wellcome

GLOPID-R

UKaid from the British people

Data sharing in public health emergencies

Learning from past outbreaks

Data sharing in public health emergencies: Learning lessons from past outbreaks

Report from a Wellcome/DfID workshop
5 December 2018

Executive summary

Data gathered during public health emergencies have the potential to enhance public health responses during an immediate infectious disease outbreak as well as preparedness for future outbreaks. However, data-sharing practices have been inconsistent across – and even within – many major outbreaks over the past two decades, limiting potential improvements to the immediate public health response, and narrowing the contribution of research to augmenting knowledge and improving disease control interventions.

A wide range of obstacles to timely data sharing are often cited, many linked to the culture and practices of academic publishing and the importance attached to publication of papers in high-profile journals for academic career security and advancement, for continued funding, and for institutional status. Funders and journal publishers have begun to take steps to ensure that these processes do not stand in the way of timely data sharing in public health emergencies

The Global Research Collaboration for Infectious Disease Preparedness (GloPID-R) group of funders has been working to support timely data sharing and has developed a set of principles to guide data-sharing practices in public health emergencies. These principles emphasise the importance of timely, ethical and equitable sharing of data.

In a bid to develop a deeper understanding of the obstacles and enablers of data sharing during emergency situations, Wellcome and the UK Department for International Development (DfID) commissioned international academic groups to develop six case studies on behalf of the GloPID-R data-sharing working group. Based on extensive literature reviews and stakeholder interviews, these cases studies explore data-sharing practices during a range of infectious disease outbreaks of varying severity, geographical exposure and public health impact, and where medical countermeasures were or were not available.

The case studies identified multiple obstacles to and enablers of data sharing in public health emergencies. While some issues were context-specific, many of the overarching obstacles and enablers were common across different outbreak scenarios.

The case studies also highlighted a range of factors influencing sharing and the use of data in public health emergencies, which add complexity to data-sharing decision-making and will need to be taken into account in order to operationalise data-sharing principles. These include heterogeneity in the nature of data being collected (e.g. epidemiological, clinical, laboratory, genetic), the varying purposes of data collection, and the

distinct data needs of different data users. Data completeness, quality and accessibility are also important factors to be considered alongside data availability.

Overcoming data-sharing barriers is likely to require action across multiple domains. These include:

- **Developing a more systems-oriented view of data sharing, with greater clarity on the purpose of data collection, the needs of data users, and the ecosystem of actors involved in data collection, management, analysis and use**, to inform the design of information management systems and processes that meet immediate as well as longer-term public health needs.
- **Developing advocacy materials** that highlight the public health, economic and other benefits of data sharing to encourage transparency in national outbreak reporting and to promote a data-sharing culture.
- **Developing international partnerships and research programmes to build local technical capacities** in epidemic preparedness and response, including laboratory capacity, data collection, management and analysis capacity, and to **encourage collaboration across research, public health and response communities**, in advance of public health emergencies.
- **Developing data-sharing platforms for diseases with epidemic potential** with robust data-governance and equitable data-ownership principles that consider the interests of all stakeholders.
- **Developing and promoting the use of standardised tools and approaches to support international research collaborations and data sharing,** including material transfer and data-sharing agreements, and streamlined and harmonised regulatory approaches.
- **Promoting funding policies that recognise research quality and public health impact**, and require and incentivise timely data sharing, rather than prioritising publication.

Introduction

Infectious disease outbreaks remain an important cause of morbidity and mortality in low- and middle-income countries (LMICs), and can inflict severe damage on countries' economies. The 2014–16 Ebola outbreak, for example, claimed more than 14,000 lives directly as well as a similar number through indirect effects on health systems, and cost the economies of the three countries most affected an estimated US$2.2bn in 2015 alone.

Infectious disease outbreaks are a common cause of public health emergencies. The causative agent may be known or a novel pathogen, and treatments and preventive interventions such as vaccines may or may not be available. As outbreaks are detected, the immediate priorities are to treat those affected and to contain the spread of infection.

In addition, public health emergencies are an opportunity to collect data for research, generating evidence to enhance responses to an ongoing outbreak as well as future outbreaks. Indeed, for some infections, such as Ebola, outbreaks are the only situations in which clinical data on intervention efficacy can be obtained. Furthermore, time is a critical factor during public health emergencies, and timely access to and analysis of data can generate a deeper understanding of an outbreak, its impact on patients, and effective methods of control – supporting more effective public health responses.

Access to data is therefore critical in emergency public health responses. Timely data can facilitate targeting of resources to control the spread on infection, promote the use of evidence-based interventions for treatment or prevention, avoid duplication of efforts and ensure the efficiency of emergency responses. In addition, analysis, and sometimes pooling of data from multiple sources, can generate a clearer picture of the epidemiology of infection, pathophysiology and intervention effectiveness.

Following a multi-stakeholder consultation in 2015, WHO issued a position statement on global norms for data sharing in public health emergencies, including the roles of scientific publishers[1] [2].In recognition of the importance of timely data sharing in public health emergencies, and in part in response to the challenges to data collection and sharing experienced during the 2014–16 Ebola outbreak[3], in 2016 funders and other stakeholders, including scientific publishers, issued a joint statement on data sharing in public health emergencies, to promote the sharing of data on the ongoing Zika outbreak[4]. These principles are enshrined in the recommendations of good scientific publishing practice published by the International Committee of Medical Journal Editors[5].

---

[1] WHO. Developing global norms for sharing data and results during public health emergencies. 2016. Available at: https://www.who.int/medicines/ebola-treatment/blueprint_phe_data-share-results/en/

[2] Modjarrad K *et al*. Developing global norms for sharing data and results during public health emergencies. *PLoS Med.* 2016;13(1):e1001935.

[3] Keusch G *et al*. Integrating Clinical Research into Epidemic Response: The Ebola Experience. 2017. Washington DC: National Academies Press

[4] Wellcome. Sharing data during Zika and other global health emergencies. 2016. Available at: https://wellcome.ac.uk/news/sharing-data-during-zika-and-other-global-health-emergencies

[5] International Committee of Medical Journal Editors. Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals. 2018. International Committee of Medical Journal Editors. Available at: http://www.icmje.org/recommendations/

LMICs face a disproportionately high infectious disease burden. Furthermore, many LMICs have limited capacity to detect, respond to and control emerging and re-emerging pathogens. In particular, such countries may lack the capacity, in terms of physical infrastructure and trained personnel, to undertake epidemiological, clinical and other research during demanding public health emergencies. Existing and new global networks can be mobilised in response to emergency situations, but the Ebola experience highlights how this can lead to a diversity of poorly coordinated actions at both national and international levels. In addition, humanitarian-driven responses typically do not prioritise data collection for research purposes, while research data may not be shared optimally to support public health responses.

Indeed, a review of the practices, policies and infrastructure for sharing data from research on WHO's 12 priority pathogens, commissioned by the GloPID-R data-sharing working group, found that most data were not being shared through formal, discoverable means[6]. Data made available through publications or data repositories are typically not in a format that can be readily be reused. In addition, considerable data sharing occurred through informal trust-based closed networks. Informal data sharing is popular and facilitates rapid dissemination of information, but lacks transparency and accountability, and excludes many potential users.

Multiple steps are now being taken to develop better-coordinated global responses to major infectious disease outbreaks. WHO's R&D Blueprint provides a global strategy and preparedness plan to ensure rapid activation of R&D activities during epidemics. The R&D Blueprint includes a Global Coordination Mechanism to facilitate dialogue among multiple stakeholders involved in research preparedness and responses. The WHO Health Emergencies Programme has been established to coordinate activities at national, regional and global levels.

In addition, multiple activities have been launched to build capacity in LMICs to detect, prevent and respond to infectious disease outbreaks, and to enhance preparedness to carry out research in outbreak situations. Most are based on regional and international networks and collaborations.

Data sharing will be crucial both within these international collaborations but also with wider research and public health communities. Past experience suggests that multiple barriers exist to timely data sharing, often linked to academic reward systems, and the need to be the first to publish, ideally in high-profile journals, in order to secure academic advancement. There are widespread, and often well-founded, concerns that those generating and collecting data do not receive due credit for their contributions, and fears that raw data rapidly made available in LMICs will be exploited by those in high-income countries with greater resources and greater capacity to undertake data analyses. Lack of trust about how data will be used and credit attributed can therefore lead to a reluctance to share data. Even when data analyses are published, underlying data are made available in only a minority of cases[7].

