Rapid Review of WHO COVID-19 Surveillance

An external review conducted by Resolve to Save Lives, an Initiative of Vital Strategies, at the request of the World Health Organization

Final Draft, with input from WHO Headquarters and Regional Offices: 27 October 2021
# Table of Contents

Executive Summary ........................................................................................................................................... 3  
Introduction and Scope of the Rapid Review ............................................................................................... 9  
Evolution of COVID-19 Surveillance by WHO ............................................................................................. 10  
   Figure 1: Timeline of WHO Surveillance Guidance and Systems Adaptations ................................ 11  
Data Use ...................................................................................................................................................... 16  
   Headquarters Perspectives ..................................................................................................................... 16  
   Regional Perspectives ............................................................................................................................. 17  
   Country Perspectives .............................................................................................................................. 17  
Missed Opportunities and Lessons Learned ............................................................................................. 18  
Reporting and Data Pipeline ....................................................................................................................... 22  
   Sources and Flows of Data, by Region .................................................................................................... 23  
   COVID-19 Platforms and Data Interoperability .................................................................................. 24  
   WHO HQ Harmonization and Output Process .................................................................................... 24  
      Figure 2: WHO Headquarters Transformation and Harmonization of Data ................................... 25  
Conclusions and Recommendations: Adapting for Scale and a Digital Future ........................................... 26  
   Summary of Key Recommendations Based on the COVID-19 Experience ...................................... 29
Executive Summary

The use of accurate, real-time data to inform decision-making is essential for infectious disease control. As countries responded to COVID-19, decision-makers needed ways to gather, manage, process, and share COVID-19 data quickly and accurately. In an effort to have an impartial view of its COVID-19 surveillance program, the World Health Organization (WHO) requested the support of Resolve to Save Lives, an Initiative of Vital Strategies, to conduct a rapid review of WHO surveillance activities, during a five-week period in March–April 2021. This review identifies technical and operational recommendations to WHO to strengthen ongoing surveillance and data systems, while anticipating the need for scalability for larger future events and the realities of real-time surveillance in the digital era.

The review draws on key informant interviews with more than 30 leaders and senior managers from WHO headquarters, regional offices and country offices in the South-East Asia region, African region, European region, Eastern Mediterranean region, and Western Pacific region; regional centers for disease control and prevention in Europe and Africa (i.e., the European Centre for Disease Prevention and Control [ECDC] and Africa Centres for Disease Control and Prevention [ACDC]); and senior managers of other COVID-19 data aggregation tools and dashboards curated and created by partners. We have not identified key informants in this review to foster candor and protect privacy. In consultation with WHO, a deep analysis of mortality surveillance, field level data collection, and systems for International Health Regulations (IHR [2005]) reporting were not conducted for this rapid review, although these topics warrant further analysis.

In this rapid review, we describe the evolution of WHO’s COVID-19 surveillance program, the use of data at all levels of the organization, and the data pipeline and reporting systems from regional offices to WHO headquarters..

Key Findings

The need to quickly characterize a new and rapidly spreading pathogen, the geographic scope of its transmission, and the volume of data to collect, store, and process, strained resources of Member States, country and regional offices, and headquarters staff. To manage the changing demands of the pandemic, WHO evolved its global guidance, case definitions, and data systems reactively. The COVID-19 pandemic required an unprecedented volume of data to be collected by the organization, including the collation of >50 million case-based records and leveraging non-traditional sources of data. These requirements placed strains on human resources and data systems at all levels.

Informants highlighted several successes of the organization’s surveillance efforts, including strong collaboration between WHO headquarters with regional offices, and regional and country offices with Member States, and the adoption of an existing data management system used for global influenza surveillance to meet the needs of the COVID-19 pandemic. Specific technical support provided by the regional offices, including harmonization of regional data systems and regional modeling, were also highlighted as successes. In addition, many country offices provided valued direct technical assistance to support Member State COVID-19 surveillance activities, although many of these outputs were not designed for public consumption and thus went unseen and unrecognized by the general public. Informants also highlighted multiple points of failure in the systematic collection, transmission and storage of data, and many missed opportunities for the timely and effective use of data which both hindered the credibility of the organization and introduced barriers for Member State reporting.
**Success areas for emulation:**

1. **Several regions were able to rapidly adjust data systems, building on existing capacity of those regions and their Member States:** Such examples included scaling of existing data sharing platforms like TESSy in the European region, the EMFLU system (for influenza surveillance) in the Eastern Mediterranean region, and DHIS2 in the Eastern Mediterranean, African, South-East Asia, and Western Pacific regions. Application programming interfaces (API) further streamlined information exchange.

2. **WHO headquarters was collaborative and communicative:** WHO regional offices appreciated the level of consultation conducted by headquarters as guidance and information sharing requirements evolved, and the high quality of the technical guidance produced. However, some did view that the extent of consultation and the broad array of stakeholder interests created bottlenecks.

3. **Technical partnerships increased the value of collected data:** The COVID-19 Modeling Consortium (CoMo) collaboration in the Eastern Mediterranean region, modeling the impact of public health and social measures (PHSMs), allowed Member States to see their data applied effectively by the regional office and incentivized data collection. Similarly, WHO partnered with several organizations to synthesize and publish the implementation of public health and social measures. However, other opportunities for formal partnerships with external partners on the collation, curation, and use of data were missed, including partners with more advanced technical capacities for data scraping and curation from publicly available sources.

4. **Investments in event-based surveillance (EBS) platforms were effectively leveraged:** Many regions and some Member States reported the effective use of EBS systems, particularly Epidemic Intelligence from Open Sources (EIOS) not only for early situational awareness and monitoring but also to facilitate identifying sources of open data to update case and death counts from existing government sources. Several informants reported on the need for WHO to implement its ability, under Article 11 of the International Health Regulations (IHR [2005]), to share data from unofficial sources publicly, even when unverified, given the COVID-19 pandemic met the requirements to do so under Article 11.

5. **Strong relationships and support to Member States before and during the COVID-19 pandemic:** The existing strong relationship between country and regional offices with Member States’ ministries of health (MOH) or national public health institutes (NPHIs) allowed for data to be shared from the majority of Member States during the pandemic. Technical assistance for implementation of guidance by WHO was appreciated, although there were human resources limitations in providing more sophisticated technical assistance such as data management, mathematical modeling, and dashboard development to Member States.

6. **Dedicated and professional personnel across all levels:** WHO staff were available and willing to go above and beyond their roles and provide technical assistance and support in very challenging and changing conditions.

**Challenge areas:**

1. **The lack of a pre-existing data architecture and data use and sharing plan with Member States hindered efforts to collect and disseminate timely and useful information:** Many informants
shared challenges with Member State (supply-side) reporting of data. Although the lack of enforcement mechanisms for Member State reporting of data under the IHR (2005) was a common theme among key informant interviews, there were many other barriers to reporting, including fear of stigmatization early in the pandemic, lack of national human resource and technical capacities, and privacy regulations. In addition to the reticence of some Member States to share data with WHO, the failure of WHO to clearly establish the conditions under which WHO would disseminate collected data significantly delayed the release of public dashboards. The self-nomination of transmission categories by Member States hindered the implementation of an evolving case definition, which was based on transmission categories during the early phase of the pandemic. Informants described the utility of establishing a data use and sharing plan with Member States before the next large event in order to facilitate reporting, collect parsimonious data, and use the data effectively to inform response decision-making.

2. **Previous lessons learned from large events were not implemented, and the WHO data architecture was not scalable to the needs of the COVID-19 pandemic**, although these weaknesses were known from experiences from the 2014–2016 Ebola virus disease epidemic and the 2009 H1N1 pandemic.

3. **Technical and operational barriers compounded political barriers to data sharing**: These barriers included the requirement for Member States to share case report data on a large number of variables that were never effectively applied, and the conflation of “front end” data entry systems with a “back end” database, resulting in inefficient data processing. Resources for the collection of detailed case-based data by the organization should have been shifted more rapidly to defining minimum data standards and data fields to Member States to publish their own data in a format to support public communications and situational awareness.

4. **The lack of compelling strategic information outputs by WHO did not incentivize reporting**: These outputs could include a clear alert level system, contextualizing data to inform the implementation or assess impact of public health and social measures, and more timely release of implementation guidance for contact tracing and case investigation. Combined with the lack of provision of timely epidemiologic data when compared to external sources, these omissions undermined the perception of WHO as the primary source of real-time COVID-19 data. The data warehousing function of WHO as the official source of global data was seen as valuable for advocacy and response option planning for Member States, but broader use of WHO data was hampered by the availability of more timely datasets.

