CE-RESEARCH LETTER TO THE EDITOR

Google trends and COVID-19 in Italy: could we brace for impact?

Jacopo Ciaffi[1] · **Riccardo Meliconi**[1,2] · **Maria Paola Landini**[3] · **Francesco Ursini**[1,2]

Dear Editor,

The first two COVID-19 cases were confirmed in Italy in January 2020 and, by the end of February, the virus had spread over northern regions of the country. During March, the toll of new cases and deaths kept surging, rapidly overtaking the numbers registered in China and making Italy the hardest-hit country in Europe and one of the epicentres of the global pandemic.

In Lombardy, the heart of the outbreak, COVID-19 was not yet labelled a pandemic, but the healthcare system had already started groaning under the strain of the invisible enemy. In the struggle to cope with the expanding threat, war-like measures were adopted by the Government, but Italy was insufficiently prepared to face this unprecedented challenge and, on the brink of collapse, weaknesses were unveiled.

As Italian citizens, and doctors, we wondered whether there was anything to be done differently, anything we could do to be more prepared to tackle the virus and allow our healthcare system to respond adequately to the crisis. We, therefore, hypothesized that digital epidemiology could provide valuable insights into the spread of the viral infection and hints useful to predict the COVID-19 outbreak in Italy.

Google Trends is a public, open-access tool (https://trends.google.com/trends/), used to monitor patterns and volumes of queries in a selected geographical area over a specified period. Absolute search volumes are not provided, but a value of 100 is assigned to the peak interest of the time range and relative search popularity in the other time points is proportionally conveyed in a data series over a 0–100 normalized scale.

Since temporal correlation between Google Trends and outbreaks of diseases different from COVID-19 has been suggested [1, 2], we imagined that a spike in searches of early symptoms characterizing the prodromal stage of the disease, as cough and fever, could precede admissions to intensive care units (ICUs) and deaths from COVID-19. Case data started to be recorded on February 24 and are provided daily by the Civil Protection Department, accessible on the Italian Government's website https://www.salute.gov.it.

We analysed official data on increase in ICU admissions and new deaths collected until April 6. A Google Trends' search was performed in Italian language using the key terms "tosse" and "febbre", respectively, meaning "cough" and "fever" in English translation, over the period between February 24 and April 6, 2020. We thus plotted the size of daily increase in ICU admissions against the synchronous proportion of interest toward "tosse" and "febbre" (Fig. 1a, b). Then, the same approach was applied to the data on new deaths (Fig. 1c, d).

Visual inspection suggested a temporal correlation, with a lag period between trends in search volumes and ICU admissions or deaths, respectively, of about 1 and 2 weeks. Hence, we repeated the Google Trends' search anticipating the search period by 1 and 2 weeks, then plotting the increase in ICU admissions against Google Trends for "tosse" and "febbre" 1 week earlier (Fig. 1e, f), and the number of new deaths against the search volumes for "tosse" and "febbre" 2 weeks earlier (Fig. 1g, h).

We subsequently analysed the findings through linear regression (Fig. 2a–d). The strength of the relationship was expressed by a coefficient of determination (R^2) of 0.580 and 0.576, respectively, between the increase in ICU admissions and Google Trends for "tosse" or "febbre" 1 week earlier. Furthermore, the R^2 was, respectively, 0.507 and 0.624 between new deaths and searches for "tosse" or

✉ Jacopo Ciaffi
  jacopo.ciaffi91@gmail.com

1 Medicine and Rheumatology Unit, IRCCS Istituto Ortopedico Rizzoli (IOR), via Pupilli 1, 40136 Bologna, Italy

2 Section of Rheumatology, Department of Biomedical and Neuromotor Sciences (DIBINEM), University of Bologna, Bologna, Italy

