medRxiv preprint doi: https://doi.org/10.1101/2020.08.09.20171181; this version posted August 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

The impact of COVID-19 on Italy web users: a quantitative analysis of regional hygiene interest and emotional response

Alessandro Rovetta[1,2], *Lucia Castaldo*[1,2]

[1]Mensana srls research and disclosure division, Via Moro Aldo 5 - 25124 Brescia, Italy

[2]Redeev technological and scientific research, Via San Pasquale a Chiaia 83, Naples, Italy

Corresponding author

Alessandro Rovetta,

Mensana srls, Research and Disclosure Division,

Via Moro Aldo 5 – 25124 Brescia, Italy.

Redeev Technological and Scientific Research,

Via San Pasquale a Chiaia 83, Naples, Italy.

Email: rovetta.mresearch@gmail.com

Phone: 39-3927112808

ORCID: 0000-0002-4634-279X

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.08.09.20171181; this version posted August 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Abstract

Background: Between the end of February and the beginning of June 2020, Italy was certainly one of the most affected countries in the world by the COVID-19 pandemic. In this period, the web interest in the novel coronavirus has undergone a drastic surge.

Objective: The aim of this study is to analyze quantitatively the impact of COVID-19 on web searches related to hygiene-preventive measures and emotional-psychological aspects, as well as to estimate the effectiveness and limits of online information during an epidemic. We looked for significant correlations between COVID-19 relative search volumes and cases per region to understand the interest of the average Italian web user in the international, national, and regional COVID-19 situation. By doing so, from the analysis of web searches, it will be possible to deduce the mental and physical health of the population.

Methods: To conduct this research we used the "Google Trends" tool, which returns normalized values, called "relative search volume" (RSV), ranging from 0 to 100 according to the web popularity of a group of queries. By comparing the $RSVs$ in periods before and after the outbreak of the novel coronavirus in Italy, we derived the impact of COVID-19 on the activity of Italian netizens towards novel coronavirus itself, hygiene, prevention, and psychological well-being. Furthermore, we calculated Pearson's correlations ρ between all these queries and COVID-19 cases for each region. We chose a p-value (p) threshold $\alpha = .1$.

Results: After the two initial spikes that occurred on February 23 and March 9, 2020, the general web interest in COVID-19 in Italy has waned, as has the correlation with the official number of cases per region ($p < .1$ only until March 14, 2020). Anyway, web interest was similarly distributed across the regions ($ASV = 92, SD = 6$). We also found that all trends depend significantly on the number of COVID-19 cases at national but not international nor regional level. Between 20 February and 10 June, 2020, web interest relating to hygiene and prevention increased by 116 and 901% respectively compared to those from 1 January to 19 February, 2020 ($95\%\ CIs$: $[115.3, 116.3]$, $[850.3, 952{,}2]$). Significant correlations between regional cumulative web searches and COVID-19 cases were found only between February 26 and March 7, 2020 ($\rho_{best} = .43$, $95\%\ CI$: $[.42, .44], p = .07$). During the COVID-19 pandemic until 10 June, 2020, national web searches of the generic terms "fear" and "anxiety" grew by 8 and 21% respectively ($95\%\ CIs$: $[8.0, 8.2], [20.4, 20.6]$) compared to those of the period 1 January, 2018 – 29 December, 2019. We found cyclically significant correlations between negative emotions related to the novel coronavirus and COVID-19 official data.

Conclusions: The Italian netizens showed a marked interest in the COVID-19 pandemic only when this became a direct national problem. In general, web searches have rarely been correlated with the number of cases per region; we conclude that the danger, once it arrived in the country, was perceived similarly in all regions. We can state that the period of maximum effectiveness of the online information, in relation to this type of situation, is limited to 3-4 days from a specific key event. If such a scenario were to occur again, we suggest that all government agencies focus their web disclosure efforts over that time. Despite this, we have found cyclical correlations with research related to negative feelings such as anxiety, depression, fear, and stress. Therefore, to identify mental and physical health problems among the population, it suffices to observe slight variations in the trend of related web queries.

Keywords: *COVID-19, Google Trends, web interests, Italy, novel coronavirus.*

medRxiv preprint doi: https://doi.org/10.1101/2020.08.09.20171181; this version posted August 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Introduction

Thanks also to introducing tools for the analysis of the queries popularity such as Google Trends, the Internet has become one of the most important platforms for obtaining valuable data on users' health and interests [1, 2]. Over time, more and more authors utilized Google Trends and the techniques of use have refined by creating tutorials articles [3]. Alongside the increasing number of users who use the web every day to obtain information on any topic, there is a growing need to control the circulation of the same to prevent fake from affecting public health and economy. For this reason, new branches of science called *infodemiology* and *infoveillance* were born [4]. Given the COVID-19 emergency, these two disciplines will provide useful data to researchers to evaluate the impact of the pandemic on people's lives and welfare, fake news and infodemic monikers spread, as well as people's attitude towards an unprecedented international issue.

Between the end of February and the beginning of April, Italy was the second most affected country by COVID-19 both in terms of total number of infected and deaths, becoming the epicenter of a shifting pandemic [5]. As evidenced in other studies, an enormous amount of information regarding the novel coronavirus has circulated in the country [6]. Although most of these were moderately infodemic, there was great interest in hygiene and prevention measures. However, no research in the scientific literature available has quantitatively analyzed the impact of COVID-19 on psychological web searches nor thoroughly investigated hygiene-related ones by looking for correlations with the number of cases per region. Since, as shown in other studies, the impact of COVID-19 on the mental and physical health of the Italian population has been devastating, we believe it is of fundamental importance to study the relationship between the trend of health-related queries and the real need for assistance [7]. In fact, this could help scientists estimate the well-being of the population from the trend of web searches.

medRxiv preprint doi: https://doi.org/10.1101/2020.08.09.20171181; this version posted August 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Methods

We used the Google Trends tool to investigate the Italian netizens' web interest in the COVID-19 pandemic. As explained in other studies of this type, Google Trends provides normalized values, called relative search volume (RSV), ranging from 0 to 100 in proportion to the popularity of the queries [3]. We searched for specific keywords that recorded high RSVs, helping us both with the items “related queries”, “related topics”, and with the results of another study on the COVID-19 infodemiology in Italy [6]. We focused on three categories of queries: generic news, hygiene, and emotional response related to stress and anxiety. After collecting the data, we looked for correlations with the official data on COVID-19 provided by the Italian Civil Protection Department regarding the number of infected, hospitalized, dead, healed, and swabs [8]. When the trends showed sufficient regularity, we looked for interpolating functions that represented them.

