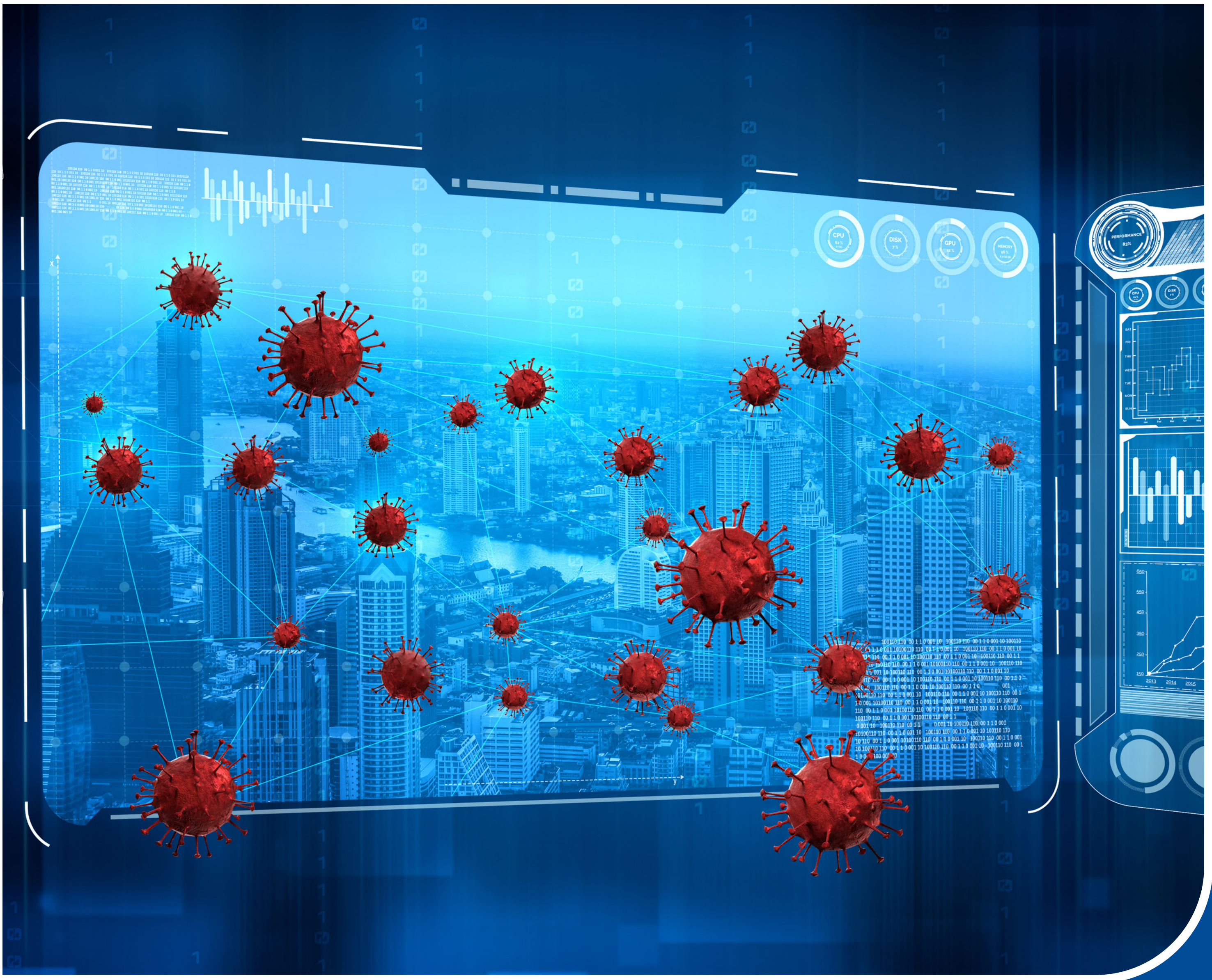
مؤسسة قطر
Qatar Foundation

مؤتمر القمة العالمي للابتكار في الرعاية الصحية
World Innovation Summit for Health

Data-enabled responses to pandemics:

Policy lessons from COVID-19

Sangeeta Dhami, Didi Thompson, Maha El Akoum,
David W Bates, Roberto Bertollini, Aziz Sheikh



Dhami S, Thompson D, El Akoum M, Bates DW, Bertollini R, Sheikh A. Data-enabled responses to pandemics: Policy lessons from COVID-19. Doha, Qatar: World Innovation Summit for Health, 2022

ISBN: 978-1-913991-18-0

Data-enabled responses to pandemics:

Policy lessons from COVID-19

TABLE OF CONTENTS

FOREWORD

The COVID-19 pandemic has made clear the pivotal importance of access to multi-dimensional, high-resolution data in near real-time to inform and enable evidence-based policy, public health and clinical decisions on how to effectively respond to various stages of the pandemic. Different data sources are particularly useful at various stages of a pandemic: initially they can help to identify and confirm the emergence of an infection with pandemic potential, understand its epidemiology, inform deliberations on risk stratification, and make decisions about travel; and later data can enable assessments of the uptake, safety and effectiveness of non-pharmacological interventions, vaccines and therapeutics, help us understand the impacts on health system functioning, and to inform deliberations on when a pandemic can be declared over.

