

RESEARCH ARTICLE

Open Access

# A systematic review of barriers to data sharing in public health

Willem G van Panhuis<sup>1\*</sup>, Proma Paul<sup>1</sup>, Claudia Emerson<sup>2</sup>, John Grefenstette<sup>1</sup>, Richard Wilder<sup>3</sup>, Abraham J Herbst<sup>4,5</sup>, David Heymann<sup>6</sup> and Donald S Burke<sup>1</sup>

## Abstract

**Background:** In the current information age, the use of data has become essential for decision making in public health at the local, national, and global level. Despite a global commitment to the use and sharing of public health data, this can be challenging in reality. No systematic framework or global operational guidelines have been created for data sharing in public health. Barriers at different levels have limited data sharing but have only been anecdotally discussed or in the context of specific case studies. Incomplete systematic evidence on the scope and variety of these barriers has limited opportunities to maximize the value and use of public health data for science and policy.

**Methods:** We conducted a systematic literature review of potential barriers to public health data sharing. Documents that described barriers to sharing of routinely collected public health data were eligible for inclusion and reviewed independently by a team of experts. We grouped identified barriers in a taxonomy for a focused international dialogue on solutions.

**Results:** Twenty potential barriers were identified and classified in six categories: technical, motivational, economic, political, legal and ethical. The first three categories are deeply rooted in well-known challenges of health information systems for which structural solutions have yet to be found; the last three have solutions that lie in an international dialogue aimed at generating consensus on policies and instruments for data sharing.

**Conclusions:** The simultaneous effect of multiple interacting barriers ranging from technical to intangible issues has greatly complicated advances in public health data sharing. A systematic framework of barriers to data sharing in public health will be essential to accelerate the use of valuable information for the global good.

**Keywords:** Data sharing, Public health, Surveillance

## Background

Public health decision making has become increasingly complex and the use of data has become essential in this information age [1]. At the local level, data are used to monitor population health and to target interventions; at the national level, data are used for resource allocation, prioritization, and planning; and at the global level for estimates on the global burden of disease, to measure progress in health and development, and to contain emerging global health threats [2-7]. In addition to their primary use by public health agencies, routinely collected

public health data have become valuable for secondary use such as academic research and technology development. Recently, global health and funding agencies have made appeals for greater availability and access to granular public health data [2,8] and have developed principles for data sharing in global health [8,9].

Benefits of data sharing have been widely recognized – transparency and cooperation, reproducibility of research, cost-efficiency and preventing redundancies, acceleration of discovery and innovation, and saving lives through more efficient and effective public health programs [5,10-12]. Despite a growing global commitment to the use and sharing of public health data, this can be challenging in reality. For example the global polio eradication initiative (GPEI) could benefit from more widely available genetic sequence

\* Correspondence: wav10@pitt.edu

<sup>1</sup>University of Pittsburgh Graduate School of Public Health, DeSoto street 130, 703 Parran Hall, Pittsburgh, PA 15261, USA

Full list of author information is available at the end of the article

data to reconstruct chains of transmission, and estimates made by the global burden of disease project (GBD) and the Millennium Development Goals (MDG) would be more accurate if better quality data would be available from parts of the world [2,5,6,13,14]. Even at the local level, the efficient use and sharing of data among different agencies can be a challenge.

The field of public health is highly interdisciplinary and includes a wide range of data sources that is always evolving in size and complexity. Much data is derived directly from populations monitored by health agencies such as from clinical records and demographic and survey data. In addition, many auxiliary data sources are used to measure determinants of health such as environmental, climate, social behavior, transport, and other types of data [4]. Although overlap exists across types of data, this paper will focus on routinely collected population derived public health data such as disease surveillance data, intervention coverage data, vital statistics and cause specific mortality data. These represent some of the most widely collected, but also some of the most underused data sources in public health science and policy.

A global policy framework or operational guidelines for data sharing in public health have not yet been developed for most types of data. For example census and survey data are increasingly shared through centralized platforms such as the International Household Survey Network (IHSN) [15] or the International Public Use Microdata Series [16], but progress in sharing of disease surveillance data or cause specific mortality data has been slow. Many potential and real barriers to sharing of public health data have been recognized such as privacy issues or legal constraints but so far have only been anecdotally discussed or presented in the context of specific examples and case studies. This has led to disjointed and incomplete evidence on the scope and variety of challenges that currently limit data sharing. Unless these barriers are better understood, solutions may remain ineffective. We conducted a systematic literature review of potential barriers to data sharing and used this evidence to group these barriers in a taxonomy that can be used as a framework to facilitate an international dialogue on solutions and instruments to advance data sharing for better population health.