The GloPID-R network of funding bodies has established a data-sharing working group to consider possible ways to enhance timely data sharing in public health emergency contexts. Following consultation, the

[6] Pisani E, Ghataure A, Merson L. Data Sharing in Public Health Emergencies: A study of current policies, practices and infrastructure supporting the sharing of data to prevent and respond to epidemic and pandemic threats. 2018. London: Wellcome

[7] Terry RF, Littler K, Olliaro PL. Sharing health research data – the role of funders in improving the impact. *F1000Research* 2018, 7:1641

working group has established a set of principles to guide timely data sharing[8], recognising the differing interests of stakeholders (Figure 1). A key challenge now is to develop mechanisms to operationalise these principles to encourage openness in data availability to enhance public health responses.

Figure 1: GloPID-R's key principles for data sharing in public health emergencies[6].

Case study project

To inform the work of the GloPID-R data-sharing working group, in 2018 Wellcome and DfID launched an open request for proposals, for research projects to generate a deeper understanding of data-sharing practices in past outbreaks. On the basis of merit, academic groups in the USA and Europe, with global collaborators were selected to undertake six case studies.

The outbreaks chosen for analysis were selected to be illustrative of different outbreak scenarios (for example, with known/unknown pathogens, where vaccines and treatments were or were not available, and in different global regions; Box 1). Annex 1 includes brief summaries of the projects; full reports will be published alongside this report.

Box: Outbreaks covered in the case studies

Infection: West African Ebola outbreak (two case studies)

Timeline: 2014–16

Main countries affected: Liberia, Sierra Leone, Guinea

Impact: 28,616 reported cases, 11,310 deaths [true impact almost certainly greater and wider]

Brief summary: Largest ever Ebola outbreak, with very high fatality rate. Ebolavirus transmitted from animal reservoirs and highly contagious. Contact tracing and isolation key to disease control. Trials of experimental vaccines initiated towards the end of the epidemic, but much controversy about appropriate study design, lack of coordination, and sample and data management practices.

Infection: Middle East Respiratory Syndrome (MERS) (two case studies)

[8] GloPID-R. Principles of data sharing in public health emergencies. 2018. Available at https://www.glopid-r.org/our-work/data-sharing/

Timeline: 2012–

Main countries affected: Saudi Arabia, United Arab Emirates, Korea

Impact: 2266 confirmed cases in 27 countries, 804 deaths

Brief summary: A novel zoonotic respiratory disease with a high fatality rate. Ultimately linked to an emerging coronavirus, MERS-CoV, transmitted to humans from camels. Continuing risk of new introductions in the Arabian Peninsula and neighbouring countries, and hospital-acquired infections there and internationally. No specific treatments or licensed vaccine.

Infection: Yellow fever

Timeline: 2016

Main countries affected: Angola, Democratic Republic of the Congo, China

Impact: 5423 suspected cases, 442 reported deaths (Angola and DRC)

Brief summary: Vaccine-preventable, mosquito-transmitted, viral disease. Low vaccine population coverage, in part linked to vaccine shortages, leads to periodic outbreaks, generally brought under control by mass vaccination campaigns. Vaccine shortages led to the use of fractional dosing – immunization with one-fifth of the standard vaccine dose (which experimental data indicated would be sufficient to provide protection, but had not been evaluated in an outbreak).

Infection: Cholera

Timeline: 2012 (Guinea), 2011, 2012, 2014-2015 (Cote d'Ivoire)

Main countries affected: Guinea, Cote d'Ivoire

Impact: 7350 cases (Guinea), 2189 (Cote d'Ivoire)

Brief summary: An endemic bacterial disease in many LMICs with potential for major outbreaks of severe diarrhoeal disease. Spread through contaminated water and person-to-person; cholera control focuses on water, sanitation and hygiene interventions, and has more recently relied on use of oral cholera vaccine (OCV). The 2012 Guinea outbreak was the first in which OCV was deployed to control the spread of infection. Factors affecting the transmission of infection are incompletely understood, hampering control efforts.

Infection: Zika [9]

Timeline: 2015–17

Main countries affected: Brazil, other Latin American countries

Impact: Estimated 1.5 million infections in Brazil and more than 3500 cases of microcephaly

Brief summary: The mosquito-transmitted Zika virus has caused multiple outbreaks globally, but the Latin American outbreak was the largest and was declared a Public Health Emergency of International Concern by WHO. Zika infections are generally mild but in pregnant women can lead to microcephaly in offspring; they also increase the risk of an inflammatory reaction, Guillain–Barré syndrome. Initial stages of the epidemic were characterised by considerable uncertainty about its epidemiology and clinical impact (including its link with microcephaly). There are no specific treatments or licensed vaccines.

[9] The zika case study was completed and self-funded by the International Severe Acute Respiratory and Emerging Infection Consortium (ISARIC)

All the projects were based on mixed methods, encompassing extensive literature reviews and interviews with key stakeholders. A key focus was on the perceived obstacles to and enablers of timely data sharing, as well as possible actions that could be taken by funders or other global stakeholders to facilitate timely data sharing in public health emergencies. A workshop was held in December 2018 at which research groups presented and discussed their initial findings. This report draws on the case studies and discussions held at the workshop.

Emergent themes

Several key themes emerged from the case studies and workshop discussions:

Data sharing is complex – operationalisation of data-sharing principles will need to take a more nuanced view of the multifaceted nature of data, the multiple purposes of data sharing, and the multiplicity of actors involved in its generation, management, analysis and use.

- **The purpose of data sharing** can vary; it can be used to guide national public health responses, to inform clinical decision-making, to provide regional and global insight into outbreaks, to improve understanding of transmission, to assess the efficacy and safety of interventions, and to assess the effectiveness of public health measures (such as contact tracing, or ring vaccination) and community engagement.
- The **data needs of different stakeholders** involved in public health responses also vary and may be inadequately understood – what data do they need, when, and in what format?
- The **type of data** also varies significantly (e.g. epidemiological, genetic, clinical, laboratory, social and behavioural data). The nature of data affects data-sharing practices; concerns about patient confidentiality, and other concerns about privacy or anonymity including national and international data protection legislation, for example, influence clinical data sharing, while academic community norms around pathogen genome sequence data tend to promote rapid data sharing. There are large areas of neglected data, specifically social science and community engagement data.
- Data flow through a **complex ecosystem**, with multiple different collectors, analysers and users of information, and multiple channels of information flow. Identifying and communicating the **value added** (public good) as data move through this ecosystem could promote data sharing.
- **Data quality** and **data completeness** are important as well as data availability. For emergency responses, quality and completeness may be less important than timely availability. For more formal meta-analyses, data quality, completeness and usability become paramount.

Many obstacles to data sharing are common across outbreak scenarios; some are context-specific.

- The case studies identified a **wide range of obstacles to and enablers of data sharing**, across multiple domains (Table 1). Relatively few related to the specific context of an outbreak.
- The case studies concluded that outbreaks are characterised by **'business as usual' behaviour with limited data sharing**, until a situation becomes very serious when behavioural norms are abandoned and data-sharing channels are opened, especially through informal networks.

What constitutes a public health emergency may not be clearcut.

- Infectious disease outbreaks form a **continuum of public health importance**, and it may be challenging to define the point at which they are considered 'public health emergencies'.
- **Inadequate detection and control can turn a minor outbreak into a public health emergency**; in addition, the epidemic potential of an emerging infection may not be clear at early stages of an infection.
- Public Health Emergencies of International Concern (as defined by International Health Regulations) are **rare and unpredictable** – 'switching on' special data-sharing measures when they are declared would set a very high bar and lead to delays before 'enhanced' data-sharing measures were initiated.
- In the long term, therefore, **data sharing should become a routine aspect of research on infections of epidemic potential** and associated public health responses.
- In the short term, **priority should be given to facilitating data sharing in broadly defined actual and potential public health emergencies**, reflecting the specific challenges associated with such circumstances, such as the need for rapid access even to unprocessed data, the likely co-existence of humanitarian, public health and research responses, and inherent uncertainties in the evolution of an outbreak.

Obstacles and enablers

Annex 2 summarises the obstacles and enablers of data sharing identified in the case studies, as well as possible actions that could be taken by funders or other international agencies.

International relationships: Historical legacies and long-standing geopolitical links can have both a positive and negative impact on data sharing. Historical connections can facilitate links between particular academic and public health institutions in the global North and global South. However, colonial legacies and economic and/or social inequalities can also foster tensions and mistrust, particularly fears of 'data exploitation'.

