COVID-19

RESOLVE TO SAVE LIVES | Vital Strategies

Guide for Analysis of Respiratory Syndromic Surveillance Data

TO SUPPLEMENT COVID-19 SURVEILLANCE

1 Identify potential sources of data.

2 Identify data that will be analyzed.

3 Extract data and conduct analysis.

4 Summarize and communicate the findings.

Visit PreventEpidemics.org for more.
Prevent Epidemics is a project of Resolve to Save Lives, an initiative of Vital Strategies.

MAY 2020

Summary

Existing data on respiratory diseases from surveillance and routine health information systems can be leveraged to identify changes in patterns of disease that could reflect the emergence of COVID-19. Data collected through daily or weekly surveillance activities can be useful where there is inadequate testing for COVID-19 or in areas that have yet to identify COVID-19 cases. These data can also provide early indications of a reemergence of COVID-19 after changes in public health and social measures.

This document provides practical guidance on using existing data from respiratory disease surveillance systems and hospital-based routine data systems to supplement COVID-19 surveillance and inform response activities. It provides step-by-step guidance on how to analyze existing data for early detection of COVID-19, and includes tools to support data analysis and communication of results. Such analysis does not replace COVID-19 case surveillance; it is a supplement to assist in monitoring the spread of COVID-19.

This guide includes practical steps and tools for:

- Identifying potential sources of data
- Identifying relevant data to conduct analysis
- Extracting the data and conducting analysis
    - Analyzing surveillance data
    - Analyzing routinely collected health services data
    - Calculating a baseline
    - Identifying triggers for action
- Summarizing and presenting the findings for decision-making

This guide is for government health officials responsible for COVID-19 surveillance.

Table of Contents

Overview

Surveillance data can signal community transmission of COVID-19 when even just a few unlinked cases have been identified, especially when symptoms are severe at the time they are diagnosed and reported. Regularly collected data from surveillance systems and hospitals can supplement COVID-19 surveillance and inform response activities. This document provides practical guidance on how existing data from systems to monitor respiratory diseases can be analyzed for early detection of COVID-19, and includes practical tools for data analysis and communication of results.

The analysis proposed in this guide is not intended to replace COVID-19 case surveillance, but rather to serve as a supplement and assist in early detection and monitoring of the emergence, reemergence, or spread of COVID-19.

Introduction

Syndromic surveillance collects, analyzes and interprets routine health-related data on a combination of symptoms and clinical signs (syndromes).[1] As a result, syndromic data can serve as an early warning system to detect unusual levels of illness before a diagnosis is confirmed clinically or with laboratory testing. Health indicators from existing syndromic surveillance systems such as influenza-like illness (ILI) and severe acute respiratory infections (SARI) can be adapted to signal a COVID-19 outbreak.[2,3] Clusters of illness detected in surveillance systems can signal changes in patterns of disease when there is inadequate testing for COVID-19, in areas where cases have not yet been detected, and in areas where existing public health and social measures have been tightened or loosened. This can be an important indicator that there is introduction or reemergence of COVID-19 in a community.

Trend analysis of data collected through daily or weekly surveillance activities is most useful, as an increase in cases can serve as an early warning system to promote action and further investigation. Ideally it forms part of a multi-pronged approach for the ongoing surveillance, response and monitoring of the impact of COVID-19. Data sources for this approach include: 1) laboratory data on confirmed cases and deaths; 2) data on probable and suspected cases and deaths; 3) respiratory disease data from surveillance systems and routine health services reporting; and 4) rapid mortality surveillance data *(Figure 1)*.

This document provides guidance on the third component—how existing respiratory disease data from surveillance systems and routine health services can be used to produce analysis in support of national and regional monitoring, risk assessment, and data-driven response actions for COVID-19.

