Overcoming Barriers to Data Sharing in Public Health
A Global Perspective
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>2</td>
</tr>
<tr>
<td>Background</td>
<td>3</td>
</tr>
<tr>
<td>Overcoming Data Sharing Barriers</td>
<td>6</td>
</tr>
<tr>
<td>Data Sharing in Public Health Emergency Situations: What is Different?</td>
<td>17</td>
</tr>
<tr>
<td>Conclusions: Data Sharing in an Ideal World</td>
<td>18</td>
</tr>
<tr>
<td>Acronyms</td>
<td>20</td>
</tr>
<tr>
<td>References</td>
<td>21</td>
</tr>
<tr>
<td>About the Authors</td>
<td>24</td>
</tr>
</tbody>
</table>
Summary

- The interaction between barriers to data sharing in public health is complex, and single solutions to single barriers are unlikely to be successful. Political, economic and legal obstacles will be the most challenging to overcome.

- Public health data sharing occurs extensively as a collection of subregional and regional surveillance networks. These existing networks have often arisen as a consequence of a specific local public health crisis, and should be integrated into any global framework.

- Data sharing in public health is successful when a perceived need is addressed, and the social, political and cultural context is taken into account.

- A global data sharing legal framework is unlikely to be successful. A global data governance or ethical framework, supplemented by local memoranda of understanding that take into account the local context, is more likely to succeed.

- The International Health Regulations (IHR) should be considered as an infrastructure for data sharing. However, their lack of enforcement mechanism, lack of minimum data sets, lack of capacity assessment mechanism, and potential impact on trade and travel following data sharing need to be addressed.

- Optimal data sharing does not equate with open access for public health data.
Background

Decision-making in public health, from routine responses to acute public health threats and long-term planning of intervention and control programmes, is increasingly reliant on the efficient use of data (1). Recent developments in information technology have created better opportunities for prompt sharing through more user-friendly systems, available in both high-income countries (HICs) and low- and middle-income countries (LMICs) (2). While the International Health Regulations (IHR) require the notification of public health events with the potential for international concern (3), there is no global systematic framework for sharing public health data (1), and there are no global norms for the systematic sharing of public health surveillance data – despite it being extensively and routinely collected in both HICs and LMICs.

In other areas of biomedical science such as genomics, data sharing is the norm, and raw sequencing data is routinely submitted to several open access databases (4–6). Proponents of complete open access to public health data have argued that publicly funded information is a public good, and that data should be available to anyone who wants to use it for advancing public health (4). Such openness would facilitate the efficient use of data and result in greater benefits for society as a whole (5). More specifically, it would lead to faster progress in improving health, better cost-efficiency and high-quality science (1, 7–8). The evidence for the public health benefits of surveillance data sharing is growing (6), with well-documented instances of an improved outcome as a result of sharing surveillance data. For example, data sharing across the Global Influenza Surveillance Network (GISN, now the Global Influenza Surveillance and Response System – GISRS), coordinated by the World Health Organization (WHO), contributed to preventing severe acute respiratory syndrome (SARS) from becoming endemic following the 2003 outbreak (9), and to a robust and timely response to the 2009 influenza A (H1N1) pandemic (10). Efficient data sharing also prompted an early response to the emergence of the H7N9 influenza virus in humans in China (11).

Conversely, reluctance to share can hinder or slow down the response. During the early phases of the 2003 SARS outbreak, it became apparent that China delayed sharing information about the situation for political reasons; this contributed to the outbreak spreading and delayed the response (12–13). Inadequate data sharing is also part of the reason why more than two years after the identification of the novel Middle East respiratory syndrome coronavirus (MERS-CoV), its origin and source remain unknown (14). More recently, limited sharing of viral sequences during the Ebola outbreak in West Africa has made it harder to evaluate the virus’s potential for mutations (15). Global outbreaks have shown that inadequate surveillance and response capacity in a single country can endanger national populations and the public health security of the entire world (9).

Among many proposals for data sharing in public health, a group of high-profile global health agencies, including the WHO, have called for increased data harmonization and sharing through capacity-building and a code of conduct for data producers and data users (16). The Bill & Melinda Gates Foundation has also proposed guiding principles for global health data access, including promotion of the common good, respect, accountability, stewardship, proportionality and reciprocity (17). Despite these proposals, public health surveillance data remain under-utilized through suboptimal sharing, with consequences for global public health. The underlying reasons for the weak implementation of data sharing principles have not been formally evaluated, but the
lack of overall guidance on how to operationalize stated data sharing principles may have been a contributory factor.

In practice, public health data sharing does occur and relies on local or regional solutions, often focusing on a single disease, with successful surveillance networks having been implemented in both HICs and LMICs. The emergence of these networks has been driven by local or regional challenges such as cross-border outbreaks, increased cross-border trade, movement of people and animals, and ecological changes related to climate change (18). These networks are organized by regional, national and supranational public health agencies, as well as by NGOs. The GISN (now GISRS), set up in 1947, is one early example, which as of 2010 incorporated 136 centres in 106 countries, sharing timely epidemiological and virological data in order to produce vaccines to match circulating influenza strains (10).