Statistical analysis

We collected all the data day by day from February 24 to June 10, 2020. In order to have a graphical representation of regional web searches trends, we exploited this stratagem: by calling RSV_i the Google Trends relative search volume of a specific keywords group for the region i and RSV_{tot} the national relative search volume for the same keywords group, we have introduced a variable x and imposed $x \cdot \sum_i^N RSV_i = RSV_{tot}$, obtaining $x = RSV_{tot}/(\sum_i^N RIV_i)$. Then, we calculated a weighted relative search volume $WRSV_i = x \cdot RSV_i$ for every region i. After that, we calculated Pearson’s correlations ρ_{ij} between the daily $RIVs$ and the number of COVID-19 active cases, total cases, new cases, home isolates, recovered, new deceased, and total deceased, for all regions. We repeated the same procedure with the cumulative values in the same time interval. The regions investigated were $N = 19$. We chose $\alpha = .1$ (10%) as p-value (p) significance threshold. Anyway, we also reported the thresholds ρ_α for $\alpha = .05$ (5%) and $\alpha = .01$ (1%): $\rho_{.01} = .58$, $\rho_{.05} = .46$, $\rho_{.1} = .39$. We calculated the ASV (Average Search Volume) values as the average

medRxiv preprint doi: https://doi.org/10.1101/2020.08.09.20171181; this version posted August 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

values of the $RSVs$ in certain time intervals (specified in the results). For each ASV we reported a Gaussian 95% confidence interval via the formula $(ASV - 2 \cdot \sigma/\sqrt{N}, ASV + 2 \cdot \sigma/\sqrt{N})$ with σ standard deviation. We denoted by ρ_{best} the arithmetic mean of specific groups of correlations shown in the results. For the estimate of the confidence interval, we used the following errors' propagation formula: $(\sum_j^n \sigma_j^2 + \sigma^2/n)^{1/2}$, with σ_j (SEM) standard deviation in the Gaussian distribution $G(ASV_j, SEM_j)$ and σ standard deviation. We used the softwares "Igor Pro 6.3.7" and "Microsoft Excel 2019" for interpolations and data processing. We calculated the percentage discrepancies through the formula $\Delta_{\%} = (y - x)/x \cdot 100$. We have calculated all errors on the functions used through the standard errors' propagation formula [9]. For all interpolations and some data, we reported the best values and the relative standard deviations using the abbreviation SD; in such cases, we have kept all possible significant figures provided by the software.

COVID-19-related significant keywords

Generic news keywords: coronavirus. Category: All.

Health keywords: amuchina, disinfectant, ffp2, ffp3, gel, gloves, mask, masks, sanitizing. Category: All.

Negative emotional response keywords: anxiety coronavirus, depression coronavirus, fear coronavirus, fear covid, stress coronavirus. Category: All.

Legend

COD: COVID-19 official data, i.e. data provided by the Ministry of Health and Civil Protection.

Critical threshold: limit of daily COD beyond which there is a sudden increase in COVID-19 related interest.

Deceased: cumulative total number of deaths from COVID-19.

medRxiv preprint doi: https://doi.org/10.1101/2020.08.09.20171181; this version posted August 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Discharged recovered / Discharged and healed: cumulative total number of patients recovered from COVID-19.

Home isolation: people currently in home-isolation by COVID-19 infection.

Hospitalized with symptoms: people currently hospitalized for COVID-19 who exhibit symptoms.

Intensive care: people currently hospitalized for COVID-19 in intensive care.

New cases / new infected: daily increase in currently active cases of COVID-19.

New deceased: daily increase in the number of deaths from COVID-19.

Swabs: cumulative total number of COVID-19 swabs.

Total active cases: total number of people currently infected with COVID-19.

Total cases: cumulative total number of COVID-19 cases.

Total hospitalized: cumulative total number of people hospitalized for COVID-19.

Δ total active cases / Δ total infected: daily increase in total cases of COVID-19.

Results

Generic news web interest

The web interest of the Italian users towards COVID-19 was similarly distributed among all regions ($ASV = 92, SD = 9$) and has dropped over time [Table 1, Figure 1]. Its trend, after the last peak at March 11th until June 10th, 2020, is well represented by the exponential function $y = y_0 + A_1 e^{-(x-x_0)/\tau_1} + A_2 e^{-(x-x_0)/\tau_2}$ with $y_0 = -5.017, SD = 23.5$; $A_1 = 56.801$, $SD = 10.6$; $\tau_1 = 0.012831, SD = 0.0122$; $A_2 = 28.665, SD = 13.6$; $\tau_2 = 0.0995501$, $SD = 0.0473$. The arithmetic mean of all the average daily COD correlations between February 24 and June 10, 2020,

medRxiv preprint doi: https://doi.org/10.1101/2020.08.09.20171181; this version posted August 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

is not statistically significant ($\rho_{best} = .17, 95\%\ CI: [.15,\ .19], p = .48$). We found significant positive daily COD correlations at the end of February with hospitalized, home isolation and recovered patients, as well as with new, active and total cases, and new and total deaths ($\rho_{best} = .45, 95\%\ CI: [.44,\ .46], p = .05$). After that, other significant positive COD correlations occurred only two times [Figure 2]. Significant negative daily COD correlations were found from 8 March onwards; in particular, the one with the number of COVID-19 swab tests reached several significant negative peaks between May and June, 2020 ($\rho_{best} = -.48, 95\%\ CI: [-.51, -.45], p = .04$). On the contrary, the cumulative COVID-19 web searches – COD correlations maintained significant values until around March 14th, although they also had a declining trend [Figure3]. In fact, excluding the items "recovered", "Δ total infected", and "swabs", a function representative this correlation (February 24 – June 10, 2020) is $y = A + B \cdot \left(1 + e^{(C-x)/D}\right)$, with $A = 0.4360,\ SD = 0.0202$; $B = -0.34409,\ SD = 0.022$; $C = 23.438, SD = 1.26$; $D = 5.8573, SD = 1.02$.