5. **WHO did not have or effectively utilize sufficient in-house capacity for data management or data science, and its internal human resources systems created substantial bottlenecks for onboarding new staff or consultants**: Almost all informants reported that there were limited staff who were mobilized internally and deployed effectively, and new staff and consultants were often junior and lacked familiarity with WHO structures, requiring significant onboarding to manage daily tasks. These bottlenecks left the organization understaffed and at times reliant on personnel deployed from partner organizations with high turnover rates and with minimal information on the function of WHO. Personnel shortages were experienced at all levels of the organization, which hindered the ability of WHO to provide technical assistance to Member States.
6. **Shifting guidance and requirements by the organization made it challenging to keep Member States engaged:** Although WHO evolved its data systems to accommodate the scale of the pandemic, including eventually upgrading its data infrastructure and collecting weekly aggregate data, these changes were challenging at times for regional offices and Member States to adapt to, emphasizing the need for stronger anticipatory planning for pandemic-prone events with scalable reporting and data systems. The large number of variables requested for case-based reporting resulted in substantial incompleteness for key data on testing rates and hospitalizations, and the number of stratifications requested for age and sex under aggregate reporting increased Member State data processing workload when line lists were available.

**Recommendations**

The COVID-19 pandemic and lessons learned provide an opportunity to identify investments in the organization to support future small and large events. Our recommendations address the key themes gathered from key informant interviews, both internal and external to WHO, which included the challenges of a novel pathogen, the challenge of scale, and the timely and effective use of data. These recommendations address systems and operations at WHO that are relevant for pandemics as well as smaller events, and address the need for developing new and scalable data solutions, streamlining and distinguishing between data useful for Member States to collect for themselves versus share with WHO and/or publicly, identifying and building capacity to create strategic information global goods, building human resource capacities and emergency hiring mechanisms, fostering partnerships, and strengthening national surveillance capacities with the support of its regional offices.

<table>
<thead>
<tr>
<th>Strategic Information Planning and Surveillance Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Identify and implement a strategic information plan that builds on existing systems and can inform decision-making across Member States to develop global public goods during major events, including data standards, data sharing, and minimum reportable data by Member States:</strong> WHO should work with Member States and IHR (2005) States Parties to develop tools, forms, and data elements required for assessing risk for new events, and coordinating outbreak response. WHO and Member States should ensure that those tools, forms and data elements are scalable for possible larger future events, with a clear data use and sharing plan in place. The WHO Outbreak Toolkit might form the basis for defining the data use plan and minimum reportable data elements. While national surveillance strategies and tools may vary based on the country context, WHO can provide thought leadership on data standards and elements to support private developers of tools that are being implemented at scale, including DHIS2, CommCare, SORMAS, Go.Data and others. In addition, WHO should provide guidance around minimum data standards for publishing COVID-19 data on country web applications and dashboards. Further, WHO should leverage its normative position to develop guidance for data standards for all public health events, which would allow for Member States to collect data using tools with common data standards that are tailored for those country contexts. In combination with consultation with Member States and support on defining and adapting tools, WHO and other stakeholders should advocate for more timely and transparent sharing of data under the IHR (2005).</td>
</tr>
<tr>
<td><strong>2. Invest in a scalable data solution that allows for relevant parties to submit data to WHO, with data management, storage, and pre-processing capabilities for small and large events:</strong> This “back-end system” should be distinguished from a data entry solution, which will likely vary by region. Enabling access to a “front-end” system for data entry for regional and country offices, as well as</td>
</tr>
</tbody>
</table>
Member States themselves, could reduce headquarters and regional office workload for data collection and entry and increase staff bandwidth for data quality and analysis. Data architecture solutions will also require clear data and metadata standards that will allow for systems operability and leverage application programming interfaces (APIs) as efficient and query-able vehicles of data exchange. This system should be used for routine reporting, and should be easy for countries to access, adopt, and obtain information from. It should incorporate reporting lessons, procedures, best practices, and best platforms from throughout WHO.

3. **Invest in national surveillance strengthening, with technical support from WHO country and regional offices:** Global surveillance will only improve when national surveillance capacities improve, and these investments should be prioritized by all stakeholders, including WHO, Member States, and development partners. WHO country and regional offices were best situated to adapt and ingest data from their Member States. Common data platforms within regions (e.g., TESSy in the European region, DHIS2 in the Eastern Mediterranean region) and integrated disease surveillance strategies (e.g., integrated disease surveillance and response strategy [IDSR] in the WHO African region) were seen as accelerators for adaptation to the COVID-19 pandemic. Catalytic investments include data management solutions at national level, and integration of outbreak response functionality into existing national surveillance systems. WHO’s SCORE technical package could provide a way for Member States to assess capacities for data collection and use and fill identified gaps, making them better prepared for health emergencies. Such efforts will require the involvement and investment of parts of the organization not limited to the Emergencies Programme or the response Incident Management Support Team (IMST). There is also a need to see a shift from WHO asking for data from countries to prioritizing assisting countries to analyze and disseminate their own data.

4. **Accelerate investments to strengthen and expand event based surveillance (EBS):** The Epidemic Intelligence from Open Sources (EIOS) platform was viewed by many stakeholders as helpful for early warning, variant tracking, and data collection. EBS might provide an additional tool to supplement formally reported data to the organization. WHO should establish a public-facing system to share curated event-based data, including the results of media scraping for new event alerts and “scraped” case data. WHO, its partners, and Member States can use unofficial data from EBS to triangulate disparities between official and unofficial data, query Member States, and initiate investigations earlier than may otherwise have been possible through traditional disease reporting systems.

5. **WHO should be empowered to share data from unofficial sources in accordance with provisions in the International Health Regulations:** IHR (2005) Article 11 permits WHO to share data with State Parties after consultation (rather than explicit approval) during a public health emergencies of international concern (PHEIC) or when: international transmission has occurred; there is evidence the control measures against the international spread are unlikely to succeed; the State Party lacks the operational capacity to carry out necessary measures; or the nature and scope of international travel and trade requires the immediate application of control measures. In instances when WHO identifies a high-risk suspected event through EBS and the relevant State Party does not respond to a verification request from WHO, WHO should make the information publicly available, noting the verification status of the event. Similarly, for larger events meeting the provisions of IHR (2005) Article 11, WHO should curate those data and make them publicly available to aid in the timeliness and completeness of its dataset.

Leveraging WHO Leadership Role and Growing Human Resources Talent and Processes
6. Revise human resource processes required for surge staffing during major events: If establishing an emergency human resources system for such events is not possible, additional human resources and finance staff should also be prioritized to accelerate the hiring process. The capacity for surge staffing was highlighted as an enabling factor at ECDC and Johns Hopkins University Center for Systems Science and Engineering.

7. Build institutional capacity for data management and data science through investment in people: Human resources in data management and data science should be prioritized before the next major event to build scalable systems and to deploy during emergency operations. A professional-level career track for data management and data science (with a focus on the effective use of data for response option planning, forecasting, and impact assessment) could incentivize recruitment and retention of talented staff across the organization, and build a pool of deployable staff during emergencies. These technical human resource gaps were noted by all levels of the organization, and best practices and resources from throughout the organization and GOARN partners should be enlisted for this purpose.

8. Leverage WHO’s convening power and international voice to foster advanced technical partnerships: WHO’s convening power is unparalleled in global public health, and its ability to build international consortia consisting of private, academic, non-governmental and/or government organizations should be further leveraged. WHO should define core technical competencies needed for data science and data management internally and establish strategic partnerships before and during emergencies for contingency-based technical needs outside of those technical competencies, which might include advanced modeling techniques. WHO can provide technical partners visibility and credibility, while maintaining its position as a global surveillance leader.

This rapid review was not a comprehensive examination of surveillance activities undertaken by Member States and was unable to assess gaps in field level data collection or systematically gather intelligence from Member States about barriers to data reporting and specific products to be developed by WHO that could provide value for future events. We would additionally recommend a Member State consultation that identifies challenges to reporting, reviews the strategic information products most relevant to government decision makers, and can pave a roadmap for addressing recommendations 1 and 2 above.
Introduction and Scope of the Rapid Review

The World Health Organization (WHO) requested the support of Resolve to Save Lives, an Initiative of Vital Strategies, to conduct a rapid review of WHO surveillance activities. The mutually agreed objective was to conduct an external review of WHO global COVID-19 surveillance approaches, identifying successes and challenges, in order to provide recommendations for improvement.

WHO HQ has identified the following four objectives for its COVID-19 surveillance activities:

1. Monitor trends in COVID-19 at national and global levels;
2. Monitor mortality caused by, and indirectly associated with, COVID-19;
3. Estimate morbidity and mortality for health care workers; and,
4. Assess the impact of control measures.