## Methods

We conducted a systematic review according to PRISMA guidelines [17] to identify documents that reported on barriers to data sharing in public health [see Additional file 1]. We defined public health data as data that were primarily collected by public health agencies for routine purposes such as disease surveillance or program monitoring without primary intention of research [4,18-20]. Barriers were defined as obstacles that could impede or delay data

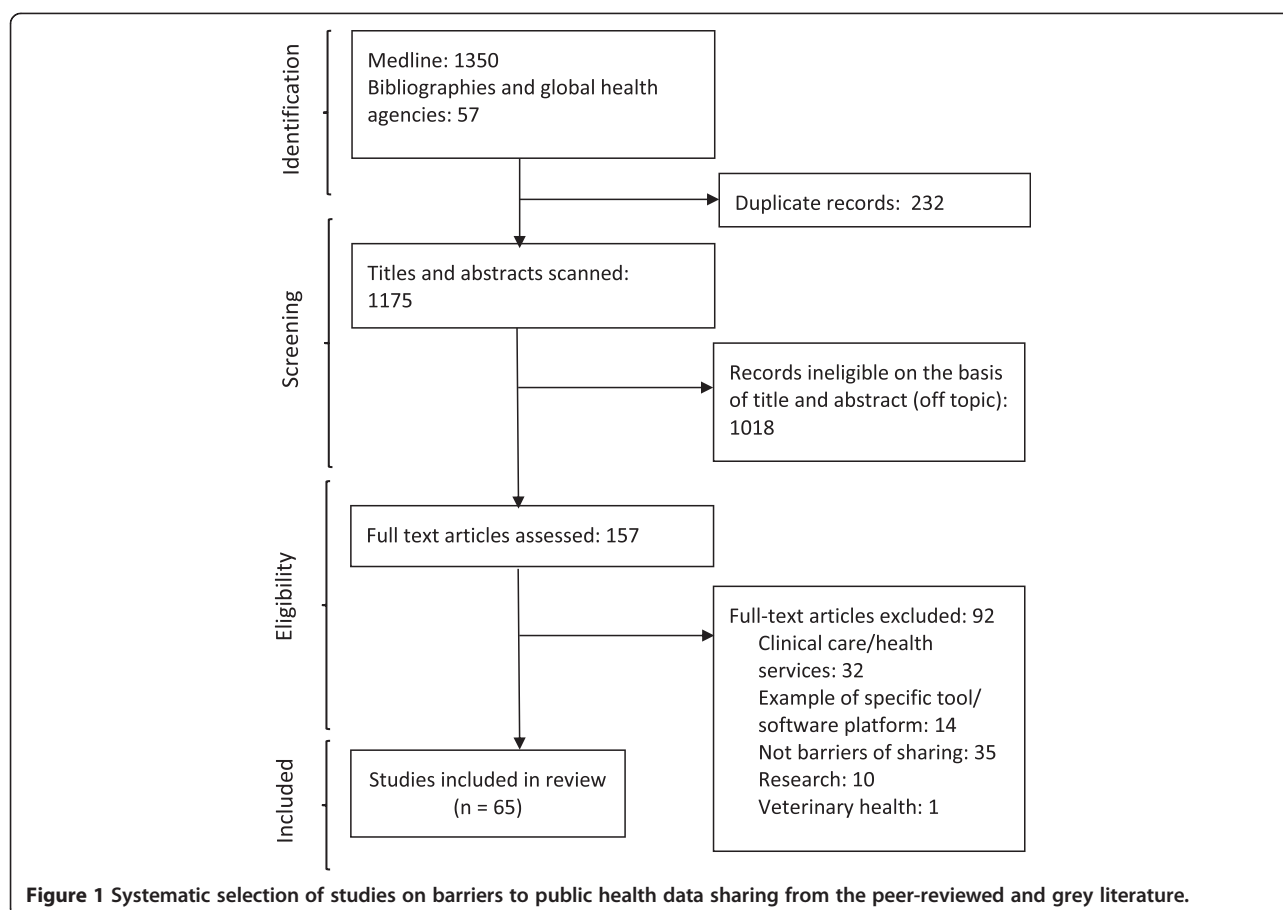
sharing or that could limit the efficiency of data sharing in public health. Studies describing barriers on clinical (patient oriented) or research data were excluded. The protocol for this review has been provided as supporting information [see Additional file 2].

We searched the MEDLINE database in August 2013 for original English-language research articles using two different queries. The first query was ["public health" OR "world health"] AND ["data sharing" OR "data access" OR "open access" OR "dissemination" OR "sharing practices"] AND ["barriers" OR "challenges"]. The second query used the following combination of key words: ["population surveillance" OR "health statistics" or "vital statistics" or "civil registry" or "health data"] AND ["data sharing" OR "data access" OR "open access" OR "dissemination" OR "sharing practices"] AND ["barriers" OR "challenges"]. Additional documentation was identified through the bibliographies of indexed papers and websites of major international agencies such as the WHO, the US Centers for Disease Control (CDC), and the Wellcome Trust.

We identified 1350 articles in MEDLINE and 57 articles from bibliographies and agency websites (Figure 1). Of these, 232 were duplicate articles and 1018 were excluded based on title and abstract review. We reviewed the full text of the remaining 157 papers. Ninety-two studies were excluded because they focused on clinical data (32), described tools for data sharing but no barriers (14), described data sharing but no barriers (35), focused on research data instead of public health data (10) or focused on animal health (1). Sixty-five studies were finally included in this review. All of these studies were initially read independently by two investigators and an initial list of barrier descriptions was extracted. This list was reviewed by domain experts among the authors and classified into preliminary categories. Experts then grouped and generalized barrier descriptions within their categories. Iteratively, a modified list of barriers was proposed and compared to the original barrier descriptions to preserve the intent of the source documents. A final taxonomy and description of barriers emerged from a series of group discussions. For each barrier, we also categorized available evidence to identify knowledge gaps. We classified studies published in peer-reviewed papers or not and presenting empirically derived evidence (through original data collection such as interviews, focus groups, etc.) or not.

