

73515_Auto_Edited.docx

Name of Journal: *World Journal of Methodology*

Manuscript NO: 73515

Manuscript Type: ORIGINAL ARTICLE

Clinical and Translational Research

COVID-19 and thyroid disease: An infodemiological pilot study

Ioannis Ilias, Charalampos Milionis, Eftychia Koukkou

Abstract

BACKGROUND

Google Trends searches for symptoms and/or diseases may reflect actual disease epidemiology. Recently, Google Trends searches for COVID-19-associated terms have been linked to the epidemiology of COVID-19. Some studies have linked COVID-19 with thyroid disease.

AIM

In this study we aimed to assess COVID-19 cases *per se* vs COVID-19-associated Google Trends searches and thyroid-associated Google Trends searches.

METHODS

We collected data on worldwide weekly GT searches regarding “COVID-19”, “SARS-COV-2”, “coronavirus”, “smell”, “taste”, “cough”, “thyroid”, “thyroiditis” and “subacute thyroiditis” for 92 wk and worldwide weekly COVID-19 cases' statistics in the same time period. The study period was split in half (approximately corresponding to the preponderance of different SARS-COV-2 virus variants) and in each time period we performed cross-correlation analysis and mediation analysis.

RESULTS

Significant positive cross-correlation function values were noted in both time periods. More in detail, COVID-19 cases per se were found to be associated with no lag with Google Trends searches for COVID-19 symptoms in the first time period and in the second time period to lead searches for symptoms, COVID-19 terms as well as thyroid terms. COVID-19 cases per se were associated with thyroid-related searches in both time periods. In the second time period, the effect of “COVID-19” searches on “thyroid” searches was significantly mediated by COVID-19 cases ($P = 0.048$).

CONCLUSION

Searches for a non-specific symptom or COVID-19 search terms mostly led Google Trends thyroid-related searches, in the second time period. This time frame/sequence particularly in the second time period (noted by the preponderance of the SARS-COV-2 delta variant), lends some credence to associations of COVID-19 cases per se with (apparent) thyroid disease (via searches for them).

4

INTRODUCTION

Digital epidemiology uses digital data which was not generated with the primary goal of serving epidemiological research ^[1]; such data are within the domain of “infodemiology” ^[2]. Google Trends (available at <https://trends.google.com>) searches may - according to some researchers - accurately reflect the epidemiology of infectious, acute or chronic diseases, including among others coronary or thyroid disease ^[2-10]. Recently, Google Trends searches for COVID-19-associated terms have been tentatively linked to the epidemiology of COVID-19 ^[11-17]. Some - but not all - clinical studies have linked COVID-19 with thyroid function abnormalities and more particularly with a form of subacute-like thyroiditis ^[18-22]. Since ² the use of Google Trends to study a wide range of medical topics is becoming more widespread and the available research on COVID-19-related thyroid disease is conflicting, with this work we aimed to look at the

issue of COVID-19-related thyroid disease from a different angle, namely that of digital epidemiology, since the latter may be a useful adjunct to classical epidemiology.

MATERIALS AND METHODS

Data and Data Collection

We collected data on worldwide weekly Google Trends searches, by means of their “relative search volumes” (RSVs). The latter are normalized internet search volume values over a given time period, with a minimum of 0 and a maximum of 100; see also <https://support.google.com/trends/>. More in detail, we used the worldwide RSVs of the search terms in the English language for “COVID-19”, “SARS-CoV-2”, “coronavirus”, “smell”, “taste”, “cough”, “thyroid”, “thyroiditis” and “subacute thyroiditis” for 92 wk, from January 26, 2020 to October 24, 2021. The search terms were chosen because of their ubiquity and uniformity in lay and medical terms. For the same time period, worldwide weekly COVID-19 cases' statistics, as provided by the Johns Hopkins University Coronavirus Resource Center (available at <https://coronavirus.jhu.edu/map.html>) were collected ^[23]. The study period was split in half: the first half corresponded to the time period with preponderance of the SARS-CoV-2 alpha variant and the second to the time period with preponderance of the delta variant. (Figure 1).

Statistical Analysis

In each of the aforementioned time periods we performed cross-correlation analysis. The threshold for statistical significance of each cross-correlation factor value at the $P = 0.05$ Level was set according to lag, thus the cross-correlation factor had to be higher than 0.290 at lag=0 and 0.324 at lag=8. A lag = 0 indicates contemporaneous correlation, a negative lag indicates that the first variable leads within a set time frame the second variable, and a positive lag indicates that the first variable follows (lags) within a set time frame the second variable. After the calculation of cross-correlation factor values, further evaluation among the variables was done with mediation analysis, implementing Sobel's test. Statistical analyses were done with Minitab v.17.1 (Minitab

Inc, State College, PA, USA, 2010) and JASP v0.15 (JASP Team, University of Amsterdam, NL, 2021).

