Journal of Infection and Public Health 18 (2025) 102621

Contents lists available at ScienceDirect

Journal of Infection and Public Health

journal homepage: www.elsevier.com/locate/jiph

Does syndromic surveillance assist public health practice in early detecting respiratory epidemics? Evidence from a wide Italian retrospective experience

Giovanni Corrao [a,b,c], Andrea Stella Bonaugurio [b,d,*], Giorgio Bagarella [c,e], Mauro Maistrello [c,f], Olivia Leoni [c], Danilo Cereda [c], Andrea Gori [g,h]

[a] University of Milano-Bicocca, Milan, Italy
[b] National Centre for Healthcare Research and Pharmacoepidemiology, University of Milan-Bicocca, Milan, Italy
[c] Welfare Department, Operative Centre for Health Data, Lombardy Region, Milan, Italy
[d] Biostatistics, Epidemiology and Public Health Unit, Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Milan, Italy
[e] Agency for Health Protection of the Metropolitan Area of Milan, Lombardy Region, Milan, Italy
[f] Local Health Unit of Melegnano and Martesana, Milan, Italy
[g] First Division of Infectious Diseases, ASST Fatebenefratelli Sacco, Luigi Sacco Hospital, Milan, Italy
[h] Centre for Multidisciplinary Research in Health Science (MACH), University of Milan, Italy

A R T I C L E I N F O

Article history:
Received 11 October 2024
Received in revised form 28 November 2024
Accepted 9 December 2024

Keywords:
COVID-19
Exponentially weighted moving average control chart
Google trends
Healthcare utilization data
Improved farrington algorithm
Syndromic surveillance

A B S T R A C T

Background: Large-scale diagnostic testing has been proven ineffective for prompt monitoring of the spread of COVID-19. Electronic resources may facilitate enhanced early detection of epidemics. Here, we aimed to retrospectively explore whether examining trends in the use of emergency and healthcare services and the Google search engine is useful in detecting Severe Acute Respiratory Syndrome Coronavirus outbreaks early compared with the currently used swab-based surveillance system.
Methods: Healthcare Utilization databases of the Italian region of Lombardy and the Google Trends website were used to measure the weekly utilization of emergency and healthcare services and determining the volume of Google searches from 2020 to 2022. Improved Farrington algorithm (IMPF) and Exponentially Weighted Moving Average (EWMA) control chart were both fitted to detect outliers in weekly searches of nine syndromic tracers. AND/OR Boolean operators were tested aimed for joint using tracers and models. Signals that occurred during periods labelled as free from epidemics were used to measure positive predictive values (PPV) and false negative values (FNV) in anticipating the epidemic wave.
Results: Out of the 156 weeks of interest, 70 (45 %) were affected by epidemic waves. Overall, 54 syndromic signals were obtained from any one of the 7 healthcare or Google tracers, generating an outlier from both the EWMA and IMPF models. PPV values of 0.95, 1.00, 0.96 admitting a delay of 0, 1, and 2 weeks, respectively, between signal and epidemic wave. The values of FNP ranged from 0.19 to 0.21.
Conclusions: High predictive power for anticipating COVID-19 epidemic waves, even two weeks ahead of the official reports, was obtained from electronic syndromic tracers of healthcare-seeking trends and Google search engine use. Following verification via a prospective approach, public health organizations are encouraged to take advantage of this free forecasting system to anticipate and effectively manage respiratory outbreaks.
© 2024 The Author(s). Published by Elsevier Ltd on behalf of King Saud Bin Abdulaziz University for Health Sciences. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

Syndromic surveillance involves tracking conditions of public health interest based on non-specific tracers (e.g., symptoms and clinical signs reported by patients and clinicians) rather than by counting cases that meet the criteria for a specific case definition, e.g., from positive laboratory assay [1]. The non-specific nature of

* Correspondence to: Dipartimento di Statistica e Metodi Quantitativi, Università degli Studi di Milano-Bicocca, Via Bicocca degli Arcimboldi, 8, Edificio U7, Milano 20126, Italy.
E-mail address: andrea.bonaugurio@unimib.it (A.S. Bonaugurio).

https://doi.org/10.1016/j.jiph.2024.102621
1876-0341/© 2024 The Author(s). Published by Elsevier Ltd on behalf of King Saud Bin Abdulaziz University for Health Sciences. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

syndromic surveillance and the rapid data collection enable makes it timely, sensitive, and flexible enough for addressing different situations/scenarios, including infectious outbreaks [2]. When syndromic surveillance systems were first established in the mid-1990s [3,4], there was a particular emphasis on monitoring influenza activity. Such surveillance systems received further impetus because of their potential utility for the early detection of bioterrorist events in the wake of the terrorist attack in the USA in September 2001 [5].