## Results

We identified 20 unique real or potential barriers to data sharing in public health and classified these in a taxonomy of six categories: technical, motivational, economic, political, legal, and ethical barriers (Table 1). These barriers and categories describe a landscape of challenges that is



highly dynamic, interconnected, and hierarchical. Although most evidence (68%) was published in peer-reviewed sources, less than a quarter (22%) of all the documents reviewed was based on empirically derived evidence, indicating that a large volume of published expert knowledge has not yet been translated into scientific evidence.

### Technical barriers

These barriers for the most part are well understood as part of resilient challenges in health information system capacity and continue to form a major obstacle to the availability and use of public health data. Solutions to these barriers have been identified but sustainable implementation and political/financial commitment have been limited.

1. *Data not collected.* As long as severe limitations persist in public health data collection, data sharing will not be considered a priority. The WHO Health Metrics Network, the CDC/USAID Data for Decision Making project (DDM) and other agencies have found significant gaps in public health data systems, in particular in low- and middle income countries [2-4,21-25]. Disease surveillance systems

in many countries cannot meet standards set by the 2005 International Health Regulations [7,25-30]. Civil registration systems in many countries are lacking as well [2,6,14,18,22,24,31].

2. *Data not preserved or 3, cannot be found.* Public health data are often collected for short-term purposes such as outbreak detection. Data preservation or archiving is often not prioritized, especially in situations of limited capacity and resources [3,18,32-35]. Even if data have been preserved, data retrieval systems may be lacking. This is amplified by relocation of offices, staff turnover, physical damage to paper or electronic files, computer viruses, computer theft, etc. [34].
4. *Language barrier.* Routinely collected public health data are often recorded in local languages, limiting the possibility to integrate and use such data together with other data sets, particularly in an international context [36].
5. *Restrictive data format.* Despite major advances in computational resources in public health, a large volume of public health data such as disease surveillance data and administrative data continue to be collected and preserved in hardcopy paper format

**Table 1 Evidence for barriers to sharing of routinely collected public health data**

Category	Barrier	Peer-reviewed		Non peer-reviewed
		Empirical data	Non-empirical*	
<b>Technical</b>	1. Data not collected	[6,21,24,31]	[2,4,7,18,22,14,26-28,30]	[3,23,25]
	2. Data not preserved		[33]	[3,32,34,35]
	3. Data not found		[45]	[3,34]
	4. Language barrier			[36]
	5. Restrictive data format		[40]	[3,34,36-39,41]
	6. Technical solutions not available		[42]	[37]
	7. Lack of metadata and standards	[21,24,43]	[40,44,45]	[1,35-37,39,41,46]
<b>Motivational</b>	8. No incentives		[27,45,49]	[35]
	9. Opportunity cost	[51,52]	[13,33,50,53]	[35]
	10. Possible criticism		[33]	[32]
	11. Disagreement on data use	[21]	[49]	
<b>Economic</b>	12. Possible economic damage		[7,26,27,30]	[55]
	13. Lack of resources	[56,21]	[13,27,28,30,42,53,57]	[3,23,34-36,39,37]
<b>Political</b>	14. Lack of trust	[19,59,60]	[33,61]	[34-37]
	15. Restrictive policies		[30]	
	16. Lack of guidelines		[45,62,65]	[37,41,63,64]
<b>Legal</b>	17. Ownership and copyright		[62,65,66,69]	[37,63,64,67]
	18. Protection of privacy	[12,19,59,73,75]	[44,57,62,66,72,74]	[36,37,64,67,68,70,71]
<b>Ethical</b>	19. Lack of proportionality			[76]
	20. Lack of reciprocity	[51,52]	[50,77,78]	
Number of unique documents (% of total)		14 (21.5%)	30 (46.2%)	21 (32.3%)

\*No or little original data presented.

or in electronic format that may be antiquated or incompatible with modern software systems [3,34,36-41].

- 6. Technical solutions not available.** Technical software solutions to collect, harmonize (transformation and recoding to enhance inter-operability), integrate (combining harmonized datasets), and share complex and heterogeneous data have been developed in the private or research sector, but have not become widely available to public health agencies [37,42].
- 7. Lack of metadata and standards.** Oftentimes, metadata that describe data content, origin, methods, etc. are lacking for public health data and standards for data format, variables, and metadata are insufficiently used, limiting secondary data use and inter-operability [1,21,24,35-37,39-41,43-46]. Some advances have been made through the development of the International Classification of Diseases (ICD) [47], the Data Documentation Initiative (DDI) and the Standard Data and Metadata eXchange (SDMX) [48]. These standards are not always used efficiently however. For example, between 1950 and 2010, up to 20% of deaths in certain countries were attributed to ill-defined ICD codes [24].

#### Motivational barriers

These include barriers based on personal or institutional motivations and beliefs that limit data sharing. Solutions for this group of barriers lie in building trust or developing transparent legal agreements.