The European Centre for Disease Prevention and Control (ECDC) has since 2008 hosted The European Surveillance System (TESSy), which collects national surveillance data from all European Union (EU) and European Economic Area (EEA) countries using standardized data sets, further disseminates these data and produces outputs for public health action (19). In the United States each state follows its own regulations to collect public health data which are then reported to the US Centers for Disease Control and Prevention (CDC), the federal agency with the mandate for infectious disease surveillance (20). The WHO also runs several disease-specific regional surveillance networks, such as the African Rotavirus Surveillance Network, established in 2006, supporting sentinel surveillance sites across member states using standardized guidelines (21). Regional networks routinely sharing surveillance data exist in several areas, including Southeast Asia (22), East Africa (23) and the Middle East (24), created locally in response to cross-border malaria, cholera and salmonella outbreaks respectively, and coordinated via an umbrella organization (25).

Conceptually, there a number of barriers facing public health data sharing. These can be classified in six broad categories – technical, motivational, economic, political, legal and ethical – and are interconnected, with complex interactions limiting the effectiveness of solutions aimed at any one barrier in isolation. Beyond restating well-known barriers, there is a knowledge gap in terms of understanding what these mean to public health professionals dealing with surveillance data as part of their work, and, more importantly, how the barriers may be overcome.

The Centre on Global Health Security (CGHS) at Chatham House examines key global health challenges, and how they manifest themselves as foreign policy and international affairs (26). The Bill & Melinda Gates Foundation is supporting CGHS in creating guidance that aims to strengthen data sharing for public health in order to create the right environment for data sharing and achieve good practice. In order to identify the policy and technical requirements for optimal data sharing in public health, and achieve the goals of this project, Chatham House is convening a series of roundtables, inviting experts in public health, policy, bioethics, data management, law, bioinformatics and other fields relevant to data. This paper aims to describe barriers in public health surveillance data sharing, and how barriers to sharing are overcome. By identifying and analysing examples – at both policy and field level – for each type of barrier, it will suggest thematic solutions to public health data sharing.
barriers that will inform the roundtable discussions, and be used as a supporting document for experts when formulating solutions to data sharing barriers. A companion paper examining data sharing practice in other sectors, and how good practice can be translated to the public health sector (27), will also inform the roundtable discussions.

**Analytical approach**

This paper relies on interviews with key experts working as high-level decision-makers, field epidemiologists, academics and individuals responsible for disease surveillance at the national level, in both HICs and LMICs. Experts were identified using a 'snowballing' approach, starting with key individuals within the CGHS network. Using a semi-structured questionnaire, participants were asked to describe how data sharing is part of their routine work; to share their experiences and expertise regarding specific barriers to public health data sharing they have encountered or are aware of; and to specify how they have overcome the barriers they have encountered, both in routine surveillance and during public health emergencies. Informants were also asked to set out their vision of public health data sharing in an ideal world. Solutions emerging from the interviews were thus examined and classified by theme; the relationships between barriers and solutions were explored using a matrix approach; and the complex interactions were captured. This paper discusses the content of the interviews in the context of the existing literature, highlighting practical solutions via quotes from informants and short case-studies.
Overcoming Data Sharing Barriers

Although the experts interviewed for this paper described a wide range of barriers to data sharing, all these fell within categories previously described in the literature, confirming that the current conceptual understanding of barriers to sharing is comprehensive. However, experts did not consider all barriers to be equal: political, economic and legal barriers were considered to be the most challenging to overcome.

These barriers hinder public health surveillance. From the local to the global level, informants described a wide variety of situations where they encountered those barriers in their daily work. The solutions they implemented were derived from their personal experience in their specific context. For this paper, based on the approaches the experts took to overcome barriers to data sharing, solutions were grouped in the following categories and subcategories (Table 1):

Table 1: Solutions to data sharing in public health

<table>
<thead>
<tr>
<th>Trust and transparency</th>
<th>Value proposition and rewards</th>
<th>Legal framework and political advocacy</th>
<th>Capacity-building and effective tools for sharing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Building trust</td>
<td>Showing the benefits and value of sharing</td>
<td>Policy development and political commitment</td>
<td>Data standardization</td>
</tr>
<tr>
<td>Communication and dialogue</td>
<td>Development of incentives</td>
<td>Data governance agreements</td>
<td>System development</td>
</tr>
</tbody>
</table>

These solutions are to a large extent interconnected, and one solution category applies to several barriers. The informants offered practical applications of these solutions for each barrier.

**Technical barriers**

Technical barriers include inadequate data collection; lack of standardization and of common protocols across agencies and surveillance sites; varying data quality; incompatibility between surveillance databases; and language barriers. A particular obstacle is the lack of an explicit minimum data set as part of the IHR notification process. As a consequence, the information shared via the IHR varies from one outbreak to the next, and the information shared during a specific outbreak is not always comprehensive or can lack important variables. Conversely, when supranational agencies such as the WHO or ECDC decide on case definitions, data formats, variables to include in a minimum data set, and validation mechanisms, these may not be compatible with data produced at the national level, and may complicate the data sharing process for national public health agencies and other data generating institutions, potentially leading to reluctance or delay in data sharing. However, the consensus among informants was that these obstacles are gradually being overcome through capacity-building and systems development.
Overcoming technical barriers