International collaborations (South–South and North–South) will undoubtedly play key roles in development of outbreak preparedness and responses capacities. Such partnerships should be based on fair and equitable agreements on data usage, and ensure that they reflect and are responsive to local public health priorities. Formal agreements, including memorandums of understanding in line with global standards and GloPID-R data-sharing principles, can facilitate the operation of such partnerships, as part of robust governance frameworks. Such partnerships and collaborations, and their governance and operating procedures, should be established in advance of emergency situations.

The importance of WHO's role was widely recognised. WHO is a trusted source of advice and guidance, and often a key entry point into global networks. It is therefore well placed to promote timely data sharing. However, WHO's accountability to national governments may lead to a reluctance to challenge data secrecy.

The involvement of commercial interests was found to have a significant impact on data sharing. Non-disclosure agreements may restrict researchers' ability to make data available. Companies may also be reluctant to share negative findings because of their potential commercial implications. Further dialogue is

required with industry associations to ensure greater transparency, and clear agreements developed for public–private product development partnerships.

National interests and capacities: Countries may be reluctant to acknowledge cases of infectious disease for fear of inflicting damage on their national interests in trade or tourism, or more generally to protect national reputations. More could be done to communicate to opinion leaders the importance of openness and the risks associated with secrecy, particularly the benefits associated with rapid detection, rapid mobilisation of an outbreak response and prevention of large-scale outbreaks. Regional health and/or economic bodies and South–South partnerships could promote more openness and coordinated responses to outbreaks.

Openness could also be enhanced by greater communication of the public health benefits of data sharing (and the risks associated with data secrecy). This could take the form of case studies of timely data sharing with public health benefit, or modelling and/or economic analyses to emphasise the value of effective outbreak control. National public health champions and civil society organisations could be identified to galvanise support for openness and data sharing.

Lack of technical capacity to detect, investigate and track outbreaks is a major challenge in many LMICs. A key goal of international partnerships and research programmes should be to build local capacity in these areas. This should include support for national infrastructure for sharing infection-related data.

Regulatory and legal frameworks for research can be significant obstacles to timely data collection. International efforts could focus on building national regulatory and ethical review capabilities, and working with regional bodies to develop harmonised, streamlined but rigorous research oversight practices. Regional approaches could also be a way to overcome tendencies towards secrecy to protect national interests.

Across all case studies, transparency was seen as a critical facilitator of data sharing – encouraging a culture of openness and collaboration.

Cultural and behavioural factors: Multiple factors were found to influence individual data-sharing behaviours. A common concern was that data would not be analysed appropriately by others. Data were often therefore shared through informal networks of 'trusted associates'. Informal networks can be important data conduits when formal channels are not functioning, but they lack transparency and accountability and can exclude important potential data users.

The case studies also suggested that a failure of 'data analysers' to recognise the investments in time and energy of 'data collectors' inhibits wider sharing. More equitable arrangements, for example support to build skills and capacity for data analysis, or even simply providing more feedback on how data have been used, could encourage more open attitudes.

Cross-sectoral collaborations (for example between veterinary and human medicine/public health communities) were often challenging, for example due to different ways of working, vocabularies and

priorities. Building partnerships in advance, for example under the umbrella of One Health initiatives, would facilitate efficient operations in emergency situations.

Similarly, pro-active investment in community engagement could promote positive public attitudes to research and clinical data sharing in outbreak situations.

Academic publishing models and academic culture: The current model of academic publishing is a well-recognised obstacle to timely data sharing. Authorship of academic papers – and being the first to publish – is seen as a critical route to academic advancement and success in research grant applications, encouraging proprietorial attitudes to data. In addition, journal policies on pre-publication communication of results have discouraged researchers from timely dissemination of data.

Several funders have recognised this issue and have developed policies promoting publication in open access publications (for example through 'plan S'[10], supported by the 'Coalition S group of funders[11]). In addition, funders are attempting to move away from publication record as the key arbiter of academic distinction, recognising other indicators of clinical or public health impact. Consistent with the San Francisco Declaration on Research Assessment[12], funders are indicating that grant-giving should be based on the quality of research not on the impact factor of the journal in which it is published. However, it is likely that, for the immediate future, researchers in LMICs are still likely to feel that career advancement will depend on publication on papers in high-impact journals, and funders need to demonstrate a firm commitment to new ways of working and to actively communicate and implement their policies on publication of data and assessment of grant applications.

For their part, academic publishers have introduced policies to enable pre-publication sharing of data of public health importance, and introduced fast-track mechanisms for publication of results of critical health importance. It will be important to ensure that these principles are not applied too restrictively, acknowledging the varying severity of public health emergencies and therefore the period of time where data has public health significance extends beyond the height of the outbreak.

International partnerships should also establish arrangements that appropriately acknowledge local contributions to research in publications. They should also ensure that they build the capacity of local researchers to generate academic outputs.

- **Technical facilitators**: The case studies suggested that technical data-sharing solutions are necessary but not sufficient for timely and effective data sharing. There are also examples of where adequate infrastructure is lacking for certain data types, such as social science and community engagement data. Data platforms and associated operating procedures (and governance frameworks) are a key infrastructure, but will need to be backed up by training and awareness

[10] Science Europe. 'Plan S' Making Open Access a Reality by 2020. 2017. Available at https://www.scienceeurope.org/making-open-access-a-reality-by-2020/

[11] https://www.coalition-s.org

[12] San Francisco Declaration on Research Assessment. Available at https://sfdora.org/read/

raising to ensure their effective use. Independent but coordinated databases following consistent data collection practices may be a possible alternative to centralised databases.

Data integration will be a key future challenge. Collection of clinical data could be enhanced by integration into national health information systems or national epidemiological data platforms. While databases for research information exist for some infectious diseases, sharing of social/behavioural rather than quantitative data may be more challenging.

The diversity of actors involved in data collection, particularly in emergency situations, is a further important challenge. Standardisation of data collection would be highly desirable, along with broad and standardised approaches for informed consent (where appropriate – this issue would have to be handled sensitively to ensure ethical practices and community support). Engagement with those involved in humanitarian responses could encourage a greater emphasis on coordinated approaches to data collection and sharing.

Formal arrangements such as material transfer agreements (MTAs) can facilitate research. More work could be done on the use of MTAs and to communicate their purpose and benefits. Building on the work carried out by WHO[10], more template MTAs could be developed although in reality the gap is likely to be the technical expertise needed to adapt these tools for the context.

Conclusions

The six case studies have provided rich insight into data-sharing practices across a range of public health emergency situations, for a range of infections, and in differing geographies. They have captured the perspectives of a diverse range of stakeholders spanning multiple sectors.

While some obstacles and enablers reflect highly context-specific issues, most are common across the different outbreak scenarios. This raises hope that solutions can be identified that will have global impact on data-sharing practices.

Importantly, however, the case studies suggest that action will be required across multiple domains – there is no silver bullet that will change practice overnight. Progress will depend on addressing cultural and behavioural factors affecting how researchers view ownership of data and the value of the journal in which they publish their findings. Funders will need to be unequivocal in their support for open access publishing and timely data sharing, and ensure that adherence to these principles does not disadvantage researchers, particularly in LMICs. Academic publishing models will need to continue to adapt to ensure they are consistent with public health goals. New platforms may be required to provide appropriate tools for data sharing.

These shifts will need to be based on a more systems-oriented view of data sharing that accommodates the needs of multiple data generators and users across public health and academic sectors with differing interests and goals. They will need to be incorporated into multiple partnerships – global, cross-sectoral, and with communities. Building capacity for research on infections in LMICs should go hand in hand with capacity

building to detect, prevent and respond to outbreaks. This will require significant investment in data management and infrastructure in the countries that bear the highest infectious disease burdens and are at greatest risk for outbreaks and public health emergencies.

Importantly, it will be key to establish at the outset the purpose of data sharing, to provide clarity on who needs data, what kind of data, for what purpose, when and in what format. This will define data needs and establish parameters for data sharing and technical specifications for shared platforms. Wherever possible, the potential for data sharing needs to become an integral part of data collection. Moreover, evaluation of the use of such platforms to inform public health responses should be used to close the cycle and feed into the development of improved tools.

