BLAVATNIK
SCHOOL OF
GOVERNMENT

UNIVERSITY OF
OXFORD

WHAT WOULD A DATA FRAMEWORK FOR POLICY RESPONSES TO PANDEMIC DISEASES LOOK LIKE?

Clare Wenham, Thomas Hale, Kaitlyn Green, Bernardo Andretti di Melo,

Rodrigo Furst, Nadezhda Kamenkovich, Anna Petherick, Toby Philips,

Will Torness, Sandhya Laping, Hao Zha and Radhika Nagesh

WHAT WOULD A DATA FRAMEWORK FOR POLICY RESPONSES TO PANDEMIC DISEASES LOOK LIKE?

Oxford COVID-19 Government Response Tracker
January 2023

This scoping paper discusses how information on government policy responses to pandemic diseases (e.g. non-pharmaceutical interventions (NPIs) and public health and social measures (PHSMs), and including behavioural rules, testing and contact tracing systems, policies to incentivise vaccination, etc.) have, can, and should be collected, analysed, and incorporated into the broader array of pandemic data (e.g. epidemiological, virological, behavioural, etc.) to build preparedness. It draws on both the academic and policy literature, as well as a series of interviews with policymakers and researchers, as well as a guided stakeholder workshop held in December 2022[1].

The objective of this scoping exercise has been to precipitate a shared vision on what a framework for policy data might look like, as part of a broader pandemic data system. The Data Framework could be a shared project of governments, international organisations, researchers, and philanthropic bodies. It could outline:

- *What information is needed - a flexible set of essential indicators that could be customised to a given situation*
- *What form that information should be reported in*
- *What actors would be well positioned to collect and report such information across the different data streams*
- *A consideration of how to build trust across the different actors contributing data streams, as well as the population overall*
- *A plan for how to rapidly deploy the protocol in the face of a new pandemic, including a plan for building up the required capacity where it is lacking*

This document outlines next steps to be considered and/or instigated if such a framework is to be developed. Our goal is to maintain momentum and capture learnings from policy trackers during COVID-19, and to make sure that there is a process for collecting and analysing such data for decision makers during the next health emergency.

---

[1] We thank all participants who took the time to be interviewed, to share relevant policies and documentation, and those who took part in the December 2022 workshop, and reviewed earlier drafts of this scoping paper.

Contents

Introduction

Three years on from the start of the COVID-19 pandemic, one of the key lessons identified has been the need for good quality, reliable, rapid, and comparable data sources which can provide much needed evidence upon which governments can make decisions as to how to respond best to the outbreak. During different stages of the pandemic life cycle, there has been a need for a range of data, including epidemiological, virological, clinical, behavioural, and policy-related data, to name but a few categories. However, at the start of the pandemic, there was no cohesive mechanism to collect and organise the data that governments needed to inform decision making. In particular, data related to government policies—such as closure and restriction measures, testing and contact tracing, economic support, policies to incentivise vaccine uptake, etc., and how these in turn impacted both on disease transmission, and on the secondary effects of the pandemic on lives and livelihoods—presented a clear gap.

During the COVID-19 pandemic, a new, emergent real-time data ecosystem developed, combining a range of data types and sources. This was a hotchpotch of different official and non-official data sources and trackers, working in real-time to try and provide pertinent data and information to decision makers, researchers, and the public to inform policy development. Some of these were a development of already existing data management systems, such as TESSy; EMLY and DHIS2 ('Rapid Review of WHO COVID-19 Surveillance: External Review, 27 October 2021' n.d.), others were created in the early stages of COVID-19 to provide additional functionality, such as tracking policy responses, and their effects on disease transmission and societal impacts[2]. Some tracked immediate responses, others focused on the longer-term policy environment in which they took place, and some combined these (Katz et al. 2022). The emergence of this data ecosystem was largely un-coordinated, driven by actors outside of governments and intergovernmental organisations, and there were few efforts to provide a central fulcrum bringing together the range of data types and sources into a unified whole

---

[2] Examples of these include (but not limited to): Covid-19 Government Tracker (OxCGRT) https://www.bsg.ox.ac.uk/research/covid-19-government-response-tracker; Coronanet https://www.coronanet-project.org/; IMF Economic response tracker https://www.imf.org/en/Topics/imf-and-covid19/Policy-Responses-to-COVID-19

to allow meaningful consideration and interaction. One example of trying to unify these activities was the Global Database of public health and social measures applied during COVID-19 established between LSHTM and WHO to try and collate data from a range of trackers into a complete dataset with a standard structure and taxonomy (WHO, n.d.).

The mixed landscape for policy data created duplication, with parallel streams and data providers leading to a relatively chaotic, unsustainable whole. This is a result of a lack of coordination for any data systems and content prior to COVID-19, and the balkanisation of health related and health adjacent data streams across disciplines, institutions, sectors and, importantly, owners and users. While different types of data had more or less coordination, in general the data ecosystem was inchoate. This in turn made it harder for decision makers to collect and analyse data points alongside each other.

As we reflect on COVID-19, one key learning must be the need for real time data across a range of topics to inform policy development. Moreover, the COVID-19 pandemic demonstrated that public engagement with data and decision making during health emergencies is ever present. Societies and the public want to know what is happening during an epidemic to inform their own individual decision making, and to understand, and in turn accept, the interventions they see their governments implementing. This, in part, is a lasting legacy of the pandemic, and our learnings on participatory governance, that data sources ideally need to be open access, in as close to real time as possible, and digestible to all in a manageable format for the layperson.

However, the data points need to be contextualised, for the public and decision makers alike. For example, when examining key data sources which emerged during COVID-19, whether from national data sets, or from aggregators, such as the Johns Hopkins Tracker, raw numbers cannot be considered in a vacuum, but need to be understood alongside the policies that different governments had implemented to mitigate the spread of disease. For example, this includes non-pharmaceutical interventions, vaccination rates, and policies determining how widespread testing may be, and broader policy positions at national and local levels. Learnings from public health and social measures (PHSM) tracking, which many governments and

international organisations introduced and developed, needs to be systematically incorporated into or alongside the epidemiological and public health data ecosystem, so as to be in place ahead of the next health emergency, to provide this much needed corollary information.

As such, we propose a data framework that seeks to harness the range of learnings from COVID-19 and prior health emergencies and develop a ready-to-use process of collecting and collating data if/when another health (or other emergency) appears. While the overall data ecosystem necessarily includes a wide variety of data types, we focus here primarily on policy data, which we view as one key underdeveloped pillar of the broader data architecture.

To do so, we propose two potential structures for this framework:

1. **A hub and spoke model with WHO at the centre** of such an arrangement, with multiple research institutions collaborating alongside. As such WHO would be the convening power, and facilitate the technical guidance that a range of research institutions would use to work together in this data ecosystem project. WHO would facilitate the institutions to be part of the network, and would, in turn, coordinate the development the technical taxonomy for the data collection and analysis. However, the actual data collection and associated analyses would be carried out by the research institutions involved, who would be the owners of the data and associated analyses. Importantly, these research institutions must be global in geography and scope, and must ensure a component of capacity development for policy tracking in research institutions in low and middle income countries, noting that they otherwise may not have resources to do so, and ideally include at least one institution from each WHO region. In time, this process could be combined into, or could mirror the WHO observatory models for health systems and policies established elsewhere in the WHO (Euro Health Observatory, n.d.).