Hygiene and prevention web interest

Web search interest related to disinfectants went from $ASV = 5.14$ ($95\%\ CI: [4.90, 5.38]$) in the period January 1 to February 19, 2020, to $ASV = 11.09$ ($95\%\ CI: [8.56, 13.63]$) in the period February 20 to June 10, 2020, i.e. it has undergone 115.8% average increase ($95\%\ CI: [115.3, 116.3]$). Considering the same two periods, web interest related to medical masks and gloves went from $ASV = 3.62$ ($95\%\ CI: [2.88, 4.36]$) to $ASV = 36.25$ ($95\%\ CI: [29.97, 42.52]$), i.e. it has undergone 901.2% average increase ($95\%\ CI: [850.3, 952,2]$). From February 20 to June 10, 2020, hygiene web interest made up about 12.6% of COVID-19 one ($95\%\ CI: [11.0, 14.2]$) but, unlike the latter, its trend remained almost constant until around May 20th [Figure 4]. Considering the daily web searches, we found significant COD correlation values only in a very few isolated cases [Figure 5]. Finally, regarding the cumulated web searches, we have highlighted significant positive COD correlations from February 26 to March 7, 2020, with all fields except "home isolation" and "swabs" ($\rho_{best} = .43,\ 95\%\ CI: [.42,\ .44], p = .07$) [Figure 6]. The differences in

medRxiv preprint doi: https://doi.org/10.1101/2020.08.09.20171181; this version posted August 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

the distribution of interest were more important than those on generic news ($ASV = 86, SD = 9$; $ASV = 85, SD = 9$) [Table 1].

Emotive-psychological well-being web interest

On the web emotional response to COVID-19, we got the following results: nationally, comparing the averages weekly RSVs of the two periods 1 January 2018 - 31 December 2019 and 1 January 2020 - 10 June 2020, we observed a decrease of 3.1% and 16.1% on generic web searches related to stress and depression respectively ($95\%\ CI: [-3.0, -2.9], 95\%\ CI: [-16.2, -16.0]$) and an increase of 8.1% ($95\%\ CI: [8.0, 8.2]$) and 20.5%($95\%\ CI: [20.4, 20.6]$) of those inherent in fear and anxiety (we point out that the trend related to anxiety had already been growing since December 2019) [Figure 7]. In the second period, the latter had a little increasing daily trend expressed by the equation $y = mx$, with $m = 0.04, 95\%\ CI: [0.00, 0.08]$; this shows that general anxiety has not diminished over time, as opposed to the number of new, active and serious cases, and new deceased. From February 20 to June 10, 2020, the web interest in negative emotions made up 4.1% of COVID-19 one ($95\%\ CI: [3.6, 4.5]$). In the same period, the web interest in negative emotions explicitly related to novel coronavirus made up about 1.4% of that related to COVID-19 ($95\%\ CI: [1.2, 1.5]$). Regarding the latter, it was only possible to look for COD correlations with the cumulative web interest since the data on the daily web interest was too uncertain. We found cyclically positive significant COD correlations throughout the investigated period [Figure 8]. To support this, web interest in negative emotions has had extremely different distributions across regions ($ASV = 55, SD = 41$).

Discussion

The quantitative investigation of web users' interests in a crisis moment is a priority aspect to test the effectiveness of online information and the perception of risk by the population. This is the first study to carry out this analysis in Italy, one country most affected by the COVID-19 pandemic in

medRxiv preprint doi: https://doi.org/10.1101/2020.08.09.20171181; this version posted August 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

the world. Our results show that interest in the novel coronavirus has exploded only in the presence of two contemporary phenomena: i) the presence of the virus in the nation; ii) a significant number of cases. Even the state of emergency declared on January 31, 2020, due to the case of the two Chinese tourists infected with COVID-19, has caused a very moderate and limited time web interest [10]. We also highlighted a disparity between online searches and the actual perception of the danger: in fact, as shown in other studies and reported by national and international newspapers, the first episodes of racism against the Chinese ethnic group and the so-called "coronavirus psychosis" [6, 11, 12]. Therefore, we believe it is plausible that the analysis of web interest in Italy shows genuine interest only if the issue exceeds a certain severity threshold. We believe that the climate of uncertainty created by the press and by the many conflicting opinions of scientists influenced negatively the perception of the risk linked to COVID-19 in the Italian population; in particular, some have compared it to an influence while others rightly denoted its dangers and critical issues also related to the lack of reliable information [13, 14]. When the problem became national instead, we recorded two peaks in the web queries (February 27, 2020 and around March 14, 2020). Significant correlations between web interest and regions with multiple cases rarely occurred only in the initial phase, i.e. web interest in the novel coronavirus has been linked only to its presence on national soil and nothing more. Since the aforementioned peaks had a maximum duration of about 7 days, we believe that online information is effective only in that time frame. The regional web interest in this topic were very similarly distributed. This leads us to plan two considerations: i) the online information on a problem in its initial phase can heavily influence the thinking of the Italian population; ergo, it is important that those who take on making information online are clear right away; ii) the online information on a problem in its initial phase can have a deep impact on its evolution.

Positive data emerge from the web interest in hygienic precautions, such as disinfectants, masks. In fact, the respective increase of about 115% and 901% compared to the periods preceding the virus signals a drastic change in the netizens' habits to face the pandemic. We must weigh such a result

medRxiv preprint doi: https://doi.org/10.1101/2020.08.09.20171181; this version posted August 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

on the fact that the mean value of these queries corresponds to approximately 13% of all the novel coronavirus-related queries. The lack of long-term correlations with the number of cases and the low variance of the data suggest that the interest was similarly distributed among all regions; this helped to avoid the spread of the epidemic nationwide [Table 1]. Even in this case we have seen two peaks in queries and then a waning interest. This does not mean that the actual interest in hygiene has decreased, since disinfectants, masks, and gloves, have experienced a great increase in sales [15].