**Standardization and automation**

Standardizing data would make sharing easier and cheaper, because it is less labour-intensive to clean up and assure the quality of data that have been standardized. Several informants mentioned the need for explicit data sharing requirements when notifying through the IHR, including minimum data set requirements. They also highlighted that automation of the data sharing process, with adequate country-level agreements, could make data sharing easier and more timely. However, several informants also saw a potential risk in standardization, particularly in LMIC settings, when standardization implied modifying functional country-specific surveillance systems, which may have required substantial human and financial investment. As one informant working for a development agency coordinating several surveillance sites across Africa explained:

> We thought, if all the countries could harmonize their disease surveillance systems reporting could be enhanced, and then we quickly discovered each of these countries, or most of them, already have pretty good systems in place and we’re not going to replace them with the neighbouring system … So we’re no longer trying to say countries should have the same systems … but we want them to be able to communicate with each other more readily.

**Workforce and systems building**

Data sharing also implies the ability to generate high-quality data in the first place. Generating these data implies a skilled workforce and good communicable disease control systems in place, including laboratory capacity. As pointed out by one informant: ‘If you don’t have the data or the competencies in your country, it’s very difficult for you to share or to fully participate in data sharing activities.’ Several experts mentioned field epidemiology and laboratory training programmes (FETPs/FELTPs) as one way to build up in-country capacity and skills. There are 57 accredited FETP/FELTPs around the world, including regional networks in Africa, South Asia and Latin America (28). Graduates acquire specific skills for setting up, managing and evaluating surveillance systems. Experts saw the IHR as the logical framework through which to build core surveillance capacity, since building local and national surveillance capacity is an IHR requirement. The specific IHR surveillance requirements are described in its Annex 1A (29). Yet this process is neither enforced nor rigorously evaluated. Countries report gaps in the resources, in particular human resources, necessary to be able to fulfil the IHR requirements (30). FETP graduates naturally fit into the IHR-mandated capacity-building (28). Regional training centres are also required to address workforce shortages and training gaps (30). Informants pointed to the need for a more rigorous surveillance capacity evaluation process as part of the IHR.

**Legal solutions to technical barriers**

Legal frameworks also have the potential to overcome technical barriers by promoting capacity-building. EU Decision No. 1082/2013/EU on serious cross-border threats (31) is one example that has been reported as having increased public health capacity by pushing member states to compliance and better standards.
Legal barriers

The main legal barriers identified by informants related to strict personal data protection laws preventing data being shared, and an ambiguous legal framework in which exceptions to data protection on health grounds were not explicit (unlike exceptions on security grounds, for example). This leads to legal ‘grey areas’ where public health institutes are not sure what they can and cannot share. In addition, lack of formal or informal data sharing agreements across borders, lack of an enforcement mechanism for the IHR, and intellectual property rights and data ownership, all hindered transparent sharing. Compliance with different national legislations was seen as particularly challenging for agencies collecting data from several countries, as one expert from a global public health agency explained:

One barrier has to do with the specific legislation in member states … for example, HIV. In some countries it’s not possible to link data from the laboratory and from epidemiology because that’s forbidden at national level, so it’s not possible for member states to share integrated data on HIV/AIDS with us.

Lack of harmonization between legal requirements and health policy can also slow down data sharing. In the United States public health laws regarding data sharing vary from state to state, and in some instances, while the law provided opportunities for data sharing, these were restricted by health department policies that had not been updated to reflect new legislation (32).

Another barrier to successful legal agreements is that such agreements are generally written by one party (generally the party receiving and using the data in an HIC); and as such the legal documents tend to be one-sided, with insufficient protection for the party providing the data. The concern is that one party (often in an LMIC) collects the data and the other (often in an HIC) analyses and uses the data, with little benefit to the party that collected the data (33). For example, in 2004 the WHO conducted a study on risk factors for non-communicable diseases in Tonga; four years later little of the data had been published, and the researchers were unwilling to share the data with the Tongan government (34).

Overcoming legal barriers

There are examples of successfully implemented legal frameworks promoting cross-border surveillance and preparedness response, such as EU Decision No. 1082/2013/EU on serious cross-border threats (31), which experts described as having improved European-level public health capacity and increased compliance with the IHR. Many informants cited the IHR as a framework already in place that can be used for data sharing, although it lacks a legal enforcement mechanism. In addition, the TESSy data access policy relies on a legal agreement approved by 29 countries (35). Different levels of data access, depending on the nature of the institute or the individual requesting it, are embedded in the legal framework, and the release of data is subject to specific conditions (35). The policy clearly outlines data ownership, data publication rights and sharing to third parties. It is unclear whether such a legal framework is transferable to other regions.