- 8. No incentives.** Data sharing requires time and resources that are chronically lacking in public health settings [27,35,37]. Personal and institutional incentives are often required to prioritize data sharing over other pressing duties [45,49], particularly if the benefit of data sharing is delayed and uncertain (e.g. possibly more efficient disease control programs) instead of immediately relevant to data providers (e.g. scientific credit or training).
- 9. Opportunity cost.** Public health officers who have invested time and effort in data collection could anticipate that scientific credit or other opportunities may be lost if data recipients with greater capacity for analysis could gain the majority of credit [13,33,35,50-52]. This is a particular challenge in low resource settings [50,53].
- 10. Possible criticism.** Data providers could be discredited by errors found during secondary use of

their data and disease control efforts may be criticized if data would reveal continued disease occurrence [32,33]. In the worst case, data sharing could reveal data fabrication or manipulation. For example, studies have shown over-reporting of vaccine coverage by country statistics compared to independent surveys after introduction of GAVI incentive funding for vaccination programs [54].

11. *Disagreement on data use.* Data providers may disagree with the intended secondary use of their data or may consider their data inappropriate for a certain use [49].

#### **Economic barriers**

These barriers concern the potential and real cost of data sharing and solutions depend on the recognition of data value and on sustainable financing mechanisms.

12. *Possible economic damage.* Data sharing in public health is challenged by the economic damage that this may cause to data providers. Public sharing of disease outbreak data, for example, can result in economic damage due to reduced tourism and trade [7,26,27,30,55]. The global SARS outbreak led to estimated economic losses of 50 billion USD between 1998 and 2004 and Foot & Mouth Disease in the UK resulted in losses of 30 billion USD between 1998 and 2003 [55]. The possibility of such significant economic implications due to (over) reactive market forces could cause great reluctance among health agencies to rapidly release disease data.
13. *Lack of resources.* The process of data sharing requires human and technical resources for data preparation, annotation, communication with recipients, computer equipment, internet connectivity, etc. [3,21,34,35,42,53]. These resources are frequently lacking in public sector agencies under economic pressure or in low income settings [3,13,21,23,25,28,30,34,36,37,56,57].

#### **Political barriers**

These are fundamental structural barriers embedded in the public health governance system that are grounded in a political or socio-cultural context. Solutions for these barriers are not clear-cut and will require global and national processes to build consensus and political will.

14. *Lack of trust.* Trust between a data provider and user greatly enables data sharing [37]. In the absence of trust, providers could anticipate potential misinterpretation, misuse or intentional abuse of the data [19,33-36,58,59]. For example the

Indonesian government refused to share H5N1 influenza samples with the international community during the 2007 pandemic due to lack of trust on the potential use of these samples for financial gain [60]. Legal arrangements were required in the absence of a trust relationship which led to the development of the Pandemic Influenza Preparedness Framework [61].

15. *Restrictive policies.* Agencies may have developed official policy guidelines that restrict data sharing, resulting from various possible underlying factors such as a general sense of distrust, negative prior experiences, or other factors [30].
16. *Lack of guidelines.* Frequently, official guidelines on data sharing simply do not exist, are unclear or inconsistent [37,41,45,62,63]. The balance between making data accessible, safeguarding privacy, and protecting intellectual, time and financial investments by public health staff is often not well regulated or standardized, resulting in protective policies on sharing of public health data in general [64,65].

#### **Legal barriers**

These barriers are legal instruments used to restrict data sharing, resulting from the underlying willingness (or not) to share data. Solutions to this group of barriers include legal instruments to facilitate data sharing and are highly dependent on solutions to underlying political barriers.

17. *Ownership and copyright.* Agencies that collect public health data are often responsible for the protection of individual and community privacy and may feel that a guardianship or ownership role is bestowed on them by the public [37,66-68]. This could result in a default of restricting access to most data [37]. Copyright can be used to restrict rather than expand access to data. In practice, it is often not well documented or known who owns public health data, resulting in inconsistent ad-hoc guidelines [37,62-65,69]. For example a project in Canada to integrate National Population Health Survey data with provincial data required a different approval process in each province [64].
18. *Protection of privacy.* Public health agencies have the mandate and authority to collect private data from the population governed by the Health Insurance Portability and Accountability Act (HIPAA) in the US or similar legislation in other countries [12,36,37,44,57,59,62,64,66-68,70-73]. A clear distinction between data containing personal identifiers and fully anonymous data may not always be possible, leading to restrictive policies on all types of data due to privacy concerns

[12,36,37,74,75]. Aggregated data without personal identifiers may not be sufficiently detailed for certain applications. Existing tools and standards for the de-identification of personal identifiers such as statistical data masking [19] may not be known or available in many contexts [12,59].

### **Ethical barriers**

These are normative barriers involving conflicts between moral principles and values. Solutions for these barriers will involve a global dialogue among all stakeholders on the ethical principles that should govern data sharing.

19. *Lack of proportionality.* The issue of proportionality, the careful deliberation in assessing the risks and benefits that derive from the amount and type of data requested compared to the potential impact of its secondary use, has been identified as a guiding ethical principle for public health data sharing [9]. Public health agencies may disagree with data requestors about the proportional risks and benefits of the secondary use of data and its impact on public health [76].
20. *Lack of reciprocity.* Data sharing practices have not always been fair, and data producers have often felt exploited in transactions where they receive little credit or benefit from their work, while data users that can rapidly analyze data and publish results benefit from academic credit and career advancement [77,78] as has happened in the past [50-52].