With a focus on these four objectives, the external team agreed with WHO to, in consultation with WHO staff at all levels: (1) produce a narrative describing how WHO approached COVID-19 surveillance and how this approach evolved over time; (2) describe the data systems used for reporting between all levels from Members States to WHO HQ, including regional variations; and (3) describe how surveillance data was used by various levels of stakeholders (WHO HQ, regional offices, country offices, Member States). Areas agreed to be within the scope of the rapid assessment included case data, the evolution of the case definition, testing data, health care worker infections, implementation of PHSMs, and data outputs including visualizations and situation reports. Several areas were excluded from the scope of the rapid review as they required a degree of Member State consultation not possible under the timeline of the project or went beyond the scope of the surveillance activities numbered above. These areas included IHR notification processes including political barriers, mortality surveillance (excess mortality surveillance, RMS), field level data collection issues experienced by Member States, and infections in schools and other congregate settings.

This rapid assessment included a review of existing technical documentation provided by the WHO team, key informant interviews with WHO staff and Ministry of Health counterparts, interviews with regional centers for disease control and prevention in Africa and Europe, and a technical review and interviews with senior managers of other COVID-19 data aggregation tools and dashboards.

The rapid assessment was conducted during a five-week period in March–April 2021. More than 30 leaders and senior managers from WHO HQ, regional offices and country offices in the South-East Asia region, African region, European region (excluding inputs from the Public Health and Social Measures team, as they are distinct from the Epidemiology Pillar in this region), Eastern Mediterranean region, and Western Pacific region participated in this review. Staff who participated included surveillance staff and incident managers based at HQ and regional offices. In addition, some WHO staff in country offices and Ministry of Health counterparts participated.

A limitation of this rapid review is the small number of interviews we were able to conduct with WHO country office and Ministry of Health staff within the five-week project period. Though we attempted to present the country office and Member State perspective, we acknowledge that additional interviews would have been needed to present a more complete range of country level perspectives and experiences, and we encourage stakeholders to gather Member State input through other processes.
Evolution of COVID-19 Surveillance by WHO

The World Health Organization (WHO) plays a critical role supporting and coordinating international health surveillance and response by working with its 194 Member States. Member States rely on WHO’s dedicated staff and strong technical capacity to direct and coordinate international health, and access to accurately and timely surveillance data are required to fulfill these functions. As COVID-19 emerged and then rapidly proliferated, WHO’s surveillance strategies evolved with the increasing challenges of an expanding pandemic. Rare events are uniquely challenging to prepare for, and the world had not experienced a pandemic of similar severity since the 1918 H1N1 influenza pandemic. In addition, unlike other past global pandemics, there was a global attempt at epidemic containment, with activities including contact tracing and travel restrictions. This led to a need for global reporting of all cases of COVID-19 and data demands that had not been anticipated and were unparalleled in WHO’s history.

There are best practices, limitations and lessons that can be learned from WHO’s approach to COVID-19 surveillance that can be used to both prepare for the next localized outbreak as well as the next pandemic. In the section below, we describe the types of data that were systematically collected by the WHO Incident Management Support Team (IMST) Health Information Pillar, the data systems that were used, and changes to data collection that occurred over time (Figure 1).
**Early Phase** (December 2019–February 2020)

After the Wuhan Municipal Health Commission, China, reported a cluster of cases of pneumonia on 31 December 2019, WHO established the IMST, which included a Health Information Pillar, on 1 January 2020. On 5 January, WHO published a Disease Outbreak News on the new virus, using available surveillance data to conduct an initial risk assessment and provide advice to Member States. In early January, WHO Health Information Pillar began to systematically collect two streams of data: case-based surveillance using a standardized case report form and count data for cases and deaths.

**Case-Based Reporting:** For case-based surveillance, the Health Information Pillar was able to capitalize on the [WHO Outbreak Toolkit](https://www.who.int/influenza/outbreaktoolkit) to develop an initial case report form through a weeklong process, and released the form on 21 January 2020. On 27 February 2020, an updated case report form was released that included additional data fields. The large number of required data fields (with no distinction between required and optional fields) posed a challenge for Member States.

WHO considered the use of two systems that were available for collection of case-based surveillance: xMart and a commercial spreadsheet software package. xMart was a platform used for influenza surveillance (FluMart), and a new instance was created by 21 January 2020 to be used to compile COVID-19 case report data (named nCov xMart). The commercial spreadsheet software package had been used during the 2014–2016 Ebola Virus Disease (EVD) outbreak for line list data. Both systems ended up being used in parallel, though the commercial spreadsheet software package served as the primary data repository because staff found it far easier to clean and modify data compared to xMart.

The large number of cases captured in the commercial spreadsheet software package resulted in systems failure of the platform within weeks of its establishment. One point of failure was that while the commercial software, with its familiar spreadsheet interface, addressed data entry user needs (a “front end” solution), it was not suited for the asynchronous updating of data and the appropriate data management and storage systems required for pre-processing and warehousing a large amount of data. It also lacked the security standards necessary to ensure protection of confidential data.

**Daily Case and Death Counts:** Early in the epidemic, the collection of daily cases and deaths presented several challenges. Firstly, there was not a pre-existing system for disseminated entry of daily case and death counts by national public health institutes (NPHIs), ministries of health (MOH), WHO country offices, or regional offices. To reduce the reporting burden on Member States, WHO accepted daily case and death counts published in any public format, rather than requiring the sending of daily updates directly to WHO. However, this left a small team at headquarters responsible for manually scraping these data from official government sources and entering the count data into a commercial spreadsheet software package. Regional offices collected and entered these data in parallel. The work was challenging due to frequently changing government platforms being used to publicly disseminate data (e.g., Facebook, Twitter, government websites), and limited human resource capacity at headquarters, which contributed to occasional data entry errors. This centralized, duplicative data collection continued until 27 March 2020.

**Launch of nCov xMart System and Weekly Aggregate Data Collection** (February – May 2020)

After the failure of the commercial spreadsheet software package, the surveillance teams switched to using xMart as their primary system for consolidating case report data from Member States. The
surveillance team attempted to facilitate country submission of data, encouraging countries to submit
data in whatever format was easiest including spreadsheet, CSV files, or via direct connection to a SQL
server. While many Member States submitted data, many others did not submit line lists to WHO.
Reasons for not submitting were variable: lack of data availability, poor data completeness that might
reflect poorly on Member State surveillance capacities, concerns around confidentiality of patient-level
data, or political sensitivities around data sharing. Technical barriers around transmission of data to
WHO was infrequently mentioned as the primary reason Member States did not report.

While the xMart system’s strength lies in its ability to merge data from a variety of sources and formats,
is it not a data entry or data management tool. A major limitation was that the system lacked the ability
to efficiently clean or correct data. To make an update, data had to be downloaded, modified and then
re-uploaded, a time-intensive process. Due to these limitations, the primary repository for daily case and
death data continued to be a commercial spreadsheet software package, though data were backed-up
in nCov xMart.

On 20 March 2020, WHO HQ released guidance requesting Member States to submit weekly aggregate
data. This data stream included aggregate reporting of cases and deaths by age and sex, as well as
testing data, hospitalization data, and self-assigned transmission classification (no cases, sporadic,
cluster or community transmission). This weekly aggregate data could be submitted either by uploading
a spreadsheet into nCov xMart or via direct data entry into a module developed using the ESRI
Survey123 platform. Setting up the ESRI platform required significant staff time, yet only nine Member
States consistently used the platform to submit data. Completeness of these data remained low and was
often delayed, leading WHO to rely on other organizations for critical data (for example, Our World in
Data’s [OWID] testing data).

On 27 March 2020 (49,659 daily cases and 2,739 daily deaths on this date), recognizing that
headquarters lacked the human resource capacity to continue collecting daily case and death counts,
WHO shifted primary responsibility for capturing these data to the regional offices. This ended the
duplicate data collection at regional and headquarters levels. From that point forward, staff at WHO
Geneva pulled data from WHO regional office dashboards, while most WHO regional offices continued
the time-consuming process of manually scraping daily cases and deaths from official public government
sources or daily situation reports sent by Member States. A consequence of this change was that the
data on the WHO global dashboard was often one day delayed. Yet, this change freed headquarters staff
from data abstraction responsibilities and allowed them to institute data quality controls, reducing
errors, ensuring consistency between regional and HQ counts, and improving data quality.

