

TITLE: Using Open-source Intelligence to Identify Early Signals of COVID-19 in Indonesia

Author: Yoser Thamtono, Aye Moa, Chandini Raina MacIntyre

Abstract

Context: Indonesia has been claiming zero cases of COVID-19 until 2 March 2020. There were doubts over the lack of reported cases through February, given high travel volume between China and Indonesia.

Aims: To identify the early signals of pneumonia of unknown cause as a proxy for COVID-19 in Indonesia.

Methods: Using open-source data of pneumonia of unknown cause in Indonesia between 1 November 2019 to 31 March 2020 obtained from *EpiWATCH*, an open-source epidemic observatory, a descriptive analysis was performed to identify the trend of pneumonia of unknown cause in Indonesia prior to official notification of COVID-19 cases.

Results: A rise in reports of pneumonia of unknown cause is identified in Indonesia starting from late January 2020. There were 304 reported cases of pneumonia of unknown cause, 30 of which occurred prior to the identification of the first COVID-19 cases on 2 March 2020. The early signals of pneumonia of unknown cause in Indonesia may indicate possible unrecognised circulation of COVID-19 prior to official detection.

Conclusions: Open source data can provide early signals of epidemics, especially in countries where surveillance, testing or reporting is low.

Keywords: *Prediction tool, Coronavirus outbreak, COVID-19, Risk Assessment, Indonesia*

Introduction

On 31 December 2019, China reported an increased occurrence of pneumonia of unknown cause in Wuhan ¹, which was later confirmed as COVID-19 (Coronavirus disease 2019), a respiratory illness due to the newly discovered coronavirus, *SARS-CoV-2* (Severe acute respiratory syndrome coronavirus 2). On 13 January 2020, Thailand became the first country to identify imported cases of COVID-19. Likewise, Singapore and Malaysia reported their first imported cases in January 2020 ^{2,3}. WHO declared COVID-19 as a public health emergency of international concern later on 31 January 2020, and by 11 March 2020, the organisation had declared a pandemic ⁴.

Unlike many other Southeast Asian countries, Indonesia has been claiming zero cases of COVID-19 until 2 March 2020, when the first two cases of COVID-19 in Indonesia were officially notified ⁵. Much concern has been raised over the country's case-detection ability during that period. One study, using a generalized linear regression of air travel volume from Wuhan, suggested that there should have been at least 5 cases in Indonesia by February 4th 2020 ⁶.

There has been growing interest in the use of open-source intelligence for rapid epidemic surveillance, especially in countries with weak surveillance and public health infrastructure, or where there may be lack of reporting or censorship of information ⁷. WHO reports that more than half of early epidemic information can be obtained through unofficial sources, including online news and social media ⁸. This study aims to use open-source data for rapid epidemic intelligence to detect the early signals of pneumonia of unknown cause as a proxy for COVID-19 circulation in Indonesia.

Methods

EpiWATCH is a semi-automated open-source epidemic observatory, which collects and analyses outbreak data from publicly available sources, such as online news items ⁹. It is established and managed by the Australian NHMRC Centre for Research Excellence, Integrated Systems for Epidemic Response (ISER) ⁹.

In this study, a database of *EpiWATCH* Outbreak Alerts was retrieved for the disease keywords in Bahasa Indonesia ‘pneumonia, lung infection, lung inflammation, severe acute respiratory infection (SARI), cough, fever, cough AND fever’. The reported cases of pneumonia of unknown cause in Indonesia dated between 1 November 2019 and 31 March 2020 were included in the search. Reports matching the keywords were analysed and were excluded if they were not reporting cases from Indonesia, not related to pneumonia or reporting pneumonia cases with a known cause.

A descriptive epidemiologic analysis was performed using Excel 2016 (Microsoft) to group cases by geolocation, and the date which they occurred. Data reported from the Ministry of Health was also reviewed and compared to the news reports.

Results

Between 1 November 2019 and 31 March 2020, *EpiWATCH* documented 211 entries related to pneumonia of unknown cause in Indonesia. **Figure 1** shows the daily counts of these entries. A steady increase in reporting of such cases is seen starting from late January 2020 until the end of February 2020, then followed by a sharp increase in the number of reports in March 2020. There were 5 entries in late January 2020, 17 entries in February 2020 and 184 entries in March 2020.

During the observation period, there were 304 reported cases of pneumonia of unknown cause in Indonesia; 30 (9.9%) of which occurred prior to the identification of the first COVID-19 cases on 2 March 2020. The number of cases started to increase in late January (6 identified

cases) to February 2020 (19 identified cases), one and a half months prior to official identification of first two COVID-19 cases in Indonesia. **Figure 2** shows the daily counts of these cases in Indonesia.