### **Discussion**

Using a systematic review of evidence from peer-reviewed and non peer-reviewed literature, we identified 20 unique real or potential barriers grouped in a taxonomy of technical, motivational, economic, political, legal, and ethical barriers. The complex interactions between tangible and intangible barriers at different levels can severely limit the effectiveness of isolated solutions. Strategies to resolve specific barriers may not advance data sharing at all if related barriers are not addressed as well in a comprehensive approach or if more fundamental barriers remain unchanged. Specific data sharing strategies should be tailored to different types of data. We focused this review on routinely collected population derived data such as disease surveillance, intervention coverage, or cause specific mortality data. These types of data are widely collected at ever growing spatio-temporal resolution and the extended use of this vast resource for research and policy making could greatly accelerate public health strategies and programs. The effective advancement of data sharing in public health will require a comprehensive understanding of all barriers

and a global consensus on the value and on the principles of data sharing.

Most technical, motivational, and economic barriers are deeply embedded in much larger challenges of health information system capacity, particularly in low- and middle income countries. Solutions are being developed as part of major international initiatives including infrastructure development, capacity building, and efficient financing [14,25,29,31,79]. For example, the need for sustainable financial mechanisms to create capacity and infrastructure for collection and sharing of public health data has been emphasized previously, especially for low income settings [14,79]. According to Global Fund estimates, 5-10% of program funds should be invested in data collection, monitoring, evaluation and operational research [2]. Global health partnerships and disease specific programs should use ongoing and additional funding to strengthen public health data systems and available data could be used more efficiently through joint use of integrated data for program monitoring and evaluation [2,3,5,6]. The increased use of standards and connected electronic data systems can accelerate the collection and integration across countries of basic longitudinal information such as counts of disease cases and deaths and coverage of interventions. Investments in such routine data systems will better position agencies to address ongoing challenges as well as new public health threats such as the current Ebola crisis in West Africa [80].

Various initiatives have successfully applied solutions for sharing of health data, such as the International Household Survey Network, the Demographic and Health Survey (DHS), the Multi-Indicator Cluster Surveys (MICS), and the International Network for the Continuous Demographic Evaluation of Populations and Their Health (INDEPTH) [2,81,82]. The solutions implemented by these initiatives should be translated to routine public health settings.

Political, legal, and ethical barriers will require a different approach. These barriers are less tangible and transparent compared to technical barriers and will need to be clearly outlined and presented for a dialogue across sectors with international agencies such as the World Intellectual Property Organization (WIPO), the World Health Organization (WHO), the World Trade Organization (WTO), countries, development and funding agencies, and experts in ethics and law [83]. This should lead to the creation of a political framework in the form of resolutions or a treaty, and operational guidelines for data sharing in public health [35,37]. A centralized mechanism such as a commission or secretariat should monitor, mediate, and facilitate data sharing among various stakeholders to ensure a fair and efficient use of data for the advancement of population health.

Most evidence for this review concentrated on disease surveillance or demographic data. More published evidence is needed on sharing of other types of public health data such as genomic data on emerging pathogens or cost data of public health interventions. Although we found the majority of evidence in the peer-reviewed literature (68%), most documents were based on experience or ideas (46%) instead of empirically derived information (22%). Levels of evidence were also different for each barrier. Lack of data collection and metadata and privacy issues were very well documented while no empirical evidence was available for other barriers such as data preservation and format or restrictive policies and data ownership. In-depth formative research is needed to expand the evidence base of these barriers. As knowledge on these barriers will increase, so will opportunities for solutions.

## Conclusion

Great opportunities have been created for global health cooperation, scientific discovery, and effective disease control programs by recent advances in public health data collection. These advancements are contrasted by real and potential barriers that limit the efficient use of these data. A global process will be essential for a more effective use of known solutions and to build consensus for new solutions to harness the potential of data towards a 21st century population health.

## Additional files

**Additional file 1:** PRISMA checklist.

**Additional file 2:** Literature search protocol.

## Competing interests

RW is employed by the Bill & Melinda Gates Foundation, however the decision to publish was solely made by the authors. The authors declare no other competing interests.

## Authors' contributions

WGVp contributed to study design, the literature review and wrote the manuscript. PP contributed to the literature review and manuscript editing. CE contributed to study design, the literature review and manuscript editing. JG contributed to manuscript editing. RW contributed to study design, the literature review and manuscript editing. AJH contributed to the literature review and manuscript editing. DH contributed to study design and manuscript editing. DSB contributed to study design and manuscript editing. All authors read and approved the final manuscript.

## Acknowledgements

The authors would like to thank Ms. Kate Colligan for her contribution in conceptualizing this study, Anita Dasu and Arvind Dabass for their assistance in the literature review. This study has been funded by the Bill & Melinda Gates Foundation Vaccine Modeling Initiative (Grant 49276) and the US National Institute of General Medical Sciences (Grant 5U54GM088491, Computational Models of Infectious Disease Threats). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Author details

<sup>1</sup>University of Pittsburgh Graduate School of Public Health, DeSoto street 130, 703 Parran Hall, Pittsburgh, PA 15261, USA. <sup>2</sup>Sandra Rotman Centre, University Health Network & University of Toronto, Toronto, Canada. <sup>3</sup>Bill & Melinda Gates Foundation, Seattle, USA. <sup>4</sup>INDEPTH Network, Accra, Ghana. <sup>5</sup>The Africa Centre for Health and Population Studies, UKZN, Somkhele, South Africa. <sup>6</sup>Chatham House London, London, UK.