Most countries have struggled to provide the required data – especially in the early stages – and fewer still have had the capacity to analyze these data in meaningful ways. These deficiencies have resulted in substantial – potentially avoidable – morbidity and social, educational and economic disruption.

However, a few countries have emerged as positive outliers with respect to their data capabilities, creating end-to-end national data infrastructures that have been and are still being used to guide their own national responses and to inform global decision-making. This paper focuses on examples of these national data infrastructures with a view to identifying potentially transferable policy lessons for other nations.

Prof Sir Aziz Sheikh
Director, Usher Institute
Dean of Data
The University of Edinburgh

N. Sultana Afdhal
Chief Executive Officer
World Innovation Summit
for Health (WISH)

SECTION 1: INTRODUCTION

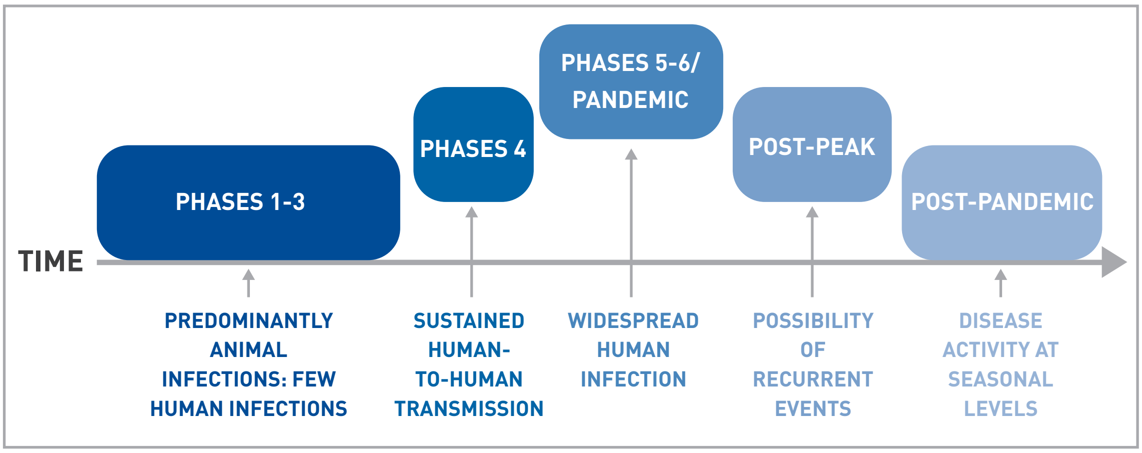
The coronavirus (COVID-19) pandemic, which emerged from Wuhan, China in December 2019 has globally resulted in at least 500 million cases and more than 6 million deaths worldwide (as of June 2022).[1] There has been considerable additional disruption, morbidity and mortality resulting from the social, economic and health system consequences that ensued as a number of countries instituted a series of national and more localized lockdowns.[2,3] The pandemic required a series of policy, public health and clinical decisions to be taken, with major consequences to societal functioning, economics and care provision. The taking of these decisions was always going to be complex, but for most countries this was exacerbated by the lack of availability of relevant data.[4,5] In contrast, a handful of countries have substantially developed their data capabilities over the course of the pandemic, generating important insights to guide their own national decisions and to inform international deliberations.

In this policy paper, we reflect on the various data sources that should ideally be available at various stages of a pandemic. Through case studies of select countries that have been positive outliers in their data capabilities, we then seek to identify potentially transferable lessons that should help ensure that countries are better equipped to generate data-enabled responses to future epidemics and pandemics. Since the COVID-19 pandemic is not yet over, the ideas contained in this paper should be seen as a work in progress.

SECTION 2: DATA REQUIREMENTS FOR DIFFERENT PHASES OF PANDEMICS

All pandemics have distinctive dimensions depending on the nature of the responsible infectious agent, the speed of national and international non-pharmacological responses, and the availability and deployment of vaccines and therapeutics. It is, however, possible to identify some core phases of pandemics and therefore consider the data sources that should ideally be available to support decision-making during these phases. The core phases of pandemics are summarized in the World Health Organization (WHO) Pandemic Phases framework, which was developed for influenza. Figure 1 shows the framework and, to illustrate, Table 1 describes the phases for influenza.

Figure 1: World Health Organization (WHO) pandemic phases



Source: Reproduced from WHO (2007)[6]