3 Scientific Direction, IRCCS Istituto Ortopedico Rizzoli (IOR), Bologna, Italy
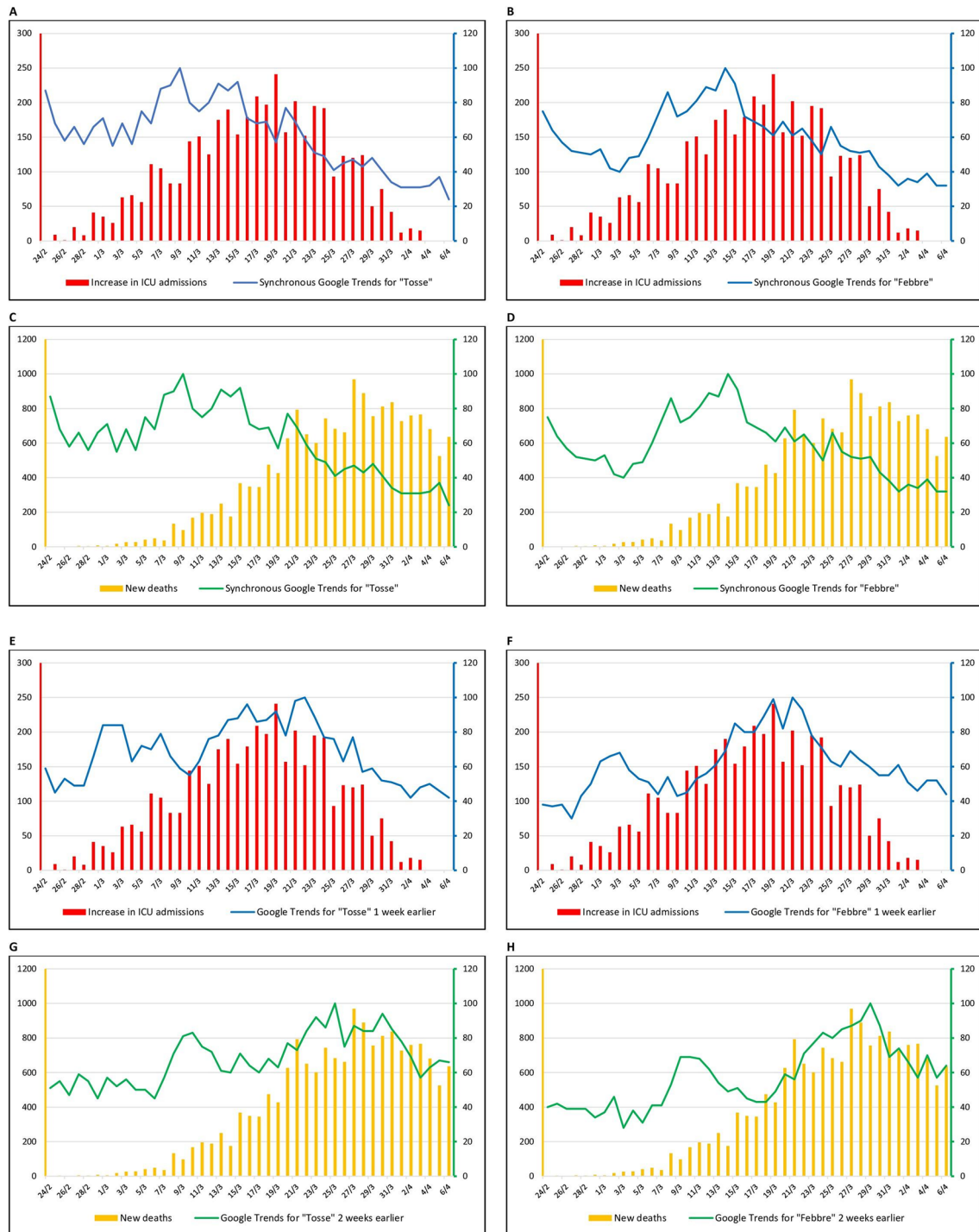
Fig. 1 Increase in ICU admissions and new deaths from COVID-19 between February 24 and April 6, 2020, plotted against Google Trends' search volumes of the Italian words "Tosse" and "Febbre", corresponding to "cough" and "fever" in English translation. **a** Google Trends for "Tosse" synchronous to the increase in ICU admissions. **b** Google Trends for "Febbre" synchronous to the increase in ICU admissions. **c** Google Trends for "Tosse" synchronous to new deaths. **d** Google Trends for "Febbre" synchronous to new deaths. **e** Google Trends for "Tosse" 1 week earlier than ICU admissions. **f** Google Trends for "Febbre" 1 week earlier than ICU admissions. **g** Google Trends for "Tosse" 2 weeks earlier than new deaths. **h** Google Trends for "Febbre" 2 weeks earlier than new deaths.

"febbre" 2 weeks earlier. All *P* values were statistically significant < 0.001.

In few days, COVID-19 stretched the Italian healthcare system beyond the limits, overloading hospitals in the northern part of the country with a quickly escalating number of critically ill patients needing ICU beds and respiratory support. As the virus piled pressure on the system, the Government responded investing resources and rushing to increase the capacity of ICUs, but in the early stages SARS-CoV-2 was spreading faster than our ability to fight it.