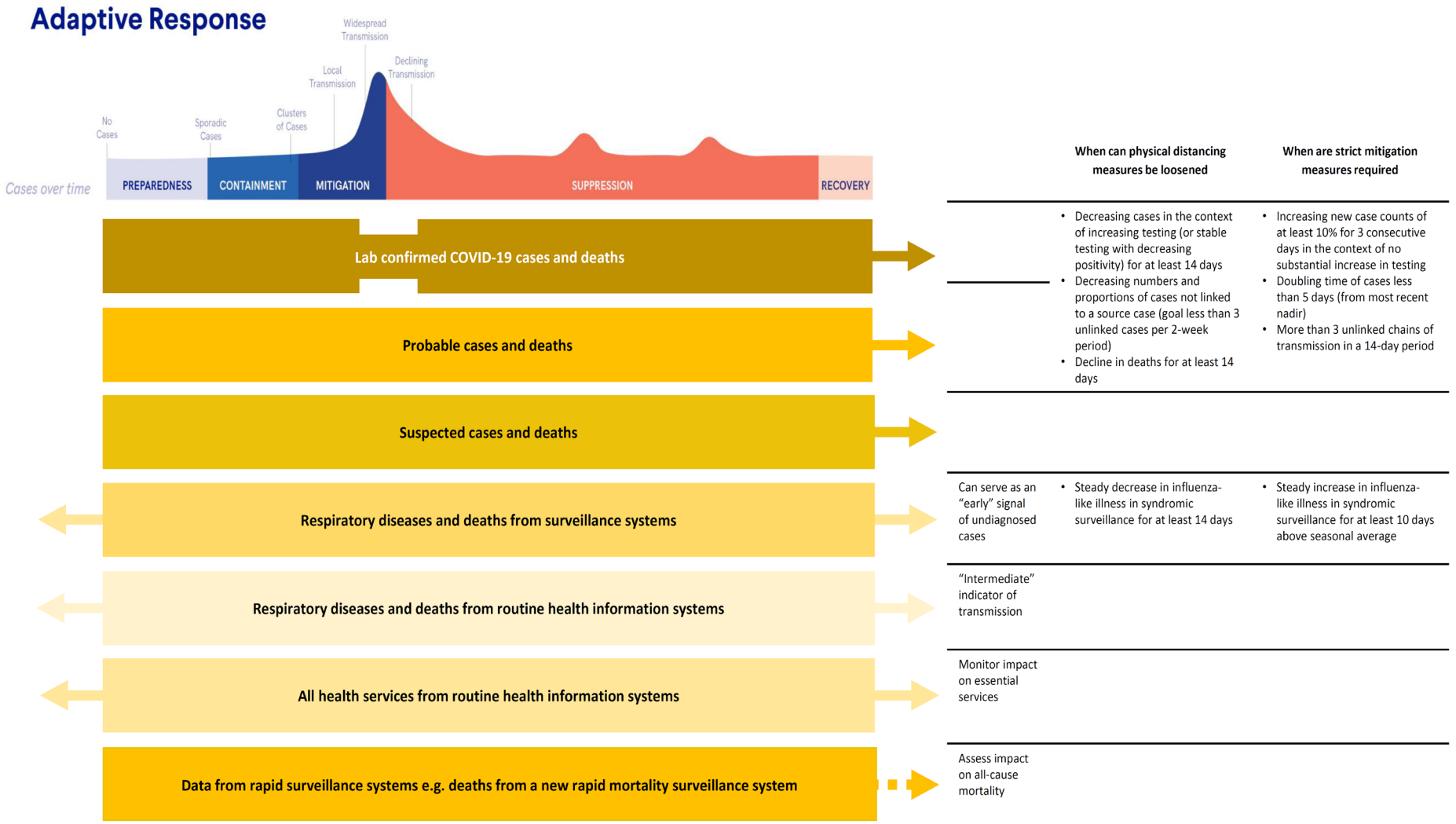

Figure 1. Data sources for surveillance, response and monitoring the impact of COVID-19 in the context of the adaptive response and tightening/loosing of measures[4]

Considerations

- **Ongoing surveillance systems vary by country.**
  Surveillance networks exist in most countries. These systems monitor specific diseases (e.g. malaria cases) and syndromes such as influenza-like illness (ILI) and severe acute respiratory infections (SARI), which are most often defined as a combination of signs and symptoms.[5] For example, the technical guidelines for Integrated Disease Surveillance and Response (IDSR) in the African region defines ILI as a sudden onset of fever (›38°C) and a cough or sore throat in a child or an adult, in the absence of other diagnoses.[6]

- **Generally, there is more than one coexisting surveillance strategy.**
  For instance, these may include passive surveillance through national health systems as well as sentinel surveillance for specific conditions. Sentinel sites for flu surveillance typically combine laboratory testing for flu subtypes in addition to ILI surveillance.