2. A group of willing / interested research institutions formulated into a **research consortium** (such as detailed above) which work cohesively, perhaps with one or two institutions which hold the leadership and management of the rest (and funding). As such, this would be a research institution led project, away

from government or member state organisation processes. This would allow for the greatest flexibility in the process of data collection, the mechanisms used, the most innovative data collection and methodological developments incorporated into projects. Such a model would have further benefits of independence (as highlighted in more detail below) but would run a risk of sustainability in the long term, and the challenge of routinely engaging with policy makers to determine what data might be most useful to collate.

Importantly, there are currently many efforts being taken elsewhere in the global health landscape to ascertain preparedness for virus emergence, or the capacity of health systems to be able to manage a health emergency, such as the National State Parties Assessment of Readiness (NSPAR), the Joint External Evaluation (JEE) and the proposed Universal Health Preparedness Review (UHPR) all occurring within the World Health Organization (WHO). Further academic assessments such as the Global Health Security Index (GHSI), aim to consider what is needed when an outbreak emerges, and what data governments and international organisations need to make decisions as the epidemic develops. Response, and in particular data on the effects of policy decisions, remains an underdeveloped part of the puzzle, so this framework could be a dynamic organism ready to provide data as soon as possible for decision-making within government responses. However, we can learn from these preparedness processes as to what works for data collection and good governance – for example, the challenges of self-reporting through the SPAR assessment tool and the unreliability of data; the top-down prescriptive considerations of the JEE, and the dispersed data challenges of the GHSI (Kentikelenis and Seabrooke 2021).

It is important to start this process of creating a policy data -ecosystem now, so that the protocol can be developed in (relative) "peace time" and trust can be built in the system between participants, with the system itself, as well as the end users, for rapid engagement during an emergency. It would be pertinent for this to be trialled or to demonstrate utility in other health emergencies beyond Covid-19 to show its value.

Role and needs of end users

We identify four main categories of users for pandemic data, including policy response data.

1. **Policymakers**. The starting point for consideration of a data framework protocol is to understand what decisions need to be made during each stage of the epidemic life cycle by decision makers, especially policymakers (primarily national decision makers, but also regional and multilateral organisations). We assume that the key decision makers to be engaged are those in leadership positions in national governments and, in some jurisdictions, at the subnational level, with support from regional and multilateral organisations. As COVID-19 demonstrated, sovereign states remained the key determinant and implementer of decisions and policies to respond to the pandemic. We did interview policymakers as part of this process, but there is a variety of data needs within different governments/ministries and thus there is no cohesive answer as to what the policy needs are.

   Understanding what questions decision makers need answers to can then reveal what data they need to make these decisions. As we saw during the pandemic, often the availability (or lack thereof) of data and analysis drove the questions asked and the decisions made, instead of the other way around.

   Importantly for this data framework, alongside the epidemiological, clinical and virological data which must play an important role in how governments make decisions, policy and related social data must be collated and coordinated more holistically. This was a secondary thought for many governments in the early phases of COVID-19, and it should be considered in the initial stages of a future health or other emergency, to be able to contextualise differing outcomes.

   Anecdotally, over the COVID-19 pandemic, we observed several use cases for OxCGRT policy data within governments and intergovernmental organisations. The primary use case we are aware

of was the use of NPI data for real-time monitoring of comparator jurisdictions. This ranged from data dashboards for ministers, to established working groups (such as the UK Cabinet Office's International Comparators Joint Unit) or the summaries published by the IMF (on economic responses) or the UNDP (on vaccination strategies). We also observed several instances where country leaders used global NPI data to communicate to their citizens and contextualise their own response (eg. prime ministers in India, New Zealand, and Spain all did this).

However, whilst many key decisions are made at the national level, national level data may in and of itself not be sufficient for decision makers to consider the options available to them in their policy toolbox. Looking at global data sources, through a global framework for data collection and sharing this would allow 1) for greater power of the data through multiple sources 2) to be able to see the effects of different interventions that governments have implemented 3) to allow for comparative research and analysis to be conducted by researchers to inform decision makers thoughts. Thus, a global data framework can serve as a public good, even if information will also need to be contextualised and adapted for many applications.

2. **Researchers**. Researchers comprise both end users of the data, and the data collectors, analysts and catalysts for the success of such an enterprise. It was research groups that developed the range of policy trackers which were established during COVID-19. Moreover, it was their contextualisation and analysis of the data which informed decision makers as to the effects of different policy interventions, from both epidemiological and socio-economic perspectives.

Policy data from the OxCGRT and other projects has found multiple use cases in helping researchers understand and interpret pandemic responses. From analysing the effectiveness of policies on improving epidemiological outcomes (eg. Mendez-Brito et al. 2021, Hale et al. 2021, and Oliu-Barton et al. 2021) to understanding the factors that

shape the responses in the first place (eg. Toshkov et al. 2022, and Phillips et al. 2021). Researchers have also used policy data to understand how pandemic policies impact a broad range of things from human behaviour (eg. Petherick et al. 2021 and Kim et al. 2021), economic outcomes (eg. Miguel & Mobarak, 2022), crime (Nivette et al. 2021), health services (eg. COVIDSurg Collaborative, 2021), and energy systems (eg. Kikstra et al, 2021).

Any protocol for a standing effort for a policy data must consider how best to engage researchers as those who will carry out the majority of the work required to facilitate such a mechanism. Researchers will need to develop the taxonomy, develop the system, collect the data, clean the data, interpret the data, offer analysis, and, where invited, provide guidance to policymakers about the significance of their data. Alongside this, researchers will want to publish the findings from the data. To do so, the data generated by such a framework or protocol must be of high quality and attain to best practice academic ethical and methodological processes. Thus, the process section below is of utmost importance to engage researchers to the best extent as possible.

3. **Publics (including advocacy groups and the private sector)**. The COVID-19 pandemic also highlighted the importance of open access information and the need to share data and analyses with the public in as close to real time as possible. Such transparency is a component of good governance, and can also support adherence to public health measures if the public feel they are able to understand the reality of any particular health emergency, and the risks that they face. Readily available data contribute to this. Advocacy groups alongside the public are also an important stakeholder as they can support public health interventions, but need data to be able to decide how best to offer support to the communities or groups they serve. Finally, trade and industry actors also find policy data helpful to support their own planning and response efforts.

Over the course of COVID-19 there was significant demand from the public for reliable and timely data about the pandemic, including policy data. Two key drivers of this were communicators (including journalists as well as educators and commentators) and businesses. In particular, policy data was used thousands of times in media coverage of the pandemic, and was integrated into popular science communication tools like Our World in Data. here was also significant usage in the private sector, where the data was used in industries as diverse as insurance, aviation, and management consulting.