Ideally, data-sharing mechanisms would apply to all infections of epidemic potential, and not restricted to public health emergencies of international concern or large-scale public health emergency situations. Indeed, one goal of timely data sharing is to prevent public health emergencies arising in the first place. Nevertheless, establishing effective mechanisms could be an important stepping stone towards more general approaches, potentially targeting the eight priority diseases (including 'disease x', an epidemic of unknown cause) identified in the R&D Blueprint. Ultimately, such mechanisms will be vital for ensuring that the impact of infections of epidemic potential is minimised, helping to keep the world safe from epidemics and other health emergencies.

Potential next steps for funders and other global stakeholders

- **Develop a clearer picture of data use, key actors, data needs:** There is a need to map out the ecosystem of actors involved in collection, analysis and use of various kinds of data, to clarify data needs and to underpin the design and development of mechanisms and processes of data sharing.
- **Fund capacity building and national/regional/international integration/collaboration across partners:** Several international partnerships have been funded to support epidemic preparedness and to facilitate clinical and epidemiological research in outbreak situations. Building on these foundations, there is a need to establish further equitable partnerships addressing capacity development in key areas such as data management and governance, data analytics and modelling, clinical research and clinical data management, and laboratory skills and infrastructure, and to strengthen connections between research, public health and response communities.
- **Support the development of sustainable platforms/IT systems and promote their use**: Data platforms need to be established in advance of emergency situations for priority infections, with agreed governance mechanisms and data security systems built in; such databases could be centralised or federated.
- **Refine templates and standardised tools**: Various tools and templates to facilitate data sharing have been developed by organisations such as Chatham House (e.g. model agreements for

data sharing[13]) and WHO (e.g. for MTAs[14]) . There is a need to identify any gaps in this toolbox and to evaluate current use of tools and templates to support their further development and utility.

- **Communicate the public health benefits of data sharing and encourage greater feedback on use of data**: Case studies and modelling studies could be developed to highlight the benefits of data sharing to political leaders, public health officials and other key stakeholders.
- **Promulgate 'new' publishing model**: Funders such as Wellcome have endorsed open access publishing[8 9] and signed up to the DORA principles, which seek to challenge the significance of journal impact factors. Funders need to be clear on their expectations of data sharing, ensure they enforce their data-sharing policies, and recognize the achievements of those who adhere to such policies.
- **WHO leadership at national, regional and global levels**: WHO Health Emergency teams and other staff clearly play key roles in epidemic preparedness and response. WHO should take leadership in promoting openness and data sharing during outbreaks among responders, researchers and national authorities.
- **Engage with additional stakeholders including industry and the humanitarian sector**: Further discussions are required to establish appropriate data-sharing principles and practices for commercially sponsored trials on interventions for infections of epidemic potential and to encourage industry to share data that are relevant from a public health perspective. Dialogue with humanitarian organisations will be required to develop common standards and approaches for data collection in emergency situations that do not compromise emergency operations but facilitate research that could benefit affected and other vulnerable communities.

[13] https://datasharing.chathamhouse.org/resources/model-agreement/

[14] http://apps.who.int/blueprint/mta-tool/

Annex 1: Case study summaries

Data sharing during the West Africa Ebola Public Health Emergency

Sam Halabi[1], Rebecca Katz[1], Alexandra Phelan[1] and Matthew Boyce[1]

[1]Georgetown University, USA, [2]Griffith University, Australia and Australian Army

Background

The 2013–15 Ebola outbreak in West Africa resulted in a reported 28,616 suspected, probable and confirmed cases of Ebola virus disease (EVD), with an attributed 11,310 deaths. Most cases occurred in Guinea, Liberia and Sierra Leone.

EVD has a case fatality rate of around 50%, yet the precise nature of virus pathogenicity and the route of zoonotic transmission is still unclear. Experimental vaccines and treatments have been approved for emergency use to prevent and control recent outbreaks of Ebola in the Democratic Republic of the Congo. To date, only China and Russia have licensed treatments and vaccines for EVD, based on limited clinical data; however, several candidates are currently being considered by the European Medicines Agency and the US Food and Drug Administration.

Ebola remains a biomedical research priority and, given the major knowledge gaps associated with Ebola, global collaboration and data sharing continue to be vital to understanding and controlling this recurrent infectious disease. Conversely, limited data sharing and communication breakdown contributed to significant delays in acknowledgement of the West African outbreak's severity and global responses.

Case study methodology

After a review of the relevant literature, 26 semi-structured phone interviews were conducted (between 11 and 67 minutes duration), while two interviewees provided responses by email. Interviewees were identified from media coverage of the Ebola outbreak, relevant scientific literature, and through the interview process. Stakeholders interviewed included public health officials, government representatives, virologists, WHO staff, public health and hospital-based researchers, and representatives from Médecins Sans Frontières.

Barriers

- **Weak health infrastructure**, resulting in multiple actors conducting health care, surveillance and response activities, including non-governmental authorities and international agencies.
- **Fractured sources of relevant data**, including a lack of standardisation and uniformity in epidemiological, operational, clinical and genetic data.
- **Lack of coordination and clear roles for data-generating sources**, including express limitations on the type of data NGOs were permitted to share and with whom.
- **Lack of infrastructure for requesting and approving assistance with research for diagnostics, therapeutics and vaccines**, including the lack of data-sharing or governance agreements.
- **Incentives for private sector researchers that penalised or prohibited sharing**, including the tension between rapid sharing and accuracy, as well as commercial interests in securing intellectual property rights to the data collected.

- **Incentives for public sector researchers that penalised or prohibited sharing**, including the tension between rapid sharing and accuracy, publication as a cornerstone of the academic reward system, and concerns that data would be used without correct attribution.
- **Ethical and legal constraints related to patient confidentiality and informed consent**, including concern over how to obtain informed consent and maintain confidentiality while sharing for the purposes of response, limited capacity for protecting sensitive health information, ambiguous procedures and tracking of informed consent, and discrimination and stigma experienced by infected persons, survivors and their families.
- **Community-level barriers to data sharing**, including the lack of community trust as a result of misinformation spread about Ebola and mishandling of information, and assumptions made regarding homogeneity and hierarchies for sources of data collection.
- **Political pressures**, including domestic political pressure to withhold data indicating a public health emergency.

Enablers

- **Centralised and standardised data sharing**, such as Sierra Leone's establishment of a single, national toll-free number to facilitate prompt investigation of Ebola cases.
- **Community engagement mechanisms**, such as multi-platform public information campaigns.
- **Fast-tracking of Ebola-related publications** by some scientific journals.
- **Establishment of data-sharing platforms and informal data-sharing networks**, although these were limited in their capacity.
- **Reviewing and updating existing reporting systems** to match the data required for a response, which prevented data-sharing bottlenecks.

Lessons learnt and potential actions

- **Ensure compliance with International Health Regulations** for a national-led response, with accompanying WHO support.
- **Develop open data platforms and databases** with standardised procedures, training and funding.
- **Develop differentiated data-sharing channels for preliminary and confirmed data.**
- **Establish minimum standards for international data-sharing protocols.**
- **Restructure the current academic reward system** to incentivise data sharing.
- **Build national public health capacity**, particularly for surveillance, reducing reliance on NGOs.
- **Establish data-sharing arrangements before public health emergencies**, including developing template data transfer agreements and enhancing data management capacity and analytic expertise in under-resourced settings.

Data sharing during the MERS-CoV outbreak

Michelle Rourke[2], Rebecca Katz[1], Alexandra Phelan[2] and Matthew Boyce[2]
[1]Griffith University, Australia and Australian Army, [2] Georgetown University, USA

Middle East Respiratory Syndrome (MERS) is caused by a coronavirus (MERS-CoV) first isolated in 2012 in Saudi Arabia. MERS is a severe acute respiratory syndrome that has caused more than 2200 laboratory-confirmed cases and over 800 deaths. The largest outbreak of MERS outside of the Eastern Mediterranean occurred in 2015 in South Korea, with 186 confirmed cases and 38 deaths. MERS has a case fatality rate of around 35%. As a severe emerging disease with the potential to cause a major global health emergency, MERS-CoV remains a global R&D priority.

The evolutionary origin and route of transmission of MERS-CoV remain unknown. There are no licensed vaccines or specific treatments for MERS. Given the major knowledge gaps associated with MERS-CoV, global collaboration and data sharing continue to be vital to understanding and controlling this novel infectious disease.

Case study methodology

After a comprehensive review of the literature, 14 semi-structured interviews were conducted (between 39 and 76 minutes duration) via teleconference. Interviewees were identified from media coverage of MERS-CoV outbreaks, relevant scientific literature, and through the interview process.