Could we brace for impact? Could we predict the outbreak? Observing Google Trends, we noted a correlation between the dynamics of key searches for symptoms referable to SARS-CoV-2 infection, the increase in ICU admissions and new deaths. Indeed, Google searches for cough and fever preceded by 1 week, the need for critical care in Italy and by 2 weeks deaths from COVID-19. Recent research attempted to explore the relationship between Google Trends and COVID-19 and to evaluate the possibility of predicting new cases from internet searches [3, 4]. However, at least in Italy, the search of keywords as "coronavirus" or

"pneumonia" during the COVID-19 crisis was broadly unspecific and mostly attributable to informative purposes, showing a linear increase temporally corresponding to the interest toward the pandemic from mass media and public opinion. Therefore, we postulated that simple terms referring to key early COVID-19 symptoms—e.g. "fever" and "cough"—or even more specific such as "loss of smell" or "loss of taste" could be more explanatory of the viral outbreak's evolution.

To strengthen our results, we assessed whether a correlation between search volumes for commonly occurring symptoms not directly related to SARS-CoV-2 infection [5] and ICU admissions existed or not. For this analysis, we selected the same period of ICU admissions from February 24 to April 6 and we searched Google trends 1 week before using the Italian words "mal di schiena", "tachicardia", "prurito", "diarrea", corresponding to "back pain", "tachycardia", "itching" and "diarrhea" in English translation. In linear regression analysis, we found an extremely weak positive correlation between searches for "diarrea" (R^2 0.122, $p = 0.023$) and increase in ICU admissions, no correlation
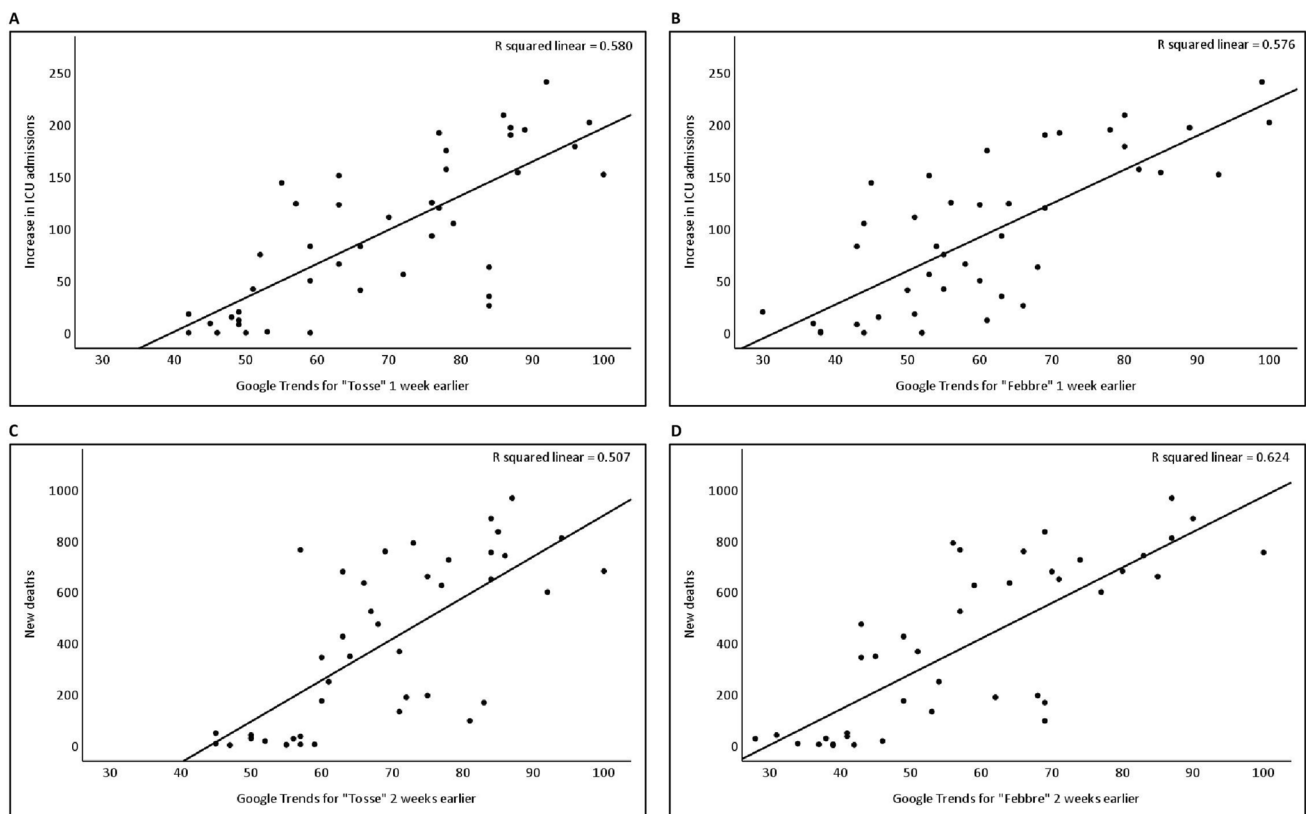
Fig. 2 Linear regression analysis showing association between Google Trends search volumes of the Italian words "Tosse" and "Febbre", corresponding to "cough" and "fever" in English translation, and increase in ICU admissions and new deaths from COVID-19 between February 24 and April 6, 2020. **a** Google Trends for "Tosse" 1 week earlier and increase in ICU admissions. **b** Google Trends for "Febbre" 1 week earlier and increase in ICU admissions. **c** Google Trends for "Tosse" 2 weeks earlier and new deaths. **d** Google Trends for "Febbre" 2 weeks earlier and new deaths.