Clarifying exceptions to data protection

In order to overcome legal barriers, legal systems must explicitly state what constitutes exception to data protection on health grounds and on intellectual property grounds: currently this is not the case, and exceptions are left to personal interpretation.
Implementing a global governance framework

Experts agreed that harmonizing a wide array of legal systems would be extremely difficult, and, as such, a global data sharing legal framework was not a realistic goal. Furthermore, a global health legal framework can have unintended negative consequences, such as petrifying principles with only contemporary relevance, imposing foreign values on less powerful countries, or forcing externally defined goals on countries (36). However, several informants saw a global data governance framework as a possible alternative, complementing existing legal systems in place. Any such framework should specify the limits of data access: what to share, when, how and with whom; and what will and will not be done with the data. Such a framework should, however, be flexible enough to address context-specific concerns, as highlighted by one informant from an academic institution:

In my experience … you have to get a clear sort of sense of what the concerns of each party [are] – usually just having a conversation – so I'll sit down and say: 'So what are you worried about?' … Then, once you know what each party is worried about, then it's important to construct an agreement or some sort of data governance document that is responsive to those concerns. If the concern is about inequity then you spell out how benefits will be equitably shared.

Complementing the law with local, context-specific agreements

As an alternative or a complement to a global framework, several experts discussed arrangements where legal requirements, often not addressing the needs of the different parties equally, were complemented by another ad hoc, context-specific document which was not legally binding but which took the consideration of all parties involved in the data sharing agreement into account. Box 1 describes an example of a data governance agreement complementing a legal framework. Surveillance experts from other countries described similar arrangements: a legal framework and a memorandum of understanding between the parties involved in sharing the data. As one informant pointed out, legal frameworks do not take into account the social, cultural or political sensitivities pertaining to a specific context where data sharing occurs, and these can be addressed by this parallel process.

Box 1: Overcoming legal and ethical concerns through data governance agreements

An informant with experience in governance models to facilitate data sharing in global health described a public health project funded by a grant from the government of an HIC. The recipients of this grant, in India and Thailand, were responsible for primary data collection of large data sets, and were paired up with investigators in an HIC who were doing some modelling work. A data sharing agreement was constructed – a requirement as part of the grant – stipulating that grantees must share data. But when it came to the time for actually exchanging the data, it didn't happen. It transpired that the grantees had entered similar agreements before and felt that they had been exploited by providing data and not being involved in the analysis or receiving any credit for their work. The solution was to prepare a parallel data governance document. In addition to the legal agreement, a document laid out the sharing principles and described the expectations of each party. It added the element of reciprocity, which legal agreements very often do not have.

Political barriers

Political barriers to data sharing include restrictive data access policies; bureaucratic hurdles; hiding data for political reasons; lack of political will and commitment to promote data sharing; lack of guidelines; and lack of trust. These barriers hinder effective sharing at the global level, but are also
apparent at the local or national level: informants from national public health institutes mentioned the difficulties in obtaining data either from regional centres or from other teams in their own institute. In LMICs the trend is often towards more restrictive, rather than more open, data policies, as highlighted by one informant working across several countries in Africa:

In Africa, we are now beginning to come across more political-type obstacles. Countries are gradually becoming aware of the data in their possession and are beginning to enact sometimes very restrictive laws about whether researchers who are working in their countries are allowed to pass on the actual data that they collect to agencies outside of their countries.

This trend was substantiated by one informant who worked as an epidemiologist in West Africa during the Ebola outbreak and was, as a matter of policy, denied access by the geographical institute of the country he was working in, to the geo-coordinates of the localities where he was operating. Restrictive data access policies can be justified by the fear of data misuse. One expert in bioethics explained that data generated for public health should only be used for public health purposes. There are also risks associated with allowing data to be shared openly with insufficient control or validation mechanisms, as illustrated in Box 2.

Box 2: When rapid data sharing can go wrong

In the early phase of an outbreak of *E. coli* O104 in Germany in 2011, a local health authority released data on a possible outbreak source, after obtaining results from environmental investigations performed on vegetables sold on the local market. They found cucumbers grown in Spain to have tested positive for the pathogen causing the outbreak, but at that stage detailed laboratory and epidemiological studies for further verification had not been conducted. This premature release of data led to other countries suspending certain agricultural imports from Spain, as well as from surrounding countries, causing serious adverse economic consequences for affected producers – yet the cucumbers grown in Spain were subsequently found not to have been associated with the outbreak (which was traced to sprouted seeds in Germany).

Overcoming political barriers

*Building trust*

Building trust among partners involved in data sharing is often mentioned as a key element to overcoming political barriers to sharing. In practice, building trust means developing one-to-one professional relationships with counterparts in other regions or countries. This can start as a ‘bottom-up’ initiative at the regional level, when public health experts realize the need to share data across regional or national borders and later bring the issue to the national ministry of health level (22). Trust-building measures can take the form of face-to-face meetings, regular regional workshops, desktop exercises or joint outbreak investigations (see Box 3). These promote the sense of working towards a common goal. Many informants reported positive experiences of developing regional and subregional networks that facilitate efficient data sharing; and sharing was more likely to occur when those requesting data were physically located in the same region, as close as possible to the data. Bottom-up, regional networks have even been created in conflict areas (24), and have remained functional even when diplomatic relations between the parties involved are tense. Such regional networks not only build trust at the regional level, but also facilitate implementation of the IHR in the countries involved (37). Such success stories have prompted some global surveillance networks to organize themselves in smaller regional and subregional networks, with a physical presence in each region.
Beyond creating technical expertise, FETP/FELTP programmes create strong links between their graduates, who train and work closely together and sometimes spend extended periods of time in one another’s company on training modules. This creates cross-border professional networks of individuals who trust each other, are in direct contact with each other, and are able to bypass bureaucracy when a public health situation warrants it. One expert in a national public health agency mentioned another setting in which technical solutions promoted the creation of trust-based networks through fostering personal relationships:

The Global Health Security Initiative … was established in 2000 among countries with the involvement of the European Commission and the WHO. All the major systems owners of the event-based surveillance systems, media scanning systems and contributor-based systems like proMED systems have agreed to put feeds into a single portal … So that’s one of the solutions technically, but then there’s the collaboration among the analysts in all the countries which see this information … You can facilitate the process of data sharing if you provide the right kind of technology.