The majority of cases were reported from provinces in Java, with East Java (39 cases/ 12.8%) and Jakarta (37 cases/ 12.1%) representing the provinces with highest number of pneumonia of unknown cause cases followed by Central Java (26 cases/ 8.5%), West Java (17 cases/ 5.6%), Yogyakarta (15 cases/ 4.9%), and Banten (7 cases/ 2.3%). Outside Java, Bali reported the highest number of cases (28 cases/ 9.2%).

Discussion

This study provides a descriptive analysis of pneumonia of unknown cause that occurred in Indonesia between November 2019 to March 2020. A rise in reports of pneumonia of unknown cause is identified in Indonesia starting from late January 2020, which may reflect the presence of COVID-19 cases in the country prior to official identification of two cases at the start of March 2020. It is also most likely that COVID-19 cases were imported to Indonesia prior to March 2nd.

Data from the Ministry of Health ¹⁰ shows that by March 31st 2020, Indonesia has confirmed 1,528 positive cases of COVID-19, with 136 documented deaths. Among these cases, almost half (48.9%) were identified in Jakarta, followed by other provinces in Java, such as West Java (11,9%), Banten (9.2%), East Java (5.9%), Central Java (5.2%) and Yogyakarta (1.2%) ¹⁰. Meanwhile, Bali only reported 19 confirmed cases of COVID-19 (1,2% of total cases) ¹⁰.

The spatial distribution of COVID-19 cases in Indonesia is similar to the distributional pattern of pneumonia of unknown cause in our observation, which may suggest the potential use of pneumonia of unknown cause as a proxy for COVID-19 cases in Indonesia. However, Jakarta

reported proportionately higher cases when compared with the spatial distribution of pneumonia of unknown cause in our study. This might suggest the geographical difference in the ability to identify and report COVID-19 cases.

Provinces in Java Island, especially Jakarta, the capital city, might have better ability in identifying cases of COVID-19. Meanwhile, other provinces, such as Bali, which reported proportionately lower confirmed COVID-19 cases when compared with the reported pneumonia of unknown cause, may have under-detection of COVID-19 cases. It is known that from the twelve COVID-19 national reference laboratories in Indonesia, only three are located outside Java Island, which may hinder the ability of provinces outside Java to rapidly identify and response to COVID-19 presence ¹¹.

This study is the first of its kind to describe the epidemiological pattern of pneumonia of unknown cause before and after the official notification of COVID-19 cases in Indonesia. It shows that monitoring trends of pneumonia of unknown cause in open-source data could provide rapid, unvalidated epidemic intelligence for early detection of COVID-19 outbreaks in Indonesia.

Several limitations, however, exist in this study. Firstly, this study relies on an online news-based surveillance system, which is unvalidated and may include other aetiologies of pneumonia ¹². Secondly, we only included pneumonia in our study; however, COVID-19 has a wide spectrum of disease, with most having a mild illness ¹³. Despite that, this study has highlighted the ability of open source data to identify early alerts before the initial confirmation of COVID-19 cases in Indonesia which makes it a promising option to enhance epidemic surveillance system.

Conclusion

In this study, we used an online news-based surveillance system using Bahasa Indonesia through *EpiWATCH* to indicate early signals of COVID-19 outbreak in Indonesia. We draw three conclusions based on our observation. Firstly, we observed the earliest signals of pneumonia of unknown cause in Indonesia in late January 2020, indicating possible unrecognised circulation of COVID-19 cases in Indonesia prior to the country's official notification of cases. Secondly, we observe a relatively similar spatial pattern of cases between our observed pneumonia of unknown cause and officially confirmed COVID-19 cases in Indonesia. Lastly, monitoring trends in open-source data could provide rapid, unvalidated epidemic intelligence for early detection of outbreaks, especially where surveillance, testing, and reporting are limited. Such signals can be used to initiate validated disease surveillance as a supplement to the traditional clinic and laboratory-based infectious or to trigger investigation and testing.