Emerging evidence has suggested that syndromic surveillance systems can predict outbreaks of COVID-19 with high spatial and temporal resolution [6–9]. The recent COVID-19 epidemic arrived in Italy earlier than in any other western country, and then spread exponentially [10]. The most severely affected part of Italy, by far, was Lombardy—a northern region in which SARS-CoV-2 infected several millions of patients and has been associated with a high rate of hospitalizations for intensive care and mortality [11]. To ensure preparedness for a new epidemic wave, the Regional Health Authority of Lombardy implemented a syndromic surveillance system of transmissible diseases based on promptly available electronic data to complement the conventional sources. High uncertainty is associated with early prediction of the start of epidemic waves using syndromic tracers and/or overburdened health services. Therefore, we gradually proceeded by verifying the properties of data generated from the out-of-hospital utilization of healthcare [12] and emergency services [13], and web-based platforms [14]. However, the predictive power obtained by combining these methods with tracers remains unknown. Furthermore, there has been considerable emphasis in the literature on the statistical methods used to underpin syndromic surveillance systems, with much of this work based on synthetic and modelled data [15–19]. However, few studies have been published on the practical experience of establishing and operating such systems—for example, by empirically investigating how tracers and models may be combined into formulating wider and prompter public health response [20].

In this study, we investigated how the syndromic data behaved in predicting the COVID-19 epidemic waves that occurred from February 2020 to December 2022. To this end, we explored the possibility of combining data on the usage of emergency and healthcare services with internet search data. Furthermore, because statistical models for early and accurately detecting the onset of epidemics are based on different assumptions, the performance of their combined use has been investigated. The principal objective was to obtain information on the best way to process syndromic data in order to maximize timeliness and predictivity of alert signals, whilst minimizing the risk of false positive signals.

Methods

Target population and data sources

Residents in Lombardy who were beneficiaries of the Regional Health Service formed the target population (just over 10 million people, around 16 % of the Italian population). Italian citizens have equal access to essential healthcare services provided by the National Health Service (NHS)— in Lombardy, this has been associated with an automated system of databases that collect a variety of information including, among others, ICD-CM-9 codes of inpatient diagnoses and services supplied by public or private hospitals as well as ATC codes of outpatient dispensation of drugs for a variety of treatments that are reimbursed to the pharmacies after filing doctors' prescriptions.

Outcomes

SARS-CoV-2 infection was the main outcome of interest. Infections were diagnosed by real-time reverse transcription-polymerase chain reaction (RT-PCR) assay of nasopharyngeal swabs processed at a laboratory accredited by Regional Health Authorities between January 2020 and December 2022 (monitored period). In addition, the weekly number of all-cause ICU admissions was considered the secondary outcome.

Potential tracers

Three families of tracers were tested. The first comprises two indicators of emergency services usage. The first of these indicators was the proportion of Emergency Department (ED) visits diagnosed as respiratory syndrome in a given week out of the total number of ED visits made in that week (note that the use of a proportion rather the absolute number overcomes the problem of the reduction of patients going to an ED during the pandemic). The second was the absolute number of calls to an Emergency Medical Communication Center recognised as respiratory or infective syndrome [21]. Both tracers are defined as conventional in Lombardy, given they were systematically used throughout the pandemic crisis period.

The second family includes tracers generated from data for the usage of out-of-hospital healthcare attributable to respiratory syndrome. Four tracers were considered. One, the proportion of chest X-rays performed in a given week among the total number of X-rays performed in that week was considered. Two, because azithromycin has both anti-inflammatory and antiviral properties and has been hypothesized to have activity against SARS-CoV-2 [22], the proportion of azithromycin dispensed in a given week among the total number of antibiotics dispensed in that week was considered. (Note that chest X-ray and azithromycin were again expressed as proportions in order to overcome the observed reduction in performing X-rays and administering antibiotics throughout the pandemic crisis period). The third and fourth healthcare syndromic tracers were the absolute number of dispensed non-steroidal anti-inflammatory drugs (NSAIDs) and corticosteroids.

Finally, Google Trends [23] was used to search for the weekly usage of a set of non-specific COVID-19-related terms that, according to a recent study by us, better capture the phenomenon under consideration [14]. The relevant terms were "cough", "fever" and "sore throat". Google Trends does not provide absolute search numbers, but instead provides a measure known as 'interest over time' that ranges from 0 to 100, with 0, 50, and 100 respectively indicating that there is insufficient use for the term, the term is half as popular, and the term is at its peak popularity [24].

Warning algorithms

To identify weekly excesses in the use of emergency, healthcare and google services, the Exponentially Weighted Moving Average (EWMA) control chart and the improved Farrington (IMPF) algorithm were applied to the nine tracers separately. Details of the current application of the two algorithms can be found in Supplementary Appendix S1, while readers interested in further details may consider the methodological literature [25,26]. Briefly, both algorithms allow the identification of unusually high values (i.e. syndromic outliers) in count time series that can present temporal trends and seasonality. This is done by comparing the counts (or the proportion) observed in the monitored period (the period between January 2020 and December 2022 in the present application) with those expected by applying a prediction model to historical data. Whenever an unusual value is detected, the algorithms generate a warning signal. While both algorithms can detect abrupt changes in the time series, EWMA was developed as an extension of cumulative sum control charts, enabling the identification of small but constant changes. In contrast, IMPF automatically reweights historical observation to reduce the influence of past outlying counts. Being the considered algorithms based on different assumptions, their concordance in generating weekly warning signal was estimated using Cohen's Kappa.