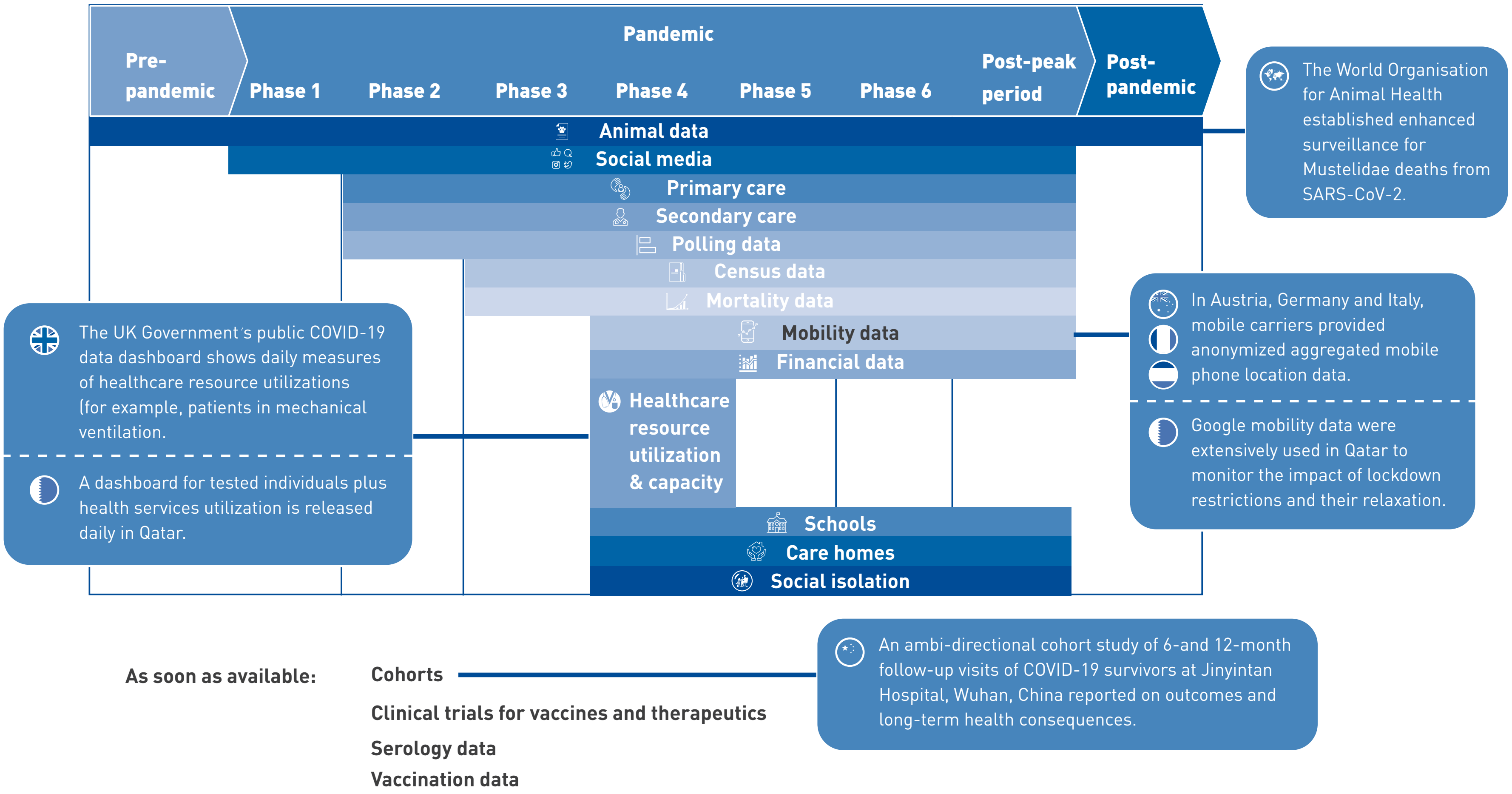
Table 1: Pandemic phase descriptions for influenza

PANDEMIC PHASE DESCRIPTIONS	
	DESCRIPTION
PHASE 1	No animal influenza virus circulating among animals has been reported to cause infection in humans.
PHASE 2	An animal influenza virus circulating in domesticated or wild animals is known to have caused infection in humans and is therefore considered a specific potential pandemic threat.
PHASE 3	An animal or human-animal influenza reassortant virus has caused sporadic cases or small clusters of disease in people, but has not resulted in human-to-human transmission sufficient to sustain community-level outbreaks.
PHASE 4	Human-to-human transmission of an animal or human-animal influenza reassortant virus able to sustain community-level outbreaks has been verified.
PHASE 5	The same identified virus has caused sustained community level outbreaks in two or more countries in one WHO region.
PHASE 6	In addition to the criteria defined in Phase 5, the same virus has caused sustained community level outbreaks in at least one other country in another WHO region.
POST-PEAK PERIOD	Levels of pandemic influenza in most countries with adequate surveillance have dropped below peak levels.
POSSIBLE NEW WAVE	Level of pandemic influenza activity in most countries with adequate surveillance rising again.
POST-PANDEMIC PERIOD	Levels of influenza activity have returned to the levels seen for seasonal influenza in most countries with adequate surveillance.

Source: Reproduced from WHO (2007)[7]

Figure 2 provides a summary of relevant data sources that can potentially be helpful to support policy responses to various phases of the COVID-19 pandemic, together with illustrative examples of how these data have been deployed.

Figure 2: Data sources and examples of their uses throughout COVID-19



Sources: Devaux et al (2021)[8], Knight et al (2020)[9], Huang et al (2021)[10], UK HSA[11]

SECTION 3: EXAMPLES OF COUNTRIES THAT HAVE DEVELOPED THEIR DATA INFRASTRUCTURES AND ANALYTIC CAPABILITIES

While most countries have, to some extent, developed their pandemic data response capabilities, a few have disproportionately contributed to the discovery of policy-relevant insights during the COVID-19 era. Examples of such countries include Iceland, Israel, Qatar, Scotland and Taiwan. Boxes 1 to 3 provide short case studies of aspects of the data infrastructure developed by some of these countries.