General interest in negative feelings such as stress and depression fell during the lockdown (around 3% and 16% respectively). The RSVs of anxiety and fear increased by approximately 21% and 8%, respectively. Therefore, it is plausible that there has been a shift of interest towards the latter. Direct associations between these symptoms of distress and the novel coronavirus made up about 1% of total COVID-19 queries. The data had an incredibly significant variance and high linear correlation values i.e. the interest was much more pronounced in some regions, such as Emilia Romagna, Piedmont, Campania, Lombardy, and Puglia [Table 1]. The inconstant trend of the correlation between cumulative web searches and COVID-19 official data shown in figure 8 suggests that the number of total searches was low, as it was easily influenced day by day. The effects of the novel coronavirus and its lockdown on the mental health of the Italian population have been serious as shown in other studies [7]. Therefore, even a small amount of these types of queries can mean a lot for mental disorders.

Limitations

Web searches provide quantitative values only for users who use the internet to obtain information on certain topics. All the queries investigated were collected from the Google search engine.

medRxiv preprint doi: https://doi.org/10.1101/2020.08.09.20171181; this version posted August 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Conclusions

The Italian netizens showed a marked interest in the COVID-19 pandemic only when this became a direct national problem. In general, web searches have rarely been correlated with the number of cases per region; we conclude that the danger, once it arrived in the country, was perceived similarly in all regions. We can state that the period of maximum effectiveness of the online information, in relation to this type of situation, is limited to 3-4 days from a specific key event. If such a scenario were to occur again, we suggest that all government agencies focus their web disclosure efforts over that time. Despite this, we have found cyclical correlations with research related to negative feelings such as anxiety, depression, fear, and stress. Therefore, to identify mental and physical health problems among the population, it suffices to observe slight variations in the trend of related web queries.

Acknowledgment: None

Conflict of Interest: None

Source of funding: None

Data availability: All the data related to this study are presented in the manuscript.

medRxiv preprint doi: https://doi.org/10.1101/2020.08.09.20171181; this version posted August 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

References

[1] Nuti VS, Wayda B, Ranasinghe I, et al, The Use of Google Trends in Health Care Research: A Systematic Review, PLoS One. 2014; 9(10): e109583, PMID: 25337815

[2] Seifter A, Schwarzwalder A, Geis K, et al, The utility of "Google Trends" for epidemiological research: Lyme disease as an example, Geospat Health, 2010 May;4(2):135-7, PMID: 20503183

[3] Mavragani A, Ochoa G, Tsagarakis KP, Assessing the Methods, Tools, and Statistical Approaches in Google Trends Research: Systematic Review, J Med Internet Res 2018;20(11):e270, PMID: 30401664

[4] Eysenbach G, Infodemiology and Infoveillance: Framework for an Emerging Set of Public Health Informatics Methods to Analyze Search, Communication and Publication Behavior on the Internet, J Med Internet Res 2009;11(1):e11, PMID: 19329408

[5] Italy, Pandemic's New Epicenter, Has Lessons for the World, The New York Times. Available at: https://www.nytimes.com/2020/03/21/world/europe/italy-coronavirus-center-lessons.html (Accessed: June 19, 2020)

[6] Rovetta A, Bhagavathula AS, COVID-19-Related Web Search Behaviors and Infodemic Attitudes in Italy: Infodemiological Study, JMIR Public Health Surveill 2020;6(2):e19374, PMID: 32338613

[7] Gualano MR, Lo Moro G, Voglino G, et al, Effects of Covid-19 Lockdown on Mental Health and Sleep Disturbances in Italy, Int. J. Environ. Res. Public Health 2020, 17(13), 4779, PMID: 32630821

[8] Italian Civil Protection Department, COVID-19 updates (Italian website). Available at: http://opendatadpc.maps.arcgis.com/apps/opsdashboard/index.html#/b0c68bce2cce478eaac82fe38d4138b1 (Accessed: June 11, 2020)

[9] J.R. Taylor, An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements, 2nd Edition, ISBN-13: 978-0935702750

[10] Reuters, UPDATE 2-Italy declares coronavirus emergency after first two cases confirmed. Available at: https://www.reuters.com/article/china-health-italy/update-1-italy-cabinet-to-discuss-declaring-coronavirus-emergency-after-first-two-cases-confirmed-idUSL8N2A02S1 (Accessed: July 08, 2020)

[11] The Guardian, Outbreaks of xenophobia in west as coronavirus spreads. Available at: https://www.theguardian.com/world/2020/jan/31/spate-of-anti-chinese-incidents-in-italy-amid-coronavirus-panic (Accessed: July 08, 2020)

[12] Huffpost, La psicosi da Coronavirus in Italia è ingiustificata: il rischio contagio è improbabile (Italian website). Available at: https://www.huffingtonpost.it/entry/la-psicosi-da-coronavirus-in-italia-e-ingiustificata-il-rischio-contagi-e-improbabile_it_5e340df5c5b6f2623327ca29 (Accessed: July 08, 2020)

medRxiv preprint doi: https://doi.org/10.1101/2020.08.09.20171181; this version posted August 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

[13] NiemanReports, Lessons from the Italian Media's Coverage of the Coronavirus. Available at: https://niemanreports.org/articles/lessons-from-the-italian-medias-coverage-of-the-coronavirus/ (Accessed: July 11, 2020).