Looking for joint use of predictors and models

The weekly incidence rate of SARS-CoV-2 infections detected by the conventional surveillance system during the monitored period was plotted. To identify the subperiods that were affected by an outbreak wave, a nine-week period was defined as being impacted by the epidemic. This period consists of four weeks preceding the peak of the epidemic, the week in which the peak occurred, and the four weeks following the peak, including the peak week itself. The remaining weeks, which fall outside of this nine-week timeframe, were classified as outbreak-free.

Syndromic outliers generated from EWMA and IMPF for each of the nine tracers were identified by including two months prior the starting observation period (i.e., from January 2020 onwards) to enable the syndromic signal to be anticipated.

We employed Boolean operators (AND/OR) to simplify the statistical process. Based on visual inspection of plots, two questions were asked. First, how may individual tracers be combined within each family of potential tracers? Second, how may statistical algorithms be combined? For example, we verified the behaviour of emergency services in generating alarm signals in the case of both visits and calls (AND operator) or at least one of the two (OR operator), to generate a syndromic outlier. Analogously, we verified the behaviour of a given tracer in generating alarm signals in the case of both EWMA and IMPF (AND operator) or at least one of two (OR operator), to generate a syndromic outlier.

Syndromic signals accuracy and timeliness

To investigate whether a syndromic signal could correctly predict an outbreak wave, we computed the proportion of weeks labelled as affected by "syndromic outliers" that fell in a week affected by an epidemic wave out of the total number of weeks affected by syndromic outliers. This measure was denoted as the positive predictive value (PPV) of a syndromic outlier. In addition, to assess whether the absence of a signal is falsely associated with an epidemic wave, we computed the proportion of weeks free from a signal that fell in a week affected by an epidemic wave out of the total number of outbreak-free weeks based on syndromic outliers. This measure (i.e., the complementary to unit of the negative predictive value) was denoted as false negative value (FNV) of the absence of a syndromic outlier.

The ability of syndromic outliers to detect the onset of an epidemic wave early was investigated by setting a time gap between the syndromic signal and the epidemic event corresponding to lags of 0, 1, and 2 weeks.

Finally, the association between combined tracers (predictors) and the rate of ICU admission (response) was investigated by multivariable Poisson regression model. Predictors were indexed as 1 or 0 according whether the syndromic outlier was detected in a given week or not. The response rate was allocated in the same week of the predictor or moved forward by 1 or 2 weeks to assess whether tracers could anticipate the onset of an epidemic wave. The number of ICU admissions expected whether (i) no syndromic outlier was detected, (ii) only one tracer was detected, or (iii) all predictors generated a syndromic outlier was estimated from the model parameters associated with the main terms (positivity to a syndromic signal alone) or the interaction term (positivity to all syndromic signal). Note that the model used ICU admission (secondary outcome) rather that SARS-CoV-2 infection (principal outcome) because only a small proportion of infections were detected at the beginning of the epidemic (a period when governments were not ready to face the emergency). Furthermore, the ability to measure the performance of syndromic tracers in predicting when essential health services—such as intensive care units—reach capacity, adds value to our approach.

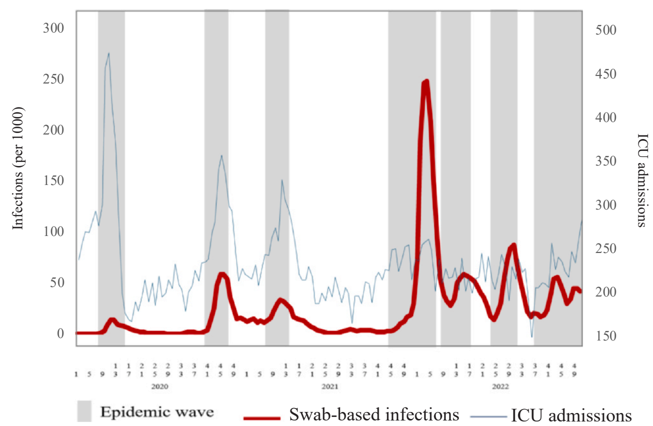

Fig. 1. Observed weekly trends in the number of confirmed SARS-CoV-2 infections and Intensive Care Unit (ICU) admissions. Grey vertical bands indicate the weeks affected by epidemic wave. Lombardy, Italy, January 2020 - December 2022.

Results

From February 23, 2020, to December 31, 2022, 4,065,829 confirmed cases of SARS-CoV-2 were detected in Lombardy (with an incidence rate of 2.6 infections per 1000 person-week), and 35,500 patients were admitted to ICUs (with an incidence rate of 2.3 admissions per 100,000 person-month).

Fig. 1 depicts the weekly trends in the number of SARS-CoV-2 infections (and epidemic waves appointed periods) and of ICU admissions. According to our criteria, of the 156 weeks of interest, 70 (45 %) were affected by epidemic waves, and seven epidemic waves occurred during the monitored period.