for "mal di schiena" (R^2 0.012, $p = 0.496$) and negative correlations for "prurito" (R^2 0.119, $p = 0.025$) and tachycardia (R^2 0.273, $p < 0.001$). No temporal trends could, therefore, be identified between search volumes of terms not closely referable to COVID-19 and increase in ICU admissions (Fig. 3a–d), confirming that symptoms suggestive of viral respiratory infection, such as "fever" and "cough", could more reliably predict the increase in ICU beds demand.

However, being aware that "fever" and "cough" cannot be considered specific of COVID-19 infection and that a surge in interest is expected during seasonal flu outbreaks, we compared Google Trends volumes for "fever" and "cough" between winter 2018–2019 and winter 2019–2020. In particular, official Italian data (https://www.epicentro.iss.it/influenza/bilancio-2018_2019) reported that, in 2018–2019, seasonal flu peaked between late January and early February 2019. We thus plotted Google Trends of key terms "fever" and "cough" from December 2018 to April 2019 against search volumes during the same months of the 2019–2020 season, outlining how the interest toward "fever" and "cough" was considerably higher during the COVID-19 pandemic compared to the peak of previous year flu outbreak

(Fig. 4a, b). Interestingly, a surge in Google Trends for "fever" and "cough" could also be observed around the end of January 2020, corresponding to the increase of seasonal flu. However, Google Trends volumes for "fever" and "cough" in 2020 seasonal flu period were, respectively, less than 50% and 65% of those referable to the COVID-19 pandemic (Fig. 4a, b). This observation suggests how COVID-19 might have a substantial impact on the occurrence of respiratory symptoms in the general population, thus corroborating the theory of a role for Google Trends in monitoring the dynamics of future COVID-19 evolution.

