Availability of open data related to COVID-19 epidemic in Italy

S. Di Federico[1], T. Filippini[1,2], C. Marchesi[3], M. Vinceti[1,4]

Abstract

Background. *Since the beginning of the COVID-19 outbreak in Italy, health authorities have released epidemiologic data about this disease. These data were the most important sources of information which were periodically updated and analyzed by researchers to predict the spread of the epidemic. However, comprehensive and timely data on the evolution of COVID-19 have not always been made available to researchers and physicians.*
Method. *The aim of our work is to investigate quality, availability and format of epidemiologic data about COVID-19 in Italy in different territorial and temporal areas. We tried to access the online resources made available by each of the 19 Italian Regions and the two autonomous Provinces, and in more detail by the Local Health Authorities of one of them, the Emilia-Romagna Region.*
We analyzed the main sources and flows of data (namely new and cumulative cases of infection, total swabs, new and cumulative COVID-19 deaths, overall and divided by sex), describing their characteristics such as accessibility, format and completeness. We eventually reviewed the data published by the Italian Ministry of Health, the National Institute of Health (ISS) and the Civil Protection Department. The Tim Berners-Lee scale was used to evaluate the open data format.
Results. *The flow of COVID-19 epidemiologic data in Italy originated from the Local Health Authorities*

[1] *CREAGEN - Environmental, Genetic and Nutritional Epidemiology Research Center, Section of Public Health, Department of Biomedical, Metabolic and Neural Sciences, University of Modena and Reggio Emilia, Modena, Italy*
[2] *School of Public Health, University of California Berkeley, Berkeley, CA, USA*
[3] *Head Office, Azienda USL-IRCCS di Reggio Emilia, Reggio Emilia, Italy*
[4] *Department of Epidemiology, Boston University School of Public Health, Boston, MA, USA*

Available ORCIDs: 0000-0003-2100-0344 (TF); 0000-0002-0551-2473 (MV)

that transmitted the data - on a daily basis - to the regional authorities, which in turn transferred them to the national authorities. We found a rather high heterogeneity in both the content and the format of the released data, both at the local and the regional level. Few Regions were releasing data in open format. ISS was the only national source of data that provided the number of COVID-19 health outcomes divided by sex and age groups since Spring 2020.

Conclusions. *Despite multiple potential useful sources for COVID-19 epidemiology are present in Italy, very few open format data were available both at a macro geographical level (e.g. per Region) and at the provincial level. The access to open format epidemiologic data should be eased, to allow researchers to adequately assess future epidemics and therefore favor timely and effective public health interventions.*

Introduction

Italy was the first country to be severely affected by the spread of the new coronavirus SARS-CoV-2 and the related disease COVID-19 (1). Italian and foreign researchers swiftly planned and conducted therefore studies on the epidemiology of the SARS-CoV-2 outbreak in the country, including the mode of transmission of the virus, the incubation period, and the effectiveness of measures to limit transmission (2-4). The study of this new epidemic was important to define, assess and support public health measures that could contain the infection: researchers demonstrated the effectiveness of lockdowns (5-10), use of protective masks, and social distancing towards of SARS-CoV-2 (11-12).

More generally, the epidemiologic characteristics of an epidemic have a very significant impact on the behavior of both the individual and of the whole community, conditioning public health choices and with consequent sanitary, psychosocial, educational and economic implications (13-16). In addition, such investigation may assess the impact of a new pandemic on access to hospital and/or home care for both infected and uninfected individuals, because these parameters have important implications on the organization of healthcare system, especially during a pandemic (17-21).

For this reason, the availability of data is a key aspect to be considered by both researchers and policy makers. In particular,

the availability of 'open' data has become of paramount relevance, and its use has widely increased, especially in the most recent years. The term 'open data' refers to certain types of data (e.g. information, numerical data) that can be freely used, modified and shared by anyone for any purpose (22). Open format is a file format with no restrictions, monetary or otherwise, placed upon its use, and can be fully processed with at least one free/libre/open-source software tool (22).

In this study, we aimed at examining the main sources and features of epidemiologic data related the COVID-19 outbreak in Italy that can be used for research purposes, also assessing the quality and completeness of the reporting and their openness, since the beginning of the pandemic up to the most recent periods.

Materials and Methods

We assessed the availability, quality and format of COVID-19 related data at two different time points, the end of 2020 and the beginning of 2022. We started our search from the most local data up to the regional and national levels. In particular, we performed a systematic screening of websites and data repositories of the subsequent authorities/agencies: the Local Health Authorities (AUSL) of the Emilia-Romagna Region (with few exceptions, the boundaries of an AUSL coincide with those of a Province); the Italian Regions and the

two Autonomous Provinces (PA) of Trento and Bolzano/Bozen (each one holding the powers of a Region); the Ministry of Health; the national Civil Protection Department; and the National Institute of Health (ISS).

For each source, we assessed the availability of open data, the type of COVID-19 related data (total and daily cases, total and daily deaths, total and daily swabs, subdivision by age groups and by sex, subdivision by municipality or province of residence), and the format of data. Starting from the most local sources, we investigated whether data were released or not, and the frequency of their publication (e.g., daily, weekly).

To assess the overall quality of retrieved data, we used the Tim Berners Lee's 5-star rating scale, which is a tool specifically developed for the evaluation of open data reporting (23). In details, the Tim Berners-Lee scale implemented a 5-star deployment scheme (Figure 1) to evaluate the quality of open data according to the following rating system: 1 star indicates the availability of data in open license (e.g. PDF or JPG); 2 stars indicate that data are structured but accessible in a proprietary format (e.g. Microsoft Office Package, like Excel datafile type); 3 stars indicate that data are accessible in non-proprietary open format (e.g. comma separated value-csv datafile type); 4 stars indicate that data are characterized by a structured open format data with Uniform Resource Identifier (URI), e.g. they are available through a unique URL (Uniform Resource Locator) that identifies and locates the data for their re-use; 5 stars level is reached when links to other data and/or online content are included to provide context and explanations, e.g. cross-links to other webpages.