Fig. 2 depicts the weeks affected by the syndromic signal according to individual tracers and models. Overall, 448 syndromic outliers were detected from EWMA (253) and IMPF (195). The PPV ranged from 70 % (corticosteroids/EWMA) to 100 % (all emergency and some healthcare tracers). The FNV ranged from 32 % (sore throat/IMPF) to 45 % (X-ray/EWMA and corticosteroids/IMPF). Between-models Cohen's Kappa coefficients ranged from 0.11 (95 % confidence interval −0.11–0.43) to 0.75 (0.57–0.93), respectively, for corticosteroids and X-ray. Cohen's Kappa quantifies the agreement between two tools when classifying data into nominal categories, considering the possibility of chance agreement, with values ranging from −1–1: a coefficient of 1 indicates perfect agreement, 0 indicates no better agreement than expected by chance (random agreement), values less than 0 suggest worse than chance agreement (systematic disagreement), and values greater than 0 indicate agreement better than chance.

Fig. 3 compares the behaviour of combining individual tracers within each setting of tracers, by AND/OR operators. Combining tracers with the AND operator: (i) generated 16 and 13 weeks interested to outliers belonging to emergency setting with EWMA and IMPF, respectively; (ii) did not generate any signal of outlier belonging to the healthcare setting; and (iii) generated 12 and 14 weeks interested to outliers belonging to the Google Trends setting with EWMA and IMPF, respectively; (iv) although all outliers fell during the epidemic waves (PPV = 100 %), FNV had values around 42–47 %. As expected, a very higher number of syndromic outliers was generated by combining tracers with the OR operator being the weeks interested 45 and 29 for emergency tracers generated with EWMA and IMPF, 68 and 35 for healthcare tracers, and 48 and 57 for Google tracers. The PPV ranged from 78 % (healthcare/EWMA) to 93 % (emergency/IMPF). The FNV ranged from 23 % (healthcare/EWMA) to 37 % (emergency/IMPF).

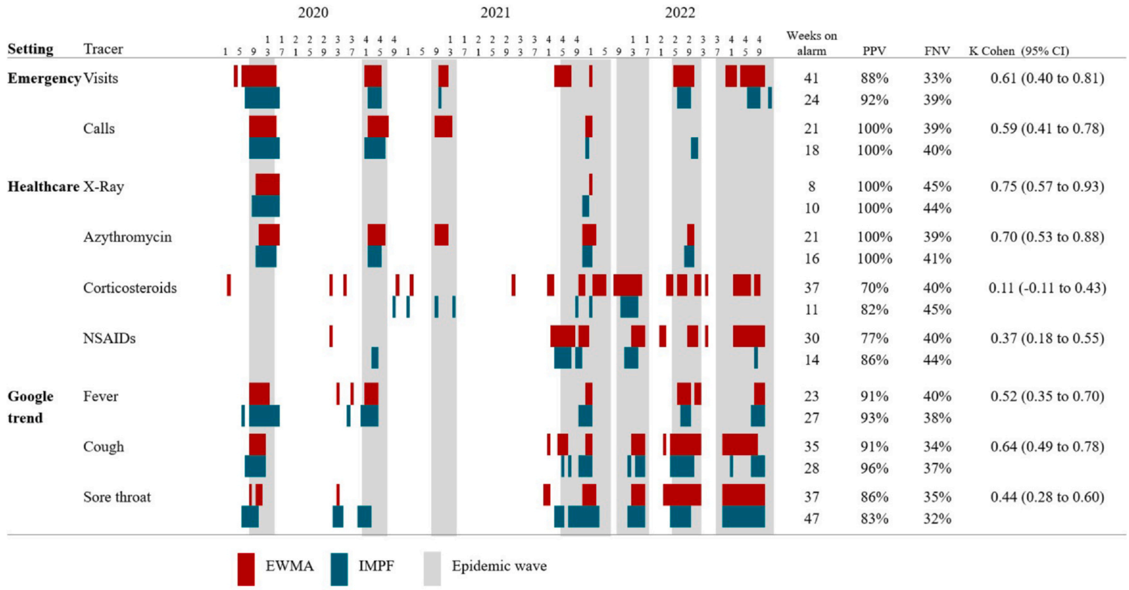

Fig. 2. Weekly syndromic outliers generated from individual emergency, healthcare and google trend tracers. Grey vertical bands indicate the weeks affected by epidemic wave. Red and green horizontal lines respectively indicate outliers generated by the EWMA chart and the IMPF model. Lombardy, Italy, January 2020 - December 2022. PPV and FNV respectively denote positive prediction and false negative values.

Fig. 4 compares the behaviour of combining EWMA and IMPF models by using AND/OR operators and further combining syndromic tracers with the OR operator. Including all syndromic tracers available, 59 and 93 outliers were observed by combining EWMA and IMPF models with the AND/OR operators respectively. Positive predictive and false negative values, respectively, were 95 % and 20 % with the AND operator and 70 % and 13 % with the OR operator. Stable figures were observed by allowing a lag of one and two weeks between syndromic signals and the onset of the outbreak wave. No substantial change in these figures was observed on combining healthcare and Google Trend syndromic tracers (i.e., omitting emergency signals). Alerts are generated in the same areas regardless of whether all three tracer groups (Emergency, Healthcare, Google Trends) are included or, for example, if the "Emergency" group is excluded. This is evident when comparing the top and bottom sections of the graph. Finally, the combination of tracers/models and lag to be preferred is the one that goes to ensure high positive predicted values and low false negative values so as to be able to anticipate any epidemic waves by avoiding false negatives.