Stakeholders interviewed included clinicians, public health officials, virologists, epidemiologists, data scientists, social scientists, WHO staff, government employees, journalists and science communicators. Interviewees were from Australia, Canada, China, the Netherlands, South Korea, Switzerland, the UK and the USA; stakeholders from Saudi Arabia did not consent to be interviewed (this is relevant as there are differing accounts relating to the initial outbreak of MERS-CoV in Saudi Arabia; this case study does not attempt to resolve these accounts).

Barriers

- **Lack of trust in official sources of information**, arising from different expectations and perceptions of countries' official data-sharing approaches, including actual or perceived delays in sharing data, incomplete data, deliberate withholding of data, or destruction of data.
- **Restrictions on news media**, including legal restrictions preventing free speech, inhibiting a potential source of data and public health information.
- **Community-level barriers to data sharing**, including different attitudes to illness and medical care, resulting in the failure to collect pertinent information in a clinical setting or failure to identify data relevant for collection.
- **Legal constraints relating to intellectual property and viral sovereignty**, including uncertainty over the scope of a patent applied for in relation to MERS-CoV and the impact on data sharing and research, and the possible impacts of international laws recognising viral sovereignty – including the Nagoya Protocol – on access to pathogens and any subsequent data sharing.
- **Lack of coordination between agencies,** including communication between the different levels of

government within countries, and between UN agencies involved in animal and human health.

- **Incentives for researchers that penalised or prohibited sharing,** including publication as a cornerstone of the academic reward system and concerns that data would be used without correct attribution.
- **Fractured quality of relevant data** without minimum standard inclusions for consistency and opportunities for informal inputs relating to the context within which the data were collected.

Enablers

- **Compliance with WHO's International Health Regulations (2005)**, including the legal obligation on countries to report information to WHO.
- **The norm and culture of open data sharing** within the scientific community.
- **Informal networks of colleagues and collaborators** within the scientific community.
- **ProMED-mail**, which enabled rapid global data sharing, even when traditional channels were unavailable.
- **Use of alternative sources for obtaining and sharing data**, such as Twitter.
- **Well-resourced bureaucratic structures** before outbreaks ensured efficient data sharing within countries during outbreaks.
- **Informal use of WHO technical teams as hubs** enabled data sharing across research teams.
- **Fast-tracking of MERS-CoV-related publications** by some scientific journals addressed delays by incentivising rapid data sharing.

Lessons learned and potential actions

- **Encouraging the use of non-traditional media in data sharing.**
- **Crowdsourcing epidemiology** and reinforcing the idea that data are a global public health good.
- **Establishing special considerations for data sharing on emerging pathogens** to address the tension between speed and accuracy of data sharing and to incentivise data sharing.
- **Ensuring compliance with International Health Regulations** for a national-led response with accompanying WHO support.
- **Establishing data-sharing arrangements before public health emergencies**, including developing template data transfer agreements and supporting development of international law that ensures unimpeded data sharing during public health emergencies.
- **Encouraging press freedom globally.**
- **Clarifying the scope of pathogen and data ownership rights.**

Data sharing in public health emergencies: Analysis of barriers and enablers from an outbreak response perspective: the case of Middle East Respiratory Syndrome (MERS)

Marion Koopmans[1], Martine van Roode[1], George Haringhuizen[2], Carolina Ribeiro[2], Mohamed Nour[3], Elmoubasher Farag[3], Minahil Ahmed[3], Aya Moustafa[3], Eric Claassen[4] and Linda van de Burgwal[4]

[1]Erasmus MC University, The Netherlands; [2]National Institute of Public Health and the Environment, The Netherlands; [3]Ministry of Public Health, Qatar; [4]Vrije Universitat Amsterdam, The Netherlands

Background

The MERS epidemic, caused by an emerging coronavirus (MERS-CoV) began in 2012. To date, 2266 confirmed cases from 27 countries and 804 MERS-CoV associated deaths have occurred. Although the disease has a zoonotic origin, large outbreaks associated with health care facilities occurred in Saudi Arabia, the United Arab Emirates and the Republic of Korea. Camels were identified as an asymptomatic reservoir for MERS-CoV and as potential source of human infections, although how this transmission occurs is not fully understood.

MERS-CoV constitutes a constant public health threat, especially in the Arabian Peninsula, where new introductions of the virus into the human population continuously take place. For example, in the past six months 45 cases of MERS have been reported, including 14 deaths.

The study focuses on the sharing of epidemiological, clinical, laboratory and surveillance data, particularly relating to primary transmissions at the animal–human interface in Qatar, while also reflecting on the wider region of the Arabian Peninsula. Data sharing relating to hospital-acquired infections was not explored. This case study explored the flow of data at different levels, including: (1) national data sharing among stakeholders in Qatar to monitor population health, target responses, and allocate resources; (2) regional data sharing among countries or a group of countries with a collaborating centre; (3) international data sharing among countries and organisations outside the region, including research centres; and (4) global data sharing among international agencies to estimate the global burden of disease and to contain emerging global health threats.

Case study methodology

A comprehensive desk study was performed using 135 sources of white and grey literature, meeting reports and media articles. This was used to reconstruct a detailed timeline of the MERS-CoV epidemic and the outbreak response, including the flow of data as described in literature. Key stakeholders relevant to the response were identified from this map and 42 stakeholders contributed to interviews (face-to-face or by phone) and a stakeholder workshop. The semi-structured interviews lasted on average 50 minutes. Stakeholders included governmental representatives, public health experts, epidemiologists, clinicians, veterinarians, virologists, and social scientists from public health or animal health institutes, academic research institutes and supranational organisations. A root-cause analysis was performed to identify causal relations between barriers.

	National	**Global**	**International**
Barriers	• Insufficient capacity for outbreak investigation and response. • Tensions in inclusive	• Delays in (official) notification and lack of data quality and completeness hampered data flow between global and national	• Strict agreements (e.g. MTAs, publication debates) and difficulties with shipment of materials.

	collaborations in a multi-stakeholder and One Health approach.	stakeholders	• Lack of data quality and completeness.
Enablers	• Establishment of a joined One Health outbreak investigation team, which helped the collection and sharing of field data. • Transparency and openness, with support from national authorities, ensuring trust and collaboration between stakeholders.	• Support from supranational organisations (advice on international partners for reference and collaboration, capacity building in terms of guidelines, protocols and training). • Improved transparency and openness in data sharing, especially by countries' authorities (e.g. by Qatar and the change in leadership at the ministry of health in Saudi Arabia). • Particular communication channels (including the institutional and government websites, and the informal and confidential sharing of information through IHR channels).	• Reciprocity of data sharing in terms of support from international institutes, for example capacity building. • Pre-existing collaborations and formal data-sharing agreements, such as memorandums of understanding.

Notable is the interconnection of these barriers at the different levels, with data flow delays at one level affecting the flow of data at or to another level.

Lessons learnt and potential actions

- **Create a globally agreed platform and mechanism for data sharing**, addressing fairness and reciprocity of data sharing, with transparency a key guiding principle for data sharing between stakeholders, and embedded within a One Health approach
- **Improve collaboration on outbreak preparedness and response between sectors**, by training and capacity building and creation of a permanent One Health working group.

Real-time barriers and solutions to data sharing during outbreaks: A case study into data-sharing practices and principles during the research response to the Zika outbreak[15]

Kajsa-Stina Longuere[1], Raul Pardinaz-Solis[1], Behnaz Schofield[2] and Gail Carson[1]

[1]University of Oxford, UK; [2]University of the West of England, UK

Background

The first locally acquired Zika virus (ZIKV) cases in Brazil were confirmed in 2015. On the evidence collected on ZIKV and its complications (i.e. Guillain-Barré syndrome and microcephaly in newborns), on 1 February 2016 WHO declared ZIKV a Public Health Emergency of international Concern (PHEIC) under the International Health Regulations (IHR 2005).

This descriptive case study examines the political, ethical, administrative, regulatory, logistical, economic and social (PEARLES) barriers and solutions to data sharing as part of the research response during the ZIKV outbreak.

Case study methodology

The research team put together a list of interviewees representing different stakeholder categories. This was based on the International Severe Acute Respiratory and Emerging Infection Consortium (ISARIC) membership and collaborators list in the Zika research response in 2015 and 2016, complemented by a list of funders' representatives from GloPID-R. The initial list included clinical researchers (virologists, epidemiologists, geneticists, and epidemiologists), social science scholars, clinicians, funders, politicians, modellers, NGOs, and public health representatives.