- **A range of health management information systems (HMIS) may be used to record surveillance information.**
  Even among countries that use the same HMIS, such as DHIS2, there are likely to be differences in the data available depending on the country-specific reporting arrangements. Further, different parts of the surveillance data may be hosted on different platforms.

- **Effective communication of the results of surveillance to decision-makers is essential.**
  To serve as an early warning system, the results and interpretation of the results should be presented in a written brief, with clear data visualizations and a standard routine format. Ministry staff can produce these briefs or presentations, highlighting patterns consistent with the presence of COVID-19, and ensuring they are easily understandable. Briefs should be produced regularly, on at least a weekly basis. Where feasible and where data are available, daily monitoring is recommended. In both instances, the data should be considered in the context of overall COVID-19 related surveillance.[7] It is expected that ILI, SARI and other syndromic data provide an early signal for COVID-19, followed by confirmed case counts, hospitalizations and deaths. The use of these data will enable rapid detection of increases in illnesses that may include COVID-19 as an underlying cause, which in turn should inform appropriate public health interventions.

- **SARI/ILI early warning surveillance should not detract from analysis of routine health data on other conditions such as immunizations and maternal and child health, so as to ensure continued delivery of essential health services.**
  Analysis of the routine health services data can also help quantify the longer term and indirect impact of COVID-19.[8] Previous research from the Ebola outbreaks in Africa has shown that there could be a decline of up to 50% in service utilization during the outbreak and that utilization of services after the outbreak does not always recover to pre-outbreak levels.[9] Women are especially vulnerable to these declines in service utilization.[10]

Using existing data on morbidity and mortality to supplement COVID-19 surveillance: a step-by-step guide

1 Identify potential sources of data.

Identify relevant data sources and systems in your country, such as the IDSR and ILI surveillance systems. In some instances, multiple sources may be required to obtain a complete picture.

Appendix A : Surveillance data inventory and metadata template

2 Identify data that will be analyzed.

Identify key data to be monitored, such as respiratory illness or syndrome. This may include data on:

- Influenza-like illness (ILI);
- acute respiratory infections or Severe Acute Respiratory Infection (SARI);
- severe acute respiratory syndrome; or
- pneumonia.

Depending on the data collected in each health ministry, authorities can include other health conditions deemed relevant.

Identify surveillance data on cases and deaths for the conditions listed. Data on outpatient (OPD) and hospitalized inpatient (IP) cases and deaths provide valuable additional information and can be used to validate syndromic data, but may only be available through routine health service data systems. In addition, the number of OPD visits and total admissions can provide

denominator data for estimating rates as the outpatient visits and admissions change over time (e.g. SARI cases per 10,000 outpatients) .

The surveillance data inventory and metadata template (Appendix A) can be used to record the sources of numerators and denominators.

Appendix A : Surveillance data inventory and metadata template

3 Extract data and conduct analysis.

It may be feasible to conduct some of the data analysis within the actual source data system. For example, pivot tables and visualization applications are available within DHIS2. However, consider downloading the data to allow more detailed analysis. An Excel sheet is provided in Appendix B with suggested analysis and templates for graphs. The analysis can be adapted depending on available data. For example, age categories can be changed if age breakdown is not available.

Appendix B: Excel sheet templates for data and visualizations

ANALYSIS OF SURVEILLANCE DATA

The analysis of surveillance data should include trends in new cases and deaths, by day or by week, compared to several years of prior data. This provides a baseline comparison to be able to identify increases in the current trend over the expected number based on prior years. Take note of any changes in case definitions and the number of facilities reporting, or changes in the utilization of health services during this time period that could provide an alternate explanation for changes in trends compared to prior years and months. Where feasible, stratify the analysis by sex, age, region and/or city.