REFERENCES

1. WHO. Coronavirus disease (COVID-19) outbreak [Internet]. 2020 [cited 2020 Apr 1]. Available from: <https://www.who.int/westernpacific/emergencies/covid-19>
2. Ministry of Health Singapore. Confirmed imported case of novel coronavirus infection in Singapore; multi-ministry taskforce ramps up precautionary measures. [Internet]. 2020 [cited 2020 Apr 1]. Available from: <https://www.moh.gov.sg/news-highlights/details/confirmed-imported-case-of-novel-coronavirus-infection-in-singapore-multi-ministry-taskforce-ramps-up-precautionary-measures>
3. Director-General of Health Malaysia. Kenyataan Akhbar KPK 26 Januari 2020 – Situasi Terkini Jangkitan Novel Coronavirus di Malaysia – (Press statement 26 January 2020- Current update on Novel Coronavirus Situation in Malaysia) [Internet]. 2020 [cited 2020 Apr 1]. Available from: <https://kpkkesihatan.com/2020/01/26/kenyataan-akhbar-kpk-26->

januari-2020-situasi-terkini-jangkitan-novel-coronavirus-di-malaysia/

4. WHO. WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020 [Internet]. 2020 [cited 2020 Apr 1]. Available from: <https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020>
5. Gorbiano MI. BREAKING: Jokowi announces Indonesia's first two confirmed COVID-19 cases - National - The Jakarta Post. Jakarta Post [Internet]. Jakarta; 2020 Mar 2 [cited 2020 Apr 1]; Available from: <https://www.thejakartapost.com/news/2020/03/02/breaking-jokowi-announces-indonesias-first-two-confirmed-covid-19-cases.html>
6. Salazar PM De, Niehus R, Taylor A, Buckee CO, Lipsitch M. Identifying Locations with Possible Undetected Imported Severe Acute Respiratory Syndrome Coronavirus 2 Cases by Using Importation Predictions. *Emerg Infect Dis J* [Internet]. 2020;26(7). Available from: https://wwwnc.cdc.gov/eid/article/26/7/20-0250_article
7. Collier N. Uncovering text mining: a survey of current work on web-based epidemic intelligence. *Glob Public Health* [Internet]. 2012/07/11. Taylor & Francis; 2012;7(7):731–749. Available from: <https://pubmed.ncbi.nlm.nih.gov/22783909>
8. WHO. Epidemic intelligence - systematic event detection [Internet]. 2016 [cited 2020 Apr 3]. Available from: <https://www.who.int/csr/alertresponse/epidemicintelligence/en/>
9. ISER. Epiwatch | NHMRC Centre for Research Excellence, Integrated Systems for Epidemic Response [Internet]. [cited 2020 Apr 3]. Available from: <https://iser.med.unsw.edu.au/epiwatch>
10. Ministry of Health Indonesia. COVID-19 update until April 1st 2020, 09.00 GMT+7 [Internet]. 2020 [cited 2020 Apr 1]. Available from: <https://infeksiemerging.kemkes.go.id/category/situasi-infeksi-emerging/#.XoQ3x4gzZyw>

11. Ministry of Health. Pemerintah Perluas Lokasi Pemeriksaan Covid-19, Berikut Daftarnya [Government decided to expand the location for COVID-19 testing, here is the list] [Internet]. 2020 [cited 2020 Apr 4]. Available from:
<http://sehatnegeriku.kemkes.go.id/baca/rilis-media/20200320/3333441/pemerintah-perluas-lokasi-pemeriksaan-covid-19-berikut-daftarnya/>
12. Aiello AE, Renson A, Zivich PN. Social Media– and Internet-Based Disease Surveillance for Public Health. *Annu Rev Public Health* [Internet]. Annual Reviews; 2020 Apr 2;41(1):101–118. Available from: <https://doi.org/10.1146/annurev-publhealth-040119-094402>
13. WHO. Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19) [Internet]. Geneva; 2020 Feb. Available from: <https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf>