The relationship between these actors is displayed in figure 1, which sketches the role that data plays in informing research, public sentiment, and decision-making.

Figure 1: The role of data in the pandemic response ecosystem
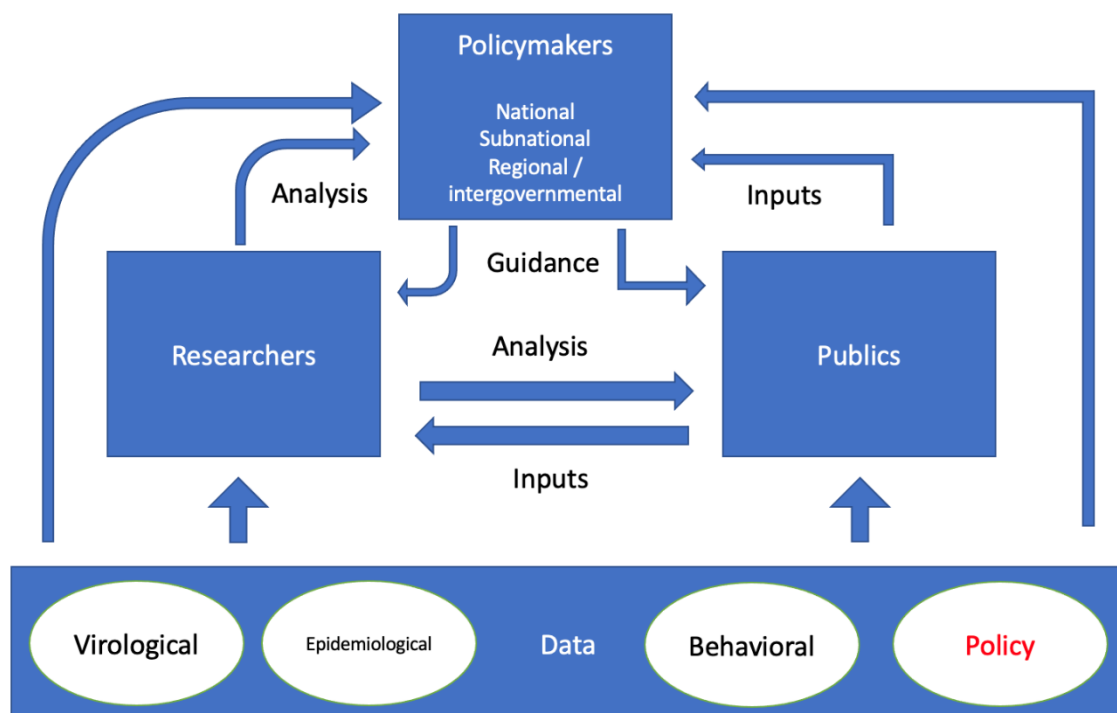
Scope of indicators

Although we are unable to answer conclusively as to what data decision makers might need – we assume that at a minimum the following data/indicators needs may be vital:

- Epidemiological
  - Diagnostic data
  - Vaccination status and distribution
- Virological
  - Spillovers
  - Variants
- Genomic
- Observational
- Animal Health Data
- Clinical + electronic health records
- Behavioural
- Climate
- Consumption
- Deforestation
- General policy data which may be pertinent for policy tracking:
  - Available financial, human and medical resources
  - Social factors (vulnerable groups, ethnicity, gender)

- Legal status of countries in terms of health emergencies legislation (Georgetown currently doing)
- Migration
- Supply chain (PPE, pharmaceuticals)
- Health system resilience
- Hospital capacity and agility

- Vaccination strategies
- Testing policies
- Mobility (domestic + international)
- School and workplace closures
- Social distancing
- Income support / fiscal measures
- Public information campaigns

The difference between policy data streams is that the more epidemiological and clinical data is routinely collected anyway as part of medical and public health practice, within most countries health systems, and thus when an outbreak emerges, the systems for data collection and analysis are in place. They might need to be

scaled up as appropriate, but they exist. For policy data this is not true. Whilst there are research groups who analyse policy and policy interventions in public health and governance[3], this is often project linked, and not systematically and routinely collected by government institutions for continuous monitoring. We think there should be routine data collection and analysis for such policy data needs, both for longer-term policy "baselines" and more immediate responses. Indeed, retrospective analysis should continue about the policy interventions undertaken during COVID-19. At the very least the process and protocols should be established for how such a process could be developed so that it can be rapidly deployed. Moreover, as the COVID-19 pandemic has highlighted, this policy data is vital to understand disease dynamics, and such should not be considered as a "nice to have" in a health emergency, or given less weight in the data needs request, but should be mainstreamed into data collection and analyses processes for pandemic preparedness and response.

Principles

A starting premise is that a global data framework which requires alignment and collaboration between multiple sources from national institutes and governments, international organisations and non-governmental sources such as academia, and civil society groups must be bound by certain principles of good governance for data sharing. Such a platform will only work with the collaboration and cooperation of those who create and generate the data, those who analyse it, and those who use it, and they will only do so if they feel at least some degree of recognition and ownership. Moreover, the process must be established within an environment of solid research ethics, that all involved can trust and believe in the data and the processes that are used to generate and collect these data points. As such, principles underpinning a data framework include:

> 1. Transparency of utilisation and process

---

[3] For example, The Global Health Security Index (GHSI) tracks policies for pandemic preparedness and response on a bi-annual basis, which includes policy relevant data https://www.ghsindex.org/; The International Disease Events Analysis (IDEA) tracks a range of pandemic preparedness policies and tools https://ghssidea.org/ ; the COVID AMP database https://covidamp.org/

2. Acknowledgement and active engagement of data creators

3. Meaningful collaboration between actors providing, managing and using the data

4. Access to all data points at the same time for all users

5. Benefits from the data to be derived/available for all user at the same time, but particularly those who shared the data in the first place (Littler et al. 2017)

6. Usability/Simplicity so as to provide useful data in as real time as possible

Process and operationalisation

There are a whole range of process challenges which will prevent or hinder the development of a data framework, and whose practicalities need to be considered. These include:

Standardisation: One key challenge identified universally by interviewed participants was the challenge of standardisation of data sources that emerged in the ad hoc manner during the COVID-19 pandemic. A global platform requires a consistent understanding of the data points that comprise it. For example, do case definitions for epidemiological data match, what does vaccinated mean (e.g. does it refer to a particular vaccine brand and/or number of boosters?), etc. In the realm of policy data, standardisation is complex as government policies, though broadly similar, often have important nuances. Policy tracking thus faces a key trade-off between precision and comparability, as well as depth and breadth. Ideally these standards should be set in the initial weeks of a health event or before, and rolled out across jurisdictions, but if this is not possible, then it is imperative that different sources provide clear definitions as to what their data refer to, and the dates of when any changes were made (and what these were).