Box 1: Scaling up data capabilities to inform the pandemic response in Qatar

Context

- Since 2011, Hamad Medical Corporation (HMC) and the Primary Health Care Corporation (PHCC), Qatar's principle public healthcare providers, have maintained a single electronic health record (EHR) across 13 hospitals and 27 primary healthcare facilities.
- Qatar has used its advanced national, centralized eHealth system to execute evidence-based public health responses to the pandemic.

Approach

- HMC – the designated national provider for patients with COVID-19 – compiles a centralized and standardized national Severe Acute Respiratory Syndrome coronavirus 2 (SARS-CoV-2) reverse transcription polymerase chain reaction (RT-PCR) testing, hospitalization, and immunization database.
- This database includes:
    - basic demographic information on all residents,
    - vaccination records for the entire population,
    - Information on RT-PCR testing in the country, including for those suspected to have a SARS-CoV-2 infection as well as traced contacts, and

- COVID-19 hospital admission data, with a WHO severity classification for each identified case.

- Scientific analyses of the surveillance and outbreak data are used to power infection transmission models that, in turn, help to monitor and predict epidemiologic trends in the country. This has allowed for the real-time estimation of key indicators, including the incidence of infection, severe and critical infections, disease mortality and basic reproduction number.

Impact

- These key indicators were reviewed and validated by a national scientific committee – the Scientific Reference and Research Task Force (SRRT) – and used to inform major public health policy decisions, such as predicting the earliest date possible for the easing of public health restrictions.
- The EHR database includes information for each resident in Qatar, which has a population of about 2.8 million people. During the COVID-19 pandemic, more than 10.5 million tests were recorded.

Lessons learned

- Policy decisions should be guided by scientific knowledge to achieve the best outcomes for any given population.
- Science-based, data-driven decision-making in Qatar during the COVID-19 pandemic facilitated social and economic stability, and helped to minimize economic losses.
- Robust data systems are essential to all health systems. They allow countries to generate new knowledge related to the epidemiology of pandemics and the efficacy of vaccines, which helps to guide effective policy responses.

Box 2: Scaling up Scotland's national pandemic data infrastructure

Context

- Scotland developed a national pandemic surveillance and evaluation platform – Early Assessment of Antiviral and Vaccine Effectiveness (EAVE) – in response to the novel influenza A (H1N1) pandemic in 2009–10.[12]
- The system linked primary care, testing, vaccination, hospitalization and mortality data on about 250,000 people (5 percent of the population) managed through Public Health Scotland, Scotland's national public health agency.
- EAVE was put into hibernation following the end of the H1N1 pandemic.

Approach

- Following the emergence of SARS-CoV-2, EAVE was repurposed to create EAVE II and scaled up to cover nearly the entire Scottish population of 5.4 million people.[13]
- EAVE II brought together primary care, testing, sequencing, vaccination, hospitalization, intensive care unit (ICU), mortality and a host of other data sets into Public Health Scotland.
- Data sets are securely linked using Scotland's unique identifier, the Community Health Index (CHI) number.

Impact

- EAVE II is one of the world's few national, end-to-end, near real-time COVID-19 data platforms.
- The EAVE II team were the first in the world to demonstrate the real-world effectiveness of first dose COVID-19 vaccines in preventing COVID-19 hospitalizations.[14]
- They have produced many other major analyses, including demonstration of increased severity of COVID-19 outcomes associated with Delta infection and reduced severity of COVID-19 outcomes associated with Omicron infection.[15,16]

- The EAVE II team were also able to demonstrate vaccine waning and characterize individuals at risk of serious outcomes post first, second and booster vaccine doses.[17]

Lessons learned

- Repurposing EAVE to create EAVE II was a time-consuming process, as information governance procedures were not fit-for-purpose in the context of a pandemic.
- The lack of trained data processors and analysts with permissions to access these data increased these challenges.
- Going forward, it is crucial that data infrastructure and associated capabilities are maintained and updated to allow more rapid responses to future pandemics.

Box 3: Reactivating Taiwan's surveillance monitoring capabilities

Context

- The Taiwan Communicable Disease Control Act (2007) was passed four years after the Severe Acute Respiratory Syndrome (SARS) outbreak; this waived data authorization and consent requirements in the context of an emerging infections disease.[18]
- Taiwan has an existing national health insurance database.[16]
- The Taiwan Centers for Disease Control (TCDC) Epidemic Intelligence Center has an existing automated early warning pandemic surveillance system (National Notifiable Disease Surveillance System), which includes automated analyses of a range of data sources.[19]

Approach

- On 20 January 2020, TCDC activated the Central Epidemic Command Center.[20]

- By 27 January 2020, the national health insurance database was linked with immigration and customs databases. Those at high risk of contracting SARS-CoV-2 infection were tracked through their mobile phones.[21]

- Patients' travel histories were made available to hospitals, clinics and pharmacies.[19]

- Passive mobile phone geolocation data was used, among other sources, for contact tracing (for example, the Diamond Princess Cruise Ship).[16]

Impact

- Taiwan was able to rapidly reactivate its pandemic plans, including provision to use mobile phone data to support surveillance efforts.
- These data were used to support Taiwan's zero-COVID policy. For the first two years of the pandemic, this policy was effective in containing transmission, leading to a low number of cases, hospitalizations, and deaths.