[14] Medical Facts, Coronavirus: Roberto Burioni parla delle notizie, dei pericoli e dei possibili sviluppi (Italian website). Available at: https://www.medicalfacts.it/2020/01/24/video-coronavirus-roberto-burioni/ (Accessed: July 11, 2020)

[15] Il Sole 24 Ore, Fase 2, dai reagenti per i tamponi alle mascherine perché tutto è esaurito. (Italian news). Available at: https://www.ilsole24ore.com/art/fase-2-mascherine-reagenti-i-tamponi-perche-tutto-e-esaurito-AD5l9rP

medRxiv preprint doi: https://doi.org/10.1101/2020.08.09.20171181; this version posted August 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Figures

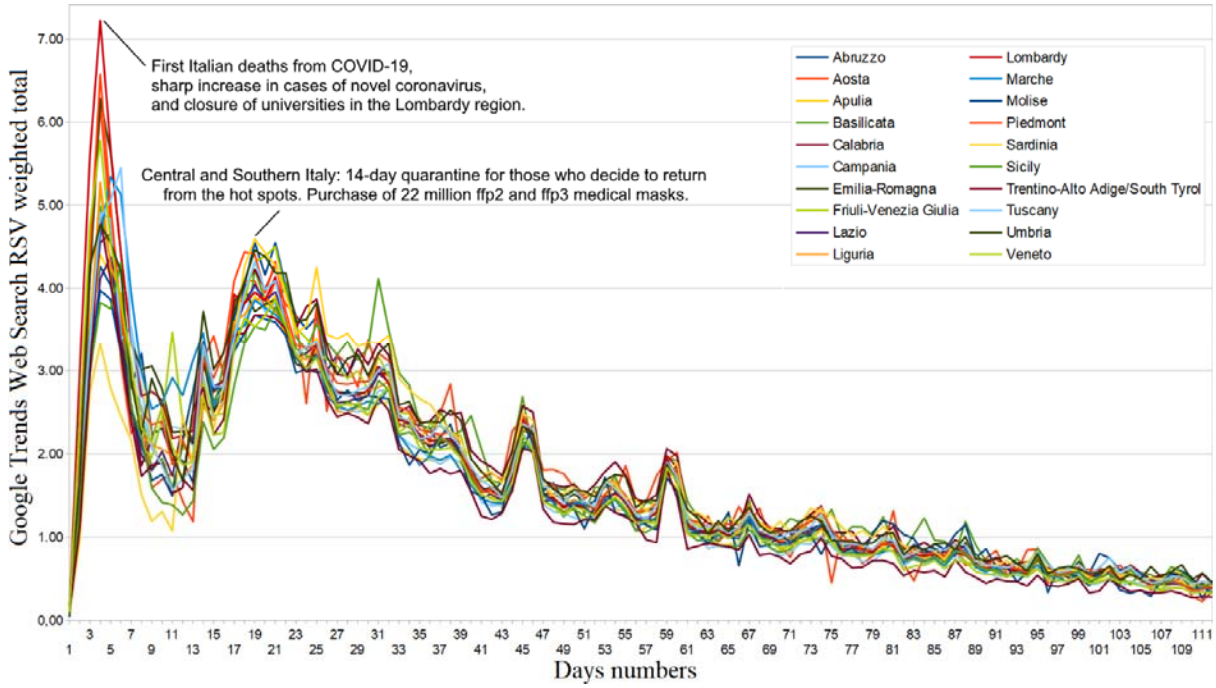

Figure 1. COVID-19 regional web interest from February 20 to June 10, 2020.

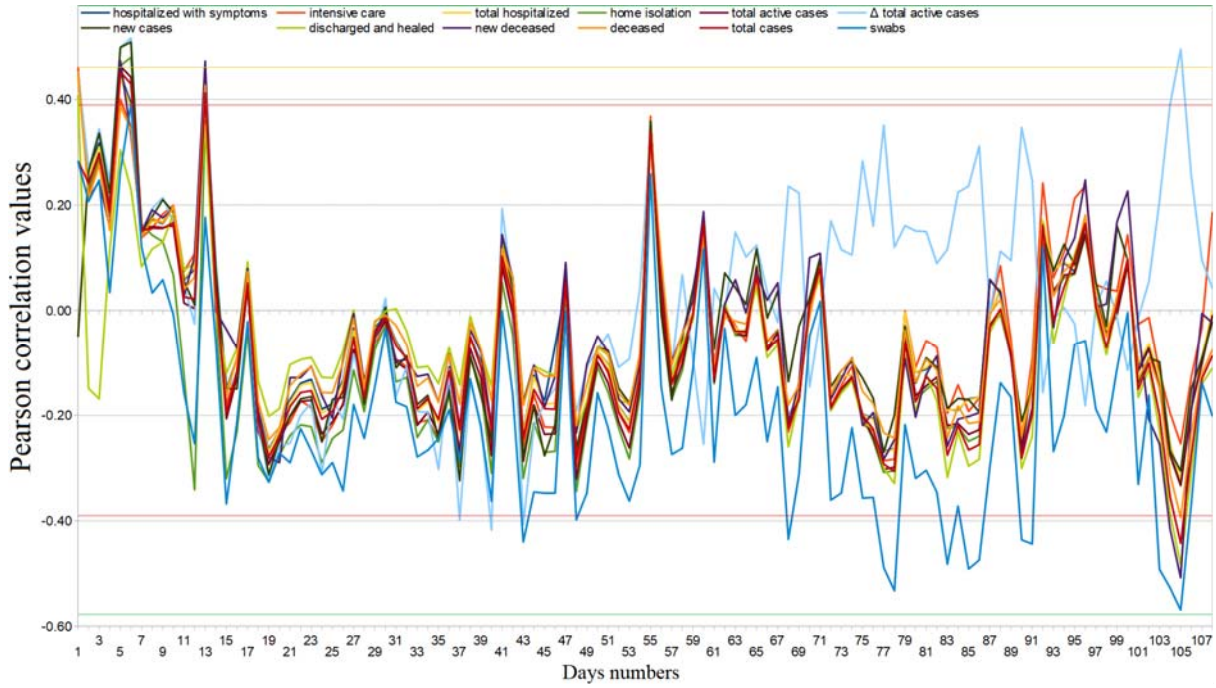

Figure 2. COVID-19 regional daily web interest – official data Pearson's correlation trends from February 24 to June 10, 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.08.09.20171181; this version posted August 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