Results

Flow of data collection about the spread of the COVID-19 outbreak are shown in Figure 2. Data are initially collected at provincial level by the Local Health
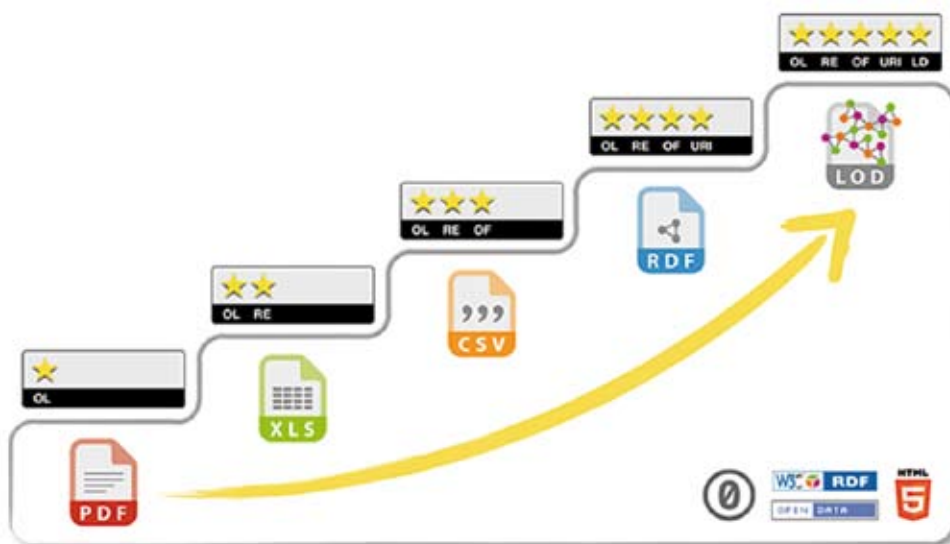

Figure 1 - 5-star scheme for open data from the simplest (1 star) to the most complex (5 stars) available on https://5stardata.info/en/ and the content on this site is freely available under the CC0 Public Domain Dedication.
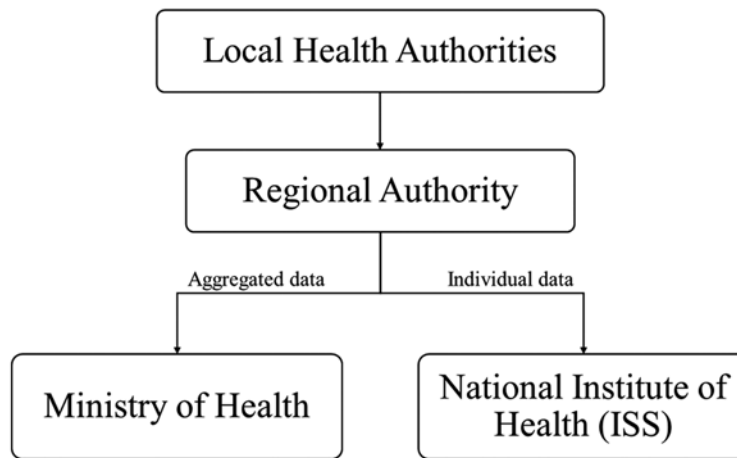
Figure 2 - COVID-19 data flow in Italy, from local to national sources.

Authorities (AUSL) working with hospitals, outpatient services and general practitioners. Each regional authority collects data from all the AUSLs of its territory and send the overall combined data to national authorities, the Ministry of Health (24) and the National Institute of Health (ISS) (25). In particular, aggregated data are sent to the Ministry of Health that releases - in collaboration with the Civil Protection Department - data of total number of positive tests, cases, deaths, hospitalizations and admissions to intensive care in each Italian Region of Italy through the GitHub platform with both daily update and historical repository from the beginning of the pandemic (26). Conversely, regional authorities send individual data to the National Institute of Health (25). These data include also demographic characteristics, presence of comorbidities, clinical status at diagnosis and evolution of the severity of the disease over time. These enriched data are fundamental for the implementation of the "integrated microbiological and epidemiologic surveillance for COVID-19" carried out by the National Institute of Health, providing essential information to support policy decisions and other public health interventions.

In Table 1 we present the availability of COVID-19 related variables for each AUSL of Emilia-Romagna Region (27-38). At the time of our first search (December 2020), only two AUSLs [Modena (27) and Parma (28)] did not release any data. No AUSL released data in open format. Moreover, some AUSL, such as Piacenza (29), Reggio Emilia (30) and Romagna (31), didn't release data at the beginning of the pandemic. After almost two years since the beginning of the pandemic in January 2022, we updated the data search (28-35) and found that Modena AUSL (33) began to release weekly reports containing overall new daily cases in the whole province and divided by municipality (Table 1). Yet, no AUSL released open format data and some did not release some data previously made available, namely "total cases by municipality" (-25% of AUSL), "new cases by municipality" (-13%), "total swabs" (-13%), and "daily swabs/weekly swabs" (-13%) and weekly cases/weekly swabs ratio (-13%).

In Table 2, we show the data available from Regions (39-66) and PAs of Trento (67, 68) and Bolzano (69, 70). In December 2020, all Regions/PAs released data on cumulative cases as well as new daily cases. However,

Table 1 - Availability of COVID-19 related variables for each AUSL of Emilia-Romagna Region. The "+" symbol indicates that data are available, the "-" symbol indicates they are not, the first one for the assessment carried out in December 2020, and the second one for the assessment carried out in January 2022. Evaluation of Tim Berners-Lee 5-star scale also reported with type of file used.

AUSL	Total cases	Daily cases	Total cases by municipality	Daily cases by municipality	Age groups	Sex division	Total swabs	Daily/weekly swabs	Weekly cases/ weekly swabs ratio	5-star scale evaluation	File type
Piacenza	+/+	+/+	+/-	+/-	+/+	-/-	-/-	+/+	+/+	★	html, pdf
Reggio Emilia	+/+	+/+	-/-	+/+	-/-	-/-	-/-	-/-	-/-	★	pdf
Bologna	+/+	*+/+	-/+	+/+	-/-	-/-	-/-	+/-	-/-	★	html, pdf
Imola	+/+	+/+	+/-	+/-	+/-	+/-	+/-	-/-	-/+	★	html, ppt##
Ferrara	+/+	+/+	+/-	+/+	-/+	-/+	-/-	+/+	-/-	★	word
Romagna	+/+	+/+	+/+	+/+	-/-	-/-	-/-	+/+	+/+	★	html, pdf
Modena	-/+	-/+	-/-	-/+	-/-	-/-	-/-	-/-	-/-	★	html##, pdf##
Parma	-/-	-/-	-/-	-/-	-/-	-/-	-/-	-/-	-/-	★	-

Notes: *After July 1 2020; ##Available in January 2022.

no Region published data on cases divided by sex, except for the Umbria and Apulia Regions (although as percentages only). Only five regional websites [Veneto (39, 40), Friuli-Venezia-Giulia (44, 45), Tuscany (48), Umbria (49-51) and Campania (58, 59) and the PAs of Trento (67,68) and Bolzano (69,70)] allowed, however, to download the data in csv format. In the subsequent assessment of data released by Regions carried out in January 2022 (41-50, 52, 55, 59-62, 64-73) we found some differences over time: in particular, the open data previously released by the Veneto Region in 2020 were no longer available in the update of January 2022 (71), and the data released by the Molise Region were no longer updated as the latest available data were published in July 10, 2021 (74).

As far as information on COVID-19 deaths is concerned, the main difference between data released by Emilia-Romagna's AUSLs in December 2020 and January 2022 (27-38) is the discontinuation of publication of data related to total deaths and subdivision by sex (Table 3). Similarly, Table 4 shows the main differences between COVID-19 deaths data published by Italian Regions/ PAs (39-73), indicating that in January 2022, Veneto and Molise Regions stopped to provide updates of the data, as already observed for cases.

At the end of 2020 and in January 2022, the Ministry of Health (24) published a daily report in pdf format and a dashboard (75) (Table 5). Another important source of data is the GitHub repository of Civil Protection Department (26, 76). The website provides the total number of cases at the provincial level without additional details such as sex or age.

