Addressing Public Health informatics patient privacy concerns

David Birnbaum
School of Population and Public Health, University of British Columbia, North Saanich, British Columbia, Canada and School of Health Information Science, University of Victoria, Victoria, British Columbia, Canada
Elizabeth Borycki
School of Health Information Science, University of Victoria, Victoria, British Columbia, Canada
Bryant Thomas Karras
Washington State Department of Health, Olympia, Washington, USA
Elizabeth Denham
Office of the Information and Privacy Commissioner for British Columbia, Victoria, British Columbia, Canada, and
Paulette Lacroix
PC Lacroix Consulting Inc., North Vancouver, British Columbia, Canada

Received 5 May 2015
Revised 5 May 2015
Accepted 27 May 2015

Abstract

Purpose – The purpose of this paper is to review stakeholder perspectives and provide a framework for improving governance in health data stewardship. Patients may wish to view their own lab results or clinical records, but others (notably academics, journalists and lawyers) tend to want scores of patient records in their search for patterns or trends. Public Health informatics capabilities are growing in scope and speed as clinical information systems, health information exchange networks and other potential database linkages enable more access to healthcare data. This change facilitates novel service improvements, but also raises new personal privacy protection issues.

Design/methodology/approach – This paper summarizes a panel session discussion from the 2015 Information Technology and Communication in Health biennial international conference. The perspectives of health service research, journalism, Public Health informatics and privacy protection were represented.

Findings – In North America, an expectation of personal privacy exists as a quasi-constitutional right. Individuals should be allowed to control the amount of information shared about them, and in particular the public expects that details of their personal healthcare data are protected. This is supported by laws, regulations and administrative structures; however, there are fundamental differences between the approaches taken in Canada and in the USA. In both countries, population and Public Health has wide powers to collect data and share it appropriately in order to accomplish a social good. A recent report issued by the British Columbia Information and Privacy Commissioner, and a recent story issued by the Bloomberg News service, highlight ways in which laws and regulations have not kept pace with advances in technology. Changes are needed to enable population and Public Health agencies to protect confidential personal information while still being able to comply with legitimate requests for data by researchers, policy makers and the public at large.

Originality/value – Similarities and differences in approach, gaps, current issues and recommendations of several countries were revealed in a conference session. Those concepts and the likelihood of ensuing legislative changes directly impact healthcare organizations' patients and leadership.

Keywords Information management, Information technology, Public health, Freedom of information, Informatics, Patient privacy

Paper type Viewpoint

Clinical Governance: An International Journal
Vol. 20 No. 2, 2015
pp. 91-100

1477-7274
DOI 10.1108/CGIJ-05-2015-0013

Public Health strives to achieve social benefit for all members of society, and public trust is fundamental in that relationship. The growing volume, velocity and vulnerabilities of "big data" and linked data in today's wired world presents good as well as bad potentials as Public Health enters a new era of information technology. In this paper, we describe the nature of health record researchers; recent privacy breaches that reveal inadequacies in current practice; proposed courses of action to improve a legal and administrative framework that balances legitimate public access to information against individual rights and expectations of privacy protection; several different approaches to establishing an effective operational framework; and an outline of what may be needed to sustain a dynamic system which undoubtedly will be challenged by future technological developments.

Patients or their designates may wish to view their own lab results or clinical records, either with their doctor or when alone; while they expect any patient portals to be secure, some of them are more willing to share other types of personal information of potential public health "crowd-sourcing" value (e.g. posting their location on Twitter during an outbreak or natural disaster; sharing shopping information through store affinity card incentive programs, etc.). Serving those interests presents legal and technical issues, which are outside the scope of this paper. This paper's focus is on the types of researchers who search across scores of patients' records. There are three major types of health record researchers (Table I) and two major places where they might seek health records (Table II).

Table I.
Who does health record research

Group	Ethical construct	What they seek
Academics	Institutional Review Board (IRB) approval, accountability through their university	Granularity, linkages across data sets, refined analysis, credibility
Journalists	Evolving standards	Details, compelling personal drama, contacts, fast access
Mainstream news media	Old: accuracy, independence, impartiality, oversight of editors	Pitching story their audience will buy
Internet/new media	New: transparency, errors corrected quickly, multiple voices and single reporters (not in a news organization under editors)	Tight story filing deadlines Primacy in 24 hour news cycle Credibility (old media) or is it "Truthiness" (new media)
Lawyers	Rules of law and professional conduct, accountability through bar associations	Case details, background context details, names, dates, etc.

Table II.
Where health record researchers can get data

Public Health agency databases	Provider organization medical records
Agency is subject to public document requests *Multiple databases can be probed to rebuild linkages* Agency has limited ability to redact anything other than personal identifiers *Even more sensitive data can be obtained by signing a data use agreement* Limited enforcement options	*Can subpoena records* Otherwise, providers operate under privacy protection laws Variation in how individual providers interpret those laws, but usually conservative Codes of conduct bar academics, journalists and lawyers from impersonating doctors to ask for patient records

Academic researchers seek as fine a level of detail ("granularity") as possible, so that they can determine optimal levels of aggregation and confirm duplications have been eliminated from their data set. These researchers are accustomed to:

- requesting detail limited to the data elements specific to a particular research question being addressed by their study;
- following national, local and institutional ethical and privacy protection laws, regulations and guidelines; and
- having ethical aspects of proposed projects reviewed by independent review boards (an IRB) to ensure that all project details (including protection of people as research subjects) conforms to well-defined constructs (such as Canada's Tri-Council Policy Statement, akin to the USA's "Common Rule" Policies for the Protection of Human Subjects).