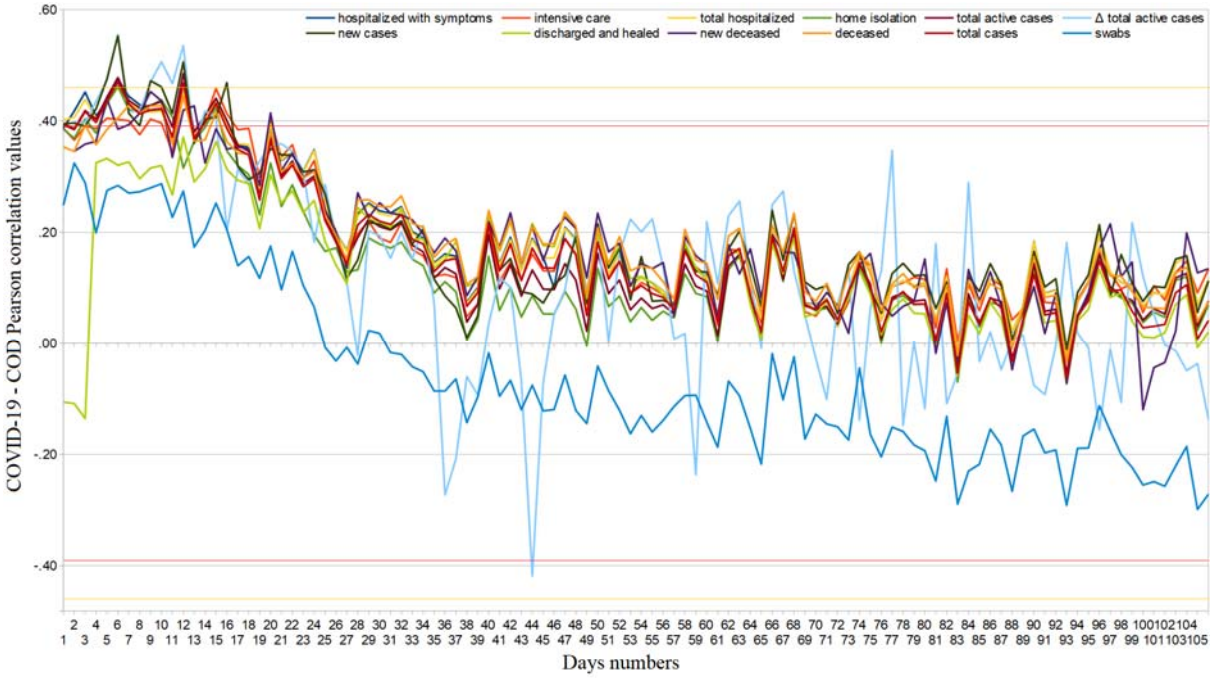

Figure 3. COVID-19 regional cumulative web interest – official data Pearson's correlation trends from February 24 to June 10, 2020.

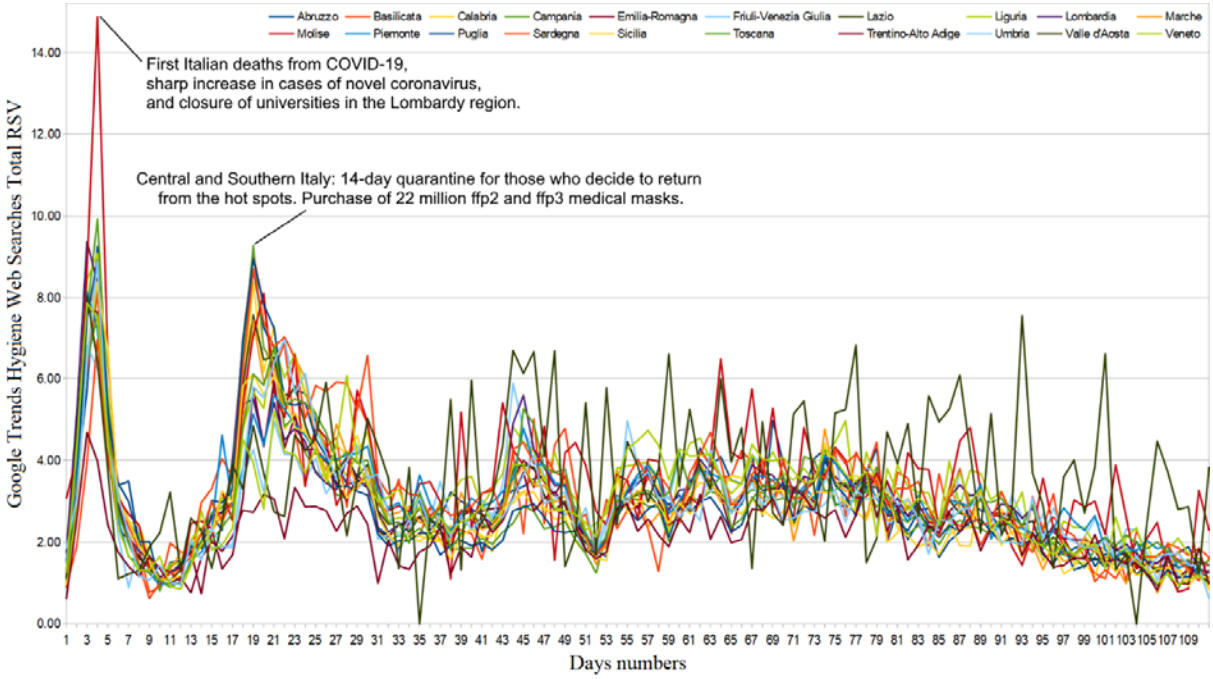

Figure 4. Hygiene and prevention regional web interest from February 21 to June 10, 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.08.09.20171181; this version posted August 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