Scalability: The challenge particularly with different case definitions and/or NPI data is the lack of capacity to scale up the data collection without considerable human resource. During COVID-19, activities such as the OxCGRT and Canadian Institute for Health Information (CIHI) relied on manual assessment of source materials generated by different jurisdictions. This makes it time consuming and resource

heavy to scale up to global levels and the scale down to a high degree of local granularity. Standardisation could support scalability, as well as novel forms of crowd-sourcing, citizen science, and algorithmic data collection methods.

Taxonomy: Particularly important for policy data, is a consistent taxonomy of what different definitions and terms mean and refer to. This qualitative description should be as detailed as possible to support interoperability across different systems and data sources. However, it also needs to be flexible to be able to expand as a health event emerges or develops, so that it does not change previous work or data if modified, such as through additional fields that allow further detail.

Metadata: Consideration must be given for:

- ○ Start date/end date of policy,
- ○ Geographical reach/scope,
- ○ Source/reference to the original location of policy,
- ○ Date of policy revision, level of governance of policy,
- ○ Proportion of population affected by policy

Each of these (and other potential metadata points) are vital analysing policy interventions so as to be able to assess the impact of any particular non-pharmaceutical intervention. Moreover, a feedback loop to encourage collection of data or analyses between end users (particularly decision makers) and those who are compiling the taxonomy and the data points collected.

Back end interoperability is vital for drawing together and compiling discrete data sources. This is only possible if/when there are meaningfully shared taxonomies and standardisations of data to be able to allow interaction. Lessons can be learned from COVID and the use of Application Programming Interfaces (APIs). Moreover, we?need to distinguish between the front-end operability of data entry from a range of locations, and the back end data platform which unifies the data streams as they arrive to develop a workable dataset from multiple sources, to reduce workload of the data platform. This will require clear data and meta data standards (WHO 2021).

Deployability: The key to mitigating the spread of an emerging infectious disease is rapid detection, rapid decision making and rapid response, to prevent a health

event from becoming a health emergency. However, this need to make quick decisions is at odds with the time it takes to develop a fully comprehensive/operable data platform across systems and data sources. The solution to this is to develop the platform during peace time, so that it can be implemented and rapidly scaled up at point of crisis. Whilst we will not be able to predict all the data points that might be needed for the next health emergency, given the type of pathogen, location of emergency, transmission routes, access to medical countermeasures etc – we can have a considerable head start with many data points. The beginning stages of developing the interoperability may be the hardest part of the process.

Timeliness: There is a fine balance to be achieved between the need for real time policy information, and for that policy data to be useful for decision makers and researchers to use. This would require the data set to be cleaned, and any contextual information provided alongside. In practice, this would realistically probably be a weekly activity, but would be reliant on the availability of individuals to undertake this labour-intensive process. Moreover, consideration needs to be given whether analysis will be done on a daily (unfeasible); weekly; or monthly (likely too long a gap for policy influence). A multi-track system in which certain data points are collected and put out rapidly, while complementary information is collected more slowly, may be optimal.

Archiving: As a living data ecosystem, it is important to ensure that changes and policy developments, as well as historical data are archived to be able to return to analysis in the future, this is particularly important for research needs compared to real-time decision making.
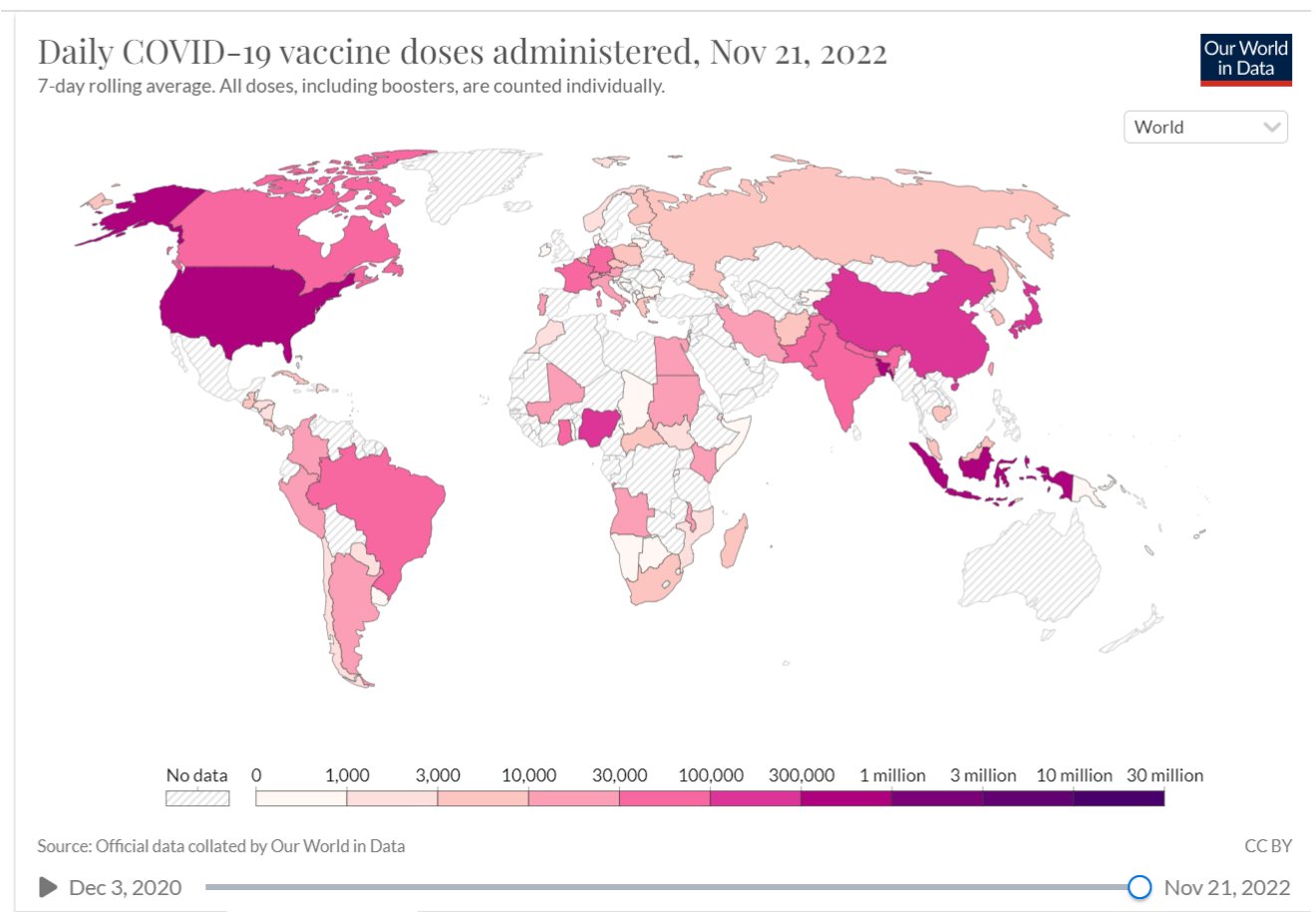
Visualisation: The reality is that those making the decisions are those who are not data experts, or even health experts. Thus, the data that is pooled needs to be presented in a format which is exceptionally user friendly, and that it allows those with no prior expertise of a particular domain to be able to interpret the data and use it to inform decision making. Although often these policy trackers are not designed for comparative analysis, such visuals do help to understand such comparisons in broad terms, which can be useful for decision makers. Several participants highlighted Our World in Data (OWID) as a great best practice example of data visualisation. One simple approach would be to ensure that there is an easily

accessible open source platform that can display data and meta data readily in an approved format, regardless of the source of that data. The importance of contextualisation and nuance grows for visual outputs that can be widely shared.