Conversely, as of December 2020, the National Institute of Health (ISS) reported a daily open data in excel format, but there were no historical data (77). The most relevant difference in the January 2022 update is the availability of an archive with complete epidemiologic data from April 23, 2020 up to the most recent period. Data were originally published on the ISS website on a daily basis only, with no availability of historical data at the beginning. Conversely, since mid-2021, an open access data format to the entire dataset is released and is continuously updated.

Table 2 - Availability of COVID-19 related variables for each Italian Region and autonomous province (PA). The "+" symbol indicates that data are available, the "-" symbol indicates they are not, the first one for the assessment carried out in December 2020, and the second one for the assessment carried out in January 2022. Evaluation of Tim Berners-Lee 5-star scale also reported with type of file used.

Region	Total cases	Daily cases	Currently positive cases	Total cases by province	Daily cases by province	Age groups	Sex division	Total swabs	Daily swabs	5-star scale evaluation	File type
Veneto	+/-	+/-	+/-	+/-	+/-	-/-	-/-	+/-	-/-	★★★	csv#
Lombardy	+/+	+/+	+/+	+/+	+/+	-/-	-/-	+/+	+/+	★	html, pdf#
Piedmont	+/+	+/+	-/-	++	-/-	-/-	-/-	+/+	+/+	★	html
Emilia Romagna	+/+	+/+	+/+	+/+	+/+	-/-	-/-	+/+	+/+	★	html
PA Trento	+/+	+/+	+/+	+/+	+/+	-/-	-/-	-/-	-/-	★★★	html##, csv
PA Bolzano	+/+	+/+	+/+	+/+	+/+	-/-	-/-	+/+	+/+	★★★	html##, csv
Friuli-Venezia-Giulia	+/+	+/+	+/+	+/+	-/-	-/+	-/-	-/-	+/+	★★★	html, csv
Aosta Valley	+/+	+/+	+/+	+/+	+/+	-/-	-/-	+/+	+/+	★	pdf
Liguria	+/+	+/+	+/+	-/-	-/-	-/-	-/-	+/+	+/+	★	html
Tuscany	+/+	+/+	-/-	+/+	+/+	-/-	-/-	-/-	+/+	★★★	html, pdf##, csv, json##
Umbria	+/+	+/+	+/+	+/+	-/-	+/+	+/+	+/+	-/-	★★★	html##, json, csv##
Marches	+/+	+/+	+/+	+/+	-/+	-/-	-/+	+/+	+/+	★	pdf
Lazio	+/+	+/+	+/+	-/-	+/+	-/+	-/-	-/-	+/+	★★★°	html, pdf##, xls##, csv##
Abruzzo	+/+	+/+	+/+	+/+	+/+	+/+	-/-	+/+	+/+	★	html
Molise	+/-	+/-	-/-	-/-	-/-	-/-	-/-	-/-	-/-	★	html#
Campania	+/+	+/+	+/+	+/-	-/-	-/-	-/-	+/+	+/+	★★★	pdf, csv
Apulia	+/+	+/+	+/+	+/+	+/+	+/-	+/-	+/+	+/+	★	html, pdf
Basilicata	+/-	+/+	+/+	+/-	+/+	-/-	-/-	+/-	+/+	★	html#, pdf
Calabria	+/+	+/+	+/+	+/+	+/+	-/-	-/-	+/+	+/+	★	html, pdf
Sicily	+/+	+/+	+/+	+/+	-/-	-/-	-/-	+/+	-/-	★	html
Sardinia	+/-	+/+	-/-	+/-	+/-	-/-	-/-	+/-	+/+	★	html

Notes: #Available in December 2020 but non available in January 2022; ##Available only in January 2022; °One star in December 2020 and three stars in January 2022.

As regards the Tim Berners-Lee scale evaluation (Tables 1-5), data released by the Emilia-Romagna Region's AUSLs were characterized by the same level, 1-star, both at the end of 2020 and in January 2022. This was the case also for 13 out of 21 Regions/PAs, while the remaining 8 Regions released the data in open format (e.g. using the csv format), reaching the 3-star rating. As regards national data sources, both at the end of 2020 and in January 2022, the Ministry of Health was rated 1-star, due to the lack of open data format. Conversley, the Civil Protection Department relased open data files (using the csv datatype) through the GitHub Repository, thus reaching a 3-star rating. Finally, the National Institute of Health was given a 2star rating, due to the use of an Excel data format for the released data.

Discussion

In the present study we assessed the availability of sources of data to conduct an epidemiologic study, and to highlight the difficulties to retrieve data about COVID-19 pandemic.

Table 3 - Availability of COVID-19 related variables about fatal cases for each AUSL of Emilia-Romagna Region. The "+" symbol indicates that data are available, the "-" symbol indicates they are not, the first one for the assessment carried out in December 2020, and the second one for the assessment carried out in January 2022. Evaluation of Tim Berners-Lee 5-star scale also reported with type of file used.

AUSL	Total deaths	New deaths	Total deaths by municipality	New deaths by municipality	Age groups	Sex division	5-star scale evaluation	File type
Piacenza	+/-	+/+	-/-	-/-	+/+	+/+	★	html, pdf
Reggio Emilia	+/+	+/+	+/+	+/+	+/+	+/+	★	pdf
Bologna	+/+	+/-	+/+	+/-	-/-	+/-	★	html, pdf
Imola	+/+	+/+	+/+	+/+	+/+	+/+	★	html, ppt[#]
Ferrara	+/+	+/+	+/+	+/+	+/+	+/+	★	word
Romagna	+/+	-/+	+/+	-/-	-/-	-/-	★	html, pdf
Modena	-/-	-/-	-/-	-/-	-/-	-/-	★	-
Parma	-/-	-/-	-/-	-/-	-/-	-/-	★	-

Notes: #Available only in January 2022.

We were able to find numerous sources of cases. Starting with the local sources, i.e. the Emilia-Romagna Region AUSLs, we noticed a large variety in the information published in the bulletins. Considerably heterogeneous information was found both among the 8 AUSLs of the Emilia-Romagna Region and in the bulletins published by the individual AUSLs in different periods (December 2020 and January 2022). The presence of such different formats increased the difficulties for the researcher to obtain a complete and clear picture of the local situation of the COVID-19 pandemic. A key and unfortunate aspect was that not a single AUSL, neither at the end of 2020 nor in January 2022, published data in open format.

Daily bulletins published in January 2022 by the Regions contained less detailed information than at the end of 2020, e.g. only 81% of Regions/PAs provided total cases and only 19 out of 21 Regions/PAs released data about daily new cases. In December 2020, every Region and PA reported total and daily cases. One possible explanation for the delays and the lack of homogeneity among the investigated sources is the absence of national guidelines to determine which data should be published in regional bulletins at the early stage of the pandemic. As a matter of fact, other researchers asked for mandatory common standards and mechanisms to collect, upload, summarize and report epidemiologic data (78). Unfortunately, each Region/PA independently decided the information to publish and the format of data. Moreover, some Regions such as Piedmont, Emilia-Romagna and Abruzzo published the data in conversational form, instead of using tables and graphs. We noticed that only few Regions decided to release data in open format. This lack of homogeneity among regional bulletins was also present in other European countries, such as Spain (79).