Trust-building among experts is not sufficient to enable systematic data sharing, and as one informant emphasized: ‘You cannot have data sharing without political buy-in.’ Support and commitment from health authorities is essential, and regional networks require connections at the technical as well as the political level, as highlighted by one regional surveillance network:

We have two meetings every year at this cross-border site. This meeting is both sides and not only for the health people, we also invited authorities like the governor and … immigration, veterinarians and some NGOs. So in the meeting we know each other, we can discuss not only health issues, but also business and tourism and trade. We have some very close relationships.

The Mekong Basin Disease Surveillance (MBDS) network was established in Southeast Asia in 2001 to enhance cross-border outbreak investigations, public health response and communication using trust-based surveillance data sharing protocols (22). The MBDS, which began as a response to a serious cross-border cholera outbreak, is built on bilateral agreements among six regional governments: Cambodia, China (Yunnan and Guangxi), Laos, Myanmar, Thailand and Vietnam. The initiative started locally, but was supported by the Rockefeller Foundation as well as several development and UN agencies. It led to the health ministers of each MBDS member country signing a memorandum of understanding to provide an agreed framework for the governing structure and processes of the consortium (22). Since the launch, experts from the network have performed several joint outbreak investigations, and have improved cross-border diseases surveillance and capacity-building. Over several years, regularly scheduled cross-border meetings and workshops significantly improved cross-border surveillance data sharing. Now, outbreak investigations in different countries are performed with standardized tools, and data from the investigations are shared in designated workshops. Challenges, achievements and future plans are discussed in annual cross-border meetings. Public health professionals, network board members and policy-makers are invited to the meetings to sit together in order to identify and address possible problems. The experiences from the network are also routinely shared with similar networks globally.

Political buy-in is an essential component to the success of any surveillance network sharing data. The context largely determines how the involvement of politicians can contribute to the success of the network.
While in the above example the bottom-up, trust-building approach was appropriate, it contrasted with another surveillance network where data sharing was enabled by strong political will at the top backed by clear guidelines and a legal obligation to share, originating from central policy-makers.

**Motivational barriers**

Motivational barriers arise from a lack of incentives to share data. At the personal level, sharing data generally implies additional time and effort (through extracting and formatting data to the required standards), with no feedback on the use of the data and no credit for the work. The consequences of sharing can even be negative for the individual collecting data and his or her institute, as it may be perceived as a lost opportunity for publication – a particularly powerful barrier in a ‘publish or perish’ academic culture that extends to public health surveillance. As a result, individuals collecting data restrict access to it until all the analyses they have planned have been published. An additional publication-related barrier is the exclusive focus on results in scientific journals. At the organizational level, the additional resources required for efficient data sharing imply that the benefits of sharing must be clear for the investment to occur. At the policy level, according to one academic expert interviewed, the lack of perceived benefits among policy-makers leads to restricting the data that goes out of a country. This tendency towards restricting data access is reinforced by a lack of knowledge about how the information is going to be used. This is a particular concern when the data are sensitive, and can lead to damaged relationships between institutes or countries.

The consequences of sharing can even be negative for the individual collecting data and his or her institute, as it may be perceived as a lost opportunity for publication – a particularly powerful barrier in a ‘publish or perish’ academic culture which extends to public health surveillance.

**Overcoming motivational barriers**

Overcoming motivational barriers requires a combination of trust, value proposition and rewarding, appropriate legal or data governance frameworks and capacity-building.

*Being explicit about secondary data use*

For an individual or an organization collecting data, confidence and guarantees regarding what will and will not happen to the data are paramount. An institute collecting data may on one level be inclined to share it onwards in order to improve global surveillance capacity, but may on another level be discouraged by seeing publications using its data without due attribution. Fostering networks of individuals who personally know one another, as described above, is one way to overcome this reservation. Legal or data governance frameworks specifying what can and cannot happen with the data (specifying, for example, that the data can be used for surveillance but not for publication without the consent and participation of the institute or the individual who generated it), with whom the data will be shared, and for what purpose, may improve the value proposition of data sharing and may alleviate fears over data misuse and missed opportunities for publication. For example, the ECDC data access policy stipulates that data from an EU member state cannot be used for publication without approval from that member state (35). The current standard practice, however, leans more towards restricting data access.
Overcoming Barriers to Data Sharing in Public Health: A Global Perspective