Our World in Data (OWID) aims to make data and knowledge on the big problems that the world faces accessible and understandable to both decision makers and the general public.

Through interactive data visualisations, OWID takes research and analysis being conducted by their team, and that of researchers across the world and works with web developers and designers to make the data open access, digestible, understandable and usable for all.

Distribution is free, and the research can be explored through a range of visual and interactive mechanisms. OWID has built and continues to build the infrastructure to make this possible as a global public good - able to download, reproduce and through open-sourced code.

Daily COVID-19 vaccine doses administered, Nov 21, 2022

7-day rolling average. All doses, including boosters, are counted individually.

Source: Official data collated by Our World in Data

CC BY

▶ Dec 3, 2020   ───────────────────────────────────────○ Nov 21, 2022

Privacy and sensitivity: Several interview participants raised the concern about privacy challenges with this sort of data platform, and how this intensifies as the data

moves from a national to international level. Whilst there may be some challenges from a legal perspective, this needs to be part of a broader conversation with the public about the trade offs which emerge through a data framework. Moreover, while policy data was in most cases not sensitive to collect (since it typically comes from open, official sources or news reports), other forms of data (e.g. epidemiological data) proved highly sensitive and prone to manipulation or secrecy.

However, there are other risks associated with the policy data and associated analysis. As with all types of data, the contextual information for interpreting the data is vital to be able to understand the data presented. However, the nature of some of this contextual data is subjective. Contextual activity also requires a lot of updating as it can change from month to month. There is a risk in this subjective analysis or interpretation which may ostracise governments with whom collaboration may be established, or trying to be so. How to manage that will be something that needs to be delicately managed. One suggestion would be to steer clear of ordinal scales or rankings of governments based on their stringency in a particular policy space as this could counter trust with the process.

Policy implementation: Any system will need to be able to demonstrate nuance between the policies launched at the governance level, and how these have been implemented in practice. This is one key challenge that policy trackers and policy analysts have experienced throughout COVID-19, which is to understand when a policy is launched, how effectively it is implemented, and thus what the actual effect of the policy is. Whilst this is a challenge experienced across the policy space more generally, it is particularly important given the need for rapid decision making during a health emergency. For example, to know if any particular social distancing policy is working, you need to know if the population have internalised and are undertaking the policy. Compliance in this way is hard to measure in real time, and often requires administrative, mobility, or poll/survey data which can take time to obtain and analyse.

Additional functions of the protocol: One potential mechanism for operationalising this protocol would be to consider what dual use such a framework could have. Is there something else that this can be used for outside of health emergency times?

Can it have a dual life to help with co-financing and longer term continuation and support. It is unclear exactly what that might be, whether to expand into other areas of health, or to consider other emergency contexts, such as humanitarian or climate, but this might provide a more robust secretariat and human resource to provide the much-needed policy context in a broader array of governmental decision-making.

Leveraging other sources and technological developments: Most trackers developed during COVID-19 relied on some form of manual assessment and scraping of data from policy documents, various official sources and media etc. However, there is much that can be leveraged from other sources of data, such as media mining, AI, remote technologies etc. This of course exists already to a certain extent with media monitoring and horizon scanning that occurs within a range of national and global public health organisations, but being able to automate and/or integrate such an approach would be beneficial to rapid data assessment abilities. Several efforts were made on this front during the COVID-19 pandemic, but to date none has been able to reach the standard of manual systems.

One suggestion would be to find a way to incorporate the need for policy data amid other conversations and policy processes/pathways that are ongoing in relation to data sharing and data collaboration in public health emergencies. For example, the Joint External Evaluation which provides monitoring and evaluation of state's activities in the implementation of the IHR have recently been updated to include data sharing principles. The IHR that are currently in the process of being amended have several proposals related to the data sharing, which at present do not specify policy data, but this might be a missing piece of the puzzle. Stakeholders must consider how to make policy data be a complicit and explicit part of this? Similarly, pilot projects are underway for the development of a Universal Health Periodic Review, aiming to demonstrate state capacity in the health sector, and in particular capacity to respond to a pandemic. Identifying where policy data can sit alongside this would be an important improvement.

Incentives: If there is a need for different data generators to share their data, there needs to be serious consideration as to the available incentives for that person to share. As there may be a conglomerate of different actors working within a protocol, there may be a range of different incentives.

For governments to have incentives to participate (e.g. by publishing their data in a particular format, or on particular areas, and even in co-finance such an initiative) they may need to be guaranteed that they would not be unfavourably compared to their neighbours, or have some control over how their policies were interpreted. Such control could destroy the objectivity and comparability of policy data.. They may further want to ensure that any compilation or analysis of their data would nor result in adverse economic impacts, such as the ceasing of trade relationships or travel routes on a consequence of domestic health policies pursued. Other governments, such as those in low and middle income states may seek to consider lessons from Access and Benefit Sharing, to ensure that they would get something in return for their active engagement, such as favourable economic treatment (e.g. no automatic travel bans), favourable engagement in other international fora, or perhaps most pertinently, access to financial resources and or health and medical countermeasures amid a health emergency.

Incentives for governments also cannot be separated from capacity to be able to process data into a particular format. This requires human capital both to establish a process to routinely report policy related data points, and to oversee its implementation. Incentives, for example if financial, could support the human resource required to implement such efforts, but that in itself may not be enough, and may not be enough to ensure a sustainable process going forward.

Researchers have different incentives for sharing. Researchers want access to data, ability to analyse, and in turn publish any associated analyses. This will require a robust data sharing agreement, publication privileges, and, where possible, cooperation and interaction for future research proposals, grants and projects. Best practice principles for academic collaboration should be undertaken, and where possible, grants should be rooted in co-creation, co-analysis and meaningful knowledge exchange as part of the activities. Funding must be shared equitably between any research consortium, noting that whilst costs for any "lead" organisation might be higher to maintain routine capacity and procedures, that costs for scaling up, and or routine provision may be more costly in low income settings where highly trained researchers may be scarce, or would need additional training to be able to take part. Mechanisms could be drawn up amongst the broader research group as to how best to analyse the data, which group is

responsible for which part, whether leading analysis is on rotation among the group, or whether particular groups have responsibility for different parts of the whole.