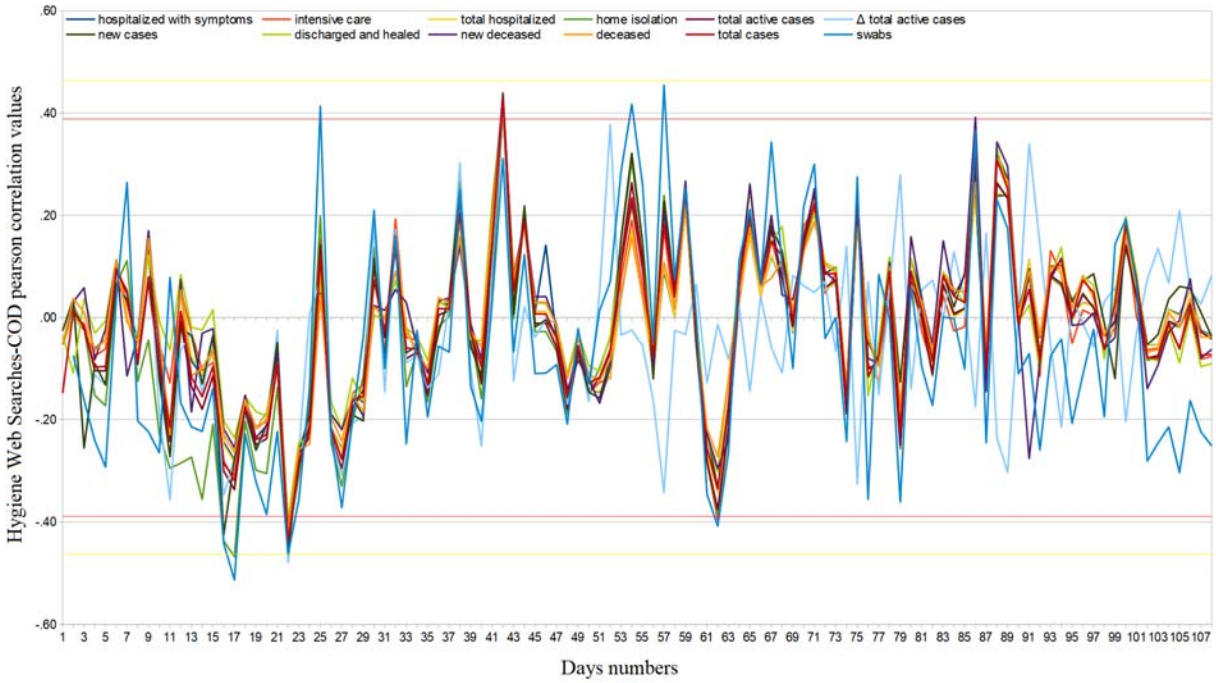

Figure 5. Hygiene daily regional web interest – official data Pearson's correlation trends from February 24 to June 10, 2020.

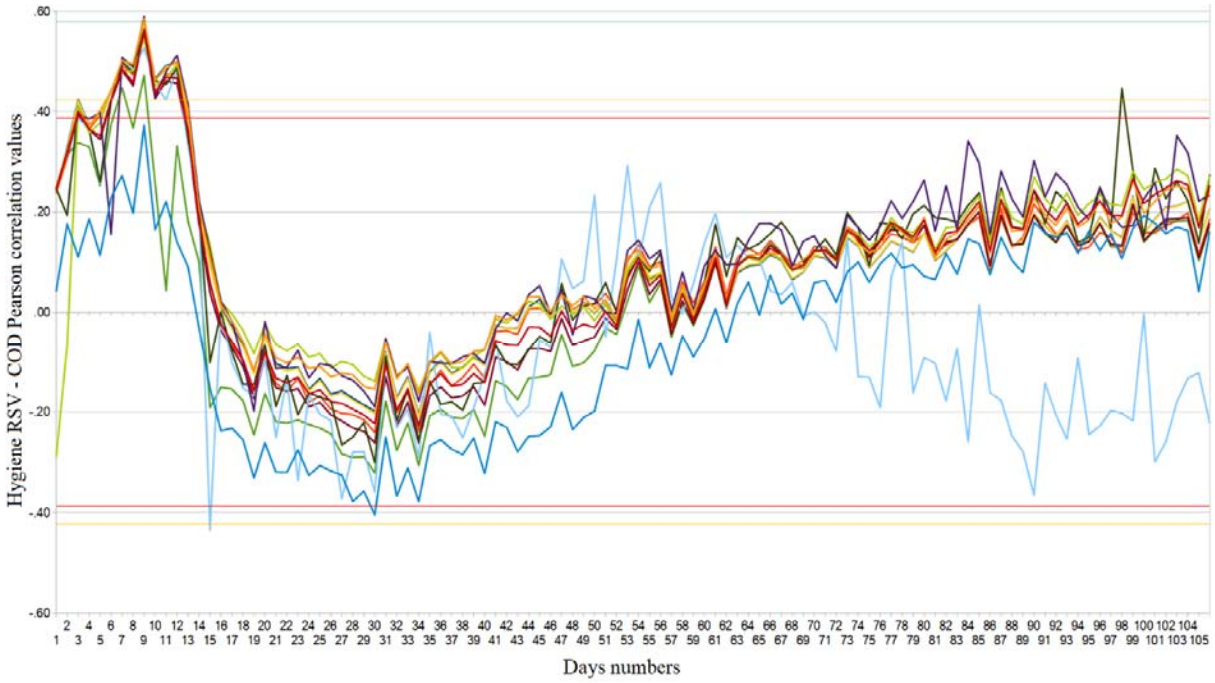

Figure 6. Hygiene regional cumulative web interest – COVID-19 official data Pearson's correlation trends from February 24 to June 10, 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.08.09.20171181; this version posted August 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

