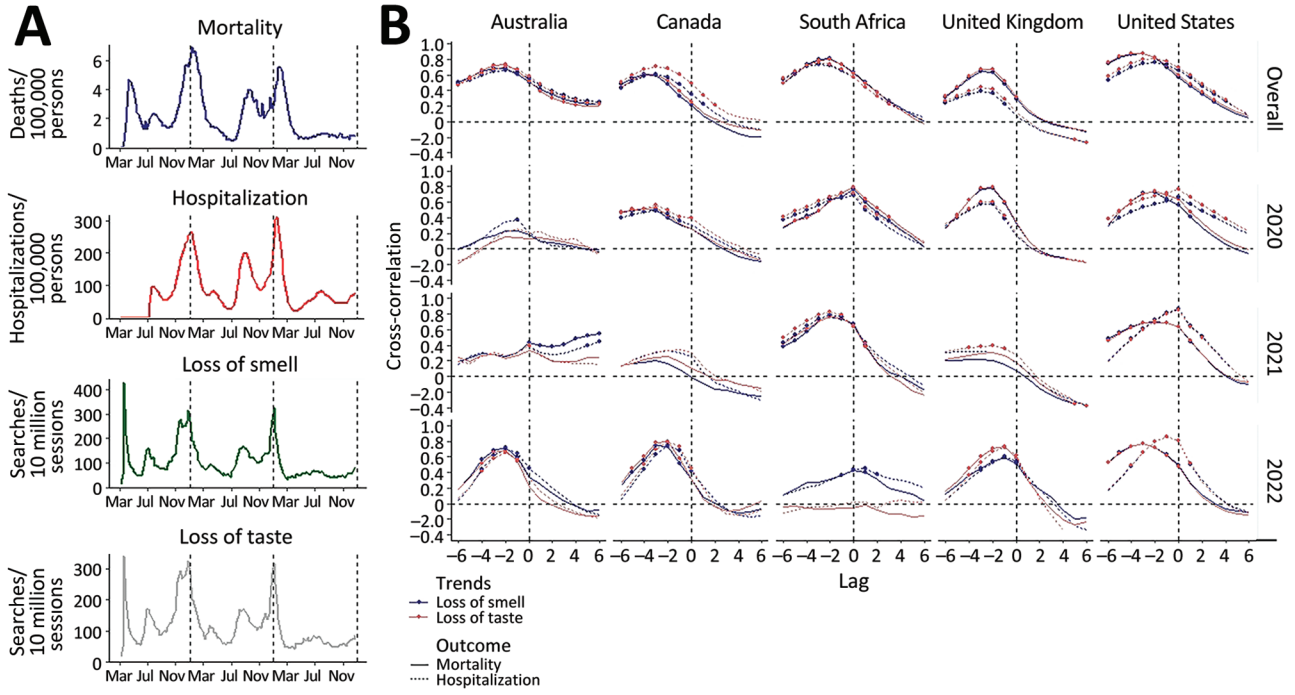

Figure. Longitudinal association of COVID-19 hospitalization and death with online search for loss of smell or taste. A) Weekly COVID-19–associated deaths (per 100,000 population), hospitalizations (per 100,000 population), and Google search trends for 'loss of smell' and 'loss of taste' (per 10 million search sessions) in the United States during March 2020–September 2022. Vertical broken lines delimit calendar years. B) Cross-correlation between Google trends of the 2 search queries, and the 2 outcomes in 5 countries (columns) over the entire COVID-19 pandemic period of March 2020–September 2022 (top row) and disaggregated by calendar year. Statistically significant correlations ($p<0.01$) are indicated by a data point. Lag between paired search trend and outcome is shown in weeks.

rates for February 2020–August 2022. For each country, we computed cross-correlation between paired search trend and outcome for each week between −6 (lead) and 6 (lag) for the study period and each calendar year (Appendix, https://wwwnc.cdc.gov/EID/article/29/8/23-0071-App1.pdf).

We observed a strong correlation in the United States between deaths, hospitalization, and searches for loss of smell or taste with surprisingly similar amplitudes for all major waves (Figure, panel A), including those associated with Omicron in December 2021. Cross-correlation was high (0.68–0.85) and significant ($p<0.01$) across all 5 countries; the peak trend for loss of smell or taste preceded hospitalization and deaths by 2–3 weeks (Figure, panel B). This correlation was seen across all years combined and was evident for most country–year combinations. The association appeared weak in years when outcome rates were low (e.g., Australia in 2020). The analysis indicates the correlation is robust over 3 years and multiple variant waves and that loss of smell or taste might give officials a useful lead indicator of the risk for COVID-19–associated hospitalizations and deaths. However, if this finding is to be used predictively, the persistence of this association would need to be closely tracked and monitored.

Strengths of this investigation are the long-duration longitudinal analysis across multiple countries, the use of simple search criteria and variable search terms, and analysis of the temporal lead-lag relationship. Limitations include potential for bias on the basis of media news cycles, the population scale of the analysis, and socioeconomic selection bias related to internet access. Future correlations will need to be monitored. Search activity might be a more useful indicator of infection levels than COVID-19–associated deaths. Despite these caveats, this accessible metric should be considered as a public health predictor.