Challenges

Trust: Trust in the data framework will be essential for its effectiveness. Decision makers and the public need to trust that the data is of good quality, well harmonised and synchronised to be of use to their decision making. Moreover, data generators need to trust in the system that they are being asked to contribute to, both that there is a value add to the process, and that the system is governed by good governance principles (as above).

Best practice for trust building for data sharing: GISAID / GISRS could be a model to build on / expand on / model – WHO at centre, many governments have common systems for influenza data sharing.

The Global Influenza Surveillance and Response System (GISRS) is an international network of influenza laboratories which conducts year around surveillance of influenza, emerging strains for alert and broader preparedness efforts. Through the mechanisms of the Pandemic Influenza Preparedness (PIP) Framework and the Influenza Virus Traceability Mechanism (IVTM), GISRS facilitates the transfer of clinical specimens, virus samples, genomic data, analysis and other contextual information, sharing these with other states, the private sector, international organisations, academia and beyond. It has been going for 70 years, and now contains over 158 National Institutions.

As a system which has detected and responded to a range of influenza viruses in its history, it has also been able to pivot to other emerging infections including SARS, MERS and COVID-19. This proven network model can offer many lessons for the development of a data framework for non-pharmaceutical / policy interventions, and for a data platform for health emergencies in general. This includes:

1. Mechanisms and best practice for sharing the whole landscape of pertinent data, including legal provisions, best practice mechanisms and practical considerations

2. Data sharing platforms for facilitating interaction between institutions

3. Public engagement and open access data

4. Fair and equitable access to data and samples

Capacity: The data comprising such a data framework is based on multiple sources, and notably would ideally need to include national sources of data. This will be hard to achieve at a national level from systems that lack capacity, but this limitation needs to be built into decision making. There is a scarcity of sources in some locations, capacity of data generators, capacity of human resource to format and share accordingly, poor national infrastructure etc. (Lal et al. 2022). This capacity issue leads to gaps in data collection, which then risks the quality of the project as a whole if you cannot develop robustness and comprehensiveness. A key challenge must be how to overcome the capacity-associated data gaps both retrospectively and prospectively. One suggestion has been to leverage the developments occurring at regional level, and through the non-state sector to support the data generation and sharing for pandemic preparedness and response (Maxmen 2021) (Wang et al. 2020). This capacity will be particularly challenged in time of crisis, and in the initial phases of an outbreak, and amid confusion and chaos it would be particularly hard to maintain quality data without considered and tested plans.

The levels of data collection is also a question which needs to be overcome. The nature of different governance systems, and its devolution of pandemic or health related decision making will mean that in some locations there will be just national level data collection, but in other locations, there will need to be sub-national policy data collection too. This variance is problematic not just from the perspective of the interoperability of the system, the need for greater contextual information to situate the data, but, moreover, multiple levels of capacity for data collection require more human resource.

Governance

Funding

Capacity for researchers to maintain and operate a policy data collection and sharing system, for governments to share data at a national level, and the capacity to develop a data protocol or framework for health emergency response will be contingent upon requisite funding to be able to support data generation, and the interoperability of the system, particularly a system which is flexible enough to reflect the evolving data needs of a health event.

The trackers developed during COVID-19 were resourced in large part by volunteers, which, while bringing certain benefits (e.g. allowing sourcing of information from many "on the ground" contexts), also faces sustainability challenges (volunteers may not be so easy to come by in the next emergency). This needs to be costed and accounted for in decision making. The success of academic trackers based on volunteers may result in this need not being identified (it wasn't needed last time) but this protocol must be based on sustainable financing from donors, philanthropy, the private sector etc. (European Observatory on Health Systems and Policies, McKee, and Greer 2021)

But the funding is not just needed for an interoperable system – there is also a need to increase funding for data collection processes within national systems. As yet, this funding for data frameworks, particularly for those of policy data does not feature as part of pandemic preparedness planning, such as the World Bank's Financial Intermediary Fund for Pandemic Preparedness and Response (Pandemic Fund) or as part of the proposed Pandemic Convention or Agreement (Treaty).

Thus, it is likely that other types of financing must be considered to establish such a framework for being able to engage in any future health emergencies. Aware of the landscape of global health security financing, we anticipate that this will need to be sourced from the philanthropic or private sectors in the first instance. Co-financing models can be introduced in an effort to develop more sustainable futures for such an initiative, whereby government or research institutions partially support the initiative, increasing the financial contribution over time as the external financing reduces. Given the uncertain time scale of when another health emergency may arise, it is vital that such mechanisms are engaged with rather than falling into the cycle of panic and neglect, and within a number of years the initiative comes to a close, potentially (hopefully) prior to its re-utilisation.

Leadership

All participants in the research have described the need for leadership to bring about a data sharing platform to be able to respond to the next pandemic. Many described a leadership vacuum during COVID-19 which meant that there were little opportunities for sharing or thinking through more meaningful sharing in real time.

As per the proposed structural **option 1**, several participants located the WHO as central to this leadership role, and the convening power and technical capacities and leadership it has within health emergencies would be catalysing to develop such a platform or process for data in future health emergencies. Moreover, with a "neutral" actor in this central position, this would hopefully overcome challenges of different trackers being in competition and the "first past the post" mentality that some participants mentioned. With the normative authority the institution has in standards setting, it could harmonise groups and data sources, as well as create a unified taxonomy, etc.

Yet the role of the WHO was also seen as problematic. Many discussed the challenges within the organisation and the broader legitimacy crisis it faces whilst discussing the role it could play. WHO has convening power, and it is accountable to member states (who are the primary end users), which can be good for governance – and could also get data from the governments, and back to them as users, through the channels that already exist such as NFPS – if they wanted to. However, this would have several limitations: Firstly, WHO have shown that they struggle to get data in real time and so this would suffer a lag time which may risk the utility of it for decision makers. Secondly, as a member state organisation, the WHO may not be in a position to share data openly, should states not want it to. Thirdly, WHO lacks the human resource and in-house capacity for data management, and it might take time to scale this up at time of need.

One way to mitigate this is for WHO to be a hub and spokes model with a range of different actors working collectively and collaboratively in this space. For example, the WHO could convene a group of research institutions, governments and other relevant stakeholders, co-design a robust taxonomy, curate the data and do preliminary checks, but then work with a range of end users, academic research institutions and non-state actors to perform more rigorous analysis and publication of data in real time, as well as enhanced visualisation. This would also require interaction with a range of other multisectoral actors to ensure diverse data sources. Importantly, whilst WHO could use its strengths and expertise to convene and offer technical guidance, and be the "at peace" hub for engagement across the academic and research sector, it would need some sort of (likely) foundation/philanthropic/private financing to be able to enact such action.