A standard set of semi-structured interview questions was developed, in English, Spanish and Portuguese. Telephone interviews of between 45 minutes and 1.5 hours were conducted. Interviews were manually transcribed and analysed using theme-oriented discourse analysis.

Barriers

- **Political situation, regulations, legal framework:** These vary across Latin American countries and contribute to the lack of unity in designing collaborative research studies.
- **Approval processes:** These are lengthy in all participating countries and contribute to the delays in starting research studies. They are also responsible for difficulties in sharing data and samples between countries.
- **Lack of resources:** This delays the timely set up of research during an epidemic outbreak. It is also responsible for local staff not being responsible for analysis of samples in some countries.
- **Ethically sensitive topic:** Research on pregnant women is classed as high risk and ethically sensitive.
- **Sharing samples:** The movement of samples is very difficult, with tedious and arduous processes.
- **Commercialisation:** Early findings may be commercially valuable, limiting collaboration with industry.

Solutions

- **International guidelines:** Guidelines on data and sample sharing and set up of biobanks would facilitate harmonisation of research processes in collaborative studies.

[15] The zika case study was self-funded by ISARIC

- **Common data capture and analysis platforms:** The creation of common protocols, consent forms, data capture forms, other templates, uniform database, data analysis platforms could facilitate data analysis across countries and studies.
- **More immediate resources:** These would allow research in public health emergencies to take place in a timely manner.
- **Creation of collaborations pre-outbreak:** The creation of such networks would enable researchers to begin working in a timely manner at the outset of an outbreak. Global interdisciplinary teams would be the most effective collaboration.
- **Streamlining of processes:** Streamlining needs to be done at local and national levels to allow collaborative work to be undertaken without delays.
- **A registry for studies conducted during an outbreak:** The creation of a registry would allow the coordination of research activities during an outbreak.
- **Broad informed consent:** Broad consent at the outset of a study, approved by research ethics committees, would facilitate future research.

Lessons learnt and potential actions

- **Ethics committees:** Ethics committee members could be provided with additional training on emergency infectious epidemic research. It would be advantageous for ethics committees to have extraordinary meetings to speed up the approval process to provide high quality but timely reviews.
- **Working in collaboration on preparedness:** Being prepared before an outbreak and creating networks and collaborations will allow teams to share good practice and lessons learnt and create trusting working environments before research during emergencies is undertaken.
- **The role of commercial companies:** Academic researchers have concerns about the role of industry during an outbreak. It may be helpful for researchers to be made aware that late-stage work needs to be done by commercial companies.
- **The role of funders:** Funders can shape when and how data and samples are shared – within the restrictions of country-specific regulations – by stipulating sharing requirements from the outset of a study.
- **Surveillance cohorts:** Surveillance cohorts provide early warning systems for outbreaks but need ongoing funding. In the absence of existing cohorts, a prospective study should be launched immediately during the outbreak.

Data sharing during cholera outbreaks

Leonard Heyerdahl[1], Berthe Marie (Betty) Njanpop-Lafourcade[2], Delphine Sauvageot[3], Isabelle Delrieu[4], Alioune Thioune[5] and Elise Guillermet[4]

[1]Independent consultant, Abidjan, Côte d'Ivoire; [2]independent consultant, Paris, France and Abidjan, Côte d'Ivoire; [3]independent consultant, Paris, France; [4]Epilinks, Geneva, Switzerland; [5]independent consultant, Conakry, Guinea.

Background

Research data sharing among stakeholders at the global, regional and field level is essential for infectious disease outbreak preparedness, particularly for pathogens causing explosive epidemics, such as *Vibrio cholerae.* Immediate priorities are to identify isolates with epidemic potential to confirm suspected cases, to make outbreaks politically visible, and to engage outbreak responses. Over the longer term, biological materials and associated data contribute to a deeper understanding of *V. cholerae* biology and cholera transmission. Identifying antibiotic susceptibilities contributes to treatment strategies beyond rehydration. In addition, the evaluation of intervention strategies (WASH, antibiotic use and vaccine campaigns) based on clinical, epidemiological and behavioural data open up new possibilities for control strategies.

Many actors contribute to data production, collection and analysis at subnational, national, regional and international levels. Before and during the last cholera outbreaks in Guinea and Côte d'Ivoire, key stakeholders involved in research and public health responses include Ministries of Health, national medical research and public health institutes, other local academic centres, NGOs, WHO country offices, and global heath research centres and public health bodies.

Case study methodology

This multi-level case study used in-depth interviews and a concise literature review to document and analyse practices and perceptions among stakeholders who were involved in research during preparedness and response to cholera outbreaks in Guinea and Côte d'Ivoire. Study participants were identified using a purposive sampling approach (mapping by literature review) and planning of interviews by convenience (availability and willingness to participate in the study). A total of 14 in-depth interviews were conducted (face-to-face or via telephone/Skype). Anthropologists analysed interview transcripts using computer-assisted qualitative data analysis software.

A concise review of 18 published articles was used to:

- Map the key research teams and stakeholders to be included in the study.
- Prepare in-depth interviews by reviewing the study questions, type of data and findings of each participant and identify their research networks.
- Develop key indicators for the protocol and the framework used in data analysis.

Timely data sharing in an emergency context

Although countries are still encouraged to notify suspected and confirmed cases, the 2005 International Health Regulations abandoned the previous (IHR 1969) mandatory notification of all cholera cases. At a country level, a cholera outbreak is supposedly declared after one laboratory-positive result (by bacterial culture). In practice,

several laboratory-positive results may be needed before action is taken, this threshold varying by country.
In some cases, countries delay or do not share their data internationally. In contrast, international stakeholders report that some countries are transparent regarding suspicion and confirmation of an outbreak, and are quick to share data on cholera for public health response and research (as occurred in Guinea during the 2012 outbreak).

Barriers for data sharing

- Competition between research teams, and even between researchers within the same institutions.
- Poor organisation of cholera surveillance and outbreak responses within countries.
- Lack of coordination between clinicians, laboratory and epidemiologists; absence of sustainable funds leads to unavailability or poor quality of data.
- Political reluctance to acknowledge outbreaks to the international community.

Enablers of data sharing

- Interpersonal relationships based on confidence in the recognition of the scientific value of all experts (with authorship attribution); this confidence can be built during collaboration in the field.
- Partnership building to promote technology transfer and to provide data management tools.
- The presence of a country public health champion who leads the partnership and research and response activities.

Lessons learnt and potential actions

- The country-driven approach has become the main framework for research and response in a context of emergency. Country authorities and national institutes for disease surveillance and control produce and manage data on their own, with the involvement of their country experts.
- Agreements for research partnerships, materials and data transfer are increasingly transparent and specifying roles and responsibilities. Fairness in authorship and recognising the contribution of all stakeholders also promotes better transparency. Formal agreements benefit from prior inter-personal trust building, rooted in face-to-face interactions and field presence. Agreements must be signed between outbreaks to allow appropriate data sharing.
- Financial support of the national surveillance system is a key condition for data availability and data sharing.
- In the case of political retention of information, an informal system of alert or data sharing exists using direct contacts or NGO local presence. A key challenge is to formalise reporting, with legal protection of national and regional public health experts and with appropriate data governance. This would provide at-risk neighbouring countries with access to appropriate information to implement preventive strategies.

Data sharing in public health emergencies: Anthropological and historical perspectives on data sharing during the 2014-2016 Ebola epidemic and the 2016 yellow fever epidemic

Sharon Abramowitz[1], Tamara Giles-Vernick[2], Jim Webb[3], Jennifer Tappan[4], Elanah Uretsky[5], Jorge Varanda-Ferreira[6], Katherine Mason[7] and Molly Beyer[8]

[1]Independent anthropologist, Boston, USA; [2]Institut Pasteur, France; [3]Colby College, USA; [4]Portland State University, USA; [5]Brandeis University, USA; [6]University of Coimbra, Portugal; [7]Brown University, USA; [8]University of North Texas, USA.

Aims

To use historical and anthropological methods to analyse data sharing during the 2014–16 West Africa Ebola outbreak and the 2016 yellow fever outbreak in Angola, the Democratic Republic of the Congo and China.

Methods

We used short-, mid- and long-term historical timeframes to consider if the Ebola outbreaks represent continuity with established practices or constituted a break with past experiences. We conducted a literature review of over 800 books, articles, working papers, websites, public statements, policy documents, software, protocols, tools, and technical guidances, to identify key actors, themes, processes, capacities, networks, incentives and policy issues around data sharing in global health emergencies. We also conducted 27 Ebola-related interviews (21 remote, six in China) and reviewed 64 existing transcripts addressing data sharing from the Ebola 100 Project.