Transition from Case-Based to Weekly Aggregate Reporting (May 2020-April 2021)

In May 2020 (86,289 daily cases and 3.37 million cumulative cases reported as of 1 May 2020), the
Health Information Pillar determined that the benefits of continuing to collect case-based surveillance
data, which in any case suffered from poor form completeness, were outweighed by the burden placed
on Member States to report through this mechanism. WHO proposed that Member States transition to
providing only weekly aggregated case data. Though less comprehensive weekly data collection had
been ongoing since March, building consensus around the exact fields to be included in this updated and
expanded weekly aggregate system delayed the transition. Official guidance informing countries that
they could stop submitting case report forms and instead submit weekly aggregate data was not
released until 7 August 2020. This weekly aggregate data could still be submitted either via upload to
nCov xMart or direct entry into ESRI Survey123. Despite countries being given the option to transition to
weekly aggregate reporting, some regions and countries elected to continue submitting case report data rather than transition to weekly aggregate reporting. Some countries perceived that the analysis required to generate the aggregated data, including multiple stratifications by age and sex, paradoxically increased their analytic workload. The Health Information Pillar has been able to combine the case-based and weekly aggregate data into a single dataset, which is now featured on the WHO COVID-19 Detailed Surveillance Data Dashboard. However, these data were not made publicly available until January 2021, and a few Member States refused to have their data included. Though a dashboard presenting aggregated analyses of the case-based data had been developed by May 2020, this was the first public access to these data because of political sensitivities around the release of data from several Member States.

By April 2021, nCov xMart contained >50 million case report forms shared by Member States, representing the largest case-based dataset WHO has ever compiled. However, completion of most of the 80+ fields on the form was extremely low, with only age, sex, date of onset, and administrative region having greater than 50% completeness. In addition, Member State case reports were often completed when patients were admitted to a health care facility, thus outcomes — including death or discharge — could not easily be updated in any system. Political sensitivities also delayed the collection of many of these case reports. For example, one regional office did not share any case-based data with HQ until November 2020 despite repeated requests. The 50 million case report forms also strained nCov xMart, which was not yet optimized to maintain such large datasets, although this challenge had already been identified during the 2009 H1N1 pandemic.

Other Data Flows

More recently, xMart has also become the repository for WHO’s vaccination data. The monitoring of the implementation of public health and social measures (PHSMs) became critical for WHO’s role supporting Member State response activities, and WHO entered a collaboration with the London School of Hygiene and Tropical Medicine (LSHTM) and other organizations who had been collecting these data to produce a publicly available dataset tracking PHSMs. Tracking and disseminating the spread of COVID-19 variants also became important later in the pandemic, and the Epidemic Intelligence from Open Sources (EIOS) initiative has helped WHO to track event-based reports on the emergence and spread of new variants of concern.

Human Resources Strains

WHO experienced human resource shortages at all levels, with managers describing challenges maintaining team morale under extremely difficult conditions. The inability of surveillance teams to quickly scale up human resources early in the pandemic led to extremely long working hours under stressful conditions that persisted for several months. Internal mobilization of staff from other program areas, particularly outside of the emergencies program, was challenging. The IMST structure allowed for mobilization of WHO staff from different parts of the organization and improved internal coordination, but with mixed results on the willingness of other departments to have staff deployed to support the COVID-19 response. When teams did receive additional staffing, many of these deployments were felt to be of too short a duration to be impactful. Many of these new staff and consultants were also unfamiliar with WHO systems, reducing their immediate utility. The large amount of paperwork required to bring on a consultant, often for a three-month contract, was described as a major barrier at a time when staff were facing a crisis. An additional challenge WHO faced in adding human resource capacity was that
during the pandemic, many of the organizations that WHO might have turned to for skilled deployments wanted to preserve their staffing to manage their own COVID-19 response.

WHO staff at all levels described mismatches between the technical skills that were available and their most pressing technical needs. At headquarters level, data scientists and data managers were perceived to be underrepresented and undervalued. Staff with strong programming and data management skills were needed to put in place systems to capture, manage and clean data. Some regional and country offices reported needing more data managers and epidemiologists. These skills were in high demand from Member States who were seeking technical support for their own internal surveillance systems. The lack of data managers in country offices emerged as a weakness, and the placement of information management officers at country level was highly valued. Regional offices described their reliance on short-term consultants in several key capacities including R programming, GIS, and dashboard development, and warned that they are at risk of losing these capacities if consultant staff are lost once the pandemic wanes.

**Evolution of the Case Definition**

WHO published the first case definition for human infection with the novel coronavirus on 11 January 2020. Subsequent clinical and surveillance findings, the global spread of the pandemic, and the development of laboratory diagnostic tests, led WHO to update the surveillance case definition five times.

Whereas the first three case definitions (published 11, 21 and 31 January 2020, respectively) required a travel history to China, the fourth case definition, published on 27 February, expanded further the geography of travel to include all countries reporting local transmission of COVID-19 cases (WHO category 3 for transmission of COVID-19) as described in the WHO daily situation report. Transmission categories were defined as: (1) no cases reported, (2) sporadic cases, (3) cluster of cases, and (4) community transmission. However, Member States defined their own case rate thresholds to determine which COVID-19 transmission category they were in. These thresholds were not uniform across countries or consistent within countries over time, which compromised not only the credibility of the transmission classification system, but the ability to confidently apply the case definitions within Member States.

Impacts of the changing case definition included unstandardized and unclear country transmission categories and hamstrung and hindered counting of cases and surveillance of COVID-19. A case definition anticipating global spread of the novel coronavirus, with clear and accurate reporting by WHO of country transmission categories could have resulted in more durability and fewer requirements for revisions of national surveillance efforts.
Data Use

The value of surveillance data collection is determined by its utility for risk assessment, situational awareness, and response option planning. WHO’s approach to COVID-19 surveillance should be designed to collect the critical data needed to inform decision-making for both WHO and its Member States. WHO’s network of global and regional offices and its established relationships and credibility with Member States provides WHO with distinct competitive advantages disseminating data and guidance to facilitate data use. WHO has critical roles to play both at the global level, as the keeper and disseminator of trusted, actionable data and at the national level utilizing its long-standing relationships with Member States to directly support analyses and evidence-based decision-making.

Within the WHO structure of offices and Member States there existed varying needs for COVID-19 surveillance data. These needs evolved over time as the COVID-19 pandemic unfolded. The section below describes both met and unmet data needs that were highlighted during discussions with WHO staff and Member States.

Headquarters Perspectives

The self-identified objectives of WHO HQ COVID-19 surveillance were to:

1. Monitor trends in COVID-19 at national and global levels;
2. Monitor mortality caused by, and indirectly associated with, COVID-19;
3. Estimate morbidity and mortality for health care workers; and,
4. Assess the impact of control measures.

As the first cases of COVID-19 were reported, data were critical for WHO to perform its key function of providing countries with situational awareness as well as key guidance documents such as case definitions. WHO headquarters began releasing daily situation reports on 22 January 2020, with early reports focusing on the geographical spread of the disease as well as indicators of disease severity including hospitalizations, deaths, and the existence of underlying comorbidities among those who had died. Reporting health care worker infections was also prioritized. A line list of cases was described to be the most critical data at this time, as WHO HQ worked to understand transmission patterns, disease severity, and the impact of comorbidities. Timely reports of cases and deaths from Member States were critical for providing situational awareness as COVID-19 cases were identified in additional countries.

Once the transmission patterns and severity of COVID-19 became understood, maintaining situational awareness by tracking daily case and death counts continued to be important, but the utility of the case report data declined. This was recognized, and countries were requested to stop submitting case report forms on 7 August 2020. However, it is not clear that the case report forms offered utility at headquarters level beyond the very early days of the pandemic, and collection likely could have stopped far sooner.

WHO headquarters produced daily situation reports for 209 consecutive days, with the last daily report released on 16 August 2020, when WHO switched to releasing weekly epidemiological and operational updates. These routine reports served as the platform for WHO to disseminate three interim analyses of COVID-19 morbidity and mortality among health care workers. Analyses relied on case report data submitted from Member States. Though less than 50% of submitted case reports indicated if the individual was a health care worker, WHO was able to use these data to demonstrate the high burden of COVID-19 among health care workers globally and advocate for health care worker protections such as PPE and vaccinations.
Other data products produced by WHO headquarters include several data dashboards. The WHO Coronavirus (COVID-19) Dashboard is a public dashboard that presents the daily case and death data collected by the Health Information Pillar, Member States’ self-assigned transmission status, and more recently, vaccine doses administered. The WHO COVID-19 Detailed Surveillance Data Dashboard presents more detailed information on cases using a dataset that merges the case report forms and weekly aggregate data, including testing data taken from Our World in Data (OWID). WHO also developed both external and internal versions of their COVID-19 Explorer, which pulls together multiple data sources to provide a more complete view of Member States’ status for situational awareness: daily case and death counts, transmission classification, population data, testing data, mobility data, and information on implemented PHSMs.

Regional Perspectives
At early stages in the pandemic, as countries were identifying their first cases, regional office data needs centered on monitoring the global spread of COVID-19. At this stage, two data streams were highlighted as critical: the collection of official reports of cases and deaths from Member States and use of EIOS media scraping to find reporting of cases from the media and identify new areas with local transmission. EIOS was highlighted as especially important for regions with limited testing capacity early in the pandemic, as the first evidence of local transmission in some countries came when travelers who had recently left those countries tested positive for COVID-19 upon return to their home country, and EIOS was helpful at identifying these incidents.