Fig. 5 shows that, at baseline—i.e., in the absence of any syndromic outlier—228 ICU admissions were expected. The occurrence of syndromic outliers generated from Google, and even more from healthcare usage figures and Google in combination, entailed a significant increase in expected ICU admissions, even with a two-week delay.

Discussion

In this paper, we explored the possibility of combining respiratory syndromic data and warning algorithms for early and accurate detection of the onset of epidemics caused by airborne

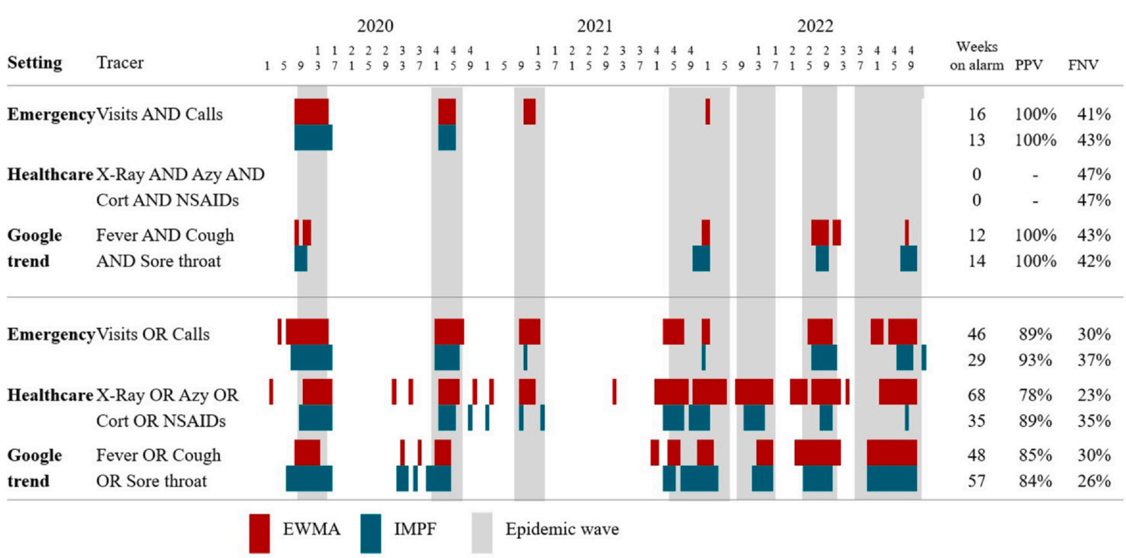

Fig. 3. Weekly syndromic outliers generated from emergency, healthcare and google trend tracers jointly considered by using the AND (upper boxes) and the OR (bottom boxes) Boolean operators. Grey vertical bands indicate the weeks affected by epidemic wave. Red and green horizontal lines respectively indicate syndromic outliers generated by the EWMA chart and the IMPF model. Lombardy, Italy, January 2020 - December 2022. PPV and FNV respectively denote positive prediction and false negative values.