This work was supported by Centers for Disease Control and Prevention contract no. 75D30122C14289.

D.T. declares competing interests as a founder of olfactory test company (u-Smell-it LLC) and for European Union patent (DM/212486) and pending US patents (29,743,100 and 29,750,313). J.S. and Columbia University disclose partial ownership of SK Analytics. J.S. discloses consulting for BNI.

About the Author

Dr. Toomre is a professor in the department of cell biology at Yale University School of Medicine. His research focus is the quantitative studies of exocytic traffic in living cells, including primary cilia, which are specialized cellular antennae that play roles in vision and smell.

References

1. Yang W, Shaman JL. COVID-19 pandemic dynamics in South Africa and epidemiological characteristics of three variants of concern (Beta, Delta, and Omicron). eLife. 2022;11:e78933. https://doi.org/10.7554/eLife.78933
2. Yang S, Santillana M, Kou SC. Accurate estimation of influenza epidemics using Google search data via ARGO. Proc Natl Acad Sci U S A. 2015;112:14473–8. https://doi.org/10.1073/pnas.1515373112
3. Payne DC, Smith-Jeffcoat SE, Nowak G, Chukwuma U, Geibe JR, Hawkins RJ, et al.; CDC COVID-19 Surge Laboratory Group. SARS-CoV-2 infections and serologic responses from a sample of U.S. Navy Service Members—USS Theodore Roosevelt, April 2020. MMWR Morb Mortal Wkly Rep. 2020;69:714–21. https://doi.org/10.15585/mmwr.mm6923e4
4. Prem B, Liu DT, Besser G, Sharma G, Dultinger LE, Hofer SV, et al. Long-lasting olfactory dysfunction in COVID-19 patients. Eur Arch Otorhinolaryngol. 2022;279:3485–92. https://doi.org/10.1007/s00405-021-07153-1
5. Pierron D, Pereda-Loth V, Mantel M, Moranges M, Bignon E, Alva O, et al. Smell and taste changes are early indicators of the COVID-19 pandemic and political decision effectiveness. Nat Commun. 2020;11:5152. https://doi.org/10.1038/s41467-020-18963-y
6. Walker A, Hopkins C, Surda P. Use of Google Trends to investigate loss-of-smell-related searches during the COVID-19 outbreak. Int Forum Allergy Rhinol. 2020;10:839–47. https://doi.org/10.1002/alr.22580
7. Jackson CB, Farzan M, Chen B, Choe H. Mechanisms of SARS-CoV-2 entry into cells. Nat Rev Mol Cell Biol. 2022;23:3–20. https://doi.org/10.1038/s41580-021-00418-x
8. Meng B, Abdullahi A, Ferreira IATM, Goonawardane N, Saito A, Kimura I, et al.; CITIID-NIHR BioResource COVID-19 Collaboration; Genotype to Phenotype Japan (G2P-Japan) Consortium; Ecuador-COVID19 Consortium. Altered TMPRSS2 usage by SARS-CoV-2 Omicron impacts infectivity and fusogenicity. Nature. 2022;603:706–14. https://doi.org/10.1038/s41586-022-04474-x
9. Nyberg T, Ferguson NM, Nash SG, Webster HH, Flaxman S, Andrews N, et al.; COVID-19 Genomics UK (COG-UK) consortium. Comparative analysis of the risks of hospitalisation and death associated with SARS-CoV-2 omicron (B.1.1.529) and delta (B.1.617.2) variants in England: a cohort study. Lancet. 2022;399:1303–12. https://doi.org/10.1016/S0140-6736(22)00462-7
10. Cardoso CC, Rossi ÁD, Galliez RM, Faffe DS, Tanuri A, Castiñeiras TMPP. Olfactory dysfunction in patients with mild COVID-19 during Gamma, Delta, and Omicron waves in Rio de Janeiro, Brazil. JAMA. 2022;328:582–3. https://doi.org/10.1001/jama.2022.11006

Address for correspondence: Derek Toomre, Yale University School of Medicine, Department of Cell Biology, SHM-C227, 333 Cedar St, PO Box 208002, New Haven, CT 06520-8002, USA; email: derek.toomre@yale.edu; Jeffrey Shaman, Columbia University Mailman School of Public Health, Department of Environmental Health Sciences, 722 W 168th St, New York, NY 10032, USA; email: jls106@cumc.columbia.edu

Article DOI: https://doi.org/10.3201/eid2908.230071

Longitudinal Association of COVID-19 Hospitalization and Death with Online Search for Loss of Smell or Taste