It should also be noted that data provided by the Ministry of Health (24) and the ISS (77) were considerably different, since

Table 4 - Availability of COVID-19 related variables about fatal cases for each Italian Region. The "+" symbol indicates that data are available, the "-" symbol indicates they are not, the first one for the assessment carried out in December 2020, and the second one for the assessment carried out in January 2022. Evaluation of Tim Berners-Lee 5-star scale also reported with type of file used.

AUSL	Total deaths	Daily deaths	Total deaths by province	Daily deaths by province	Age groups	Sex division	5-star scale evaluation	File type
Veneto	+/-	-/-	-/-	-/-	-/-	-/-	★★★	csv[#]
Lombardy	+/+	+/+	+/+	-/-	+/+	-/-	★	html, pdf[#]
Piedmont	+/+	+/+	+/+	-/-	-/-	-/-	★	html
Emilia Romagna	+/+	+/+	-/-	+/+	+/+	+/-	★	html
PA Trento	+/+	+/+	+/+	+/+	-/-	-/-	★★★	html[##], csv
PA Bolzano	+/+	+/+	+/+	+/+	-/-	-/-	★★★	html[##], csv
Friuli-Venezia-Giulia	+/+	+/+	+/+	+/+	+/+	+/+	★★★	html, csv
Aosta Valley	+/+	-/+	+/+	-/+	+/-	+/-	★	pdf
Liguria	+/+	+/+	-/-	+/+	+/+	+/+	★	html
Tuscany	+/+	+/+	+/+	+/+	-/-	-/-	★★★	html, pdf[##], csv, json[##]
Umbria	+/+	-/-	+/+	-/-	-/-	-/-	★★★	html[##], csv[##], json
Marche	+/+	+/+	+/+	+/+	+/+	+/+	★	pdf
Lazio	+/+	+/+	-/-	+/+	+/+	+/+	★★★	html, pdf[##], xls[##], csv[##]
Abruzzo	+/+	+/+	-/-	+/+	+/+	+/+	★	html
Molise	-/-	-/-	+/-[°]	-/-	+/-	+/-	★	html[#]
Campania	+/+	+/+	-/-	-/-	-/-	-/-	★★★	pdf, csv
Apulia	+/+	+/+	-/-	+/-	+/-	-/-	★	html, pdf
Basilicata	+/-	-/+	+/-	-/+	-/-	-/-	★	html[#], pdf
Calabria	+/+	+/+	+/+	-/+	-/-	-/-	★	html, pdf
Sicily	+/+	+/+	-/-	-/-	-/-	-/-	★	Html
Sardinia	+/+	+/+	-/-	+/+	+/+	-/+	★	Html

Notes: # Available in December 2020 but not available in January 2022; ##Available only in January 2022; °Available by municipality.

while the Ministry releases daily aggregated data, the ISS provided daily individual cases which also include demographic characteristics, comorbidities, clinical status, and evolution over time of the disease which are helpful for more accurate analysis (80). The aggregated data were referred to the date of notification by Regions/PAs, while the data of the ISS individual case flow were expressed according to the date of onset of symptoms, diagnosis, hospitalization, and death, thus adding some discrepancies when comparing aggregated and individual data. In January 2022, the most important and useful new feature was the presence of an archive containing daily open data. This archive can be downloaded and accessed by any citizen. On the contrary, at the end of 2020, ISS open data were replaced at each update.

The advantage of aggregated data is the speed and simplicity of collection to generate hypothesis about possible risk factors (81-83) and to compare the waves of the outbreak (84, 85). Nonetheless, such findings need to be confirmed by other studies as individual data take more time to be collected and analyzed,

Table 5 - Availability of COVID-19 related variables for Ministry of Health (MH), National Institute of Health (ISS) and Civil Protection Department (CPD). The "+" symbol indicates that data are available, the "-" symbol indicates they are not, the first one for the assessment carried out in December 2020, and the second one for the assessment carried out in January 2022. Evaluation of Tim Berners-Lee 5-star scale also reported with type of file used.

	Total cases	Daily cases	Currently positive	Total cases by province	Daily cases by province	Age groups	Sex division	Total swabs	Daily swabs	Total deaths	Daily deaths	Total deaths by province	Daily deaths by province	5-star scale evaluation	File type
MH	+/+	+/+	+/+	+/+	+/+	-/-	-/-	+/+	+/+	+/+	+/+	+/+	-/-	★	html, pdf
ISS	-/-	+/+	-/-	+/+	-/-	+/+	+/+	-/-	-/-	+/+	+/+	-/-	-/-	★★	xls
CPD	+/+	+/+	+/+	+/+	-/-	-/-	-/-	+/+	-/-	+/+	-/-	+/+	-/-	★★★	csv, json, xml

but allow a more detailed and insightful analysis. For this reason, the two types of data differ. Individual data are important both for research purposes and for public health interventions. For example, timely reporting of individual data showed that the elderly population was at higher risk of mortality for COVID-19 (86). For this reason, senior citizens were prioritised at the beginning of the vaccination campaign in Italy.

As for the publication frequency, local, regional, and national sources generally release data reports daily. However, there are marked differences between weekday and weekends, since fewer people are swabbed on weekends, particularly on Sunday. Since data are released about 24 hours later, on Mondays the figures about the numbers of swabs performed in the previous day are systematically lower compared to the other weekdays, and the positivity rate tends to be higher. In order to addressed this issue, three- or seven-day averages are generally used to estimate positivity rate (87).

Interestingly, the Ministry of Health published a daily bulletin in pdf format which is a one-star level format and an interactive dashboard that provides a link to access the GitHub repository (26). The GitHub platform (26) provided data in csv format at national, regional and provincial level. The main

limitation of this repository was that at the provincial level (which generally coincides with the territory of one AUSL) only the total number of cases was reported and more detailed information, such as new cases or deaths, was not present. Also Germany and Spain use the GitHub repository (88, 89) for data sharing on COVID-19. In Germany the number of cases and deaths were reported in a file providing information aggregated by nation, state or province. Data were accessible in csv format. The European Centre for Disease Prevention and Control (ECDC) created a daily reporting data in open format, containing cases and deaths (90) and the information are summarized in a dashboard (91). Data are notified to ECDC by national health authorities.

Open data are important for scientific research, particularly in the field of epidemiology (92), because the availably of data permits a rapid response to pandemic (93, 94). In addition, if researchers do not have open format data available, the database must be created manually and this greatly increases both the time needed for the research activity and the possibility of mistakes and inaccuracies.

COVID-19 is the first pandemic occurred in the 'social-media era'. Social media can be used to publish rapidly correct information

such as the number of cases and deaths. However, they can also be used to convey incorrect information (95). The sharing of open data on social media can simplify correct communication of risk to the entire population, in particular to young people, and this can be in turn associated with a more adequate effectiveness of pandemic control measures (96).