Our proposed structural **option 2,** seeks to break away from the aforementioned limitations of WHO, through the development of a rigorous consortium approach to this policy data framework. In this scenario, leadership would be comprised of either one "lead" institution which manages the day to day running of a data framework, and engages with other groups in the consortium as and when are required. This would have strategic purpose for managerial and operational concerns, and yet the academic development could be shared equally between institutions. In this scenario, consideration would need to be given as to the ownership of the policy tracker, the data produced by it and the analyses undertaken. It is important that if it is research teams that are doing the lion's share of the work to sustain such a tracker, and scale it up during a time of crisis, that there is a sense of ownership by the consortium (see principles). This will facilitate greater engagement with the process, and hopefully close alignment with the principles and improved outcomes.

Which option of these two becomes dominant will depend on how the community of individuals involved in such tracking already are able to organise themselves and work together, the demand which emerges from end users for such routine policy data, and ultimately, whether either option is able to obtain sustainable financing.

Community engagement and trust

Although the end user identified are decision makers in governments, a secondary user is that of the public, as a tool for health communication, information sharing about an outbreak and, if necessary, community engagement. Several participants noted how important it was to have a user-friendly interface [see visibility] so that the public can understand the data. There is also potential for citizen science in the data collection / collation (Chu et al. 2020). Engagement in this way with the public also helps to engender trust both in the protocol itself, and in the government decision makers who are using the protocol, and in turn the response to a health emergency. To do this, and more broadly across any policy tracking framework, ethical consideration of how the data is used, and how individuals are involved must be carefully considered and implemented.
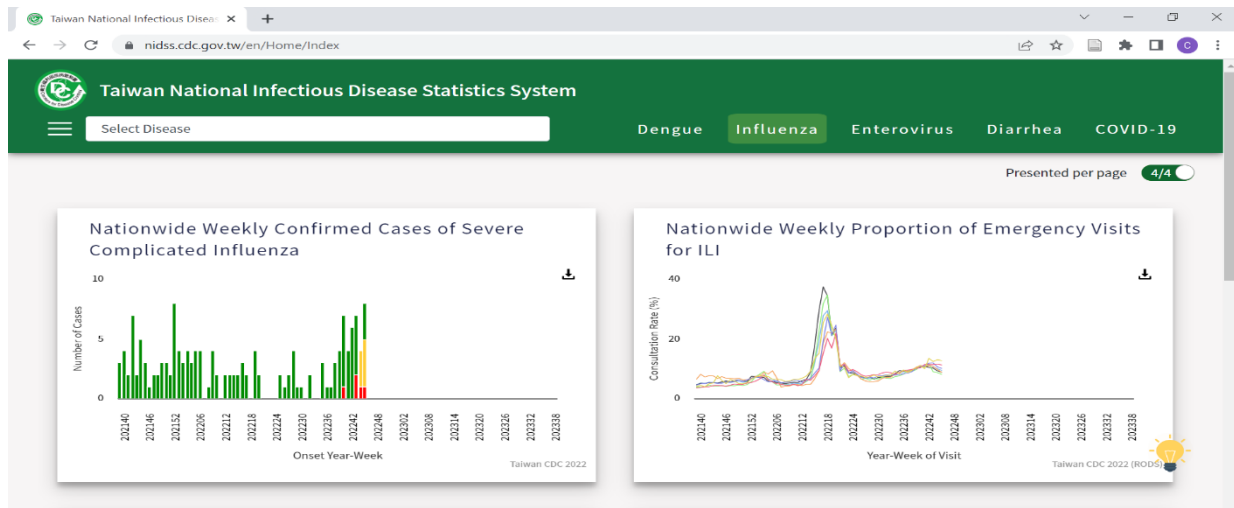
One pertinent example of data transparency for improved community engagement has been the success of Taiwan's National Infectious Disease Statistics System which is updated on a daily basis and open access for all. It allows the public, media, and international decision makers to browse real time data on a host of infectious diseases. This offers not just case data, but hospitalisation, clinical and other pertinent data points for analysis.

Building toward a fit for purpose data framework for policy data: proposed next steps

Below we tentatively outline a few potential next steps for the purposes of pushing this agenda forward for a working policy data framework. These are non-exhaustive, and also are not sequential.

1. It is important for the community of data collectors, researchers, and users (especially policymakers) to work toward a stronger approach for future pandemics. To do this, all stakeholders need to reach a broad consensus on the need for such a framework and protocol, and be willing to be part of a more permanent policy data landscape. Using this scoping paper as a starting point, we propose that this should be circulated to any and all institutions who may have interest / desire to be part of any proposed framework for discussion and consensus building as to the principles, process, structure and approaches to take this framework forward.

2. In this scoping paper we propose two different structures for such a data framework - stakeholders involved will need to weigh up the pros and cons of both options and make a decision about which to pursue:
    a. Model 1: WHO as hub, researcher institutions as spoke, philanthropy as supporter
    b. Model 2: Research network with secretariat amid consortium, philanthropy as supporter
    c. Consider other models which might be suitable

From a practical point of view, we propose that either or any option begins with a small secretariat or standing committee function to iron out the operationalisation and be the driving force for implementation.

3. Key tasks for either structure and for operationalising the data framework for policy data:
    a. Developing a standardised template (taxonomy) for policy data that can quickly be adapted and deployed. This can use the process level considerations from this scoping paper as a starting point, and build on them to create a working framework to pilot. As this is likely to be the biggest task for any group, we propose that this should start with designing a template of the kinds of policies we can expect in future pandemics, rendered in a flexible way so that they can be adapted to the specific requirements of future diseases. As this may well be too abstract to identify, and not wanting to base all contents on the last/ongoing crisis, an alternative option would be for the group to agree to a process for making the general template more specific once a new potential pandemic is identified. For example, a standing committee that can quickly mobilise to agree a standardised approach.
    b. Building and maintaining a data collection infrastructure to be the legacy of the range of trackers established during COVID-19, and ensure that there is a starting point for a framework, and scaling up for the next health emergency. This data collection infrastructure should, as this scoping paper has started to do, seek to ascertain the learnings from the ad hoc projects that arose in response to COVID-19. These learnings should be codified into working knowledge and practice, to

be used for the tracker and its future developments. Those stakeholders and individuals that have been involved in trackers across COVID-19 should be engaged/retained or their institutional history be preserved to as great extent as possible.

c. Funding should be a priority to ensure the continuity of this data framework protocol. This will need to be considered now in the establishment / consolidation phase of the data framework, but also to ensure ongoing, sustainable financing (even if to a lesser degree) to keep the minimal infrastructure in place to reactivate if/when needed. For example, where funding allows, those staff / researchers who have been involved in different policy trackers to date should continue to devote some time to this ongoing project. A next step will be to fundraise for innovative/flexible financing to support the reactivation of policy data tracking for the next health event

d. Mainstreaming policy data in the wider pandemic data architecture. Whilst many decision makers and end users of the policy trackers established during COVID-19 have highlighted the importance and utility of policy data in their decision making, it is far from embedded within routine public health practice at present. A further next step must be to create a more systematic role for such policy data within pandemic preparedness and response activities, to ensure it is not a secondary thought in any future health emergency. This will require trust building and awareness raising activities by a Secretariat / Standing Committee. For example, this might include awareness raising / promotion amongst more traditional pandemic preparedness data collectors and analysts, such as epidemiology, modellers and virology groups. Institutions such as WHO, GOARN, and university based research centres would be a good starting point to facilitate integration with policy data into more traditional public health settings. Due consideration should also be given for the legal and policy developments occurring within the pandemic preparedness and response landscape, such as the IHR Amendment Process, the Intergovernmental Negotiating Bureau for the proposed Pandemic Accord, the Universal Health Preparedness Review, the Pandemic

Fund, the High - Level Meeting on Pandemic Preparedness and Response etc. Each of these are considering the role of data, data needs for decision makers, and data sharing. Any proposed consortium or Secretariat for a policy data framework must try to advocate for the inclusion of policy data within the formal structures of governance.