FIGURES

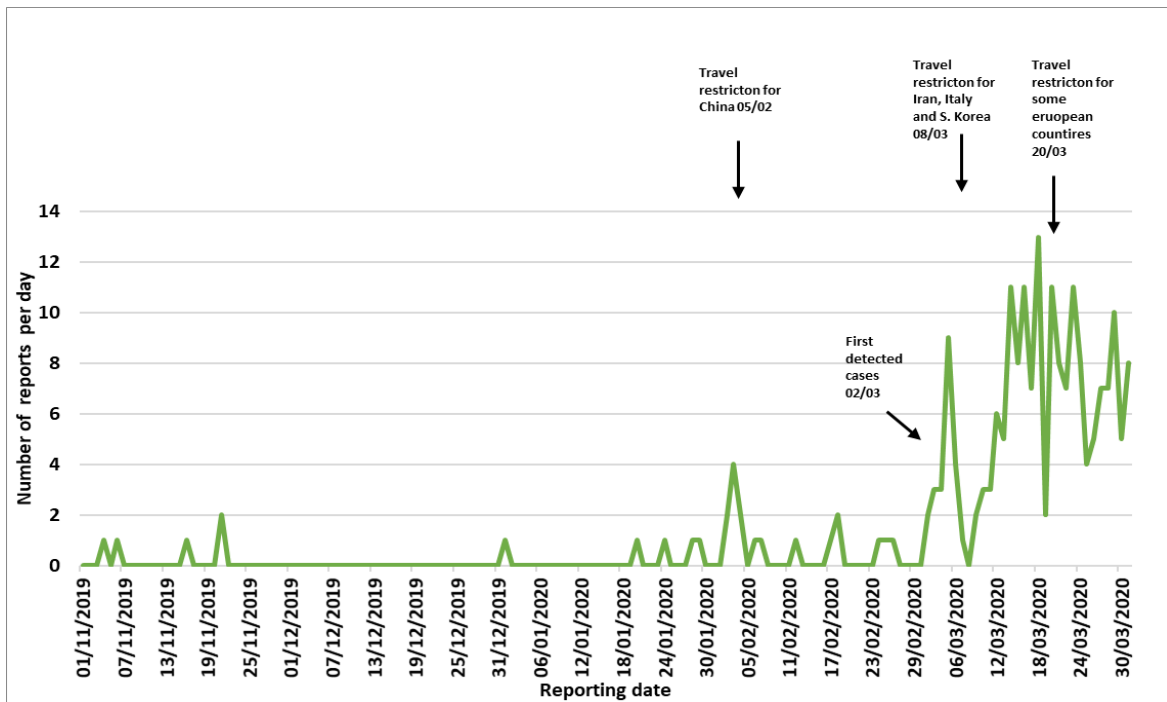


Figure 1: Number of Reports of pneumonia of unknown cause in Indonesia between 1 November 2019 and 31 March 2020.

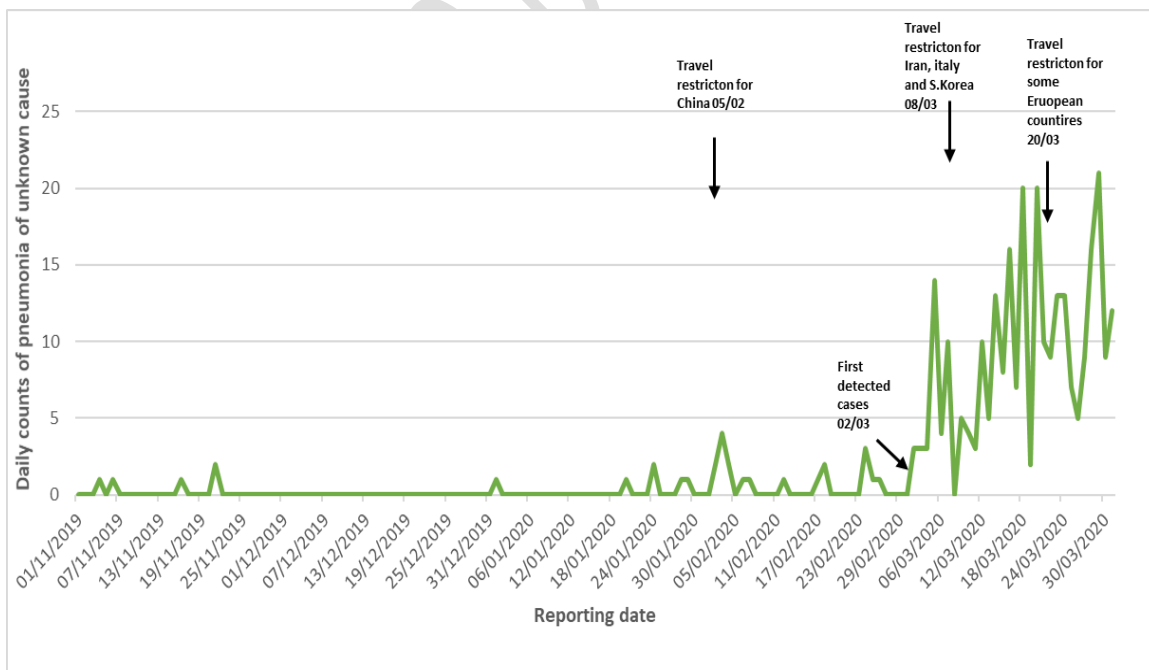


Figure 2. Number of Cases reported by *EpiWATCH* between 1 November 2019 and 31 March 2020.