Conclusions

In conclusion, there appears to be multiple data sources available to investigate COVID-19 epidemiology in Italy at both the local, regional, and national level, but such data are generally difficult to access, as well as often incomplete and heterogeneous with reference to their format and frequency of update. Unfortunately, most Regions either do not publish open format data or provide aggregate or out-of-date data, further hampering the implementation of epidemiologic research on an outbreak such as that of COVID-19. The availability of up-to-date, accessible and complete data is important for both communication to the general population, who should be timely informed about the evolution of the outbreak and with adequate details, and scientific research, because accessible open data allow fast transmission of information that facilitates research in the clinical, diagnostic and pharmacological fields. Therefore, to improve and facilitate epidemiologic research, it might perhaps be useful to have a unified management system on a national basis, to which each Region would daily upload the new data, and from which all data useful for epidemiologic and scientific research could be freely downloaded. All the data, both local and regional, should then be comprehensive, of good methodological quality, and made easily and freely accessible to all researchers, in the best interest of both epidemiologic research and public health.

Riassunto

Disponibilità di open data sull'epidemia di COVID in Italia

Introduzione. Fin dall'insorgenza della pandemia COVID-19 in Italia, le autorità sanitarie hanno pubblicato dati epidemiologici. Questi dati hanno costituito fondamentali fonti informative per la ricerca epidemiologica, al fine di prevedere e contenere l'evoluzione dell'epidemia. Non sempre, tuttavia, sono stati resi disponibili ai ricercatori e agli operatori di sanità pubblica dati esaustivi e tempestivi sull'evoluzione del COVID-19, sia in ambito locale che nazionale.

Metodo. L'obiettivo del nostro lavoro è indagare la qualità, la disponibilità e il formato dei dati epidemiologici sul COVID-19 in Italia in diversi ambiti territoriali e temporali. Abbiamo tal fine preso in esame le risorse online messe a disposizione da ciascuna delle 19 Regioni italiane, dalle due Province autonome e, più in dettaglio, dalle AUSL della Regione Emilia-Romagna. Abbiamo analizzato le principali fonti e i flussi di dati (ossia i casi di infezione nuovi e cumulativi, i tamponi totali, i decessi COVID-19 nuovi e cumulativi, complessivi e suddivisi per sesso). Abbiamo descritto le loro caratteristiche, tra cui l'accessibilità, il formato e la completezza. Inoltre, abbiamo esaminato i dati pubblicati dal Ministero della Salute, dall'Istituto Superiore di Sanità e dal Dipartimento della Protezione Civile. Per valutare il formato dei dati aperti è stata utilizzata la scala Tim Berners-Lee.

Risultati. I dati epidemiologici COVID-19 sono stati sistematicamente trasmessi su base giornaliera dalle AUSL alle autorità regionali, le quali a loro volta li hanno inviati alle autorità nazionali. Abbiamo riscontrato un'eterogeneità piuttosto elevata sia nel contenuto che nel formato dei dati rilasciati, sia a livello locale che regionale. Poche Regioni hanno pubblicato i dati in formato aperto. L'ISS è la fonte di dati che ha fornito il numero di esiti sanitari riferiti al COVID-19 suddivisi per sesso e fasce d'età, sin dall'inizio dell'epidemia nella primavera 2020.

Conclusioni. Nonostante in Italia siano presenti molteplici fonti potenzialmente utili per l'epidemiologia del COVID-19, pochi dati in formato aperto sono stati resi disponibili sia a livello regionale che in ambito provinciale. L'accesso ai dati epidemiologici in formato aperto dovrebbe essere sensibilmente facilitato, al fine di consentire ai ricercatori di svolgere tempestivamente un'efficace attività di ricerca sullo sviluppo di future epidemie, favorendo in tal modo l'individuazione di interventi di sanità pubblica tempestivi ed efficaci.

References

1. Filippini T, Zagnoli F, Bosi M, Giannone ME, Marchesi C, Vinceti M. An assessment of case-fatality and infection-fatality rates of first and second COVID-19 waves in Italy. Acta Biomed. 2021 Oct 1; **92**(S6): e2021420. doi: 10.23750/abm.v92iS6.12241.

2. Krumkamp R, Duerr HP, Reintjes R, Ahmad A, Kassen A, Eichner M. Impact of public health interventions in controlling the spread of SARS: modelling of intervention scenarios. Int J Hyg Environ Health. 2009 Jan; **212**(1): 67-75. doi: 10.1016/j.ijheh.2008.01.004. Epub 2008 May 6.

3. Liang WN, Huang Y, Zhou WX, Qiao L, Huang JH, Wu ZL. Epidemiological characteristics of an outbreak of severe acute respiratory syndrome in Dongcheng District of Beijing from March to May 2003. Biomed Environ Sci. 2003 Dec; **16**(4): 305-13.

4. Gandhi M, Yokoe DS, Havlir DV. Asymptomatic Transmission, the Achilles' Heel of Current Strategies to Control Covid-19. N Engl J Med. 2020 May 28; **382**(22): 2158-60. doi: 10.1056/NEJMe2009758. Epub 2020 Apr 24.

5. Lau H, Khosrawipour V, Kocbach P, et al. The positive impact of lockdown in Wuhan on containing the COVID-19 outbreak in China. J Travel Med. 2020 May 18; **27**(3): taaa037. doi: 10.1093/jtm/taaa037.

6. Kraemer MUG, Yang CH, Gutierrez B, et al. The effect of human mobility and control measures on the COVID-19 epidemic in China. Science. 2020 May 1; **368**(6490): 493-7. doi: 10.1126/science.abb4218. Epub 2020 Mar 25.

7. Vinceti M, Filippini T, Rothman KJ, et al. Lockdown timing and efficacy in controlling COVID-19 using mobile phone tracking. EClinicalMedicine. 2020 Aug; **25**: 100457. doi: 10.1016/j.eclinm.2020.100457. Epub 2020 Jul 13.

8. Guzzetta G, Riccardo F, Marziano V, et al. Impact of a Nationwide Lockdown on SARS-CoV-2 Transmissibility, Italy. Emerg Infect Dis. 2021 Jan; **27**(1): 267-70. doi: 10.3201/eid2701.202114. Epub 2020 Oct 20.

9. Sjödin H, Wilder-Smith A, Osman S, Farooq Z, Rocklov J. Only strict quarantine measures can curb the coronavirus disease (COVID-19) outbreak in Italy, 2020. Euro Surveill. 2020 Apr; **25**(13): 2000280. doi: 10.2807/1560-7917.ES.2020.25.13.2000280.

10. Vinceti M, Balboni E, Rothman KJ, et al. Substantial impact of mobility restrictions on reducing COVID-19 incidence in Italy in 2020. J Travel Med. 2022 Jul 24: taac081. doi: 10.1093/jtm/taac081. Online ahead of print.