Once COVID-19 was characterized as a pandemic, the focus shifted to using data to provide situational awareness to regional Member States to help inform decision-making and guide response activities. Evaluation of the impact of response measures including PHSMs was also critical; one success story was in the Eastern Mediterranean region, when regional PHSM impact modeling (in a collaboration with the CoMo Consortium) was found to be useful by Member States and incentivized the collection of case-based data to the region. Contextual information important for correctly interpreting surveillance data, for example the national testing strategy and case definitions being used, were also highlighted as essential to be able to provide support and guidance to Member States. Hospital system indicators (e.g., hospital bed occupancy, oxygen supply, ventilators, personal protective equipment [PPE]) were also highlighted for their importance guiding response activities, although regional availability of these data varied, and in most cases, were limited.

Publicly disseminated outputs were quite similar across regional offices. All regions produced their own regional COVID-19 dashboards and produced weekly or biweekly situation reports. Some regions also produced daily updates, country-specific situation reports, or separate epidemiologic updates that contained more detailed analyses of the available epidemiologic data. Regions also invested significant time providing direct technical assistance to Member States that sought their support in areas including information management, data analysis, and contact tracing. This direct technical assistance did not frequently involve the direct use of the data being collated by WHO HQ, although the importance of guidance documents produced by HQ was emphasized during interviews.

Country Perspectives
Discussions with staff at WHO regional offices, country offices and national public health institutes (NPHIs) indicated that many countries did not view WHO reporting as a priority because they did not see a strong direct benefit from reporting. Countries did desire having situational awareness of COVID-19 among their neighbors, and they reported accessing the WHO COVID-19 dashboard, but this was not
sufficient motivation for many countries. Country-level informants reported seeing limited benefits from reporting case-based data to WHO. It was perceived that many of the Member States that did reliably share data did so because they had a longstanding positive relationship with their respective WHO country offices and felt that WHO added value to their in-country surveillance activities. For example, in one country, WHO had closely supported the implementation of surveillance systems, and close collaboration and data sharing with their WHO country offices was routine.

Data quality and completeness was hindered both by the variability of the strength of national and subnational surveillance systems, and the technical capacities for national authorities to compile data, update case status with outcomes, and report to WHO. Requests to WHO around data use varied according to country capacity. Some countries requested support in data analysis; however, they often wanted support to analyze the data themselves and were not interested in sending data to external partners for analysis. NPHIs also requested data management support from WHO for their own surveillance systems. Support for modeling was also highly in demand and needed to be delivered in a timely fashion. These requests varied by the stage of the epidemic, with early requests often focused on modeling epidemic trajectory and later requests often focused on modeling the impact of different PHSMs to inform policy decisions. It is important to note that many of the analyses supported by WHO for its Member States were considered politically sensitive and not for public consumption. Thus, this added value provided by WHO to Member States was not released into the public domain and went unseen and unrecognized by the wider public.

A shared frustration was a lack of coordination of data requests from HQ and ROs, with various teams requesting similar data. A thorough inventory of data needs across WHO offices and better coordination may have reduced the reporting burden on overwhelmed Member State and country office surveillance teams. Providing the reasoning and rationale behind data requests and explaining the potential benefits of data sharing to Member States was described as critical.

**Missed Opportunities and Lessons Learned**

In our interviews, informants described several missed opportunities and lessons learned that could have facilitated improved data use and helped inform decision-making among Member States in the context of the COVID-19 pandemic. Some interviewees felt that much of the focus of data collection and dissemination at headquarters, and to a lesser extent at regional level, was on providing situational awareness to WHO leadership. However, more advanced planning and better data systems would have enabled WHO HQ to produce more timely strategic information outputs that would have added value to Member States as they coordinated their response. WHO Country and Regional Offices have provided significant direct technical support to Member States throughout the COVID-19 pandemic, and this support was highly valued by many MOHs and NPHIs. However, much of this support was done on an *ad hoc* basis, scrambling to find consultants and relying on personal connections, and a more systematic approach to supporting Member State data use through established partnerships and consortiums could have increased WHO’s impact and perceived usefulness to Member States.

In the following sections, we described four areas where changes to the collection or dissemination of data by WHO could have facilitated Member State data use:

1. The need for prompt dissemination of data to Member States;
2. The need for timely dissemination of guidance documents to inform data use;
3. Data that were collected but went unused or were underutilized;
4. Data that were not systematically collected that could have added value.
Prompt dissemination of data to Member States

When WHO requested case report data early in the pandemic, these data including age and sex breakdowns were not made routinely accessible to all Member States or to the public for nearly a year. WHO had developed an internal dashboard in May 2020, shared it with Regional Offices, and had planned to disseminate it to all Member States. However, some regional offices chose not to share the dashboard with their Member States. Though these data were periodically analyzed at global level, with results disseminated in WHO situation reports, aggregate country-level case report data were not made available on a public facing dashboard until January 2021. Two barriers to sharing these data were highlighted. First, the lack of prior agreement among Member States and WHO on data sharing within the scope of IHR created political impediments to dissemination of the case report data. De-identified data available as a “pseudo line list” would provide researchers a valuable source of data to inform modeling efforts. Second, in early 2020 when WHO first requested case report data, WHO lacked a data platform to analyze the data and provide access and utility to Member States.

This lack of a feedback loop was widely perceived to have reduced Member States’ motivation for reporting case report data to WHO, especially at a time when many Member States’ human resources were stretched. Member States did not see an immediate benefit from reporting case report data to WHO, and in some Member States there was general fatigue from partner requests for data. Interview participants described a need to see a shift from WHO asking for data from countries to prioritizing assisting countries to analyze and disseminate their own data. This approach would have required WHO to have adequate human resources, data sharing agreements, processes and platforms in place prior to the pandemic, but would have increased the dissemination of important information to the public.

Consensus around Member State data sharing agreements, minimum datasets, processes, and platforms should be reached in advance of an outbreak to align WHO HQ and regional office approaches to data collection and avoid long stakeholder processes that restrict and delay data use during a crisis. However, the effective use of these systems requires Member States to develop capacities as described under the IHR (2005).

Timely dissemination of guidance documents to inform data use

In a crisis, balancing the tradeoffs between building consensus through solicitation of stakeholder feedback versus the time constraints of providing timely guidance is extremely difficult. WHO headquarters was praised by informants for its consultative approach and excellent communication, as regional offices felt empowered to provide input and felt that their input was incorporated into HQ technical products. In addition, regional offices appreciated the high quality of the technical guidance produced by WHO HQ.

However, staff at both regional and headquarters levels indicated that at times headquarters needed to prepare guidance faster for it to be relevant to Member States. The consultation and approval processes at times delayed the development process, preventing WHO from adequately fulfilling its leadership role. Some informants from headquarters felt that WHO needed a mandate during an emergency to limit consultations, rather than seek the input of external experts that slowed the release of critical documents. A few of the guidance documents that were mentioned as having come too late include recommendations around international travel, contact tracing, and updated guidance on implementing and adjusting PHSMs.

Data that were unused or underused
WHO invested significant time and resources in the collection of over 50 million case report forms. However, there is limited evidence that the case report data were widely used at headquarters or regional levels after the initial stages of the pandemic. Collection of detailed case report forms likely could have been discontinued far earlier in the pandemic. Once the pandemic became established, it was reports and analyses done at country level that generally added value, with the potential exception of global health care worker infection data that could have been collected in aggregate formats.

Additionally, the revised case report form released on 27 February 2020 contained 83 fields. Compiling these data quickly became untenable in many countries as case counts increased. Some Member States found themselves with line lists with largely incomplete data, and this led some Member States and even some levels of WHO to resist sharing their line list with higher levels of WHO. A clearly defined minimum set of variables and a case report form with fewer fields may have helped encourage case reporting.

WHO attempted to collect testing data and hospitalizations through the weekly aggregate reporting process. However, these data were widely perceived by WHO staff as being too delayed and too incomplete to be useful. Many Member States did not provide these data to WHO, so WHO relied on testing data from OWID, which scrapes these data daily from official public sources. Additionally, the collection of Member States’ self-determined transmission classification through weekly aggregate reporting, and displaying this data prominently on the WHO dashboard, was perceived to diminish WHO’s credibility as some countries with large numbers of cases refused to classify themselves as having community transmission. This highlights a major political challenge faced by WHO, which is forced to balance the value-add of making independent evidence-based judgements that may conflict with the official position of Member States against the risk of fracturing relationships with Member States whom WHO rely on to voluntarily share data. WHO should be able to have more autonomous authority for interpreting and publishing Member State data, as stipulated under IHR (2005) Article 11.