ANALYSIS OF ROUTINE DATA

If available, additional data such as outpatient (OPD) cases, Inpatient (IP) cases, total OPD visits and total IP admissions, can be used to conduct secondary analyses. These can include:

- Number of cases or deaths due to specific respiratory illnesses or syndromes
- Rate of OPD cases or deaths due to specific respiratory illnesses or syndromes per 10,0000 OPD visits (more commonly used for ILI)
- Rate of IP cases or deaths due to specific respiratory illnesses or syndromes per 10,000 IP admissions (more commonly used for SARI)

- Rate of cases or deaths due to specific respiratory illnesses or syndromes per 10,000 patients (sum of OPD visits and IP admissions)

If denominator data is unavailable, calculate proportional morbidity, by comparing the percentage of cases due to a specific respiratory illness or syndrome with other illnesses or syndromes that are monitored. This disease-based data should be compared to syndromic data for validation or to identify discrepancies.

DETERMINING THE BASELINE

The number or proportion of cases or deaths in previous years can be used to calculate the average level of activity in a surveillance system; this is referred to as the expected number or baseline.[11] The World Health Organization (WHO)'s "Global Epidemiological Surveillance Standards for Influenza" provides guidance in its Appendix 8 on the simplest way of producing a baseline. This is done by creating an average curve centered on the median week of peak transmission for several years.[11] This analysis can be done at the national level, but would be most useful at the geographic level of observation such as subnational or district level. Guidance on this calculation is provided here in Appendix C.

For routine health data the average for every month can be calculated using the previous years' data. It is ideal to have three to five years worth of data, but if not available even one year of prior data can be used. This information can then be used to set triggers for action. Historical facility data is also useful for comparison to observed data in rapid surveillance systems established during the COVID-19 response, as can be seen from its use in facility-based rapid mortality surveillance.

Appendix C: Excel sheet for calculating the baseline and alert threshold

TRIGGERS FOR ACTION

Analysis of syndromic surveillance data should be used as one of the indicators to inform the implementation and timing of mitigation measures. For example, a steady increase in ILI for at least 10 days above the seasonal average, or baseline, can indicate that mitigation measures are needed, whereas a steady decrease in ILI for at least 14 days can serve as one of the indicators that physical distancing measures can be loosened.[12,13]

Disaggregating data by demographic groups (e.g. children, elderly, etc.) and geographic areas (e.g. districts) can help to identify particular groups at risk. Statistical measures of clustering can also be applied if the available data allows for this analysis.

For analysis using routine data, thresholds can be set using the upper limit of either 90% or 95% confidence interval of the average for every month. Countries can decide on the limit depending on how conservative they wish to be in their identification of potential cases. There is merit in keeping a low threshold to serve as an early warning signal, which could then trigger further investigations.

WHO standards provide a simple statistical calculation of variance to define an alert threshold to detect an unusual pattern of illness.[11] Detailed steps on this calculation are provided in Appendix C. The baseline data are also used to define seasonal thresholds indicating the start and end of a seasonal increase in disease burden, such as influenza season, which requires more complex methods for calculation, such as the moving epidemic method.[11] These approaches are useful for illustrating what seasonal data (e.g. for ILI) look like and how seasonal trends and year-to-year variation in timing and size can influence baseline data. However, these approaches may be less useful for guiding alerts for an individual country or city in relation to COVID-19.

Appendix C: Excel sheet for calculating the baseline and alert threshold

4 Summarize and communicate the findings.

The results of the analysis should be compiled in an easy-to-understand format for decision-makers. This can be in the form of a brief report or presentation.

A data brief is an essential tool for communicating the data to a decision-maker. It should be concise and clearly convey the key findings and recommendations. A template for a weekly brief is available in Appendix D. It can be adapted to the country's needs based on the available data and analysis. The brief should provide the needed information in a format that is easy to understand, interpret and apply, and that is consistent from week to week. Seek feedback from stakeholders to ensure it meets their needs and expectations. Agree on the frequency of updates. Consider the quality, timeliness, collection frequency and completeness of the data sources when interpreting the data. The results could also be presented as a weekly slide deck for the emergency operations center or displayed in the emergency operations center.