Received: 11 June 2014 Accepted: 7 October 2014

Published: 5 November 2014

## References

1. Thacker SB, Qualters JR, Lee LM, Centers for Disease C, Prevention: **Public health surveillance in the United States: evolution and challenges.** *Morb Mortal Wkly Rep Surveill Summ* 2012, **61**(Suppl):3–9.
2. Chan M, Kazatchkine M, Lob-Levyt J, Obaid T, Schweizer J, Sidibe M, Veneman A, Yamada T: **Meeting the demand for results and accountability: a call for action on health data from eight global health agencies.** *PLoS Med* 2010, **7**(1):e1000223.
3. Health Metrics Network: **Framework and Standards for Country Health Information Systems.** 2nd edition. Edited by Organization WH. Geneva: World Health Organization; 2008.
4. AbouZahr C, Adjei S, Kanchanachitra C: **From data to policy: good practices and cautionary tales.** *Lancet* 2007, **369**(9566):1039–1046.
5. Murray CJ: **Towards good practice for health statistics: lessons from the Millennium Development Goal health indicators.** *Lancet* 2007, **369**(9564):862–873.
6. Boerma JT, Stansfield SK: **Health statistics now: are we making the right investments?** *Lancet* 2007, **369**(9563):779–786.
7. Heymann DL, Rodier GR: **Hot spots in a wired world: WHO surveillance of emerging and re-emerging infectious diseases.** *Lancet Infect Dis* 2001, **1**(5):345–353.
8. Walport M, Brest P: **Sharing research data to improve public health.** *Lancet* 2011, **377**(9765):537–539.
9. **Global health data access principles.** <http://www.gatesfoundation.org/global-health/Documents/data-access-principles.pdf>.
10. Office of E-government and IT, Office of Management and Budget: **Data.gov Concept of Operations Draft.** Washington DC: Executive office of the President of the United States; 2009.
11. Laine C, Goodman SN, Griswold ME, Sox HC: **Reproducible research: moving toward research the public can really trust.** *Ann Intern Med* 2007, **146**(6):450–453.
12. Wartenberg D, Thompson WD: **Privacy versus public health: the impact of current confidentiality rules.** *Am J Public Health* 2010, **100**(3):407–412.
13. Chandramohan D, Shibuya K, Setel P, Cairncross S, Lopez AD, Murray CJ, Zaba B, Snow RW, Binka F: **Should data from demographic surveillance systems be made more widely available to researchers?** *PLoS Med* 2008, **5**(2):e57.
14. Editorial: **Vital statistics.** *Nature* 2013, **494**(7437):281.
15. **International household survey network.** <http://www.ihsn.org/home/index.php?>
16. Center MP: **Integrated Public Use Microdata Series, International: Version 6.3.** Minneapolis: University of Minnesota; 2014.
17. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P: **Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement.** *Int J Surg* 2010, **8**(5):336–341.
18. Byass P: **The unequal world of health data.** *PLoS Med* 2009, **6**(11):e1000155.
19. Anderson M, Seltzer W: **Federal statistical confidentiality and business data: twentieth century challenges and continuing issues.** *J Privacy and Confidentiality* 2009, **1**(1):7–52.
20. Georgia Department of Human Resources (DHR) Division of Public Health: **HIPAA Fact Sheet: Public Health Practice Versus Research.** Atlanta GA: Georgia Department of Human Resources.
21. Wilkins K, Nsubuga P, Mendlein J, Mercer D, Pappaioanou M: **The data for decision making project: assessment of surveillance systems in developing countries to improve access to public health information.** *Public Health* 2008, **122**(9):914–922.
22. Vogel G: **How Do You count the dead?** *Science* 2012, **336**(6087):1372–1374.
23. Happel Lewis S, Chretien JP: **The potential utility of electronic disease surveillance systems in resource-poor settings.** *Jhns Hopkins APL Technical Digest* 2008, **27**(4):366–373.