**Data that were not systematically collected that could have added value**

WHO staff highlighted testing data, hospitalizations, and hospital bed capacity as critical data for risk assessment and response coordination. Though WHO HQ attempted to capture testing data as part of the weekly aggregate reporting, Member State completeness was low, and regional offices either compiled testing data from their Member States’ public sources or used data from OWID. The availability and quality of hospitalization data varied by region, and it is an area of surveillance that should be strengthened at country level. It was also emphasized that hospital capacity data are critical for interpreting hospitalization data, so response managers can determine where resources are needed. One region emphasized the need to collect top-line data – cases, tests, hospitalizations, and deaths – at subnational levels. The WHO Regional Office for Europe displays subnational data on their [public dashboard](https://www.who.int), and subnational data exist in many Member State DHIS2 systems. However, some WHO regions had no system to ingest subnationally reported daily data (cases, hospitalizations, deaths, etc.) and use it for risk assessment and response planning, while others had a system for data capture but did not receive subnational data from Member States. This capacity might not be the purview of WHO headquarters, and is potentially an area of development for country and regional offices.

Systematic collection of metadata around case definitions, testing strategies, and contact tracing strategies used by Member States were repeatedly highlighted as critical to contextualize the epidemiologic data that were being collected and helping avoid misinterpretation of data. Though WHO
requested in their surveillance guidance that Member States submit these metadata via email, many Member States did not provide them.

Although outside of the scope of this review, stronger and more timely implementation of excess mortality surveillance systems at national level and reporting to WHO would allow Member States to better understand the impacts of the pandemic in Member States with stronger surveillance systems. Additionally, mortality surveillance would provide improved situational awareness and inform response activities via a sentinel site approach in Member States with more limited surveillance and diagnostic capacities. Though WHO released a technical package for rapid mortality surveillance in May 2020, this is an area that would benefit from additional investment and support at country level.
Reporting and Data Pipeline

WHO Surveillance Systems

Prior to COVID-19, WHO did not routinely collect and analyze case or patient level data. Systems in place were not designed to monitor national and global burdens of disease and generally collected aggregate data on a monthly or weekly basis.

Several systems were in place at all levels of WHO to support reporting of COVID-19 surveillance data through the various levels of the organization. However, the depth and width of data quickly became a challenge to capture and manage through existing systems. In other instances, data were minimally or inadequately reported from subnational to national and to regional offices. As these challenges emerged, staff at headquarters and regional offices looked for additional ways to support data collection efforts. This effort was collaborative, and regions reported that the ability to work with Member States and WHO HQ occurred through frequent meetings and other communication methods.

WHO regional offices that initially or continually received limited Member State data heavily relied on teams to use media and web aggregators such as the EIOS platform. From their scans, data were collated into local data collection systems to produce data summaries of basic daily case and death counts as well as surface additional contextual information within their regions.

Attempts to support collection and management of increasing data volume and complexity occurred at all levels, leading to the implementation of nCov xMart to harmonize data from disparate sources. Member States with high participation in existing regional surveillance platforms (TESSy, EMFLU) for influenza and other diseases were able to alleviate some of their data reporting challenges, and these regions had an advantage in being able to quickly expand upon their platforms to share data into nCov xMart.

When data were able to be collected, they were often incomplete or required mapping and transformation, thereby impacting the timeliness and utility of the data. Through the course of the pandemic, WHO issued interim guidance that posed additional clarification for Member States on case definition and surveillance methods (Timeline Figure 1). However, some posed challenges with adapting data collection systems to the guidance. For instance, when WHO requested weekly aggregate data, several Member States that had been collecting and reporting case-based data either opted not to report weekly or needed additional support to transition to weekly reporting. These challenges underscore the need to identify system variables that shall be minimally collected, their value sets, defining how to append or deprecate variables and value sets if needed, level of data detail (individual vs aggregated) and the method(s) the data can be reported (e.g., file upload, user interface, direct connectivity via application programming interface) at each level prior to the next event.

Member State Surveillance Platforms

As COVID-19 spread across regions, NPHIs and local public health agencies raced to enable existing surveillance and case management platforms or adopt systems to support rapid response. Many regional offices reported that Member States soon utilized data collection systems, however system sophistication varied. Some Member States started with paper-based case report forms while others were able to compile electronically reported case report forms and lab reports.

In settings where platforms were in place, initial cases were reported to these systems as required by those Member States who were willing to share data. However, the pace and volume of cases soon
became overwhelming. Delays and incompleteness of data were most apparent when subnational areas were not yet integrated into national platforms. In addition, the ability to report data from these systems typically required time consuming manual manipulations and was compounded at times by data sharing sensitivities. Inevitably, these data supply chain issues impacted the ability to quickly identify target areas and respond with public health interventions. National systems were not designed to scale for pandemics. As platforms began to struggle, some were reinforced with increased data processing and database capacity or separate COVID-19 specific platforms were implemented.

Reliance on Spreadsheets

In settings with minimal surveillance and case management systems in place, spreadsheets were often relied upon to capture case data and were also soon found to be inadequate. Although these tools are simple to implement and familiar to users, the use of spreadsheets, especially those that can be manipulated and modified by multiple users, poses several risks, including difficulty maintaining data validation and formulas, data scalability, stability and integrity. For example, Google Sheets has the ability to support up to 5 million cells. If a line list consists of 50 variables, 100,000 cases is the maximum data set for a sheet. Excel can support approximately 1.05 million rows and 16,000 columns but can be difficult to share and poses similar risks. The limited ability to restrict user privileges in a simple spreadsheet often allows any user the possibility to inadvertently corrupt the entire dataset.

Sources and Flows of Data, by Region

Until March 2020, public data were scraped manually by HQ for daily case and death counts. Data were then collected from Regional Dashboards.

Member States communicated case-based data (case reporting forms, line lists, and formal reports) through several mechanisms. Some used IHR (2005) reporting systems and email via their country office or IHR (2005) National Focal Point, while others used other tools made available by WHO (file, data input into nCov xMart). Member States also used connectivity available to them with existing regional surveillance networks such as WHO Regional Office for Europe/ECDC TESSy and Eastern Mediterranean Regional Office EMFLU networks.

As with case-based data reporting, a number of methods were used to share weekly aggregate counts. Member States shared the data directly via the nCoV xMart portal, weekly aggregate spreadsheet upload or the aggregate counts were compiled from case-based data by the regional office or headquarters. DHIS2 is another system used by many Member States and as well as at three regional offices to help support detailed and aggregate reporting.

Digital platforms play a critical role in the management of cases and collection of data at the community level and by NPHIs. Several platforms have been developed and/or adapted to support COVID-19 case-based surveillance and aggregate reporting. For instance, WHO in partnership with the Global Outbreak Alert and Response Network (GOARN), has enabled Go.Data, an outbreak investigation tool for field data collection during public health emergencies, to support COVID-19 case management and surveillance efforts. WHO has facilitated its implementation in many Member States; however, use of Go.Data for surveillance reporting was minimal. Data from some platforms can be exported for use in aggregate reporting.
COVID-19 Platforms and Data Interoperability

Interoperability and data standards can play a critical role in data management and reporting within and between COVID-19 digital platforms. Standards such as open application programming interfaces (API) for data input and access, HL7 FHIR, HL7 laboratory messaging and integration with laboratory management information systems (LMIS) enable the transmission of data between local, national and regional levels. In countries where electronic health records are widely available, additional efforts are being made to enable these clinical systems to electronically communicate clinical data on cases and transmit electronic case report forms using HL7 standards. Though these technologies are complex and have had their share of issues, building upon this work can alleviate manual and duplicative data processing efforts, especially if data governance and sharing agreements are well developed and properly implemented.

Several platforms developed or adapted for COVID-19 have implemented interoperability standards; these include Go.Data, EWARS, SORMAS, DHIS2 and CommCare. Each platform has had varying degrees of national adoption. WHO has facilitated Go.Data implementation in approximately 80 instances across 60 Member States. It was also mentioned that 23 Africa Region Member States participated in Go.Data awareness training and that eight Member States implemented the system. Use of Go.Data was primarily used for Case Management and Contact Tracing and was not used to report surveillance data to nCov xMart. SORMAS was in use in some Member States and DHIS2 and CommCare have had broader implementation. In the South-East Asia Region, the majority of Member States in the region are utilizing DHIS2 and the region has implemented DHIS2 to facilitate case-based data sharing. In regard to surveillance guidance, many of the systems mentioned above developed COVID-19 modules built upon WHO guidance (case reporting variables).

In addition to the adoption of these platforms, it is important that WHO recognize its role in driving data and interoperability standards in collaboration with the needs and abilities of Member States. This will provide the informatics framework from which WHO can develop a strategic information plan and definition of data sets and transmission standards that are appropriate and applicable to the surveillance needs at each level. Finally, WHO should also continue to support subnational and national agencies with technical expertise on standards and together work with platform developers to implement features and requirements to support and enhance interoperability efficacy.