We also conducted 21 yellow fever-related interviews with participants recruited through purposive and snowball sampling strategies and reviewed 13 publicly available interviews from the US Centers for Disease Control and Prevention (CDC). We used historical and anthropological qualitative data analysis methods to evaluate the degree of data sharing, the quality of data sharing, and the experiences of data sharing in order to identify the barriers to and facilitators of data sharing.

Summary

The Ebola and yellow fever epidemics illustrated two radically different models for data sharing during epidemic response – a routine data-sharing environment and a paradigm shift in sharing practices during a Public Health Emergency of International Concern (PHEIC).

Business as usual vs PHEIC

The yellow fever epidemic of 2016 constituted a case of 'business as usual' epidemic response, involving the stakeholders typically engaged in disease outbreaks in LMICs (Table 1). During the yellow fever outbreak, data access was strictly regulated by government authorities. WHO actors used informal means to gain access to data, but they had little authority to circulate that data beyond an inner circle of data stakeholders (e.g. WHO Collaborating Centres, response partners), and even within that circle, implementing partners had inconsistent access to data. Countries with close bilateral relationships with the national governments – such as Cuba and Angola - were able to gain access to data by leveraging long histories of public health response partnerships.

During the West Africa Ebola epidemic, data sharing prior to the PHEIC declaration on 8 August 2014 was typical of routine epidemic responses. Data sharing was restricted to stakeholders integral to operational responses, and

access to data was deeply restricted and shaped by pre-existing governmental and institutional arrangements. However, with the PHEIC declaration, expectations and demand for data led to a radical shift in data-sharing practices. The public failure of conventional approaches to epidemic containment, visible on the front pages of newspapers around the world, challenged conventional standards, norms and agreements. The legitimacy of data stakeholdership was challenged by a range of insurgent data claimants and data sharers, both globally and on the ground in West Africa. With these drivers, insurgent actors within and outside conventional stakeholder relationships moved to share data informally with 'outsiders', creating open-source mechanisms to capture and share data, and putting pressure on core response actors to increase data access and transparency.

Table 1: Key stakeholders

Stakeholders	**Yellow fever**	**Ebola**
National governments in affected countries	X	X
WHO	X	X
WHO Collaborating Centers	X	X
WHO/epidemic response NGO implementers (IFRC, MSF)	X	X
National research partners	X	X
International research community (not WHO CCs)	X/O	X
International global health community		X
Multilateral agencies and funders (e.g. World Bank, IMF)		X
Foreign governments		X
Private donors/philanthropists		X
Commercial/for-profit entities		X
Humanitarian response actors		X

In all this, the role of the nation-state as a focal point in arbitrating data access was subject to continuous testing and was managed quite differently by different national governments.

Data: What was important?

During epidemics, specific types of data are routinely prioritised for collection and circulation to inform response activities. During the yellow epidemic, these included epidemiological, clinical data, laboratory/diagnostic, genetic and vaccination data. There were capacities created to conduct research during this epidemic due to the need to use fractional dosages of the yellow fever vaccine, due to a global vaccine shortage.

During the Ebola epidemic, the first eight months of the outbreak were marked by a dependence on these routine sources of data collection. However, as the epidemic grew in scale, visibility and risk to high-income countries, it exceeded the capacity of responders to contain the disease. When the PHEIC was declared, conventions regarding needed data were tested. While routine actors such as WHO Collaborating Centres were subject to strict limitations on the kinds of data they could access and circulate, external actors – from 'star' scientists to philanthropists – were able to access sensitive data, on informal or formal bases.

The chaotic conditions were unfavourable to structured data sharing (and even encouraged data hoarding), but were remarkably facilitative of informal and open source data sharing. There was also an enormous surge of data sharing innovation from unconventional data stakeholders demanding access to and input into the Ebola response. Anthropological, geospatial, clinical, and political and economic data were in high demand – and were generally unavailable due to a lack of platforms for collection and sharing. People working within the region felt morally compelled to share data outside of institutional norms and conventions in order to raise international

awareness of the impact of the epidemic. And researchers outside of West Africa worked intensively on the limited data emerging from the region to make it accessible and useful to unrecognised, informal 'response actors', from the US White House Office of Science and Technology Policy to college classrooms.

Was there a therapeutic or vaccine?

In the absence of a rapid diagnostic, therapeutic or vaccine, governments around the world mobilised vast research networks and resources to target responses to the Ebola outbreak. Scientific journals tried to facilitate the flow of data and analyses by opening up rules around prepublication dissemination of data, removing paywalls, and increasing expectations regarding shared attribution.

Barriers

Different types of data were associated with different barriers to sharing. Often these barriers existed on a country-by-country basis and changed over time.

Yellow fever

Factors	Enablers	Barriers

Political-social-historical	• International Health Regulations
	• 'Trust', a marker that represented longer-term relations, shared experiences, confidence that data would not be misused.
	  • Circularity: trust facilitates sharing; sharing to cultivate trust to secure material or other support
	  • Shared experiences of prior outbreak responses
	  • Shared histories (e.g. Angola–Cuba) in communities of responders and researchers facilitate credibility
		• Angolan restrictions on data circulation contributed to strained relations with WHO, other partners
		• Post-colonial relations: sense of neo-colonial relations and hierarchies
		• Social tensions (political undercurrents) between teams
		• State image vis-à-vis press, population, neighbouring countries, multinational industries

Logistics	• Emergency operations centers (IMS). • Key actors from UNICEF, the US CDC, MSF, the Cuba Cooperation, and the WHO • Formal and informal exchanges of data and analyses	• Insufficient technical capacity and funding • Insufficient laboratory capacity to diagnose yellow fever, collect data • Gaps between response needs and research needs • Procedural problems: Unclear procedures, strict procedures, repeat requests, delays
Ethics	• Moral obligation facilitated data sharing • Data sharing is seen by some as a moral good. We share to respond to an epidemic (although what	• Political sovereignty: also based on ethics, even in an emergency. • Epidemic management and data produced: linked to moral claims of sovereign control over data and labour of employees

	"response" was, and how it relates to research, remains undefined).	• Moral obligations to share data broke down in the post-epidemic phase: "response" is over, and research doesn't matter

Ebola

Factors	Enablers	Barriers
Political-social-historical	• Climate of 'emergency' encouraged sharing, including breaking norms and rules • Chaotic response environment • High visibility in the media	• Elites discouraged inclusion • Perceptions that the situation was 'under control' • Lack of power (e.g. local community leaders)
Economic	• Declining costs of research • Countries used data as leverage for assistance • Donor mandates to share data	• High costs • Institutional pressure to gain consolidated access to data
Administrative	• IMS/emergency operations centres • Data-sharing agreements between WHO and collaborating partners expedited funding and data sharing for rapid analysis • Strong leadership in some governments • Funding conditions • Willingness to engage non-traditional sharing pathways	• Post-epidemic failures to maintain data • Lack of regional data coordination • Non-standardisation of data • Restricted entry to data portals
Regulatory	• Private sector risks: legal and regulatory	• Climate of 'emergency' encouraged breaking norms and rules
Logistics	• Mass mobilisation of resources	• Insufficient technical, funding, logistical, bureaucratic, laboratory capacity
Ethics	• Moral obligation facilitated data sharing	• Privacy of individual health data • Non-disclosure agreements in clinical trials • The lack of understanding around discipline-specific ethics and norms around data sharing
Social	• A culture of curiosity and learning • Small, specialised research networks with consensus around key actors, norms, values • Strong informal relationships with governments prior to PHEIC • Recognition of expertise	• Lack of trust

Lessons learned

A comparison of responses during the yellow fever and Ebola outbreaks highlights the shared characteristics of data-sharing activities during 'routine' epidemics. Only after the PHEIC was declared in West Africa were conventional approaches overthrown, leading to a profound change in the entire approach to public health emergency response, research and learning. Unless specific measures are taken to change the overall approach to epidemic response, this business as usual model – with limited data sharing – will remain the dominant model during epidemic responses.