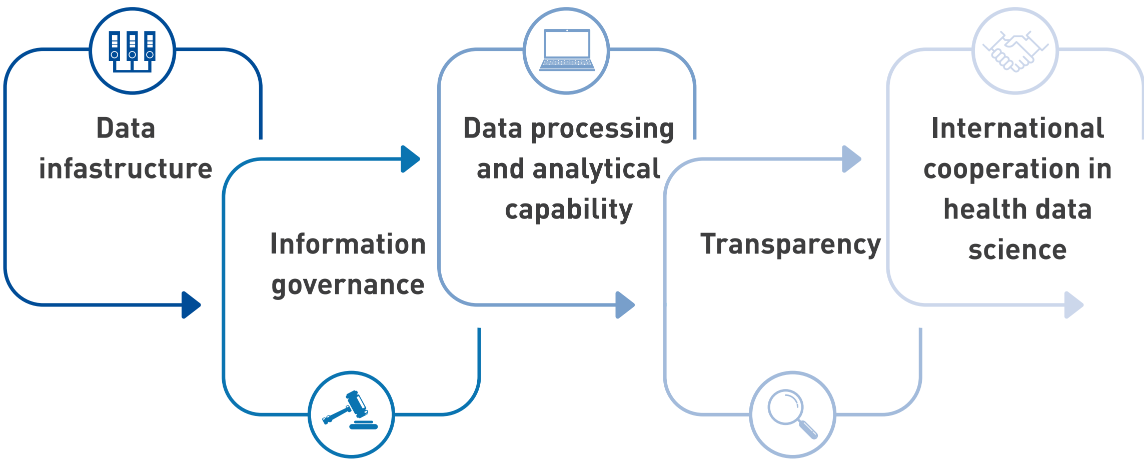
Lessons learned

- It is important to enact legislative changes that may prove helpful in the context of epidemics/pandemics as part of national pandemic plans.
- Safeguards are needed to ensure that these data are not used outside of exceptional circumstances.

SECTION 4: CATALYZING DATA-ENABLED RESPONSES

Having availability of data sets along the lines considered above is fundamental, but insufficient to ensuring capacity for data-enabled policy responses to pandemics. Also needed are permissions to access data by different stakeholders – ideally co-ordinated and granted by a national scientific committee – and the ability to curate, link, analyze, visualize, interpret and communicate these data to government bodies, policymakers, health system leaders and other audiences, often across national boundaries (see Figure 3). These are each time-consuming steps, but time is one luxury not available in the context of the exponential growth of infections seen in pandemics. It is therefore crucial that due attention is given to the data infrastructure and pipeline as part of countries' national pandemic preparedness plans.

Figure 3: Factors to facilitate data-enabled responses



Data infrastructure

There is a need to access disparate data, including EHRs, travel, and other health-related data on every person in the country, in as close to real-time as possible. Key data sets could be stored in a central secure warehouse – for example, as is the case for Qatar (see Box 1). In turn, this requires the need for adequate computational power, which can be substantive when dealing with millions of rows of data. Bringing together these disparate data sets can be done through deterministic or probabilistic approaches; where possible, this is most efficiently achieved using unique identifiers.[22,23] An alternative approach is to leave data *in situ* and deploy a service orientated architecture (SOA) approach, which creates interfaces between disparate data sets through application programming interfaces (APIs). This requires upfront engineering costs, but offers the potential for periodic synchronized updates and accompanying substantial reductions in downstream resource demands.

Information governance

Access to health and other sensitive data needs to be carefully regulated.[24] Countries will therefore have a variety of processes in place to ensure that data are not inappropriately used. These checks are typically extensive and time-consuming. However, the risk balance in providing access to these data needs to be shifted in the context of national and global emergencies such as pandemics. It is therefore important that the countries have agreed policies and plans in place. This may involve the need for special legislation. For example, Taiwan passed legislation to allow access to mobile phone data (see Box 3). The Control of Patient Information (COPI) notice was issued by the UK Government's Secretary of State for Health and Social Care to allow sharing of confidential patient information among healthcare organizations and other relevant bodies to safeguard public health.[25]

Data processing and analytical capability

Another key rate-limiting step in the ability to provide data-enabled insights is the lack of data processing and analytical capability. There is a need for trained staff who are ideally familiar with the data sets in question who can, at pace, check, clean, link, analyze and help visualize data for policy (and other) audiences. These steps require staff with a range of skills working in concert with each other.[26] Taking the time to develop, for example, a data dictionary and the sharing of source code can greatly increase efficiency of analysis and transparency of methods (see below).

Transparency

As ever, it is important that analyses are undertaken in transparent ways.[27] Key issues that need to be addressed include the opportunity for exploratory analyses. For example, it was unclear during the early stages of the pandemic which variables would be most useful to identify patients at greatest risk of poor COVID-19 outcomes. Other key issues include: reporting meta-data; specifying a statistical analysis plan (SAP) in advance and making this publicly available; making source code available through a repository such as GitHub; and where possible, making actual or synthetic data available to facilitate replication and validation studies. While the immediate need is to provide insights to policymakers, there is considerable merit in also publishing analyses in preprints and peer-reviewed journals to allow independent verification of methods and share insights with the global community.