The results of the data analysis can also be used as an opportunity to improve the quality of the data by reinforcing and ensuring that facilities understand case definitions and laboratory testing procedures. This includes the ongoing training of epidemiology teams and health care workers on reporting procedures, and job aids to remind facilities to report. When facilities do not report or report late, there should be follow-up to troubleshoot and provide additional support, as needed.

Appendix D: Word template for weekly brief

Appendices and additional resources

Tools to support analysis and reporting:

Appendix A : Surveillance data inventory and metadata template

Appendix B: Excel sheet templates for data and visualizations

Appendix C: Excel sheet for calculating the baseline and alert threshold

Appendix D: Word template for weekly brief

This guide does not address the establishment of new programs for rapid surveillance of mortality or the establishment of community-based surveillance, nor does it provide advice on response and mitigation measures. For guidance on these measures, the following resources are recommended:

1. Rapid Mortality Surveillance in Low-Resource Settings[14]
2. COVID-19 Playbook, Adaptive Response[4]

References

1 Smith GE, Elliot AJ, Lake I, Edeghere O, Morbey R, Catchpole M, et al. Syndromic surveillance: two decadesexpe ence of sustainable systems - it's people not just data! Epidemiol Infect 2019;147:e101.

2 Prevent Epidemics. COVID-19 Weekly Science Review March 28–April 3, 2020. [Internet]. [cited 2020 Apr 10]; Available from: https://preventepidemics.org/coronavirus/science-review/march-28-april-3-2020/.

3 World Health Organization. Operational considerations for COVID-19 surveillance using GISRS. Geneva: World Health Organization; 2020.

4 Prevent Epidemics. COVID-19 Playbook [Internet]. [cited 2020 Apr 21]; Available from: https://preventepidemics.org/coronavirus/playbook/

5 Henning KJ. What is syndromic surveillance? MMWR Suppl 2004;53:5–11.

6 World Health Organization, Centers for Disease Control and Prevention. Technical Guidelines for Integrated Disease Surveillance and Response in the African Region. Brazzaville, Republic of Congo and Atlanta, USA: 2010.

7 Centers for Disease Control and Prevention. COVID View Week 15, ending April 11, 2020 [Internet]. Centers for Disease Control and Prevention; 2020. Available from: https://www.cdc.gov/coronavirus/2019-ncov/covid-data/pdf/covidview-04-17-2020.pdf

8 McKay G, Black B, Mbambu Kahamba S, Wheeler E, Mearns S, Janvrin A. Not All That Bleeds Is Ebola: How has the DRC Ebola outbreak impacted Sexual and Reproductive Health in North-Kivu? New York, USA: The International Rescue Committee; 2019.

9 Camara BS, Delamou A, Diro E, Béavogui AH, El Ayadi AM, Sidibé S, et al. Effect of the 2014/2015 Ebola outbreak on reproductive health services in a rural district of Guinea: an ecological study. Trans R Soc Trop Med Hyg 2017;111(1):22–9.

10 Chattu VK, Yaya S. Emerging infectious diseases and outbreaks: implications for women's reproductive health and rights in resource-poor settings. Reprod Health 2020;17(1):43.

11 World Health Organization. Global epidemiological surveillance standards for influenza. [Internet]. Geneva: World Health Organization; 2013 [cited 2020 Apr 13]. Available from: https://www.who.int/influenza/resources/documents/WHO_Epidemiological_Influenza_Surveillance_Standards_2014.pdf

12 Resolve to Save Lives. When and how to reopen after COVID 19 [Internet]. Vital Strategies; 2020 [cited 2020 Apr 20]. Available from: https://preventepidemics.org/wp-content/uploads/2020/04/COV020_WhenHowLoosenFaucet_v4.pdf

13 Resolve to Save Lives. When and how to close due to COVID-19 spread [Internet]. Vital Strategies; 2020. Available from: https://preventepidemics.org/wp-content/uploads/2020/04/COV020_WhenHowTightenFaucet_v3.pdf

14 Vital Strategies, et al (2020). Rapid Mortality Surveillance in Low-Resource Settings: Data for epidemic awareness. New York: Vital Strategies