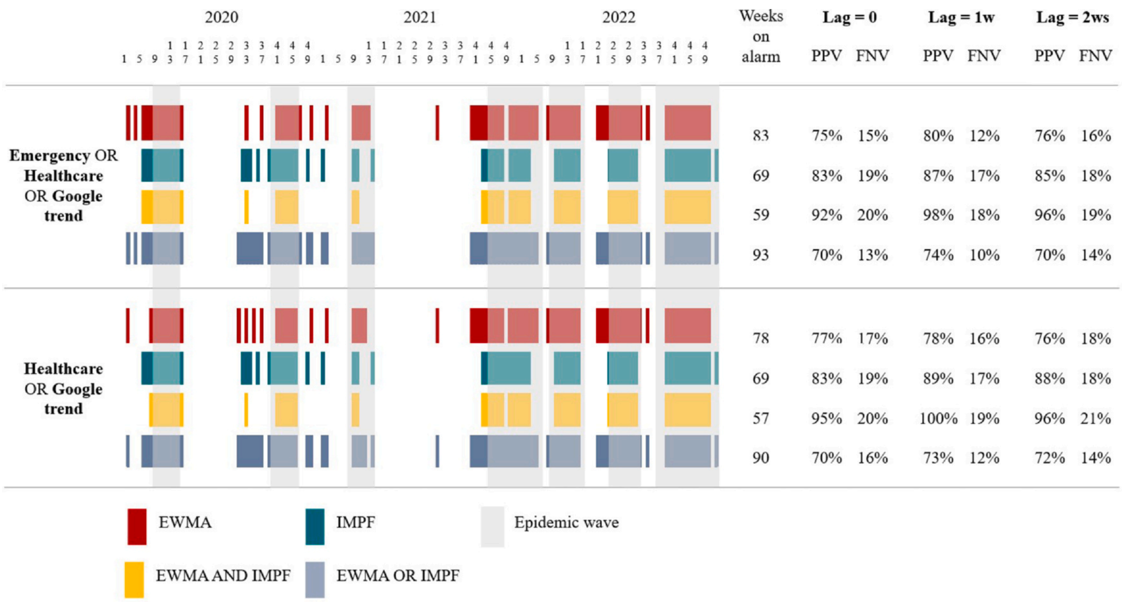

Fig. 4. Weekly syndromic outliers generated by jointly considering emergency, healthcare and google tracers (upper boxes) and healthcare and google tracers (bottom boxes) by using the OR Boolean operator. Red, green, orange and blue horizontal lines respectively indicate syndromic outliers generated EWMA chart, IMPF model, EWMA AND IMPF combined, and EWMA OR IMPF combined. Grey vertical bands indicate the weeks affected by epidemic wave. Lombardy, Italy, January 2020 - December 2022. PPV and FNV respectively denote positive prediction and false negative values. Lag time denotes the number of weeks between that of outlier and epidemic wave was denoted as "lag time".

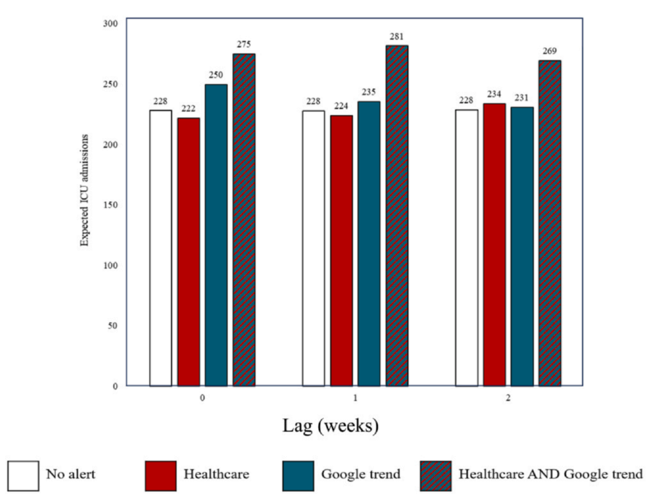

Fig. 5. Expected number of all-cause intensive care unit admissions associated with no syndromic tracers (baseline), syndromic outliers belonging to either healthcare seeking or search google engines, or syndromic outliers belonging to both healthcare seeking and search google engines. Lombardy, Italy, January 2020 - December 2022.

pathogens. Our findings suggest that a surveillance system 'infrastructure' that is fed, in real time, with real-world data on the usage of healthcare, as well as on keyword searches in Google, can anticipate the signal of an epidemic wave even two weeks in advance. The system showed acceptable detection capability, i.e., good positive predictivity (a syndromic outlier was followed by an epidemic wave in 96 % of cases) and tolerable false positive proportion (the absence of a syndromic outlier was followed by an epidemic wave in 21 % of cases).

Our study shows that despite being widely used during the recent COVID-19 pandemics in Lombardy, the performance of conventional syndromic tracers, such as those encompassing emergency services usage data, were rather disappointing. These findings are consistent with those from a recent paper showing that a syndromic surveillance system encompassing emergency data from 12 European countries successfully identified only one in 147 outbreaks [27]. Similar findings were reported even more recently from a UK study [28]. Consistent with other reports, we observed that tracers generated from both healthcare-seeking data [12,29] and use of Google search engines [14,30] exhibited better performance than tracers based on emergency services usage.

As a novel and original message, our study showed that it is sufficient that one of the seven healthcare/google tracers turns on with from both the EWMA chart and the IMPF model for generating an alarm signal. In the absence of consolidated evidence explaining these results, our speculative reasoning is as follows. First, because every individual syndromic tracer is the response of the system (emergency), doctors (healthcare), and citizens (Google) to the spread of the epidemic, and because responses vary over time and between countries, the inclusion of any outlier should allow the identification of any emerging anomalous situation. Second, because the two algorithms for detecting outliers are based on different assumptions [25,26], the imposition of the AND operator allows greater caution against the risk of generating unjustified alarms.

Implications for public health authorities

In this paper, we provide an assessment of the detection capability of algorithms to help decision makers perform weekly surveillance based on three non-hierarchical needs (i) timeliness: anticipating, as far as possible, the warning signal, (ii) detection capability: probability that the triggering of a syndromic alarm signal is suggestive of the imminent occurrence of an epidemic wave, for early initiation of preventive actions aimed at controlling its spread (and to promptly adapt treatment services to ensure the provision of appropriate medical assistance), and (iii) avoiding false positive syndromic alarms that could cause the system to lose credibility.

Limitations and future work

The retrospective design as well as the lack of a validation setting are the main limitations of our study. These limitations notwithstanding, implementation of the present prototype infrastructure for the early detection of respiratory outbreaks should enable prospective validation of the syndromic surveillance algorithm.