Appendix

```
        ## Aggregate daily data to weekly resolutions
aggregateToWeek <- function(x){
  x <- merge(x, date.week.map, by = 'date')
  x <- x[, j = .(deaths = sum(deaths,na.rm = T),
                 hosp = sum(hosp,na.rm = T),
                 pop = mean(pop)),
         by = .(region, year, week)]

  x <- merge(x,
             date.week.map[day =  = 1, j = .(date, year, week)],
             by = c('year', 'week'))

  x[,j = .(date,
           region,
           deaths = deaths*1E5/pop, # per 100000 population
           hosp = hosp*1E5/pop, # per 100000 population
           pop)]
}
```

```
## Function to find for correlations for outcomes vs trends in a county
doCountry <- function(req.region = 'United States', req.nrands = 10){
  x <- merge(d[region =  = req.region and variable %in% c('smell'),
               j = .(date, smell = value)],
             mort[region =  = req.region, j = .(date, deaths)],
             by = 'date')
  x <- merge(x, d[region =  = req.region and variable %in% c('taste'),
                  j = .(date, taste = value)],
             by = 'date')
  x <- merge(x, mort[region =  = req.region, j = .(date, hosp)],
             by = 'date', all.x = T)



  smell <- cbind(x$date, scale(x$smell))
  taste <- cbind(x$date, scale(x$taste))
  mort <- cbind(x$date, scale(x$deaths))
  hosp <- cbind(x$date, scale(x$hosp))



  # Calculate cross-correlation for each pair and stack
  rbind(
    data.table(var1 = 'smell', var.2 = 'mort',
               pairSignalWrapper(smell, mort)),
    data.table(var1 = 'taste', var.2 = 'mort',
               pairSignalWrapper(taste, mort)),
    data.table(var1 = 'smell', var.2 = 'hosp',
               pairSignalWrapper(smell, hosp)),
    data.table(var1 = 'taste', var.2 = 'hosp',
               pairSignalWrapper(taste, hosp))
  )
}
```

```
### For the two time series, find correlations over the entire period and
### by calendar year
pairSignalWrapper <- function(x, y){
  overall <- pairSignal(x[,2], y[,2]) # over the entire study period
  y20 <- pairSignal(x[checkDate(x, 2020),2],
                    y[checkDate(y, 2020),2]) # calendar year 2020 only
  y21 <- pairSignal(x[checkDate(x, 2021),2],
                    y[checkDate(y, 2021),2]) # calendar year 2021 only
  y22 <- pairSignal(x[checkDate(x, 2022),2],
                    y[checkDate(y, 2022),2]) # calendar year 2022 only

  ret <- rbind(
    data.table(period = 'Overall', overall),
    data.table(period = '2020', y20),
    data.table(period = '2021', y21),
    data.table(period = '2022', y22))

  ret$period <- factor(ret$period,
                       levels = c('Overall', as.character(2020:2022)))

  ret
}


## Function to calculate cross-correlation between the two signals
pairSignal <- function(x, y){
  obj <- ccf(x, y, lag.max = 6, plot = F, type = 'correlation')
  p.value <- 2* (1 - pnorm(abs(obj$acf),
                           mean = 0,
```

```
                    sd = 1/sqrt(obj$n.used))) %>%
   round(4)
  ret <- data.frame(lag = −6:6, coeff = round(obj$acf, 4), p.value)

  ret
}


## Check if date is in a calendar year
checkDate <- function(x, req.year = 2020){
  year(as.Date(x[, 1], '1970–01–01')) =  = req.year
}
## Load mortality and hospitalization data from Our World in Data and
## aggregate daily data to week
<- readRDS(paste0(baseDir, 'owid-covid-data_v2.Rds'))
mort <- mort[date < = '2022–12–31',
          j = .(region = location, date, deaths = new_deaths, hosp = hosp_patients,
pop = population)] %>%
  aggregateToWeek()



## Load trends from Google Health Trends API
d <- readRDS(paste0(baseDir, 'Trends_v2.Rds'))
setnames(d, c('region', 'date', 'smell', 'taste'))
d$date <- as.Date(d$date)
d <- melt(d[date> = '2020–01–01' and date < = '2022–12–31'], id.vars = c('region', 'date'),
variable.factor = F)


##  Stack datasets
```

```
temp <- rbind(mort[,.(variable = 'mort', region, date, value = deaths)],
        mort[,.(variable = 'hosp', region, date, value = hosp)], d, use.names = T)
temp$variable <- factor(temp$variable,
                levels = c('mort', 'hosp', 'smell', 'taste'),
                labels = c('Mortality', 'Hospitalization', 'Loss of smell', 'Loss of taste'))
```

Calculate correlations

```
ret <- lapply(unique(temp$region),
        doCountry)
```

post process to clean outcomes and trend labels

```
names(ret) <- unique(temp$region)
ccf.pairs <- rbindlist(ret, idcol = 'region')

ccf.pairs$var1 <- factor(ccf.pairs$var1,
                 levels = c('smell', 'taste'),
                 labels = c('loss of smell', 'loss of taste'))
ccf.pairs$var2 <- factor(ccf.pairs$var2,
                 levels = c('mort', 'hosp'),
                 labels = c('Mortality', 'Hospitalization'))
```