While many local and national public health institutes are using these platforms and are able to produce case-based aggregate reports and line lists for local use, regional offices reported that data were not readily shared at the WHO regional level for a variety of reasons such as sensitivity in sharing data and IT and personnel capacity to produce data outputs. Given that these platforms support interoperability and data access, it remains a priority for WHO to encourage and incentivize states to leverage this capacity to support surveillance.

WHO HQ Harmonization and Output Process

Although WHO published guidance on data collection tools and suggested data sets, WHO welcomed data from Member States and regions in a multitude of formats and methods. This flexibility was welcomed by the regions and enabled them to share data with WHO headquarters more easily. Member States submitted a range of surveillance data to their regions; similarly, regions opted to report varied data to WHO. At WHO headquarters, data transformation processes were established to harmonize data for various data products and outputs. These processes included means to validate, de-duplicate and
enrich surveillance data via nCov xMart, a SQL database that supports xml transformation of reported data.

As the data grew, so did the complexity of data processing. Regions were often unable to report incremental data because of varying identification numbers (IDs) or date/time stamps, which impacted the performance of the nCov xMart system. It typically takes several hours to process a single region’s data. It was reported that on several weekends, the incidence of cases was so high that the nCov xMart system slowed substantially or made uploads impossible. The API service that loads data into the R-Shiny application also lost functionality. Availability of after-hours support to correct these issues was limited; it was not until October 2020 that an on-call roster system from data management and IT was in place.

Once processed, the data are used by data visualization tools (PowerBI, R-Shiny, and a custom system developed by Sprinklr) to produce WHO dashboards, situation reports and other in-depth analyses (Figure 2). These were developed in close partnerships with regional offices. In one region, data are synched with WHO nCov xMart to produce regional reports and dashboards. While more recent traffic to WHO dashboards is less than that of some third-party sites (e.g., in April 2021, WHO dashboards had 3.65M visits, Johns Hopkins had 8.85M visits and OWID had 12.9M visits), it was shared that regions and Member States are deeply dependent on these publications to support official state-based guidance and interventions.

Figure 2: WHO Headquarters Transformation and Harmonization of Data
Conclusions and Recommendations: Adapting for Scale and a Digital Future

WHO’s Role in Disease Surveillance

WHO plays a unique and critical surveillance role before and during public health emergencies, and has several distinct advantages compared to other organizations: its existing relationships with Member States, its global network of regional and country offices, and its credibility as a source for both surveillance data and guidance documents that inform how Member States conduct their own surveillance activities. However, WHO also faces unique organizational and political challenges that directly impact the timeliness with which the organization can disseminate data. There are no enforcement mechanisms or reprisals when countries choose not to report in accordance with the International Health Regulations (2005) for public health emergencies of international concern (PHEICs), and no formal mechanism or regulation exists for data sharing for disease events that fall outside of PHEIC criteria as described in IHR (2005) Annex 2. As described earlier, for the purposes of this rapid review, the review team was unable to speak with a representative sample of Member States and we do recommend that further consultations on the emergent themes identified in this report occur in order to proactively and collaboratively develop solutions.

Early in a public health emergency, WHO must fulfil several critical surveillance roles, including providing technical guidance on surveillance and case definitions: supporting countries in data collection and analysis to understand transmission patterns and to inform severity and risk assessments; and when relevant, serving as a trusted source of information to other Member States about potential risks. During emergencies, WHO can support its Member States by providing technical support for the various pillars of response efforts and disseminating and sharing information through situational updates by means such as dashboards or formal situation reports.

The quality and completeness of the data on which WHO relies depends on the capacity of its Member States for surveillance and their willingness to share surveillance data, as WHO is often a secondary consumer of data generated by Member States. During and before emergencies, WHO must have resources to support Member States on strengthening their surveillance and reporting systems that can quickly be adapted to scale. These needs include providing guidance around minimum surveillance datasets and tools for situational awareness. Just as pre-pandemic work developing the WHO Outbreak Toolkit helped WHO quickly release an initial case report form, anticipatory planning for data elements and standards for public-facing dashboards and sample templates could allow WHO to provide timely and helpful guidance to its Member States and ensure that key stakeholders including HQ and regional WHO offices agree on what minimum set of data elements should be collected. Guidance around minimum dataset and data standards for publishing COVID-19 data on country web applications and dashboards, disseminated by the eHealth Network with support from ECDC, could serve as a model resource to facilitate timely data sharing and usage.

WHO’s strength also lies in supporting countries to conduct risk management. This includes advising countries on the use of PHSMs, clinical guidance, and vaccine policy. While data often inform these decisions, given challenges in data sharing, WHO might not necessarily be best suited to serve as the repository of detailed case information. In Member States, WHO can support governments on data collection and analysis to understand transmission patterns and severity assessments, but rigorous clinical transmission and clinical data are often analyzed and promulgated by the countries themselves, or their hospitals or academic institutions, as was the case during the COVID-19 pandemic.
WHO’s convening power is unparalleled in global public health, and its ability to build international consortia consisting of private, academic, non-governmental and/or government organizations should be further leveraged. An example of this potential is the consortium WHO established to create a publicly available dataset tracking PHSMs. As described below, WHO could benefit from establishing formal collaborations or consortia which could be useful when responding to both smaller outbreaks and pandemics, either via GOARN or other partnership models. Other organizations may have a technological advantage collecting certain types of data or have more flexible human resource structures better able to respond to a crisis, allowing WHO staff to better maintain focus on strengthening national and subnational surveillance capacities, transmitting and supporting the implementation of new guidance, providing risk assessments, and supporting evidence-based response option development and implementation.

A Novel Pathogen and Evolution to a Global Pandemic

The novel coronavirus identified in Wuhan presented dual challenges in the early phase: relatively little was known about the pathogen itself, and the surveillance and data systems would inevitably need to evolve from a focus on disease characterization to providing adequate situational awareness to inform response efforts. WHO prepared and released guidance on surveillance for the novel coronavirus, including a detailed case investigation form and first few cases and contacts (FFX) protocol. These releases were necessary and relevant for all Member States as virus transmission was rapidly detected in all WHO regions. However, the initial clinical characterization of COVID-19 was based on data produced and published by Member States and their health care and academic institutions, and was not derived from case investigation or surveillance data. As the pandemic progressed, Member States were still asked by WHO to submit detailed case data via line lists on more than 80 variables, the vast majority of which were incompletely reported. Further, there were few publicly available analytic applications of those data to indicate that they were truly required by WHO. As the pandemic progressed, WHO recommended weekly aggregate data be submitted to the organization, although the requirement for stratified data (including stratifications by age and sex) did not necessarily diminish the complexity of the data submission process, as countries would inevitably draw from line list data to conduct the required analyses for submission to WHO. These challenges point to the need for a more parsimonious case-based dataset, including case and death data, age, sex, health care worker status, hospitalizations, in addition to national testing and excess mortality rates.

WHO’s Outbreak Toolkit does provide recommendations for required data fields, which are likely of value in future large events. For future events with pandemic potential, WHO might find it most efficient to define a minimum data set and data standards for ministries of health or national public health institutes to publish, and allow for the aggregation of those publicly available data. WHO might also serve a valuable function as a repository of the surveillance “metadata”, including case definitions and testing strategies used by Member States, which further contextualize the public-facing data.

COVID-19: The Challenges of Scale

The scale of the COVID-19 pandemic, affecting all WHO regions and with a nearly unprecedented case count, resulted in several operational bottlenecks. Firstly, human resources capacities were limited. Although WHO has skilled epidemiologist and event managers, there was a perception among staff at all levels that there was not an adequately staffed workforce to design big data systems, perform the type of data management and processing needed for the pandemic scale, or provide technical support for Member State surveillance activities. When resources were not available within the organization,
human resources processes resulted in substantial delays in hiring consultants, whereas high staff turnover of GOARN deployed and embedded staff resulted in operational discontinuity of the organization. If an emergency hiring mechanism cannot be established, human resources and finances surge staff should be prioritized in addition to technical staff.

Secondly, the high number of cases meant that *ad hoc* data systems failed to meet the demands of the system. While all outbreaks do not require big data management capacities, a scalable data solution for the organization that can be used for both smaller outbreaks and for pandemic-size events needs to be established. A key point of failure in the design of the data systems was the initial conflation of user needs for data entry (where spreadsheet formats are often preferred), with a technologically capable data management and storage system; there was little distinction between the “front end” and the “back end” requirements for data systems. There was no pre-existing established network of interoperable data entry systems that could facilitate decentralized data entry across WHO regions and Member States, resulting in a high and at times duplicative workload for data collection and entry by headquarters staff, although the use of APIs and leveraging of existing capacities (e.g., DHIS2, TESSy, xMart) increased throughout the pandemic. The data management and storage system should provide key features, including data ingestion from an input system (a highlight of WHO’s response was flexibility in accepting data in various formats to accommodate regional reporting systems), data storage, and data cleaning and processing to allow for read-only access for output generation, including dashboards and situation reports. While the xMart system allowed for data integration from multiple sources, its limited ability to perform data management and cleaning meant that modifying data required exporting data to a spreadsheet format (e.g., Microsoft Excel), making the corrections, and then re-importing the data, leading to inefficiencies and limiting usefulness.