RESULTS

Worldwide, COVID-19 wkly cases per se gradually increased over time and showed wide fluctuations during the second half of the study period (Figure 1). The RSVs of the studied search terms also showed fluctuations (Figure 1 and Supplemental Figures 1 and 2). Significant positive cross-correlation factor values were noted in both time periods. More in detail, significant cross-correlation factors for “COVID-19” and “SARS-CoV-s” were mostly found in the second time period (Table 1), whereas COVID-19 cases per se were associated with “thyroid” searches in both time periods. In the second time period, the effect of “COVID-19” searches on “thyroid” searches was significantly mediated by COVID-19 cases (Sobel test statistic $P = 0.048$).

DISCUSSION

COVID-19 cases per se were found to be associated with no lag with Google Trends searches for COVID-19 symptoms in the first time period and in the second time period to lead searches for symptoms, COVID-19 terms as well as thyroid terms. Searches for a non-specific symptom or COVID-19 search terms mostly led Google Trends “thyroid” searches, in the second time period. This time frame/sequence particularly in the second time period, which was noted by the preponderance of the SARS-CoV-2 delta variant, lends some credence to associations of COVID-19 cases per se with (apparent) thyroid disease (via searches for them). Moreover, this finding, points to a possible higher probability of thyroid disease with SARS-CoV-2 delta variant compared to the alpha variant (and may also explain discrepancies regarding COVID-19 *vs* thyroid disease among previous relevant studies). Digital health is in the spotlight as the COVID-19 crisis progresses [24,25]. At the same time, digital epidemiology is emerging at a very fast pace [25]. More and more of what we do and say - including epidemiologically relevant behaviors - is stored

electronically, often in an accessible form. Internet data mining has ³ a revolutionary impact on the way we monitor global health and health behaviors. Infectious and chronic disease data can be collected and disseminated in almost real time through a number of online sources. Google Trends provides a powerful measure of public interest in a topic, being a proxy of internet searches for it. The frequency of internet searches for disease terms may not reflect directly the epidemiological characteristics of a given disease, which is related and/or described by such search terms. Media coverage may skew subsequent internet searches. Nevertheless, the frequency of internet queries for various diseases' symptoms are correlated to a degree with physician visits for these diseases [26,27]. Google Trends has been used – despite its shortcomings – to monitor the yearly influenza epidemics [24,28]. Another source that has provided health data is Twitter. A smartphone application can be used to assess COVID-19 symptoms and may indicate future disease hotspots within 5–7 days [29]. The collection and classification of data ranging from the detection of suspected cases to the monitoring and assessment of pandemic risk is crucial. However, as this is a very evolving field, validation of digital health measures vis-à-vis input data, tentative associations or predictive models is still needed. Regarding COVID-19, the influence of media on Google Trends RSVs has been studied and was found to be maximal after a week [30], whereas the effect of COVID-19 cases on “COVID-19” searches has been studied [31], and has been found to be most notable after 11.5 days [32]. Thus, with the lags in the observed cross-correlation factors we believe that the Google Trends searches for COVID-19 and/or thyroid-related items may reflect personal interest fuelled by probable real disease (COVID-19 or thyroid disease). Receptors for the SARS-CoV-2 virus are found in tissues beyond the respiratory system, such as the thyroid, thus an effect of COVID-19 on the thyroid is plausible [33]. Indeed, there is some evidence of thyroid dysfunction in patients with COVID-19, characterized by changes in hormone levels (low triiodothyronine or low thyrotropin levels) or laboratory results compatible with the presence of subacute thyroiditis [20,34]. Italian researchers observed that in the spring of 2020, 15% of COVID-19 patients ($n = 93$)

admitted to the intensive care unit (ICU) at a hospital in Milan had changes in thyroid hormones. By comparison, only 1% of patients in the same period of 2019 ($n = 101$) had changes in thyroid hormones [19]. Considering the fact that viral infections can cause thyroiditis, the researchers began a monitoring program to look at thyroid function 3 mo after COVID-19 treatment. The researchers found that thyroiditis, in patients with moderate to severe COVID-19, was different from common subacute thyroiditis: many patients had mild dysfunction and the rate of thyroid disease was higher in men. Thyroid dysfunction appeared to be associated with more severe COVID-19 disease. After 3 mo, thyroid function was normal in all followed-up patients ($n = 53$), with persistence of ultrasound findings of thyroiditis in one third of them [35]. Another study from Greece was based on the premise that the interpretation of thyroid tests in ill patients is hampered by changes that ensue in the context of non-thyroidal illness syndrome and studied thyroid function in cohorts of COVID-19 positive ($n = 102$, 46 in the ICU) and COVID-19-negative patients ($n = 94$, 41 in the ICU) [18]. The researchers noted a non-thyroidal illness syndrome pattern in 60% of ICU and 36% of ward patients (with no significant differences between COVID-19 positive or negative patients) [18]. The thyroid laboratory work-up was compatible with thyrotoxicosis in 14.6% of SARS-CoV-2 ICU patients vs 7.7% in ICU negative ($P = NS$) and, overall in 8.8% of SARS-CoV-2 positive vs 7.4% of negative patients. Thus, the authors concluded that a non-thyroidal illness syndrome pattern is common in COVID-19 but it relates to the severity of disease rather than SARS-CoV-2 infection, whereas a thyrotoxicosis pattern was less frequently observed and was no different between patients with and without COVID-19 [18]. Our study has several limitations and its caveats have to be considered. We collected only Google Trends data for English-language searches; however, we have shown in an older study that searches in this language dwarf searches in all other languages [4]. Additionally Northern hemisphere internet searches dwarf Southern hemisphere searches [4]. Analyses were done on a weekly worldwide basis since Google Trends searches for extended time periods are provided as such. From the literature, worldwide and weekly or monthly Google Trends data are considered to be more