11. Ford N, Holmer HK, Chou R, et al. Mask use in community settings in the context of COVID-19: A systematic review of ecological data. EClinicalMedicine. 2021 Aug; **38**: 101024. doi: 10.1016/j.eclinm.2021.101024. Epub 2021 Jul 19.

12. Coclite D, Napoletano A, Gianola S, et al. Face Mask Use in the Community for Reducing the Spread of COVID-19: A Systematic Review. Front Med (Lausanne). 2020 Jan 12; **7**: 594269. doi: 10.3389/fmed.2020.594269.

13. Dubey S, Biswas P, Ghosh R, et al. Psychosocial impact of COVID-19. Diabetes Metab Syndr. 2020 Sep-Oct; **14**(5): 779-88. doi: 10.1016/j.dsx.2020.05.035. Epub 2020 May 27.

14. Singh S, Roy D, Sinha K, Parveen S, Sharma G, Joshi G. Impact of COVID-19 and lockdown on mental health of children and adolescents: A narrative review with recommendations. Psychiatry Res. 2020 Nov; **293**: 113429. doi: 10.1016/j.psychres.2020.113429. Epub 2020 Aug 24.

15. MacDonald M, Hill C. The educational impact of the Covid-19 rapid response on teachers, students, and families: Insights from British Columbia, Canada. Prospects (Paris). 2021 Jan 15: 1-15. doi: 10.1007/s11125-020-09527-5. Epub ahead of print.

16. Nicola M, Alsafi Z, Sohrabi C,et al. The socio-economic implications of the coronavirus pandemic (COVID-19): A review. Int J Surg. 2020 Jun; **78**: 185-93. doi: 10.1016/j.ijsu.2020.04.018. Epub 2020 Apr 17.

17. Di Bidino R, Cicchetti A. Impact of SARS-CoV-2 on Provided Healthcare. Evidence From the Emergency Phase in Italy. Front Public Health. 2020 Nov; **8**: 583583. doi: 10.3389/fpubh.2020.583583.

18. Maccarone MC, Masiero S. The Important Impact of COVID-19 Pandemic on the Reorganization of a Rehabilitation Unit in a National Healthcare System Hospital in Italy: Lessons From Our Experience. Am J Phys Med Rehabil. 2021 Apr 1; **100**(4): 327-30. doi: 10.1097/PHM.0000000000001707.

19. Bonalumi G, Giambuzzi I, Barbone A, et al. A call to action becomes practice: cardiac and vascular surgery during the COVID-19 pandemic

based on the Lombardy emergency guidelines. Eur J Cardiothorac Surg. 2020 Aug 1; **58**(2): 319-27. doi: 10.1093/ejcts/ezaa204.

20. Ruggieri P, Trovarelli G, Angelini A, Pala E, Berizzi A, Donato D. COVID-19 strategy in organizing and planning orthopedic surgery in a major orthopedic referral center in an area of Italy severely affected by the pandemic: experience of the Department of Orthopedics, University of Padova. J Orthop Surg Res. 2020 Jul 23; **15**(1): 279. doi: 10.1186/s13018-020-01740-4.

21. Deana C, Rovida S, Orso D, et al. Learning from the Italian experience during COVID-19 pandemic waves: be prepared and mind some crucial aspects. Acta Biomed. 2021 May 12; **92**(2): e2021097. doi: 10.23750/abm.v92i2.11159.

22. Foundation Open Knowlege. The Open Definition 2015. Available on: http://opendefinition.org/ [Last accessed: 2022 June 23].

23. Berners-Lee T. 5-star scheme for open data. Available on: https://5stardata.info/en/ [Last accessed: 2022 June 23].

24. Ministry of Health. Coronavirus daily bulletin 2020. Available on: https://www.salute.gov.it/portale/nuovocoronavirus/homeNuovoCoronavirus.jsp [Last accessed: 2022 June 23].

25. National Institute of Health. COVID-19 integrated surveillance system 2020. Available on: https://www.epicentro.iss.it/coronavirus/sars-cov-2-sorveglianza [Last accessed: 2022 June 23].

26. CPD. Italian Civil Protection Department. COVID-19 data 2020 [Available on: https://github.com/pcm-dpc/COVID-19 [Last accessed: 2022 June 23].

27. Local Health Authority of Modena. Web site of Local Health Authority of Modena 2020. Available on: https://www.ausl.mo.it/coronavirus [Last accessed: 2022 June 23].

28. Local Health Authority of Parma. Web site of Local Health Authority of Parma 2020. Available on: https://www.ausl.pr.it [Last accessed: 2022 June 23].

29. Local Health Authority of Piacenza. Web site of Local Health Authority of Piacenza 2020. Available on: https://www.covidpiacenza.it/situazione-covid-19-a-piacenza/ [Last accessed: 2022 June 23].

30. Local Health Authority of Reggio Emilia. Web site of Local Health Authority of Reggio Emilia 2020. Available on: https://www.ausl.re.it/covid-19-aggiornamento-quotidiano-dei-dati [Last accessed: 2022 June 23].

31. Local Health Authority of Romagna. Web site of Local Health Authority of Romagna 2020. Available on: https://www.auslromagna.it/quadro-epidemiologico-covid-19-ausl-romagna [Last accessed: 2022 June 23].

32. DSP AUSL Bologna. COVID-19 data 2020. Available on: https://public.tableau.com/app/profile/dsp.ausl.bologna [Last accessed: 2022 June 23].

33. Modena Local Health Authority. COVID-19 Emergency-Joint weekly update 2020. Available on: https://www.ausl.mo.it/coronavirus-andamento-casi [Last accessed: 2022 June 23].

34. Local Health Unit of Ferrara. Daily bulletins cases COVID19 epidemic trend 2020 [Available from: https://www.ausl.fe.it/ [Last accessed: 2022 June 23].

35. Local Health Unit of Imola. Daily updates on the COVID-19 epidemic in the imolese territory January 2022 2020. Available on: https://www.ausl.imola.bo.it/gennaio22 [Last accessed: 2022 June 23].

36. Local Health Unit of Imola. daily updates on the COVID-19 epidemic in the imolese territory 2020. Available on: https://www.ausl.imola.bo.it/flex/cm/pages/ServeBLOB.php/L/IT/IDPagina/10285 [Last accessed: 2022 June 23].

37. Local Health Unity of Ferrara. Daily bulletins cases COVID-19 epidemic trend 2020. Available on:https://salute.regione.emilia-romagna.it/tutto-sul-coronavirus [Last accessed: 2022 June 23].

38. Local Health Unit of Bologna. Report archive 2020. Available on: https://www.ausl.bologna.it/per-i-cittadini/coronavirus/rcps/archivio-report-settimanali/archivio-report-2020 [Last accessed: 2022 June 23].

39. Veneto Region. Veneto situation monitoring daily update 2020. Available on: https://dati.veneto.it/opendata/COVID_19_Monitoraggio_situazione_Veneto_aggiornamento_quotidiano [Last accessed: 2020 December 10].

40. Veneto Region. Monitoring situation data in Veneto 2020 [Available on: https://dati.veneto.it/opendata/COVID_19_Monitoraggio_situazione_Dati_di_dettaglio_nel_Veneto [Last accessed: 2020 December 10].