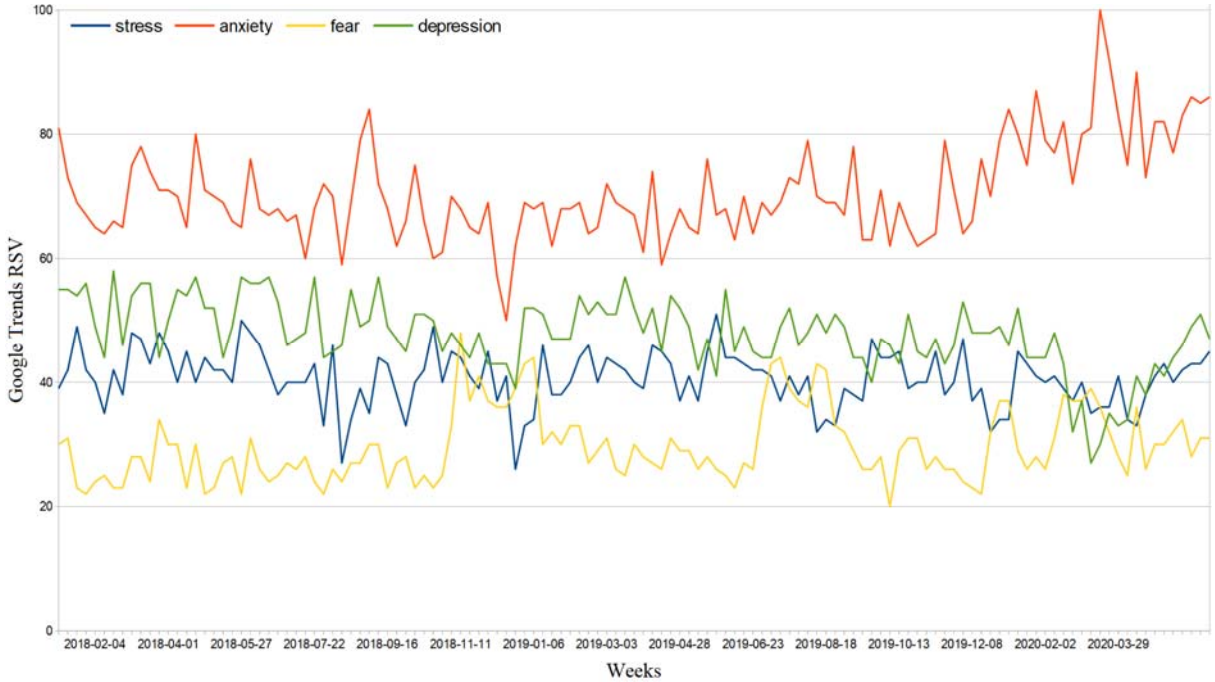

Figure 7. Italy web interest in negative emotions from January 1, 2018, to June 10, 2020.

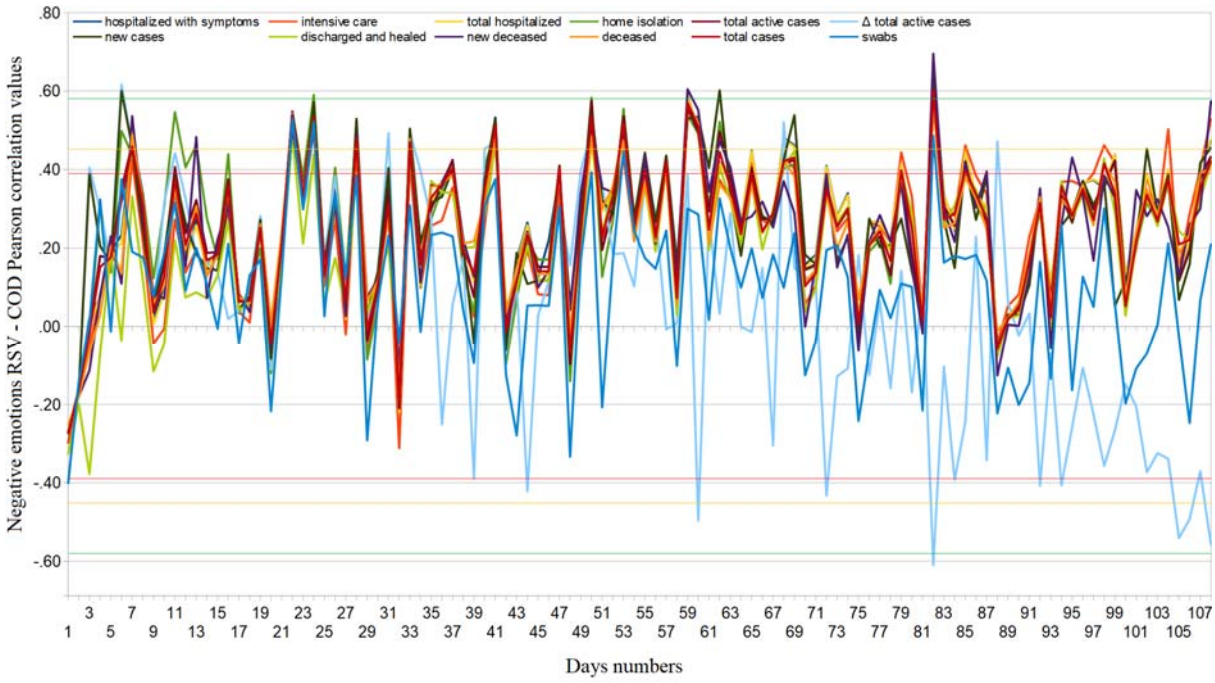

Figure 8. Regional cumulative web interest in negative emotions related to novel coronavirus – COVID-19 official data Pearson's correlation trends from February 24 to June 10, 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.08.09.20171181; this version posted August 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Tables

Region	Generic News RSV	Hygiene RSV	Prevention RSV	Negative Emotions RSV
Abruzzo	95	91	90	88
Aosta	99	68	83	0
Apulia	87	96	76	92
Basilicata	98	100	85	0
Calabria	97	93	82	0
Campania	83	93	75	94
Emilia-Romagna	96	81	85	100
Friuli-Venezia Giulia	95	81	85	0
Lazio	90	86	84	86
Liguria	92	94	98	67
Lombardy	96	81	92	93
Marche	91	85	86	79
Molise	87	97	90	0
Piedmont	98	79	93	96
Sardinia	97	96	100	88
Sicily	82	89	72	81
Trentino-Alto Adige*	80	68	59	0
Tuscany	95	83	90	70
Umbria	100	84	84	0
Veneto	86	78	88	75
ASV	92	86	85	55
SD	6	9	9	41

Table 1. COVID-19-related generic news, hygiene, prevention, and negative emotions RSVs.

RSV = relative search volume, ASV = average search volume, SD = standard deviation. *South Tyrol is included.