Protection of individual privacy is inherent in those constructs, which gain a further layer of oversight by policies and offices of the academic institution through which its researchers must enter into data use agreements. Those contractual agreements between organizations detail how data will be collected, accessed, used, stored and ultimately destroyed when no longer needed.

Privacy has never been part of the ethical construct in journalism. Consistent with vital freedom of the press, it has been said that good journalism invades individual privacy (so long as revealing personal details is critical to the story rather than solely salacious, and there is absence of malice). Traditional press values have centered on accuracy (getting a story out first is good, but getting it right is critical), independence (for which consolidation of media ownership raises challenges) and impartiality (which does not simply imply giving equal time to people expressing opposing views regardless of factual evidence). However, the rise of blogs, Twitter or Facebook postings, YouTube and other internet broadcasts have enabled anyone to call themselves a reporter. The business model of "old media" mainstream news channels has been disrupted by these technological innovations to the point that established news agencies also attempt to give their subscribers more avenues to comment on-line along with their employed reporters. In this new model, "trust me because I'm Edward R. Murrow or Walter Cronkite" is replaced by transparency (showing raw data so people can judge conclusions for themselves); "watch the evening news" is replaced by 24-hour-a-day-7-day-a-week instant on-demand news cycle in which being first with the story, and quick with corrections, becomes paramount. The so-called Fourth Estate free press, as well as its newer internet-empowered offshoot, is an important asset for an informed democracy; however, its newer offshoot also presents more challenges for those trying to balance access to information vs fundamental privacy protection rights. As outlined in Table I, it also is important to recognize that journalism achieves influence by conveying pictures of powerful human drama – personal stories that have stronger public resonance than the more abstract scientific reports of academics.

The third group is lawyers, who have obvious vested interests in obtaining as much data and as detailed data as possible. They have greater knowledge of freedom of information laws (also known as request for public documents), and greater ability to subpoena records. They, like the academic researchers and journalists, have come to recognize that Public Health departments could be a richer and easier source of information than actual health service providers like doctors or hospitals (Table II). The crux of the problem is that legislative provisions have not kept pace with advances

in technology and a growing scope of surveillance data is being collected that can enable increasing effectiveness of health promotion, disease outbreak management and bio-terrorism response.

The list of surveillance data sources is long, including:

- vital statistics (e.g. birth and death records);
- hospital discharge abstracts (e.g. CIHI in Canada, CHARS in USA);
- clinical laboratories;
- emergency rooms (e.g. Syndromic Surveillance);
- doctors' offices, hospitals, clinics (e.g. Reportable and Notifiable Diseases);
- sentinel events (e.g. geographic reporting of a single occurrence of major concern);
- immunization registries;
- occupational health risks (e.g. disease exposures, school attendance records, etc.);
- case registries (e.g. cancer, cardiovascular disease, diabetes);
- health surveys (e.g. habits of nutrition, exercise, smoking, etc.);
- environmental data (e.g. transportation access, weather, water and air quality, etc.); and
- social media (e.g. Facebook, Twitter) and other internet sites (e.g. Google Flu Trends, Healthmap.org).

Public Health data also crosses jurisdictional boundaries because countries, states and provinces share information to deal with global health problems as well as build Public Health surveillance capacity around the world. For example, the Centers for Disease Control and Prevention (CDC) Global Public Health Informatics Program:

- provides informatics assistance to CDC supported countries (e.g. electronic integrated disease surveillance systems);
- established/administers a World Health Organization Collaborating Center for Public Health Informatics;
- participates in development of International Health Regulations; and
- supports Health Metrics Network and Open Architectures, Standards and Information Systems (OASIS) for development of personal health information architecture.

Data collected and stored by Public Health agencies has come from hospital discharge summaries, from laboratory or physician notifications when they recognize a case of reportable contagious disease, from field monitoring studies of air or water pollution levels … from traditional sources. Other data are now coming from novel sources – for example faster than usually possible control of a food-borne outbreak of Hepatitis A by using grocery store loyalty card data (www.cbc.ca/news/health/hep-a-food-outbreak-traced-with-grocery-store-loyalty-card-clue-1.2637273).