Owing to the specific cultural, social, and behavioural factors relevant to this study, our findings likely reflect the Italian (and perhaps Mediterranean) setting; this must be considered when designing syndromic surveillance systems. In general, surveillance based on healthcare seeking and Google Trends search is circumscribed to the health system and to the internet user behaviour typical of the investigated population. Thus, the wider generalizability of its utility in syndromic surveillance, spatially and temporally, is questionable.

Conclusions

The use of syndromic tracers to identify outliers of healthcare seeking and Google searches has high predictive power for anticipating COVID-19 epidemic waves, even two weeks ahead of the official reports. Although the period considered was almost exclusively affected by the spread of the COVID-19 epidemic due to the SARS-CoV-2 virus, because of the non-specific nature of the syndromic tracers used, the results could be extended to the total respiratory virus detections (RVDs) in Lombardy, including influenza, respiratory syncytial virus (RSV), severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), and other respiratory viruses [31]. The system described herein can provide valuable buffer time to take early preventive actions and allocate the necessary supplies and personnel to hospitals. Upon verification by prospective approaches, public health organizations are encouraged to take advantage of this free forecasting system to anticipate and effectively manage COVID-19 outbreaks.

Ethical statement

Ethical approval was not required for the study based on registry data in accordance with the local legislation and institutional requirements. All procedures were in accordance with the Ethical standards of the institutional and national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Data Availability

In this study, publicly available datasets were analysed in addition to administrative data held by Regione Lombardia. Publicly datasets are available here: Google trend data: https://trends.google.it/trends/explore; Covid-19 confirmed-case data: http://github.com/pcm-dpc/COVID-19/tree/master/dati-regioni

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: GC received research support from the European Community (EC), the Italian Agency of Drug (AIFA), and the Italian Ministry for University and Research (MIUR). He took part to a variety of projects that were funded by pharmaceutical companies (i.e., Novartis, GSK, Roche, AMGEN and BMS). He also received honoraria as member of Advisory Board from Roche.

Acknowledgements

This research was funded by Lombardy Region, by means of an agreement aimed at supporting regional health policies, practices, and decision-making processes (project code H45F21003340002).

Use of artificial intelligence tools

'None declared'.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jiph.2024.102621.

References

[1] Cocoros NM, Willis SJ, Eberhardt K, Morrison M, Randall LM, DeMaria A, et al. Syndromic surveillance for COVID-19, Massachusetts, February 2020-November 2022: the impact of fever and severity on algorithm performance. Public Health Rep 2023;138:756–62. https://doi.org/10.1177/00333549231186574
[2] Hughes HE, Edeghere O, O'Brien SJ, Vivancos R, Elliot AJ. Emergency department syndromic surveillance systems: a systematic review. BMC Public Health 2020;20:1891. https://doi.org/10.1186/s12889-020-09949-y
[3] Lazarus R, et al. Using automated medical records for rapid identification of illness syndromes (syndromic surveillance): the example of lower respiratory infection. BMC Public Health 2001;1:9. https://doi.org/10.1186/1471-2458-1-9
[4] Heffernan R, et al. New York city syndromic surveillance systems. MMWR Suppl 2004;53:23–7.
[5] Steiner-Sichel L, et al. Field investigations of emergency department syndromic surveillance signals – New York city. MMWR Suppl 2004;53:184–9.
[6] Desjardins MR. Syndromic surveillance of COVID-19 using crowdsourced data. Lancet Reg Health West Pac 2020;4. https://doi.org/10.1016/j.lanwpc.2020.100024
[7] Chan AT, Brownstein JS. Putting the public back in public health—surveying symptoms of Covid-19. N Engl J Med 2020;383:e45. https://doi.org/10.1056/NEJMp2016259
[8] Nomura S, Yoneoka D, Shi S, et al. An assessment of self-reported COVID-19 related symptoms of 227,898 users of a social networking service in Japan: has the regional risk changed after the declaration of the state of emergency? Lancet Reg Health West Pac 2020;1. https://doi.org/10.1016/j.lanwpc.2020.100011
[9] Maharaj AS, Parker J, Hopkins JP, Gournis E, Bogoch II, Rader B, et al. The effect of seasonal respiratory virus transmission on syndromic surveillance for COVID-19 in Ontario, Canada. Lancet Infect Dis 2021;21:593–4. https://doi.org/10.1016/S1473-3099(21)00151-1
[10] Alteri C, Cento V, Piralla A, Costabile V, Tallarita M, Colagrossi L, et al. Genomic epidemiology of SARS-CoV-2 reveals multiple lineages and early spread of SARS-CoV-2 infections in Lombardy, Italy. Nat Commun 2021;12:434. https://doi.org/10.1038/s41598-022-14426-0
[11] Onder G, Rezza G, Brusaferro S. Case-fatality rate and characteristics of patients dying in relation to COVID-19 in Italy. JAMA 2020;323:1775–6. https://doi.org/10.1001/jama.2020.4683
[12] Merlo I, Crea M, Berta P, Ieva F, Carle F, Rea F, et al. Italian Alert_CoV Project group. Detecting early signals of COVID-19 outbreaks in 2020 in small areas by monitoring healthcare utilisation databases: first lessons learned from the Italian Alert_CoV project. Eur Surveill 2023;28:2200366. https://doi.org/10.2807/1560-7917.ES.2023.28.1.2200366
[13] Bagarella G, Maistrello M, Minoja M, Leoni O, Bortolan F, Cereda D, et al. Early Detection of SARS-CoV-2 epidemic waves: lessons from the syndromic surveillance in Lombardy, Italy. Int J Environ Res Public Health 2022;19:12375. https://doi.org/10.3390/ijerph191912375
[14] Porcu G, Chen YX, Bonaugurio AS, Villa S, Riva L, Messina V, et al. Web-based surveillance of respiratory infection outbreaks: retrospective analysis of Italian COVID-19 epidemic waves using Google Trends. Front Public Health 2023 May 18;11:1141688. https://doi.org/10.3389/fpubh.2023.1141688
[15] Jafarpour N, et al. Quantifying the determinants of outbreak detection performance through simulation and machine learning. J Biomed Inform 2015;53:180–7. https://doi.org/10.1016/j.jbi.2014.10.009
[16] Jafarpour N, et al. Using hierarchical mixture of experts model for fusion of outbreak detection methods. AMIA Annu Symp Proc 2013;2013:663–9.
[17] Spreco A, et al. Integrated detection and prediction of influenza activity for real-time surveillance: algorithm design. J Med Internet Res 2017;19:e211. https://doi.org/10.2196/jmir.7101
[18] Unkel S, Farrington CP, Garthwaite PH, Robertson C, Andrews N. Statistical methods for the prospective detection of infectious disease outbreaks: a review. J R Stat Soc Ser A (Stat Soc) 2012;175:49–82. https://doi.org/10.1111/j.1467-985X.2011.00714.x
[19] Yan P, Chen H, Zeng D. Syndromic surveillance systems. Annu Rev Inf Sci Technol 2008 2008;42:425–95. https://doi.org/10.1002/aris.2008.1440420117
[20] Smith GE, Elliot AJ, Lake I, Edeghere O, Morbey R, Catchpole M, et al. Public Health England real-time syndromic surveillance team. Syndromic surveillance: two decades experience of sustainable systems - its people not just data!. Epidemiol Infect 2019;147:e101. https://doi.org/10.1017/S0950268819000074
[21] Perego E, Balzarini F, Botteri M, Favetti S, Zoli A, Sechi GM, et al. Emergency treatment in Lombardy: a new methodology for the pre-hospital drugs management on advanced rescue vehicles. Acta Biomed 2020;91(3-S):111–8. https://doi.org/10.23750/abm.v91i3-s.9421
[22] Oliver ME, Hinks TSC. Azithromycin in viral infections. Rev Med Virol 2021;31(2):e2163. https://doi.org/10.1002/rmv.2163