Bibliography

Chu, Isaac Yen-Hao, Prima Alam, Heidi J Larson, and Leesa Lin. 2020. 'Social Consequences of Mass Quarantine during Epidemics: A Systematic Review with Implications for the COVID-19 Response'. *Journal of Travel Medicine* 27 (7): taaa192. https://doi.org/10.1093/jtm/taaa192.

COVIDSurg Collaborative., 2021. Effect of COVID-19 pandemic lockdowns on planned cancer surgery for 15 tumour types in 61 countries: an international, prospective, cohort study. *Lancet Oncol.*, 22, pp.1507-17.

European Observatory on Health Systems and Policies, Martin McKee, and Scott L. Greer. 2021. 'Effective Pandemic Preparedness Demands That Governments Confront Hard Questions about Sovereignty'. *Eurohealth* 27 (1): 32–35.

European Observatory on Health Systems and Policies, About us. https://eurohealthobservatory.who.int/about-us/overview

Hale, T., Angrist, N., Hale, A.J., Kira, B., Majumdar, S., Petherick, A., Phillips, T., Sridhar, D., Thompson, R.N., Webster, S. and Zhang, Y., 2021. Government responses and COVID-19 deaths: Global evidence across multiple pandemic waves. *PLoS One*, 16(7), p.e0253116.

Katz, Rebecca, et al. 2022. COVID Analysis and Mapping of Policies (COVID AMP). Available: https://covidamp.org/

Kentikelenis, Alexander, and Leonard Seabrooke. 2021. 'Organising Knowledge to Prevent Global Health Crises: A Comparative Analysis of Pandemic Preparedness Indicators'. *BMJ Global Health* 6 (8): e006864. https://doi.org/10.1136/bmjgh-2021-006864.

Kikstra, J.S., Vinca, A., Lovat, F., Boza-Kiss, B., van Ruijven, B., Wilson, C., Rogelj, J., Zakeri, B., Fricko, O. and Riahi, K., 2021. Climate mitigation scenarios with persistent COVID-19-related energy demand changes. *Nature Energy*, 6(12), pp.1114-1123.

Kim, J. and Kwan, M.P., 2021. The impact of the COVID-19 pandemic on people's mobility: A longitudinal study of the US from March to September of 2020. *Journal of Transport Geography*, 93, p.103039.

Lal, Arush, Henry C. Ashworth, Sara Dada, Laura Hoemeke, and Ernest Tambo. 2022. 'Optimizing Pandemic Preparedness and Response Through Health Information Systems: Lessons Learned From Ebola to COVID-19'. *Disaster Medicine and Public Health Preparedness* 16 (1): 333–40. https://doi.org/10.1017/dmp.2020.361.

Littler, Katherine, Wee-Ming Boon, Gail Carson, Evelyn Depoortere, Sophie Mathewson, Daniel Mietchen, Vasee S Moorthy, Denise O'Connor, Cathy Roth, and Carlos Segovia. 2017. 'Progress in Promoting Data Sharing in Public Health Emergencies'. *Bulletin of the World Health Organization* 95 (4): 243. https://doi.org/10.2471/BLT.17.192096.

Maxmen, Amy. 2021. 'Has COVID Taught Us Anything about Pandemic Preparedness?' *Nature* 596 (7872): 332–35. https://doi.org/10.1038/d41586-021-02217-y.

Mendez-Brito, A., El Bcheraoui, C. and Pozo-Martin, F., 2021. Systematic review of empirical studies comparing the effectiveness of non-pharmaceutical interventions against COVID-19. *Journal of Infection*, 83(3), pp.281-293.

Miguel, E. and Mobarak, A.M., 2022. The economics of the COVID-19 pandemic in poor countries. *Annual Review of Economics*, 14, pp.253-285.

Nivette, A.E., Zahnow, R., Aguilar, R., Ahven, A., Amram, S., Ariel, B., Burbano, M.J.A., Astolfi, R., Baier, D., Bark, H.M. and Beijers, J.E., 2021. A global analysis of the impact of COVID-19 stay-at-home restrictions on crime. *Nature Human Behaviour*, 5(7), pp.868-877.

Oliu-Barton, M., Pradelski, B.S., Aghion, P., Artus, P., Kickbusch, I., Lazarus, J.V., Sridhar, D. and Vanderslott, S., 2021. SARS-CoV-2 elimination, not mitigation, creates best outcomes for health, the economy, and civil liberties. *The Lancet*, 397(10291), pp.2234-2236.

Petherick, A., Goldszmidt, R., Andrade, E.B., Furst, R., Hale, T., Pott, A. and Wood, A., 2021. A worldwide assessment of changes in adherence to COVID-19 protective behaviours and hypothesized pandemic fatigue. *Nature Human Behaviour*, 5(9), pp.1145-1160.

Phillips, T., Zhang, Y. and Petherick, A., 2021. A year of living distantly: global trends in the use of stay-at-home orders over the first 12 months of the COVID-19 pandemic. *Interface Focus*, 11(6), p.20210041.

'Rapid Review of WHO COVID-19 Surveillance: External Review, 27 October 2021'. n.d. Accessed 28 November 2022.

https://www.who.int/publications/m/item/rapid-review-of-who-covid-19-surveillance-external-review-27-october-2021.

Toshkov, D., Carroll, B. and Yesilkagit, K., 2022. Government capacity, societal trust or party preferences: what accounts for the variety of national policy responses to the COVID-19 pandemic in Europe?. *Journal of European Public Policy*, 29(7), pp.1009-1028.

Wang, Zhebin, Yuqi Duan, Yinzi Jin, and Zhi-Jie Zheng. 2020. 'Coronavirus Disease 2019 (COVID-19) Pandemic: How Countries Should Build More Resilient Health Systems for Preparedness and Response'. *Global Health Journal*, Special Issue on Global Responses to Covid-19 Pandemic: Challenges and Opportunities, 4 (4): 139–45. https://doi.org/10.1016/j.glohj.2020.12.001.

World Health Organization, Tracking Public Health and Social Measures, n.d. Accessed 20 January 2023 https://www.who.int/emergencies/diseases/novel-coronavirus-2019/phsm