International co-operation in health data science

There are numerous instances where it is important to be able to run analyses across countries, regions or globally.[28] This is, however, difficult to achieve in practice because it is seldom possible to move sovereign data sets across national boundaries. This therefore requires federated analyses to be undertaken with some form of data synthesis. The most prominent example has been the Johns Hopkins Coronavirus Resource Center COVID-19 Testing Dashboard (see Case study 4).

Box 4: Johns Hopkins COVID-19 Testing Dashboard

Context

- In December 2019, a new coronavirus, SARS-CoV-2, was detected as spreading in Wuhan, China.
- Ensheng Dong – a Chinese graduate student at Johns Hopkins University, Baltimore, USA, who had studied epidemics – was concerned about the pandemic's effect on his home country and family.[29]
- He consulted with his advisor, Professor Lauren Gardner, and they decided this virus needed "a closer look".

Approach

- Dong had previous experience in creating a dashboard for visualizing measles risk in the US using a geographic information system (GIS), and so used the same model for COVID-19.
- Initially the dashboard was manually updated twice a day with data on China, but quickly expanded to include worldwide data.[30]
- The dashboard was interactive and provided near real-time data to track and visualize the location of cases of COVID-19, deaths and recovery.
- In February 2020, Esri's Living Atlas of the Worlds team automated the task of importing data from China (that is, data scraping), and a team of volunteers was recruited from Johns Hopkins University to update and maintain the site.

Impact

- The dashboard has now evolved to a multi-layered resource that provides expert analyses and graphics.
- The dashboard map was first shared publicly on 22 January 2020, and provided researchers, public health and the general public near real-time information on the unfolding pandemic.[27]
- The dashboard reports cases at the province level in China; at the city level in the US, Australia, and Canada; and at the country level in other regions.
- The Johns Hopkins team identified most newly infected countries ahead of WHO.[32]
- *TIME* magazine recognized this as the "go-to data source" for COVID-19 and named it in the Top 100 Inventions of 2020.[31]

Lessons learned

- The quick thinking and prompt action of a graduate student led to the creation of an invaluable global data resource.
- A team science-based approach was essential to the rapid scaling-up of this effort.

Other examples include analyses of data across UK nations investigating the impact of lockdown measures on health system functioning, and investigation of rare vaccine safety signals (for example, cerebral venous sinus thrombosis).[33,34] These have also been examples of cross-continental analyses investigating the impact of variants of concern (Gamma in Brazil and Delta in Scotland) on disease severity and waning of vaccine effectiveness[35] and work undertaken across more than 40 countries through the International COVID-19 Data Alliance (ICODA) to investigate the impact of lockdown measures on perinatal outcomes.[36,37]

CONCLUSIONS

Ready access to high-quality data is fundamental to generating effective evidence-informed policy responses to pandemics. Most countries have struggled with this. A few countries have, however, excelled in offering pointers to how this capacity and capabilities might be developed and deployed in other countries. We have identified a set of policy recommendations, which should help countries develop their capabilities in this respect. There is also the need to enable effective analyses that extend across international boundaries, which is most likely to be achieved through federated analytical approaches.

In Box 5, we offer some key policy recommendations to catalyze the ability for countries to develop their data science capabilities.

Box 5: Key policy recommendations

1. Development of the underlying data infrastructure, governance, and analytic capacity to provide data for policymaking and respond to pandemics should be a core component of countries' national pandemic preparedness plans.
2. Most countries have enhanced their data capabilities in some ways over the course of the COVID-19 pandemic. It is crucial that this capacity is not allowed to regress; rather, nations should seek to further develop their capabilities. Repurposing COVID-19 data capabilities to help to respond to other major health concerns, such as influenza, pneumonia, cancer, cardiovascular disease and mental health, will help ensure that capabilities are maintained, allowing for rapid redeployment in the context of any future epidemics/pandemics.
3. Through identifying the range of data sources that can prove useful during various stages of an epidemic/pandemic, nation states should now reflect on key data gaps and prioritize and proactively develop strategies to plug these gaps as part of their data roadmaps.
4. Being able to securely link data sets greatly increases the range of questions that data can help answer; countries with unique identifiers are able most efficiently to link disparate data sets, and therefore, the maintenance and strategic extension of unique identifiers should be prioritized. In countries where unique identifiers do not exist, these should be considered a policy priority.

5. Bringing together disparate data sets on entire populations is challenging from the perspectives of computational ability, information security and governance, and the human capacity needed to process, link, analyze and interpret these data. Developing capacity and capabilities should be a central component of national data and workforce strategies.
6. Countries should proactively review their legislative frameworks governing the use of health data and should have provisions in place to expedite permissions for the use of health and health-related data in exceptional circumstances such as pandemics.
7. It is vitally important that public trust is maintained. A national commitment to transparency about access to and uses of data is crucial. The 'Five Safes Framework' – that is, safe people, safe projects, safe settings, safe data and safe outputs – is an example of a potentially effective approach.
8. There are areas where a handful of countries have been able to make substantive progress in their capacity to generate data-enabled responses to the COVID-19 pandemic. These are the easily achieved goals that many other countries can examine over the coming years. Direct dialogue with the policy teams in these data frontier countries is to be encouraged with a view to identifying potentially transferable lessons.
9. There has thus far been relatively little policy attention as to how to enable citizen data science initiatives in the context of the pandemic. Given the abundance of publicly available data, opportunities to generate data at scale through smartphones and other devices, and considerable public interest to support pandemic responses, this is something of a missed opportunity.
10. Finally, pandemics are, by their nature, international. It is vital therefore that there are effective mechanisms for the efficient sharing of data and data-enabled insights between countries and regions. Given the sovereign nature of national data assets, this will most likely be achieved through federated approaches to data analysis where data remain within national jurisdictions. There is an urgent need to develop global capacity in undertaking federated analyses.