reliable than country-wide and daily data ^[36,37]. No periodicity in the data was assessed since the total time duration of data collection was rather short. As stated above, the datasets were split in half given the vast differences in COVID-19 epidemiology in 2020-2021 due to the preponderance of different SARS-CoV-2 variants. Finally we have to bear in mind the fact that Google Trends searches are limited to internet-literate persons, who are easily influenced by media items, although few (medical) research articles are reported by news outlets (targeting diverse audiences) and generate public interest ^[38].

CONCLUSION

+ADw-html+AD4APA-p+AD4Given the relatively recent onset of SARS-CoV-2 virus infection, the available monitoring data are limited in time and therefore long-term studies are needed to evaluate even longer-term effects on the endocrine glands. Research into the virus continues to grow, shedding more light on the real health risks posed by COVID-19. Ideally, it would be interesting to assess time and localization-delimited Google Trends searches with the corresponding thyroid disease incidence, as reported by +ACY-ldquo+ADs-sentinel+ACY-rdquo+ADs- physicians or as recorded in healthcare databases, to verify the associations observed. Understanding the nature of a pandemic of this magnitude means saving human lives and proper knowledge of ways to prevent further infection.+ADw-/p+AD4APA-/html+AD4-

ARTICLE HIGHLIGHTS

Research background

Google Trends searches for symptoms and/or diseases may reflect actual disease epidemiology. Recently, Google Trends searches for COVID-19-associated terms have been linked to the epidemiology of COVID-19. Some studies have linked COVID-19 with thyroid disease.

Research motivation

Since the use of Google Trends to study a wide range of medical topics is becoming more widespread and the available research on COVID-19-related thyroid disease is conflicting, with this work we aimed to look at the issue of COVID-19-related thyroid disease from a different angle, namely that of digital epidemiology, since the latter may be a useful adjunct to classical epidemiology.

Research objectives

We assessed worldwide COVID-19 cases per se *vs* COVID-19-associated Google Trends searches and thyroid-associated Google Trends searches for 92 wk.

Research methods

We collected data on worldwide weekly GT searches regarding “COVID-19”, “SARS-COV-2”, “coronavirus”, “smell”, “taste”, “cough”, “thyroid”, “thyroiditis” and “subacute thyroiditis” for 92 wk and worldwide weekly COVID-19 cases' statistics in the same time period. The study period was split in half (approximately corresponding to the preponderance of different SARS-COV-2 virus variants) and in each time period we performed cross-correlation analysis and mediation analysis.

Research results

Significant positive cross-correlation function values were noted in both time periods. More in detail, COVID-19 cases per se were found to be associated with no lag with Google Trends searches for COVID-19 symptoms in the first time period and in the second time period to lead searches for symptoms, COVID-19 terms as well as thyroid terms.

Research conclusions

Searches for a non-specific symptom or COVID-19 search terms mostly led Google Trends thyroid-related searches, in the second time period. This time frame/sequence particularly in the second time period (noted by the preponderance of the SARS-COV-2

delta variant), lends some credence to associations of COVID-19 cases per se with (apparent) thyroid disease (via searches for them).

Research perspectives

Given the relatively recent onset of SARS-CoV-2 virus infection, the available monitoring data are limited in time and therefore long-term studies are needed to evaluate even longer-term effects on the endocrine glands. Research into the virus continues to grow, shedding more light on the real health risks posed by COVID-19. Ideally, it would be interesting to assess time and localization-delimited Google Trends searches with the corresponding thyroid disease incidence, as reported by “sentinel” physicians or as recorded in healthcare databases, to verify the associations observed. Understanding the nature of a pandemic of this magnitude means saving human lives and proper knowledge of ways to prevent further infection.

6%

SIMILARITY INDEX

PRIMARY SOURCES

1	ec.bioscientifica.com Internet	80 words — 3%
2	en.m.wikipedia.org Internet	34 words — 1%
3	Salathé, Marcel, Linus Bengtsson, Todd J. Bodnar, Devon D. Brewer, John S. Brownstein, Caroline Buckee, Ellsworth M. Campbell, Ciro Cattuto, Shashank Khandelwal, Patricia L. Mabry, and Alessandro Vespignani. "Digital Epidemiology", PLoS Computational Biology, 2012. Crossref	28 words — 1%
4	www.endocrine-abstracts.org Internet	18 words — 1%

EXCLUDE QUOTES ON
EXCLUDE BIBLIOGRAPHY ON

EXCLUDE MATCHES OFF