Changing the publishing culture

While public health surveillance data are primarily collected for monitoring and for public health action, routinely collected data generates extensive data sets that are often underutilized. When surveillance systems detect an unusual or novel event, the publication potential and value of these data becomes high, and the distinction between research and surveillance data for publication purposes becomes theoretical. For example, two articles – one describing the epidemiology of the first 47 reported cases of MERS-CoV in Saudi Arabia (38), and one describing the epidemiology of the first nine months of the Ebola outbreak in West Africa (39), both relying on routine surveillance data – were published in high-impact journals and respectively cited 120 and 228 times. Changing the publishing culture among scientific journals could therefore lead to more timely data sharing. Changing the ‘publish or perish’ message to one of ‘publish data or perish’ (5) would allow researchers to share data without fear of negative effects on their career development. Researchers would be able to share data immediately for the best interest of the public, and journals would allow this sharing without devaluing the subsequent peer-review publication. Several initiatives have encouraged data sharing as part of the publication process: requiring a data management plan as part of grant proposals (4); requiring or encouraging authors to make data publicly available when publishing (40); and publishing data sets in citable data journals (41). Furthermore, rewarding researchers who engage in sharing data, by means of professional bonuses, funding or promotions, can help interrupt the current norms influencing sharing in academia (13).

Economic barriers

Economic barriers restrict data sharing in two ways. First, a country or an institution may lack the financial resources for disease surveillance, and may also lack skilled human resources because of limited training capacity, difficulties in retaining staff or high staff turnover. Second, sharing data may have negative economic consequences resulting from decreased travel and trade.

Developing training capacity at the local and regional level is a step towards addressing the lack of human resources, but subsequently retaining staff raises an additional barrier, especially in LMICs.

The direct costs of data sharing need to be addressed. Sharing implies locating and extracting data, cleaning the data and preparing them in the required format. This process requires skilled staff and dedicated time. An academic expert described how in East and Southern Africa, apart from in South Africa, it is difficult to get high-quality IT staff who maintain high data management standards. Developing training capacity at the local and regional level is a step towards addressing the lack of human resources, but subsequently retaining staff raises an additional barrier, especially in LMICs.

The negative impact of outbreaks on travel and trade can lead to countries' reluctance to share surveillance data openly. The 2003 SARS outbreak cost an estimated $40–$80 billion, with the tourism and travel industry severely affected (43). In Europe in 2011, the food-borne outbreak caused by the contamination of sprouted seeds by *E. coli* O104 (described in Box 2) resulted in losses to farmers in the fruit and vegetable sector in the EU of more than €800 million in the first two weeks of the outbreak, and led to a temporary ban on agricultural exports to Russia (44). As a senior member of an international public health institute explained: ‘Countries economically dependent on tourism would be very hesitant to disclose any kind of outbreaks in their countries or some of the food and
Overcoming Barriers to Data Sharing in Public Health: A Global Perspective

Overcoming economic barriers

Learning from successful capacity-building initiatives

If sharing is to be systematic and standardized, the time and skills required should be planned for when hiring and training staff. Beyond providing funding to help countries lacking resources to build up their public health capacity, offering incentives for sharing is another strategy for overcoming economic barriers, at the personal or organizational level, as described by an informant from an international disease reporting network:

> There are costs to gathering the data, and there are costs to sharing the data. Both actual costs of just the process, but also secondary costs, of trade barriers and political and so forth. So a way of compensating those providing the information, incentivizing the sharing of information maybe should go hand-in-hand with regulatory reform or legal reform.

Staff retention remains a problem. However, more than 80 per cent of FETP graduates stay in their home country, many becoming leaders within their public health systems, and there may be lessons to learn from these programmes for the broader public health workforce. Financial incentives or programme-specific regulations such as obligatory service in the region after the training can also bolster retention rates. Dedicated funding for data sharing is dependent not only on resources available, but also on the political will to fund a quality surveillance system, and engaging at the political level in data sharing discussions may encourage such investments (see Political Barriers).

Box 4 describes an alternative solution to funding data sharing during public health emergencies.

Mitigating the economic impact

One expert from an international public health agency suggested a mechanism to mitigate the economic impact of outbreaks to countries sharing their data, via the IHR:

> Many people in infectious disease prevention and control only see one side of the IHR, that countries should be reporting outbreaks internationally, but there’s a flipside to that coin and that is that countries should not be punished when they report and that’s almost always forgotten by experts: the trade, the traffic … So I think the rich countries have to realize that when they sign the IHR, it’s not just for poor countries to report their problems, but also for rich countries to take the responsibility when it comes to measures … how you do that I have no idea, but it’s really at the heart of the IHR if it’s going to work.

Influencing the private sector to mitigate the economic impact of an outbreak – such as by encouraging airlines to maintain flights when the WHO does not recommend travel restrictions – remains a challenge. For example, in August 2014, at the height of the West African Ebola outbreak, only one airline had maintained its scheduled flights to Liberia, compared with the normal 12, even though the WHO and other UN agencies insisted that the risk of Ebola transmission during a flight was low and urged airlines to keep their schedules in place (40). Such economic consequences may delay data sharing or prevent affected countries from releasing epidemiological information.
Overcoming Barriers to Data Sharing in Public Health: A Global Perspective

Ethical barriers

Two issues dominate the ethical debate on data sharing: lack of reciprocity and lack of shared benefits between those providing and those requesting data; and privacy and confidentiality. The benefits can be academic credit, where inequity between HICs and LMICs is frequent, as one academic expert explained:

Once others have data that you're holding onto they will be able to analyse it. Publishers will quickly use it to make all sorts of decisions that perhaps don't come back as benefits to the originators of the data. I see this very often in agreements between stakeholders in high-income countries and low-income countries, with those in low-income countries usually being the generators of data (and then feeling) like they've entered into agreements which do not equitably distribute the benefits.