ACKNOWLEDGMENTS

We would like to thank Sultana Afdhal and Zakariya Sheikh for their thoughtful comments and contributions to earlier drafts of this paper.

REFERENCES

1. World Health Organization. WHO Coronavirus (COVID-19) Dashboard. Available at: https://covid19.who.int/ [Accessed: 12 July 2022].

2. UNDP. COVID-19: Socio-Economic Impact. United Nations Development Programme (UNDP). https://www.undp.org/coronavirus/socio-economic-impact-covid-19?utm_source=EN&utm_medium=GSR&utm_content=US_UNDP_PaidSearch_Brand_English&utm_campaign=CENTRAL&c_src=CENTRAL&c_src2=GSR&gclid=Cj0KCQjwwJuVBhCAARIsAOPwGASe-Gr_rRBY1RZ15TGMOww7dfzOWxh8SCjyFqiSrX5svHWelo37H1caAphgEALw_wcB [Accessed 12 July 2022].

3. Wellcome. From equality to global poverty: The Covid-19 effects on societies and economies. Wellcome 29 June 2021 https://wellcome.org/news/equality-global-poverty-how-covid-19-affecting-societies-and-economies [Accessed 12 July 2022].

4. Ioannidis JPA. A fiasco in the making? As the coronavirus pandemic takes hold, we are making decisions without reliable data. Stat News, 17 March 2020. Available at: https://www.statnews.com/2020/03/17/a-fiasco-in-the-making-as-the-coronavirus-pandemic-takes-hold-we-are-making-decisions-without-reliable-data/ [Accessed 12 July 2022].

5. Downey A. Poor data flows 'throttled' timely response to Covid-19. digitalhealth, 14 January 2021. Available at: https://www.digitalhealth.net/2021/01/poor-data-flows-throttled-timely-response-to-covid-19/ [Accessed 12 July 2022].

6. World Health Organization. *Pandemic Influenza Preparedness and Response: A WHO Guidance Document;* 2010. Available at: https://apps.who.int/iris/bitstream/handle/10665/44123/9789241547680_eng.pdf?sequence=1&isAllowed=y [Accessed 12 July 2022].

7. WHO. *Pandemic Influenza Preparedness and Response: A WHO Guidance Document.* World Health Organization (WHO) Global Influenza Programme. Available at: https://apps.who.int/iris/bitstream/handle/10665/44123/9789241547680_eng.pdf?sequence=1&isAllowed=y [Accessed 12 July 2022].

8. Devaux C et al. Spread of mink SARS-CoV-2 variants in humans: A model of sarbecovirus interspecies evolution. *Frontiers in Microbiology,* 12; 2021.

9. Knight SR et al. . Risk stratification of patients admitted to hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: Development and validation of the 4C Mortality Score. *BMJ.* 2020; Sep 9;370:m3339.

10. Huang L et al. 1-year outcomes in hospital survivors with COVID-19: A longitudinal cohort study. *The Lancet*. 2021; Volume 398, Issue 10302, P747-758.

11. UK Health Security Agency. Healthcare in United Kingdom. Available at: https://coronavirus.data.gov.uk/details/healthcare [Accessed 12 July 2022].

12. Simpson C. Early estimation of pandemic influenza Antiviral and Vaccine Effectiveness (EAVE): Use of a unique community and laboratory national data-linked cohort study. *Health Technology Assessments*, 2015;19(79). Available at: https://www.journalslibrary.nihr.ac.uk/hta/hta19790 [Accessed 12 July 2022].

13. Simpson C et al. Early pandemic evaluation and enhanced surveillance of COVID-19 (EAVE II): Protocol for an observational study using linked Scottish national data. *BMJ Open*. Volume 10, Issue 6. Available at: https://bmjopen.bmj.com/content/10/6/e039097 [Accessed 12 July 2022].

14. Vasileiou E et al. Interim findings from first-dose mass COVID-19 vaccination roll-out and COVID-19 hospital admissions in Scotland: A national prospective cohort study. *The Lancet*. 2021; May 1;397(10285):1646-1657.

15. Sheikh A et al. Public Health Scotland and the EAVE II Collaborators. SARS-CoV-2 Delta VOC in Scotland: Demographics, risk of hospital admission, and vaccine effectiveness. *The Lancet.* 2021; Jun 26;397(10293):2461-2462.

16. Sheikh A et al. Severity of omicron variant of concern and effectiveness of vaccine boosters against symptomatic disease in Scotland (EAVE II): A national cohort study with nested test-negative design. *Lancet Infectious Diseases*. 2022; Jul;22(7):959-966.

17. Katikireddi SV et al. Two-dose ChAdOx1 nCoV-19 vaccine protection against COVID-19 hospital admissions and deaths over time: a retrospective, population-based cohort study in Scotland and Brazil. *The Lancet*. 2022; Jan 1;399(10319):25-35.