41. Lombardy Region. Dashboard Covid-19 Lombardy Region 2020. Available on: https://www.

regione.lombardia.it/wps/portal/istituzionale/ HP/servizi-e-%20informazioni/cittadini/salute-e-prevenzione/coronavirus/dashboard-covid19 [Last accessed: 2022 June 23].

42. Piedmont Region. Covid: gli aggiornamenti su contagi e vaccini 2020. [Available on: https://www.regione.piemonte.it/web/pinforma/notizie/covid-gli-aggiornamenti-contagi-vaccini [Last accessed: 2022 June 23].

43. Emilia-Romagna Region. Emilia-Romagna news 2021 2020 [Available on: https://www.regione.emilia-romagna.it/notizie/raccolte/coronavirus?b_start:int=870 [Last accessed: 2022 June 23].

44. Friuli-Venezia-Giulia Region. FVG Civil Protection map 2020 [Available on: https://covid19-map.protezionecivile.fvg.it [Last accessed: 2022 June 23].

45. Friuli-Venezia-Giulia Region. FVG report 2020. Available on: https://www.regione.fvg.it/rafvg/comunicati/comunicati/comunicati/comunicati/comunicati/comunicati/comunicati/ricercacom.act?txtMese=10&txtAnno=2020&btnCerca=vai [Last accessed: 2022 June 23].

46. Aosta Valley Region. Aosta Valley daily bulletin 2020. Available on: https://www.regione.vda.it/pressevda/Eventi/bollettini_i.aspx [Last accessed: 2022 June 23].

47. Liguria Region. Liguria Homepage Coronavirus 2020. Available on: https://www.regione.liguria.it//homepage/salute-e-sociale/homepage-coronavirus.html [Last accessed: 2022 June 23].

48. Tuscany Region. Data Covid daily update 2020. Available on: https://www.ars.toscana.it/banche-dati/dati-sintesi-sintcovid-aggiornamenti-e-novita-sul-numero-dei-casi-deceduti-tamponi-per-provincia-e-per-asl-della-regione-toscana-e-confronto-con-italia-con-quanti-sono-i-decessi-per-comune?provenienza=home_ricerca&dettaglio=ric_geo_covid&par_top_geografia=090 [Last accessed: 2022 June 23].

49. Umbria Region. Coronavirus Umbria Region 2020. Available on: https://coronavirus.regione.umbria.it/casi [Last accessed: 2022 June 23].

50. Umbria Region. Dataset Covid-19 Municipalities 2020. Available on: http://dati.umbria.it/dataset/covid_comuni [Last accessed: 2022 June 23].

51. Umbria Region. Umbria Region's API 2020. Available on: https://apistore.regione.umbria.it/store/apis/info?name=COVID-19-rsion=1.0.0&provider=admin&tag=Agenda%20

digitale-group [Last accessed: 2022 June 23].

52. Marche Region. Coronavirus daily report 2020. Available on: https://www.regione.marche.it/News-ed-Eventi/Categorie/Term/2440/Coronavirus [Last accessed: 2022 June 23].

53. Lazio Region. Twitter daily bulletin 2020. Available on: https://twitter.com/SaluteLazio?ref_src=twsrc%5Egoogle%7Ctwcamp%5Eserp%7Ctwgr%5Eauthor[Last accessed: 2022 June 23].

54. Lazio Region. Lazio Coronavirus data 2020. Available on: https://www.salutelazio.it/ [Last accessed: 2022 June 23].

55. Abruzzo Region. Coronavirus daily bulletin 2020. Available on: https://www.regione.abruzzo.it/notizie_stampa [Last accessed: 2022 June 23].

56. Molise Region. Covid-19 Monitoring 2020. Available on: http://www3.regione.molise.it/flex/cm/pages/ServeBLOB.php/L/IT/IDPagina/17206 [Last accessed: 2022 June 23].

57. Campania Region. Campania Region daily bulletin 2020. Available on: http://www.regione.campania.it/regione/it/la-tua-campania/coronavirus-kyxz/aggiornamenti?page=1 [Last accessed: 2022 June 23].

58. Civil Protection Department. COVID-19 Monitoring: total cases by province in Campania Region 2020. Available on: https://dati.regione.campania.it/catalogo/datasetdetail/covid-19-monitoraggio-situazione-totale-casi-per-provincia-nella-regione-campania [Last accessed: 2022 June 23].

59. Campania Region. COVID-19 Monitoring: Campania Region data 2020. Available on: https://dati.regione.campania.it/catalogo/datasetdetail/covid-19-monitoraggio-situazione-dati-di-dettaglio-relativi-alla-regione-campania [Last accessed: 2022 June 23].

60. Apulia Region. Apulia Region daily bulletin 2020. Available on: https://www.regione.puglia.it/web/press-regione/ricerca?delta=75&q=archivio%20bollettini%20covid&category=181787 [Last accessed: 2022 June 23].

61. Basilicata Region. COVID-19 Emergency daily update 2020. Available on: https://www.regione.basilicata.it/giunta/site/giunta/section.jsp?otype=1012&value=regione [Last accessed: 2022 June 23].

62. Calabria Region. Calabria daily bulletin 2020. Available on: https://portale.regione.calabria.it/website/ [Last accessed: 2022 June 23].

63. Sicily Civil Protection department. Sicily CO-

VID-19 Dashboard 2020. Available on: https://drpcsicilia.maps.arcgis.com/apps/opsdashboard/index.html#/e9ba8a03150f486a8ff06a1ca7-b6278e [Last accessed: 2022 June 23].

64. Sicily Region. COVID-19 Monitoring 2020. Available on: https://www.protezionecivilesicilia.it/it/219-monitoraggio-e-mappatura-dinamica-dei-contagi-da-covid-19-nel-territorio-della-regione-sicilia.asp [Last accessed: 2022 June 23].

65. Sicily Region. Special COVID-19 Emergency 2020. Available on: https://www.protezionecivilesicilia.it/it/archivio/speciale-covid19.asp [Last accessed: 2022 June 23].

66. Sardinia Region. COVID-19 daily bulletin 2020. Available on: https://www.regione.sardegna.it/j/v/2571?&s=1&v=9&c=94255&nodesc=1 [Last accessed: 2022 June 23].

67. PA Trento. COVID-19 Spread 2020. Available on: https://patn.maps.arcgis.com/apps/opsdashboard/index.html#/30d39ef2c83a47d285512a4685235843 [Last accessed: 2022 June 23].

68. PA Trento. COVID-19 Emergency Health Situation 2020. Available on: https://covid19trentino.fbk.eu/#download [Last accessed: 2022 June 23].

69. PA Bolzano. COVID-19 News 2020. Available on: http://www.provincia.bz.it/news/it/news.asp?news_page=1 [Last accessed: 2022 June 23].

70. PA Bolzano. Current Coronavirus data 2020. Available on: https://www.provincia.bz.it/sicurezza-protezione-civile/protezione-civile/dati-attuali-sul-coronavirus.asp [Last accessed: 2022 June 23].