[23] Google Trends, ⟨https://trends.google.com/trends/?geo=IT⟩.
[24] Yang S, Santillana M, Kou SC. Accurate estimation of influenza epidemics using Google search data via ARGO. Proc Natl Acad Sci USA 2015;112:14473–8. https://doi.org/10.1073/pnas.1515373112
[25] Noufaily A, Enki DG, Farrington P, Garthwaite P, Andrews N, Charlett A. An improved algorithm for outbreak detection in multiple surveillance systems. Stat Med 2013;32:1206–22.
[26] Höhle M, Paul M. Count data regression charts for the monitoring of surveillance time series. Comput Stat Data Anal 2008;52:4357–68. https://doi.org/10.1016/j.csda.2008.02.015
[27] Ziemann A, et al. A concept for routine emergency-care data-based syndromic surveillance in Europe. Epidemiol Infect 2014;142:2433–46. https://doi.org/10.1017/S0950268813003452
[28] Todkill D, Elliot AJ, Morbey R, Harris J, Hawker J, Edeghere O, et al. What is the utility of using syndromic surveillance systems during large subnational infectious gastrointestinal disease outbreaks? An observational study using case studies from the past 5 years in England. Epidemiol Infect 2016;144:2241–50. https://doi.org/10.1017/S0950268816000480
[29] Nikhab A, Morbey R, Todkill D, Elliot AJ. Using a novel 'difference-in-differences' method and syndromic surveillance to estimate the change in local healthcare utilisation during periods of media reporting in the early stages of the COVID-19 pandemic in England. Public Health 2024;232:132–7. https://doi.org/10.1016/j.puhe.2024.04.022
[30] Kurian SJ, Bhatti AUR, Alvi MA, Ting HH, Storlie C, Wilson PM, et al. Correlations Between COVID-19 cases and google trends data in the United States: a state-by-state analysis. Mayo Clin Proc 2020;95:2370–81. https://doi.org/10.1016/j.mayocp.2020.08.022
[31] Ben Moussa M, Rahal A, Lee L, Mukhi S. Syndromic surveillance performance in Canada throughout the COVID-19 pandemic, March 1, 2020 to March 4, 2023. Can Commun Dis Rep 2023;49:501–9. https://doi.org/10.14745/ccdr.v49i1112a06