24. Mahapatra P, Shibuya K, Lopez AD, Coullare F, Notzon FC, Rao C, Szreter S: **Civil registration systems and vital statistics: successes and missed opportunities.** *Lancet* 2007, **370**(9599):1653–1663.
25. States U, General Accounting Office: *Global Health: Challenges in Improving Infectious Disease Surveillance Systems : Report to Congressional Requesters.* Washington, D.C: The Office; 2001.
26. Cash RA, Narasimhan V: **Impediments to global surveillance of infectious diseases: consequences of open reporting in a global economy.** *Bull World Health Organ* 2000, **78**(11):1358–1367.
27. Morse SS: **Global infectious disease surveillance and health intelligence.** *Health Aff (Millwood)* 2007, **26**(4):1069–1077.
28. Chretien JP, Burkom HS, Sedyaningsih ER, Larasati RP, Lescano AG, Mundaca CC, Blazes DL, Munayco CV, Coberly JS, Ashar RJ, Lewis SH: **Syndromic surveillance: adapting innovations to developing settings.** *PLoS Med* 2008, **5**(3):e72.
29. World Health Organization: **International Health Regulations.** 2nd edition. Geneva: World Health Organization; 2005.
30. Keller M, Blench M, Tolentino H, Freifeld CC, Mandl KD, Mawudeku A, Eysenbach G, Brownstein JS: **Use of unstructured event-based reports for global infectious disease surveillance.** *Emerg Infect Dis* 2009, **15**(5).
31. Setel PW, Macfarlane SB, Szreter S, Mikkelsen L, Jha P, Stout S, AbouZahr C: **A scandal of invisibility: making everyone count by counting everyone.** *Lancet* 2007, **370**(9598):1569–1577.
32. Brest P: **President's Statement 2007: The Importance of Data.** Edited by Foundation TWaFH: The William and Florea Hewlett Foundation. 2007.
33. Pisani E, AbouZahr C: **Sharing health data: good intentions are not enough.** *Bull World Health Organ* 2010, **88**(6):462–466.
34. Baldwin W, Diers J: **Demographic Data for Development in Sub-Saharan Africa.** New York: Population Council, New York; 2009.
35. Pisani E: **OpenEpi – A new Culture for Public Health Data?** London: Ternyata; 2012.
36. Wright G, Prakash P, Abraham S, Shah N: **Open Government Data Study: India.** London: The Centre for Internet and Society; 2010.
37. Rudolph R, Davis R: **Administrative Data and Disease Surveillance: An Integration Toolkit.** In *Public Health Data Dissemination Guidelines: NAHDO Working Technical Paper Series.* Edited by Shah GH. Salt Lake City, Utah: The National Association of Health Data Organizations; 2005:51–79.
38. National Center for Vital Statistics: **Multiple Cause of Death Public use Tape, ICDA-8 (1968–1978).** Hyattsville, MD: National Center for Vital Statistics; 1980.
39. Lungo JH, Nhampossa JL: **The Impacts of Legacy Information Systems in Reporting Routine Health Delivery Services: Case Studies from Mozambique and Tanzania.** In *International ICT Workshop 2004.* Dar es Salaam, United Republic of Tanzania: University of Oslo; 2004.
40. Wu M, Zhao T, Wu C: **Public health data collection and sharing using HIPAA messages.** *J Med Syst* 2005, **29**(4):303–316.
41. Heeks R: **Information Systems for Public Sector Management.** In *Government Data: Understanding the Barriers to Citizen Access and Use.* Manchester, UK: Institute for Development Policy and Management; 2000.
42. McNabb SJ: **Comprehensive effective and efficient global public health surveillance.** *BMC Public Health* 2010, **10**(Suppl 1):S3.
43. Friedman DJ, Anderka M, Krieger JW, Land G, Solet D: **Accessing population health information through interactive systems: lessons learned and future directions.** *Public Health Rep* 2001, **116**(2):132–147.
44. Gao S, Mioc D, Anton F, Yi X, Coleman DJ: **Online GIS services for mapping and sharing disease information.** *Int J Health Geogr* 2008, **7**:8.
45. Fan X, Yu P: **A discussion about the importance of laws and policies for data sharing for public health in the People's Republic of China.** *Stud Health Technol Inform* 2007, **129**(Pt 1):316–319.
46. Centers for Disease C, Prevention: **State electronic disease surveillance systems — United States, 2007 and 2010.** *MMWR Morb Mortal Wkly Rep* 2011, **60**(41):1421–1423.
47. **International Classification of Diseases (ICD).** <http://www.who.int/classifications/icd/en/index.html>.
48. Arofan G, Heus P: **DDI and SDMX: Complementary, not Competing, Standards.** Tucson, AZ: Open Data Foundation; 2007.
49. Lopez AD: **Sharing data for public health: where is the vision?** *Bull World Health Organ* 2010, **88**(6):467.
50. Tangcharoensathien V, Boonperm J, Jongudomsuk P: **Sharing health data: developing country perspectives.** *Bull World Health Organ* 2010, **88**(6):468–469.
51. Butler D, Cyranoski D: **Flu papers spark row over credit for data.** *Nature* 2013, **497**(7447):14–15.
52. Irwin R: **Indonesia, H5N1 and Global Health Diplomacy.** *Global Health Governance* 2010, **3**(2).
53. Lang T: **Advancing global health research through digital technology and sharing data.** *Science* 2011, **331**(6018):714–717.
54. Lim SS, Stein DB, Charrow A, Murray CJ: **Tracking progress towards universal childhood immunisation and the impact of global initiatives: a systematic analysis of three-dose diphtheria, tetanus, and pertussis immunisation coverage.** *Lancet* 2008, **372**(9655):2031–2046.
55. MARSH Inc: **The Economic and Social Impact of Emerging Infectious Diseases.** New York, NY: Marsh Inc; 2008.
56. Chen L, Evans T, Anand S, Boufford JI, Brown H, Chowdhury M, Cueto M, Dare L, Dussault G, Elzinga G, Fee E, Habte D, Hanvoravongchai P, Jacobs M, Kurowski C, Michael S, Pablos-Mendez A, Sewankambo N, Solimano G, Stilwell B, de Waal A, Wibulpolprasert S: **Human resources for health: overcoming the crisis.** *Lancet* 2004, **364**(9449):1984–1990.
57. Fountain JE: **Digital government and public health.** *Prev Chronic Dis* 2004, **1**(4):A2.
58. Boussard E, Flahault A, Vibert JF, Valleron AJ, Noah N, Williams J, LaPorte RE: **Sentiweb: French communicable disease surveillance on the world wide Web.** *BMJ* 1996, **313**(7069):1381–1384.
59. El Emam K, Mercer J, Moreau K, Grava-Gubins I, Buckeridge D, Jonker E: **Physician privacy concerns when disclosing patient data for public health purposes during a pandemic influenza outbreak.** *BMC Public Health* 2011, **11**:454.
60. Fidler DP: **Influenza virus samples, international law, and global health diplomacy.** *Emerg Infect Dis* 2008, **14**(1):88–94.
61. Fidler DP, Gostin LO: **The WHO pandemic influenza preparedness framework: a milestone in global governance for health.** *JAMA* 2011, **306**(2):200–201.
62. Safran C, Bloomrosen M, Hammond WE, Labkoff S, Markel-Fox S, Tang PC, Detmer DE: **Toward a national framework for the secondary use of health data: an American Medical Informatics Association White Paper.** *J Am Med Inform Assoc* 2007, **14**(1):1–9.
63. Centers for Disease Control and Prevention: **Data Security and Confidentiality Guidelines for HIV, Viral Hepatitis, Sexually Transmitted Disease, and Tuberculosis Programs: Standards to Facilitate Sharing and Use of Surveillance Data for Public Health Action.** Atlanta (GA): U.S. Department of Health and Human Services, Centers for Disease Control and Prevention; 2011.
64. Kephart G: **Barriers to Accessing & Analyzing Health Information in Canada.** Ottawa: Dalhousie University; 2002.
65. Strobl J, Cave E, Walley T: **Data protection legislation: interpretation and barriers to research.** *Br Med J* 2000, **321**:890–892.
66. Lee LM, Gostin LO: **Ethical collection, storage, and use of public health data: a proposal for a national privacy protection.** *JAMA* 2009, **302**(1).
67. CDC/ATSDR: **CDC-CSTE Intergovernmental Data Release Guidelines Working Group (DRGWG) Report: CDC-ATSDR Data Release Guidelines and Procedures for Re-Release of State-Provided Data.** Centers for Disease Control and Prevention; 2005:1–22.
68. Centre for Disease Control and Prevention: **HIPAA Privacy Rule and public health: Guidance from CDC and the U.S. Department of health and Human Services.** *Morbidity And Mortality Weekly Report* 2003, **52**(Supplement):1–20.
69. Stansfield S: **Who owns the information? Who has the power?** *Bull World Health Organ* 2008, **86**(3):170–171.
70. US Department of Health and Human Services: **Summary of the HIPAA Privacy Rule.** Washington, DC: Department of Health and Human Services; 2003.
71. Gostin LO: **Health information privacy.** *Cornell Law Rev* 1995, **80**(3):451–528.
72. McGraw D: **Building public trust in uses of Health Insurance Portability and Accountability Act de-identified data.** *JAMIA* 2013, **20**(1):29–34.
73. Flaherty DH: **Privacy, confidentiality, and the use of Canadian health information for research and statistics.** *Canadian Public Administration: Administration Publique du Canada* 1992, **35**(1):75–93.
74. Croner CM: **Public health, GIS and the Internet.** *Annu Rev Public Health* 2003, **24**(1):57–82.
75. Lane J, Schur C: **Balancing access to health data and privacy: a review of the issues and approaches for the future.** *Health Serv Res* 2010, **45**(5p2):1456–1467.