18. Chen CM et al. Containing COVID-19 among 627,386 persons in contact with the Diamond Princess cruise ship passengers who disembarked in Taiwan: Big Data Analytics. Journal of Medical Internet Research. Vol 22, No 5 (2020): May. Available at: https://www.jmir.org/2020/5/e19540 [Accessed 12 July 2022].

19. Jian SW et al. Real-time surveillance of infectious diseases: Taiwan's experience. *Health Securities*. Mar/Apr 2017;15(2):144-153.

20. Taiwan Centers for Disease Control. Taiwan CDC announces activation of Central Epidemic Command Center (CECC) for Severe Special Infectious Pneumonia to comprehensively prevent novel coronavirus pneumonia outbreak in China and ensure health of Taiwanese public. Available at: https://www.cdc.gov.tw/En/Category/ListContent/tov1jahKUv8RGSbvmzLwFg?uaid=vmv22PiH7-k3K-yh6FkmKw [Accessed 12 July 2022].

21. Wang CJ et al. Response to COVID-19 in Taiwan Big Data Analytics, New Technology, and Proactive Testing. *JAMA*. 2020;323(14):1341-1342. Available at: https://jamanetwork.com/journals/jama/fullarticle/2762689 [Accessed 12 July 2022].

22. Lewis S. deterministic/probabilistic data definition. TechTarget. August 2019. Available at: https://www.techtarget.com/searchdatamanagement/definition/deterministic-probabilistic-data#:~:text=While%20deterministic%20data%20is%20consistent,get%20the%20most%20valuable%20insights [Accessed 12 July 2022].

23. DbBee. What is unique identifier and how to use it? Available at: https://www.dbbee.com/what-is-unique-identifier-and-how-to-use-it- [Accessed 12 July 2022].

24. NHS Research Authority. Data protection and information governance. 30 June 2021. Available at: https://www.hra.nhs.uk/planning-and-improving-research/policies-standards-legislation/data-protection-and-information-governance/ [Accessed 12 July 2022].

25. NHS Digital. Control of patient information (COPI) notice. Available from: https://digital.nhs.uk/coronavirus/coronavirus-covid-19-response-information-governance-hub/control-of-patient-information-copi-notice [Accessed 12 July 2022].

26. Allsopp R et al. Data-driven responses to COVID-19 lessons learned. *OMDDAC Research Compendium*, October 2021. Available at:https://www.researchgate.net/publication/355493241_Data-Driven_Responses_to_COVID-19_Lessons_Learned_OMDDAC_Research_Compendium [Accessed 12 July 2022].

27. NHS. How data is supporting the COVID-19 response. NHS. October 2021. Available at: https://www.nhsx.nhs.uk/covid-19-response/data-and-covid-19/how-data-supporting-covid-19-response/ [Accessed 12 July 2022].

28. Yang L et al. Big data driven COVID-19 pandemic crisis management: Potential approach for global health *Archives of Medical Science.* 2021; 17(3): 829–837.

29. Milner G. COVID-19: Inside look at the Johns Hopkins Dashboard, keeping tabs on the virus. Esri Blog, 16 July, 2020. Available at: https://www.esri.com/about/newsroom/blog/how-researchers-built-johns-hopkins-dashboard/ [Accessed 12 July 2022].

30. Center for Systems Science and Engineering, Johns Hopkins University. COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University. Available at: https://github.com/CSSEGISandData/COVID-19 [Accessed 12 July 2022].

31. Rosen, J. TIME names Johns Hopkins Coronavirus Resource Center a top invention of 2020. Johns Hopkins HUB. Available at: https://hub.jhu.edu/2020/11/19/time-coronavirus-resource-center-top-100 [Accessed 12 July 2022].

32. Dong E et al. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infectious Diseases.* 2020 May;20(5):533-534. Available at: https://pubmed.ncbi.nlm.nih.gov/32087114 [Accessed 12 July 2022].

33. Shah SA et al. Impact of first UK COVID-19 lockdown on hospital admissions: Interrupted time series study of 32 million people. *eClinicalMedicine.* 2022; Jul;49:101462.

34. Kerr S et al. First dose ChAdOx1 and BNT162b2 COVID-19 vaccinations and cerebral venous sinus thrombosis: A pooled self-controlled case series study of 11.6 million individuals in England, Scotland, and Wales. *PLOS Medicine*. 2022; Feb 22;19(2):e1003927.

35. Katikireddi SV et al. Two-dose ChAdOx1 nCoV-19 vaccine protection against COVID-19 hospital admissions and deaths over time: A retrospective, population-based cohort study in Scotland and Brazil. *The Lancet.* 2022; Jan 1;399(10319):25-35.

36. ICODA. A globally coordinated, health data-led research response to tackle the COVID-19 pandemic. International COVID-19 Data Alliance (ICODA). Available at: https://icoda-research.org/#homepageContent [Accessed 12 July 2022].

37. Stock S et al. The international Perinatal Outcomes in the Pandemic (iPOP) study: Protocol. Wellcome Open Research; 2021. Available at: https://wellcomeopenresearch.org/articles/6-21 [Accessed 12 July 2022].

ISBN 978-1-913991-18-0

9 781913 991180

www.wish.org.qa