71. Veneto Region. Coronavirus emergency 2020. Available on: https://www.regione.veneto.it/web/guest/emergenza-coronavirus [Last accessed: 2022 June 23].

72. Lazio Region. Dashboard Covid-19 Lazio Region 2021. Available on: https://www.dep.lazio.it/covid/dashboard/dashboard_COVID19_Lazio.php [Last accessed: 2022 June 23].

73. Molise Region. Covid-19-useful information 2020. Available on: https://www.regione.molise.it/flex/cm/pages/ServeBLOB.php/L/IT/IDPagina/1 [Last accessed: 2022 June 23].

74. ASREM. Coronavirus vaccinations in the Molise Region 2020. Available on: https://adesionivaccinazionicovid.regione.molise.it/app/f?p=300:92 [Last accessed: 2022 June 23].

75. Ministry of Health. COVID-19 Dashboard 2020. Available on: https://opendatadpc.maps.arcgis.com/apps/dashboards/b0c68bce2cce478eaac82-fe38d4138b1 [Last accessed: 2022 June 23].

76. Italian Civil Protection D, Morettini M, Sbrollini A, Marcantoni I, Burattini L. COVID-19 in Italy: Dataset of the Italian Civil Protection Department. Data Brief. 2020 Apr 10; **30**: 105526. doi: 10.1016/j.dib.2020.105526.

77. National Institute of Health. COVID-19 integrated surveillance: key national data 2020. Available on: https://www.epicentro.iss.it/en/coronavirus/sars-cov-2-integrated-surveillance-data [Last accessed: 2022 June 23].

78. Gardner L, Ratcliff J, Dong E, Katz A. A need for open public data standards and sharing in light of COVID-19. Lancet Infect Dis. 2021 Apr; **21**(4): e80. doi: 10.1016/S1473-3099(20)30635-6. Epub 2020 Aug 10.

79. Martín Fernandez A, Alfaro Latorre M, Crespo Sanchez-Eznarriaga B, Fidalgo Garcia L. [Availability of open data related to COVID-19 in the web portals of the Spanish regions and the Ministry of Health]. Rev Esp Salud Publica. 2021 Dec 2; **95**: e202112190. Spanish.

80. Ministry oh Health. The Italian surveillance network 2020. Available on: https://www.salute.gov.it/portale/nuovocoronavirus/dettaglioContenutiNuovoCoronavirus.jsp?lingua=italiano&id=5351&area=nuovoCoronavirus&menu=vuoto [Last accessed: 2022 June 23].

81. Filippini T, Rothman KJ, Goffi A, et al. Satellite-detected tropospheric nitrogen dioxide and spread of SARS-CoV-2 infection in Northern Italy. Sci Total Environ. 2020 Oct 15; **739**: 140278. doi: 10.1016/j.scitotenv.2020.140278. Epub 2020 Jun 16.

82. Copat C, Cristaldi A, Fiore M, et al. The role of air pollution (PM and NO2) in COVID-19 spread and lethality: A systematic review. Environ Res. 2020 Dec; **191**: 110129. doi: 10.1016/j.envres.2020.110129. Epub 2020 Aug 24.

83. Filippini T, Rothman KJ, Cocchio S, et al. Associations between mortality from COVID-19 in two Italian regions and outdoor air pollution as assessed through tropospheric nitrogen dioxide. Sci Total Environ. 2021 Mar 15; **760**: 143355. doi: 10.1016/j.scitotenv.2020.143355. Epub 2020 Nov 4.

84. Vinceti M, Filippini T, Rothman KJ, et al. SARS-CoV-2 infection incidence during the first and second COVID-19 waves in Italy. Environ Res. 2021 Jun; **197**: 111097. doi: 10.1016/j.envres.2021.111097. Epub 2021 Apr 1.

85. Vinceti M, Filippini T, Rothman KJ, et al. The association between first and second wave COVID-19 mortality in Italy. BMC Public Health. 2021 Nov 11; **21**(1): 2069. doi: 10.1186/s12889-021-12126-4.

86. Characteristics of COVID-19 patients dying in Italy. Available on: https://www.epicentro.iss.it/coronavirus/sars-cov-2-decessi-italia [Last accessed: 2022 June 23].

87. Sciannameo V, Goffi A, Maffeis G, et al. A deep learning approach for Spatio-Temporal forecasting of new cases and new hospital admissions of COVID-19 spread in Reggio Emilia, Northern Italy. J Biomed Inform. 2022 Jul 11; **132**: 104132. doi: 10.1016/j.jbi.2022.104132. Online ahead of print.

88. Robert Koch Institute (RKI). Covid-19-germany-gae 2020. Available on: https://github.com/jgehrcke/covid-19-germany-gae [Last accessed: 2022 June 23].

89. Ministerio de Sanidad. Datasets relacionados con la incidencia de la COVID-19 en España 2020. Available on: https://github.com/datadista/datasets/tree/master/COVID%2019 [Last accessed: 2022 June 23].

90. European Centre for Disease Prevention and Control (ECDC). Data on the daily number of new reported COVID-19 cases and deaths by EU/EEA country 2020. Available on: https://www.ecdc.europa.eu/en/publications-data/data-daily-new-cases-covid-19-eueea-country [Last accessed: 2022 June 23].

91. European Centre for Disease Prevention and Control (ECDC). COVID-19 Situation Dashboard 2021. Available on: https://qap.ecdc.europa.eu/public/extensions/COVID-19/COVID-19.html#global-overview-tab [Last accessed: 2022 June 23].

92. Sarli L, Monaco F. Does open data tell us everything? Acta Biomed. 2021 Dec 22; **92**(S2): e2021337. doi: 10.23750/abm.v92iS2.12715.

93. Data Europa Academy. The Power of Data in the Everlasting Battle: Humanity vs. Virus (1/2) 2020. Available on: https://data.europa.eu/en/impact-studies/covid-19/power-data-everlasting-battle-humanity-vs-virus-12 [Last accessed: 2022 June 23].

94. Moorthy V, Henao Restrepo AM, Preziosi MP, Swaminathan S. Data sharing for novel coronavirus (COVID-19). Bull World Health Organ. 2020 Mar 1; **98**(3): 150. doi: 10.2471/BLT.20.251561.

95. González-Padilla DA, Tortolero-Blanco L. Social media influence in the COVID-19 Pandemic. Int Braz J Urol. 2020 Jul; **46**(Suppl 1): 120-4. doi: 10.1590/S1677-5538.IBJU.2020.S121.

96. Kobayashi S, Falcon L, Fraser H, et al. Using Open Source, Open Data, and Civic Technology to Address the COVID-19 Pandemic and Infodemic. Yearb Med Inform. 2021 Aug; **30**(1): 38-43. doi: 10.1055/s-0041-1726488. Epub 2021 Apr 21.

Corresponding author: Prof. Marco Vinceti, Department of Biomedical, Metabolic, and Neural Sciences, University of Modena and Reggio Emilia, Via Campi 287, 41125 Modena, Italy
e-mail: marco.vinceti@unimore.it