The benefits in some instances are understood as broader public health gains, such as access to public health interventions. One highly publicized event was Indonesia’s refusal to share H5N1 influenza viral sequences during the 2007 avian influenza outbreak, out of concern that samples provided freely by LMICs are used to develop vaccines that LMICs themselves cannot afford (43). This greatly precipitated the emergence of the Pandemic Influenza Preparedness Framework, which encourages the WHO member states to share influenza virus sequences in exchange for increased access to vaccines (45).

Overcoming ethical barriers

Anonymization

Confidentiality issues can often be solved by anonymization, although in some instances even anonymized data enable the identification of individuals, particularly when accessing small area statistics, and given that linking different databases is becoming increasingly feasible (45). One expert gave the example of the notification of a case of HIV in an individual of African origin in a sparsely populated area of a Nordic country, which only had one resident of African origin. Such possibilities must be taken into account when developing anonymization standards. An academic expert also pointed out that anonymization standards need to be enforced with the same efficiency in LMICs as in HICs, which is not currently the case. Another informant also pointed out that confidentiality was not only about anonymizing data, but also about the public feeling secure with regard to the use of their data.

Box 4: Outbreak response contingency fund

The US CDC command centre for monitoring and coordinating emergency response has set up an outbreak response contingency fund. The aim of this fund is to promote information sharing within the agency that consists of several programmes responsible for responding to events within their disease area. If the aetiology of the disease event is unclear or unknown, no programme is directly responsible for the response, and programmes do not necessarily have the funds for immediate information gathering and investigation. When the programmes detect particular disease events or become aware of an event through different channels, they turn to the command centre and request access to the contingency fund and share their information with the centre. Thus, the fund is a means for the CDC to have the ability rapidly to gather and disseminate information in situations where timeliness is important.
Global ethical framework

A global ethical framework may be a realistic approach to overcoming inequity in data sharing. The value of data may be seen differently by the involved stakeholders, and it is important to identify and acknowledge the different values and interests that people have in the sharing process. The principles should address the risks and benefits of sharing in a reciprocal manner, clearly informing the stakeholders about mutual expectations. Reciprocity also helps build goodwill and trust. Global data governance principles that are responsive to the concerns and needs of stakeholders, and to some extent adaptable on a case-by-case basis, could be effective in overcoming ethical barriers. This would entail inclusion in the academic process to a degree commensurate with involvement, if that is valued by the data provider, as well as access to the public health benefits deriving from the data provided. It would also ensure that data provided are used for specific purposes agreed at the outset, and that any other use would involve a dialogue with data providers.
Data Sharing in Public Health Emergency Situations: What is Different?

There are a number of differences when public health data are shared during emergencies. The sense of urgency may make sharing easier; however, economic interests may impede sharing. Another delay to sharing during emergencies occurs when the situation is managed locally and stakeholders at the national level are not informed in detail at the outset. When the data are eventually shared, this is often too late for an intervention. Political barriers are exacerbated in public health emergency situations, and the situation rapidly becomes very complex when different data sources emanating from different agencies need to be integrated, as a public health specialist from a national public health agency described:

During a major natural disaster, health data were being collected in emergency shelters in multiple jurisdictions and in multiple formats (different variables, some individual data, some aggregate data, different time periods, sometimes with denominators, sometimes only numerator data). Data sharing resulted from multiple conversations with multiple jurisdictions to clarify the type of data that would be useful to collect so that it could be interpreted and compared across jurisdictions. Bidirectional communication, providing opportunities for input and feedback, and time for the process to work out the discrepancies allowed useful data to be captured. But, there were delays in generating and accessing useful data.

Data ownership sometimes becomes unclear in emergencies, and there are some ethical issues around the use of data after the emergency situation is over. There are often no plans for the long-term use of data collected in these situations. Trust and pre-existing relationships are key to achieving well-functioning data sharing during emergencies. When responding to an emergency, public health professionals tend to be more trusting in relation to those who can help them, and where the help is more immediate. The application of the data is also more apparent, and therefore increases motivation to share. In the words of one informant: ‘Parsimony can be difficult when exploring a new situation to establish intervention measures, and over-collection is a common problem.’
Conclusions: Data Sharing in an Ideal World

Global surveillance could ideally be, as one academic informant saw it: ‘A globally networked system that is constantly monitored by a dedicated institution of highly trained individuals that can detect and investigate anomalies with local authorities in real-time and that is linked to a well-funded international response system that can direct resources where and when needed.’

When shared, data should be accessible, real-time, interdisciplinary, robustly anonymized and properly documented, to allow easy secondary use. This is underpinned by the principle that data collected by publicly funded institutions should be publicly accessible. This would also promote shared benefits among HICs and LMICs and encourage scientific innovation: ‘You don’t know where the next great idea is going to come from. So a doctoral student or a postdoc in Nigeria or in Vancouver who has access to this data may have ideas about different things that could be done. That would be useful for decision-making, better policy decisions around how you handle some of these diseases.’