Our idea is indeed that the interpretation of Google Trends data could have been useful to warn the healthcare system concerning what we were about to experience 1 week ahead. This time we were close to breaking point, but second waves of COVID-19 are expected [6] and monitoring Google Trends might help us to anticipate them. One week might be enough to strengthen and re-organize the healthcare system and activate critical care facilities. Next time, 1 week in advance could be crucial to react faster and be more aggressive in the effort to prevent and contain the pandemic, hoping treatments will be available to fight COVID-19 [7, 8].
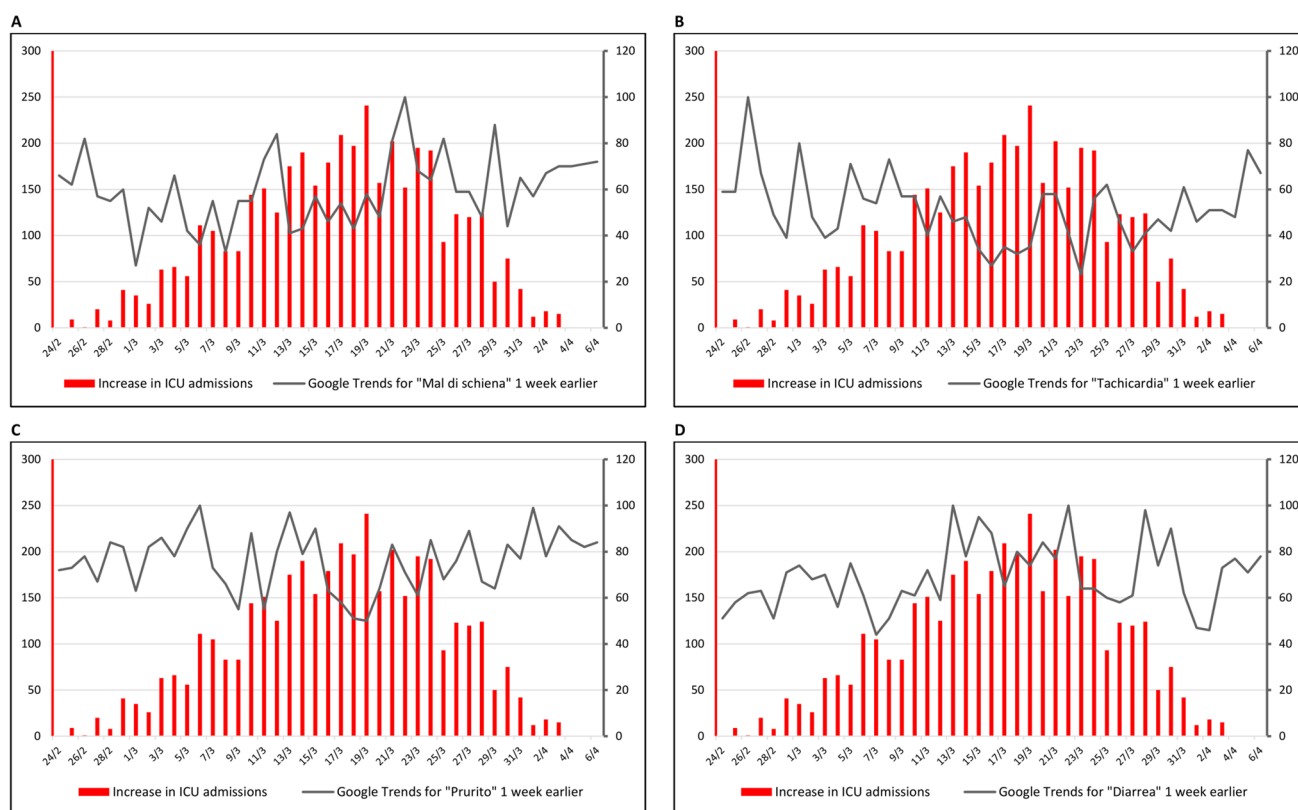

Fig. 3 Increase in ICU admissions COVID-19 between February 24 and April 6, 2020, plotted against Google Trends' search volumes of the Italian words "Mal di schiena", "Tachicardia", "Prurito" and "Diarrea", corresponding to "back pain", "tachicardia", "itching" "diarrhea" in English translation. **a** Google Trends for "Mal di schiena" 1 week earlier than ICU admissions. **b** Google Trends for "Tachicardia" 1 week earlier than ICU admissions. **c** Google Trends for "Prurito" 1 week earlier than ICU admissions. Google Trends for "Diarrea" 1 week earlier than ICU admissions.
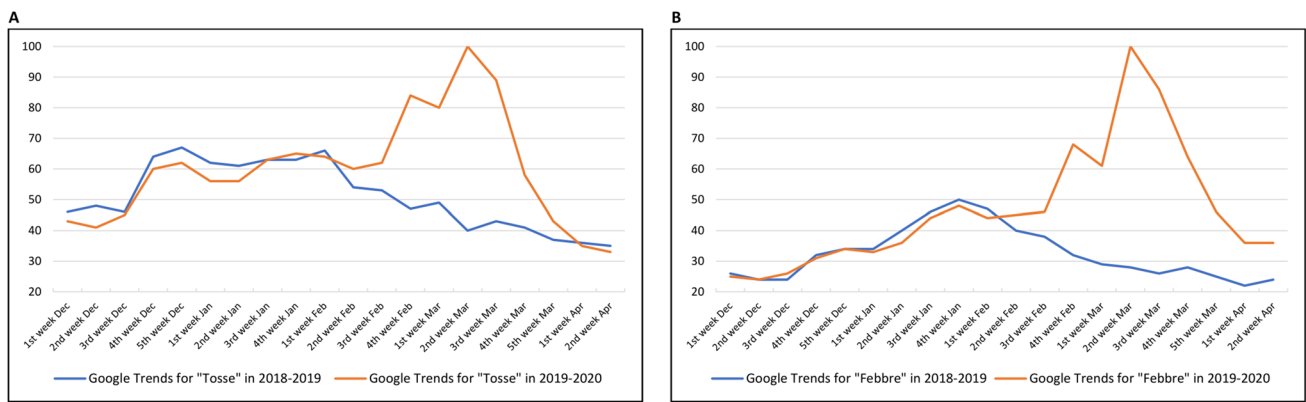
Fig. 4 Google Trends for "Tosse" (**a**) and "Febbre" (**b**) from December 2018 to April 2019 compared to same searches between December 2019 and April 2020