Annex 2: Data sharing in public health emergencies (PHEs): Barriers, enablers and potential actions by funders and other international agencies

Theme	Barrier	Enabler	Action
INTERNATIONAL RELATIONSHIPS			
Political history	North–South tensions and power imbalances; perceptions of 'data exploitation'	Well-established academic and other links between particular LMICs and HICs and institutions within them Equitable two-way data sharing and technology/knowledge transfer, capacity building	*Build equitable international partnerships (North–South and South–South) in advance of PHEs, where the purpose of data sharing is clear and those involved understand each other's expectations.* *Enhance international funders' coordination with national priorities and international donor agencies, so research is more in line with country needs*
Global agencies	WHO may be reluctant to challenge national practices	Global agencies (e.g. WHO) have crucial global leadership role WHO can promote effective and appropriate national data sharing, especially in its role to support countries with outbreaks and coordinating research response; WHO can also advise data producers on what data is needed and useful to public health	*Enhance visibility of the leadership role that WHO has at global, regional and national levels, both in data sharing in PHEs and in promoting data sharing by others, such as responders, national authorities and researchers*
Memorandums of understanding/material transfer agreements		Formal agreements based on common principles developed in advance can expedite data sharing; they should be sufficiently informed by the rights and interests of all stakeholders	*Promote use of memorandums of understanding and material transfer agreements in international partnerships, in line with GloPID-R data-sharing principles*
Commercial studies	Non-disclosure agreements, commercial interests may inhibit timely sharing of data Reluctance to share negative data because of possible commercial implications		*Discuss with industry associations mechanisms to ensure timely access to data (including negative findings)* *Agree data-sharing procedures in advance for product development partnership projects and other public–private partnerships*
NATIONAL INTERESTS AND CAPACITIES			

National reputations	Reluctance of political leaders, especially non-health actors, to release information related to outbreaks in order to protect national reputation and/or economic interests		*Communicate national and global importance of openness to opinion leaders and decision makers, e.g. through advocacy case studies* *Expose senior national decision-makers to the realities of outbreak response (e.g. simulation exercises)*
Transparency		Openness increases awareness of data availability and encourages culture of sharing	*Build promotion of transparency into partnership activities*
Advocacy		Showcasing the public health and other benefits of timely data sharing (e.g. through case studies, economic analyses, modelling) Public health champions can drive data sharing and use of data in public health responses	*Develop case studies and mobilise other evidence to support advocacy activities* *Identify potential national public health champions and/or civil society organisations and incorporate into international partnerships*
Lack of local technical capacity	Data slow to be generated or of poor quality due to lack of local technical capacity (laboratory, surveillance, outbreak investigation etc.)		*Build local capacity for surveillance, outbreak detection and response, for example through international networking and research programmes*
Formal bureaucratic structures		Provide formal mechanisms for data sharing among key stakeholders, which take into account the rights and interests of those involved	*Consider as part of local capacity building*
Regulatory frameworks	Variation in national regulatory and legal frameworks inhibits international collaboration; lengthy approval processes slow research		*Work with regional and international bodies to encourage harmonisation of practices, and develop guidance on how to navigate different regulatory systems*
International Health Regulations (IHR)	Changes to notification rules may discourage countries from reporting cases	Formal regulations can help to promote global sharing	*Consider the responsibility of the research community with respect to IHR requirements*
Nagoya Protocol	Emphasis on national ownership of non-human genetic resources and complexity of different requirements across different countries may discourage sharing		*Provide guidance on ownership rights for pathogen samples and data and how to promote international collaboration in the context of the Convention on Biological Diversity*
Media	Restrictions on news media may reduce access to information of public health importance	General media may have key role in highlighting early cases and promoting accountability; social media provide alternative sources of data	*Consider how a communications strategy can promote greater awareness of and support for timely data sharing*

ATTITUDINAL/BEHAVIOURAL FACTORS			
Trust	Fears of inappropriate use of data inhibit sharing	Existing strong relationships foster collaborative practice and open sharing	*Build and nurture international networks and cross-sectoral relationships (e.g. academic–public health) in advance of PHEs*
Lack of awareness of public health needs	Researchers may not focus on issues of most relevance to public health response		*Clarify key public health data needs for different infections in different settings* *Promote interactions between academic researchers and public health professionals*
Informal networks		When 'official' channels are not functioning, informal channels provide alternative routes of information flow	*Consider how informal networks can contribute to information flows when formal channels are functioning inadequately, while recognising implications for transparency and equity of data sharing*
'Data as labour'	Failure to recognise time and energy put into data gathering can inhibit desire to share	Feedback on use of data signals value of data gathering	*Provide support and develop the skills of those contributing to data collection, and encourage greater feedback of data analyses to data providers* *Articulate the potential benefits and risks of sharing, so that data collectors can assess the value of collecting and sharing data*
Community attitudes	Lack of trust reduces willingness to provide data or to allow it to be shared		*Promote community engagement activities that build trust and promote data sharing for the public good, in advance of PHEs* *Encourage greater feedback to the communities that data originate from*
Cross-sectoral dialogue	Differing terminologies, ways of working, priorities inhibit collaboration	Cross-sectoral structures/groups (e.g. One Health bodies)	*Build and nurture cross-sectoral relationships and networks (e.g. One Health) in advance of PHEs*
ACADEMIC PUBLISHING PRACTICES			
Academic incentives I	Desire for publication credit to advance career encourages data hoarding	Clear agreements on accreditation can promote timely sharing	*Promote collaborative agreements that ensure recognition of LMIC research contributions*

Academic incentives II	LMIC researchers may be strongly influenced by journal impact factor; high-impact journals are often not open access	Funders' policies can promote sharing in open-access publications	*Communicate and enforce relevant funding policies, stressing importance of quality of research rather than where findings are published* *Enforce compliance with funding policies and recognise timely sharing of data in the public interest*
Academic incentives III	Using journal publication impact factor as a proxy for quality of research encourages publishing in venues which are often not open access	Broader conceptualisation of impact, by research funders and academic institutions when recognising the contributions of researchers, can encourage more proactive dissemination of data	*Recognising this kind of impact when awarding grants*
Academic capacity	Academics in LMICs may not have the resources or capacity to write academic papers		*Build local academic capacity, especially around data management, data modelling and analytics, and authorship*
Journal practices		Fast-track publication, open-access publishing can accelerate data access	*Encourage scientific publishers to reflect on their roles in facilitating timely and effective public health responses to infectious disease outbreaks*
TECHNICAL/PRACTICAL			
Platforms		Data platforms (or linked federated databases) with standard operating procedures and agreed governance frameworks	*Develop and promote the use of shared data platforms for prioritised pathogens to build consensus around governance, standards with contributors and likely end users before outbreaks.*
Social/behavioural data	Few platforms or mechanisms for efficient sharing of social/ behavioural data	Social science evidence and methods used to inform responses	*Develop sustainable platforms for sharing of social/behavioural data*
Health information systems		Embedded data collection systems enable data collection in routine clinical practice	*Work with technical partners to incorporate appropriate functionality into commonly used systems (e.g. DHIS2)*
Standards		Developing prescriptive protocols for sharing when planning research, potentially before outbreaks	*Build requirement into grant applications where appropriate*
Material transfer agreements (MTAs)	Confusion about purpose and use can delay/inhibit sharing of samples/data	Use of 'off-the-shelf' agreements can expedite sharing; seen as integral in research	*Undertake activities to promote deeper understanding of the nature and use of MTAs* *Develop and evaluate template MTAs for use by international collaborations*

Informed consent	Differing attitudes to informed consent lead to inconsistencies in sharing practices	Broad informed consent at point of capture	*Develop and promote use of standardised consent forms where appropriate*
Data collection	Diversity of actors involved in data collection	Centralised data collection mechanisms and training in data collection and management	*Explore potential for innovative data collection mechanisms and linkage to national/global data platforms*
NGO attitudes	Focus on emergency response rather than research; differing attitudes to data collection and sharing		*Engage with NGOs to share best practices and promote consistent approaches to data sharing*

Wellcome exists to improve health by helping great ideas to thrive.

We support researchers, we take on big health challenges, we campaign for better science, and we help everyone get involved with science and health research.

We are a politically and financially independent foundation.

Wellcome Trust, 215 Euston Road, London NW1 2BE, United Kingdom
T +44 (0)20 7611 8888, E contact@wellcome.ac.uk, wellcome.ac.uk

The Wellcome Trust is a charity registered in England and Wales, no. 210183.
Its sole trustee is The Wellcome Trust Limited, a company registered in England and Wales, no. 2711000
(whose registered office is at 215 Euston Road, London NW1 2BE, UK). MS-7170/03-2019/RK