A creative commons-type licence for data sets that allow creators to communicate which rights they reserve, and which rights they waive, for the benefit of recipients or other creators, would alleviate fears of inequity and would enable sharing to improve public health while retaining publishing privileges. Even in an ideal world, public health data sharing is not ‘everything available to everyone as long as the person’s name is not on it’. With the risk of de-anonymization and data misuse, this approach could eventually backfire and undermine public trust. Open access public health data may be neither feasible nor desirable, and it remains debatable whether the Panton Principles, whereby data in science should be freely available for downloading, copying, reprocessing or analysis, without any financial, legal or technical barriers (47), apply to public health surveillance.

In the real world, data sharing in public health is a complex landscape, made up of a range of data sharing arrangements based on local needs in a local context and dependent on the specific type of data. One recurring pattern is that successful surveillance networks sharing data as a matter of routine have emerged from local or regional initiatives responding to a local public health crisis. More globally, the few data sharing frameworks that exist in public health have arisen from crises, rather than from forward planning. Data sharing in public health is successful when a perceived need is addressed, and when the social, political and cultural context is taken into account. This can translate into an informal, trust-based local network or a nationally enforced mandatory sharing policy, but a single global model is unlikely to be successful: the wide range of very different legal systems worldwide makes the prospect of a global enforceable data sharing policy unlikely. However, a global ethical framework, or ethical principles, accompanied by local governance agreements that are flexible enough to take the local context into account, may be more conducive to success. Thus, perhaps a collection of local networks feeding into a gradually more global structure is a model to consider.

In order to achieve this, legal, political, economic, technical, motivational and ethical barriers have to be overcome, with the first three considered to be the greatest challenges. A complex interaction of trust, value proposition, political advocacy and capacity-building is required (Figure 1).
Generating a global framework for disease surveillance does not occur in a vacuum, and the global health community already has tools in place. The IHR are widely seen as the backbone of any future global framework, but several issues need to be addressed: first, the lack of an enforcement mechanism; second, the lack of detailed specifications on the minimum data sets for sharing; third, the lack of a capacity-assessment mechanism; and fourth, potential negative trade and travel consequences for notifying countries. Regardless of the mechanisms of sharing, the first step is always that the individual or institution collecting the data is able and willing to share it. In other words, as one informant stated: ‘The solution needs to be something which everybody, all the stakeholders, see as improvements to their own work.’
## Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMC</td>
<td>British Medical Council</td>
</tr>
<tr>
<td>CDC</td>
<td>[US] Centers for Disease Control and Prevention</td>
</tr>
<tr>
<td>CGHS</td>
<td>Centre on Global Health Security</td>
</tr>
<tr>
<td>E. coli</td>
<td>Escherichia coli</td>
</tr>
<tr>
<td>ECDC</td>
<td>European Centre for Disease Prevention and Control</td>
</tr>
<tr>
<td>EEA</td>
<td>European Economic Area</td>
</tr>
<tr>
<td>EU</td>
<td>European Union</td>
</tr>
<tr>
<td>FETP</td>
<td>Field Epidemiology Training Program</td>
</tr>
<tr>
<td>FELTP</td>
<td>Field Epidemiology and Laboratory Training Program</td>
</tr>
<tr>
<td>GISN</td>
<td>Global Influenza Surveillance Network</td>
</tr>
<tr>
<td>GISRS</td>
<td>Global Influenza Surveillance and Response System</td>
</tr>
<tr>
<td>HICs</td>
<td>high-income countries</td>
</tr>
<tr>
<td>HIV/AIDS</td>
<td>human immunodeficiency virus/acquired immune deficiency syndrome</td>
</tr>
<tr>
<td>IHR</td>
<td>International Health Regulations</td>
</tr>
<tr>
<td>LMICs</td>
<td>low- and middle-income countries</td>
</tr>
<tr>
<td>MBDS</td>
<td>Mekong Basin Disease Surveillance</td>
</tr>
<tr>
<td>MERS-CoV</td>
<td>Middle East respiratory syndrome coronavirus</td>
</tr>
<tr>
<td>proMED</td>
<td>Program for Monitoring Emerging Diseases</td>
</tr>
<tr>
<td>SARS</td>
<td>severe acute respiratory syndrome</td>
</tr>
<tr>
<td>TESSy</td>
<td>The European Surveillance System</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
</tbody>
</table>
References


About the Authors

**Michael Edelstein** is a medical doctor specializing in field epidemiology, with expertise in surveillance systems, outbreak investigations and global health policy, in both routine and emergency contexts. He is currently a consultant research fellow at the Centre on Global Health Security at Chatham House. He has worked for national public health agencies and international public health organizations in Europe, Africa and Southeast Asia.

**Dr Jussi Sane** is a senior epidemiologist at the National Institute for Health and Welfare, Helsinki, Finland. His work includes infectious disease surveillance, preparedness and research. He is also active in Finland's contribution to the ongoing Global Health Security Agenda. He has a master's degree in medical science from Karolinska Institute, Stockholm, and a PhD in medical virology/epidemiology from the University of Helsinki, and is a graduate of the European Programme for Intervention Epidemiology Training (EPIET).