76. Willison D, Ondrusek N, Dawson A, Emerson C, Ferris L, Saginur R, Sampson H, Upshur R: **A Framework for the Ethical Conduct of Public Health Initiatives**. Toronto, Canada: Public Health Ontario; 2012.
77. Pisani E, Whitworth J, Zaba B, Abou-Zahr C: **Time for fair trade in research data**. *Lancet* 2010, **375**(9716):703–705.
78. Editorial: **Sharing public health data: necessary and now**. *Lancet* 2010, **375**(9730):1940.
79. AbouZahr C, Cleland J, Coullare F, Macfarlane SB, Notzon FC, Setel P, Szreter S, Monitoring of Vital Events (MoVE) writing group, Anderson RN, Bawah AA, Betrán AP, Binka F, Bundhamcharoen K, Castro R, Evans T, Figueroa X, George CK, Gollogly L, Gonzalez R, Grzebién DR, Hill K, Huang Z, Hull TH, Inoue M, Jakob R, Jha P, Jiang Y, Laurenti R, Li X, Lievesley D, *et al*: **The way forward**. *Lancet* 2007, **370**(9601):1791–1799.
80. Kupferschmidt K: **Infectious disease. Estimating the Ebola epidemic**. *Science* 2014, **345**(6201):1108.
81. Dupriez O, Boyko E: **Dissemination of Microdata Files: Principles, Procedures and Practices**. International Household Survey Network; 2010.
82. Sankoh O, Herbst A, Juvekar J, Tollman S, Byass P, Tanner M: **INDEPTH launches a data repository and INDEPTHStats**. *Lancet Global Health* 2013, **1**:e69.
83. Whitworth J: **Data sharing: reaching consensus**. *Bulletin- World Health Organization* 2010, **88**(6):468.

doi:10.1186/1471-2458-14-1144

Cite this article as: van Panhuis *et al.*: A systematic review of barriers to data sharing in public health. *BMC Public Health* 2014 **14**:1144.

Submit your next manuscript to BioMed Central  
and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