Intrigued by the potential of Google Trends data, we used intuitive statistics to explore their possible value in the management of COVID-19 outbreak. Big data analytics is beyond our competence, but integrating information of top online health search requests from different countries, languages and time periods, might in our opinion shed light on the dynamics of COVID-19 global spread.

In conclusion, although appropriate use and reliability of the tool have not been thoroughly defined and data should be calibrated for future use, tracking public health information from online search engines, which is called "infodemiology" and "infoveillance" [9], might have a role in the prediction of future COVID-19 waves, complementarily to traditional public health surveillance systems.

Compliance with ethical standards

Conflict of interest The authors declare they have no conflict of interest.

Statement of human and animal rights No human participant or animal was involved in the present research.

Informed consent The authors used open-access data and information available on official websites of the Italian government. Only aggregate, anonymous data were used for the purpose of the present research. Therefore, informed consent was not required.

References

1. Verma M, Kishore K, Kumar M, Sondh AR, Aggarwal G, Kathirvel S (2018) Google search trends predicting disease outbreaks: an analysis from India. Healthc Inform Res 24(4):300–308. https://doi.org/10.4258/hir.2018.24.4.300**PMID: 30443418**

2. Husnayain A, Fuad A, Lazuardi L (2019) Correlation between Google Trends on dengue fever and national surveillance report in Indonesia. Glob Health Action 12(1):1552652. https://doi.org/10.1080/16549716.2018.1552652**PMID: 31154985**

3. Li C, Chen LJ, Chen X, Zhang M, Pang CP, Chen H (2020) Retrospective analysis of the possibility of predicting the COVID-19 outbreak from Internet searches and social media data, China, 2020. Euro Surveill 25(10):2000199. https://doi.org/10.2807/1560-7917.ES.2020.25.10.2000199

4. Ayyoubzadeh SM, Ayyoubzadeh SM, Zahedi H, Ahmadi M, Niakan R, Kalhori S (2019) Predicting COVID-19 incidence through analysis of google trends data in iran: data mining and deep learning pilot study. JMIR Public Health Surveill. 6(2):e18828. https://doi.org/10.2196/18828

5. Zhai P, Ding Y, Wu X, Long J, Zhong Y, Li Y (2020) The epidemiology, diagnosis and treatment of COVID-19 [published online ahead of print, 2020 Mar 28]. Int J Antimicrob Agents. https://doi.org/10.1016/j.ijantimicag.2020.105955**(PMID 32234468)**

6. Leung K, Wu JT, Liu D, Leung GM (2020) First-wave COVID-19 transmissibility and severity in China outside Hubei after control measures, and second-wave scenario planning: a modelling impact assessment published online ahead of print, 2020 Apr 8]. Lancet. https://doi.org/10.1016/S0140-6736(20)30746-7**(PMID 32277878)**

7. Sanders JM, Monogue ML, Jodlowski TZ, Cutrell JB (2019) Pharmacologic treatments for coronavirus disease 2019 (COVID-19): a review [published online ahead of print, 2020 Apr 13]. JAMA. https://doi.org/10.1001/jama.2020.6019

8. Testa S, Paoletti O, Giorgi-Pierfranceschi M, Pan A (2019) Switch from oral anticoagulants to parenteral heparin in SARS-CoV-2 hospitalized patients [published online ahead of print, 2020 Apr 15]. Intern Emerg Med. https://doi.org/10.1007/s11739-020-02331-1

9. Eysenbach G (2011) Infodemiology and infoveillance tracking online health information and cyberbehavior for public health. Am J Prev Med 40(5 Suppl 2):S154–S158. https://doi.org/10.1016/j.amepre.2011.02.006

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.