**Timely and Effective Use of Data**

The COVID-19 pandemic demonstrated the need to use data to inform response options including the implementation of PHSMs; targeting risk communications messaging, contact tracing, and vaccinations; and informing recommendations around travel. The data elements required to inform these response options include case incidence, testing rates (to contextualize case incidence), hospitalization and intensive-care unit admissions (to understand when national capacity is at risk of being exceeded), all-cause mortality, and country transmission categories. Defining transmission patterns within countries was critical in the early stages of the pandemic for Member States to apply WHO’s case definitions, and important to countries as they considered implications for global travel. However, self-nomination of transmission categories resulted in a loss of credibility of these categorizations.

There was limited public application and use of many data fields included in the case-based reporting forms, which jeopardized the perceived utility of sharing these data with WHO. The Eastern Mediterranean Regional office was able to use surveillance data submitted by Member States to model the impact of PHSMs, which was seen as valuable by Member States and provided an incentive for data sharing. Providing services to Member States that are not routinely available might increase the perceived utility of the global data collection process.

IHR (2005) Article 11 permits WHO to share data with State Parties after consultation (rather than explicit approval) during a PHEIC or when: international transmission has occurred; there is evidence the control measures against the international spread are unlikely to succeed; the State Party lacks the operational capacity to carry out necessary measures; or the nature and scope of international travel and trade requires the immediate application of control measures. Although these criteria were all in
place during the COVID-19 pandemic, the lack of a coherent and anticipatory data use and sharing plan hindered the response.

During pandemics, event-based surveillance that included both media scraping for situational awareness (e.g., EIOS) and other data scraping efforts from publicly available sources (as done by JHU, ECDC, some regional offices) to supplement data formally reported to WHO was seen as particularly important. Event-based surveillance systems can also support early warning systems, and the EIOS platform allowed user communities to share information between countries, which was seen as valuable. Key enabling factors for organizations that were able to effectively use web scraping technologies to create time series datasets were their ability to rapidly surge staff capacity, making code and data publicly available, and having a mechanism for anomaly detection and user feedback for concerns about data quality. However, as the pandemic evolved, minor differences in timeliness and completeness of WHO data versus other publicly available data were likely less relevant, except for the perception by users that there were other data sources to inform situational awareness. WHO’s competitive advantage in this space was as an official data source with multiple forms of validation and retrospective data adjustment, allowing it to be a trusted source of data by Member States. However, county-level aggregate data from more than 50 million case report forms were not made available to Member States or the public until January 2021, compromising the perceived utility of real-time data collection efforts.

Summary of Key Recommendations Based on the COVID-19 Experience

---

**Strategic Information Planning and Surveillance Systems**

1. **Identify and implement a strategic information plan that builds on existing systems and can inform decision-making across Member States to develop global public goods during major events, including data standards, data sharing, and minimum reportable data by Member States:** WHO should work with Member States and IHR (2005) States Parties to develop tools, forms, and data elements required for assessing risk for new events, and coordinating outbreak response. WHO and Member States should ensure that those tools, forms and data elements are scalable for possible larger future events, with a clear data use and sharing plan in place. The WHO Outbreak Toolkit might form the basis for defining the data use plan and minimum reportable data elements. While national surveillance strategies and tools may vary based on the country context, WHO can provide thought leadership on data standards and elements to support private developers of tools that are being implemented at scale, including DHIS2, CommCare, SORMAS, Go.Data and others. In addition, WHO should provide guidance around minimum data standards for publishing COVID-19 data on country web applications and dashboards. Further, WHO should leverage its normative position to develop guidance for data standards for all public health events, which would allow for Member States to collect data using tools with common data standards that are tailored for those country contexts. In combination with consultation with Member States and support on defining and adapting tools, WHO and other stakeholders should advocate for more timely and transparent sharing of data under the IHR (2005).

2. **Invest in a scalable data solution that allows for relevant parties to submit data to WHO, with data management, storage, and pre-processing capabilities for small and large events:** This “back-end system” should be distinguished from a data entry solution, which will likely vary by region. Enabling access to a “front-end” system for data entry for regional and country offices, as well as
Member States themselves, could reduce headquarters and regional office workload for data collection and entry and increase staff bandwidth for data quality and analysis. Data architecture solutions will also require clear data and metadata standards that will allow for systems operability and leverage application programming interfaces (APIs) as efficient and query-able vehicles of data exchange. This system should be used for routine reporting, and should be easy for countries to access, adopt, and obtain information from. It should incorporate reporting lessons, procedures, best practices, and best platforms from throughout WHO.

3. **Invest in national surveillance strengthening, with technical support from WHO country and regional offices:** Global surveillance will only improve when national surveillance capacities improve, and these investments should be prioritized by all stakeholders, including WHO, Member States, and development partners. WHO country and regional offices were best situated to adapt and ingest data from their Member States. Common data platforms within regions (e.g., TESSy in the European region, DHIS2 in the Eastern Mediterranean region) and integrated disease surveillance strategies (e.g., integrated disease surveillance and response strategy [IDSR] in the WHO African region) were seen as accelerators for adaptation to the COVID-19 pandemic. Catalytic investments include data management solutions at national level, and integration of outbreak response functionality into existing national surveillance systems. WHO’s SCORE technical package could provide a way for Member States to assess capacities for data collection and use and fill identified gaps, making them better prepared for health emergencies. Such efforts will require the involvement and investment of parts of the organization not limited to the Emergencies Programme or the response IMST. There is also a need to see a shift from WHO asking for data from countries to prioritizing assisting countries to analyze and disseminate their own data.

4. **Accelerate investments to strengthen and expand event based surveillance (EBS):** The Epidemic Intelligence from Open Sources (EIOS) platform was viewed by many stakeholders as helpful for early warning, variant tracking, and data collection. EBS might provide an additional tool to supplement formally reported data to the organization. WHO should establish a public-facing system to share curated event-based data, including the results of media scraping for new event alerts and “scraped” case data. WHO, its partners, and Member States can use unofficial data from EBS to triangulate disparities between official and unofficial data, query Member States, and initiate investigations earlier than may otherwise have been possible through traditional disease reporting systems.

5. **WHO should be empowered to share data from unofficial sources in accordance with provisions in the International Health Regulations:** IHR (2005) Article 11 permits WHO to share data with State Parties after consultation (rather than explicit approval) during a PHEIC or when: international transmission has occurred; there is evidence the control measures against the international spread are unlikely to succeed; the State Party lacks the operational capacity to carry out necessary measures; or the nature and scope of international travel and trade requires the immediate application of control measures. In instances when WHO identifies a high-risk suspected event through EBS and the relevant State Party does not respond to a verification request from WHO, WHO should make the information publicly available, noting the verification status of the event. Similarly, for larger events meeting the provisions of IHR (2005) Article 11, WHO should curate those data and make them publicly available to aid in the timeliness and completeness of its dataset.

*Leveraging WHO Leadership Role and Growing Human Resources Talent and Processes*
6. **Revise human resource processes required for surge staffing during major events:** If establishing an emergency human resources system for such events is not possible, additional human resources and finance staff should also be prioritized to accelerate the hiring process. The capacity for surge staffing was highlighted as an enabling factor at ECDC and Johns Hopkins University Center for Systems Science and Engineering.

7. **Build institutional capacity for data management and data science through investment in people:** Human resources in data management and data science should be prioritized before the next major event to build scalable systems and to deploy during emergency operations. A professional-level career track for data management and data science (with a focus on the effective use of data for response option planning, forecasting, and impact assessment) could incentivize recruitment and retention of talented staff across the organization, and build a pool of deployable staff during emergencies. These technical human resource gaps were noted by all levels of the organization, and best practices and resources from throughout the organization and GOARN partners should be enlisted for this purpose.

8. **Leverage WHO’s convening power and international voice to foster advanced technical partnerships:** WHO’s convening power is unparalleled in global public health, and its ability to build international consortia consisting of private, academic, non-governmental and/or government organizations should be further leveraged. WHO should define core technical competencies needed for data science and data management internally and establish strategic partnerships before and during emergencies for contingency-based technical needs outside of those technical competencies, which might include advanced modeling techniques. WHO can provide technical partners visibility and credibility, while maintaining its position as a global surveillance leader.