Different countries have taken different approaches to protecting individual privacy while enabling legitimate data access. The USA has taken a sectorial approach, with laws specific to each setting (e.g. the Health Insurance Portability and Accountability Act (HIPAA) to cover healthcare data). While statutory authority in acts like HIPAA delineate privacy provisions, other acts ("sunshine laws") delineate as much open access to internal

government documents as possible (http://ballotpedia.org/State_sunshine_laws). Other federal laws also need to be considered if federal agency computer systems are involved (Federal Information Security Management Act of 2002). It then becomes the role of in-house counsel to interpret the law and the courts to provide redress when data access is denied or a privacy breach causes alleged damage. Canada has taken an umbrella approach, with public-sector and private-sector laws overarching all settings to promote a trust model. This is closer to, but not identical with, the European Union Directive approach. More like Europe and unlike the USA, Canada also has established administration of its information access and privacy framework under the offices of independent regulators who report to their respective legislative assemblies (rather than to the government political party in power). For example, as an officer of the provincial legislature, the British Columbia Information and Privacy Commissioner has regulatory and quasi-judicial authority to:

- conduct public education and outreach activities;
- provide confidential consultations;
- make public comments on programs;
- conduct investigations and audits; and
- issue binding orders enforceable in a Court of Law.

Decisions have a legal basis in British Columbia's Freedom of Information and Protection of Privacy Act (FIPPA), Personal Information Protection Act (PIPA) and a dozen other federal and provincial laws. While FIPPA, PIPA, federal PIPEDA and other laws do present a jigsaw puzzle of laws to navigate, their overarching nature probably leaves fewer gaps than the American HIPPA statute. HIPPA, for example, prohibits "covered" entities (identified in the Act, including doctors and hospitals) from releasing patient data to others without a patient's permission yet other entities are not so "covered" (e.g. HIPPA has a special provision for Public Health that exempts the patient permission requirement before it can receive patient data required for a specific purpose; HIPPA contains no coverage of companies that provide direct-to-consumer testing such as those now offering to map anyone's genome). Also, in the USA any redress is available only through the courts which involve costs for legal representation; conversely, complaints in Canada at both the provincial and federal levels can be brought directly to the Information and Privacy Commissioner without cost.

Historically, there are a number of different ways in which Public Health data sets can be shared. Data stripped of individual identifiers (data elements that directly identify an individual, such as name, social security number, etc.) in accord with America's Safe Harbor convention was presumed safe for unrestricted access. Data sets containing elements that might indirectly facilitate identification of individuals have been available to third parties after those parties sign a data use agreement. Particularly sensitive data sets have been shared between trusted government agencies (e.g. collected by state Public Health but shared with local Public Health departments, data shared between state and federal Public Health agencies, data required by law to be shared with other governmental agencies). However, as illustrated by a recent Bloomberg News item (Robertson, 2013), current legislated privacy protections in the USA have been overcome by technology-enabled challenges in the relationships between Public Health, non-governmental organizations, industry, public requestors and various inter-governmental agency levels (Table III). News reporters, using Public Health public use data sets that contain hospital

Table III. Current American challenges to sharing data while protecting personal identifiable information

Challenge	NGO partners	Industry	Public	Government
Statutory requirements that are intended to protect patient confidentiality but are not adequate to meet the intent	√	√	√	√
Data-sharing agreements cannot be enforced – Misuse of de-identified data to re-identify people (outside of IRB approval or other statutory authority)	√	√	√	
Lack of statutory authority or statutory restrictions re: sharing confidential data with other governmental entities				√
Lack of statutory authority of receiving governmental entity to protect confidentiality of data from re-disclosure (from/to a department of health)				√
Sophistication of data matching ability even with use of HIPAA "safe-harbor" criteria	√	√	√	
Targeted public records requests (identifies one person or facility)			√	

discharge record abstracts stripped of direct personal identifiers in accord with Safe Harbor specifications, were able to reconstruct links by comparing those data sets against other publically available information sources to identify patients by name and home address. At the same time, an increase in sophisticated records requests by other parties for multiple data sets that, when combined, could facilitate identification of individuals was noticed within the Washington State Health Department. It was decided that Safe Harbor did not de-identify public use data sets adequately; so much detail needed to be removed or masked that the utility of public use data sets is compromised; and current laws are not keeping up with technology available to make linkages in analysis of "big data." Present technology offers a half dozen different ways to share data electronically, but none are free from all potential cybersecurity, stewardship, functionality and related risk issues (Table IV). As Public Health moves farther and farther into use of geographic information system (GIS) technology, that path will open even broader potential benefits, privacy threats, as well as potential to be overwhelmed by the sheer volume of data unless an underlying data strategy guides Public Health into the future.

The big gaps in today's wired world are not unique to health data, but health data is among the most sensitive of personal information. Gaps facing today's balance between legitimate data access and personal privacy expectations include:

(1) Deficient meaningful remedies:

- tangible and intangible personal harm is done by breaches, along with serious social harm in resulting loss of trust; and
- data use agreements are only civil contracts, so do not insulate health departments from legal recompense by offended third parties.

(2) Remedial rather than punitive authority under law:

- new authority to make findings and issue significant fines warrants consideration.

(3) Variable accountability in management practices:

- sound data management should be expected at a level no less than security of accurate transactions in bank deposits and withdrawals.

Issue	Data transfer	Hosted analytics	Algorithm sharing	Application sharing	Distributed application	Directory service
Cybersecurity Data access	managed	*exposed*	avoided	avoided	managed	avoided
Cybersecurity Malicious code	avoided	managed	managed	*exposed*	*exposed*	avoided
Stewardship Intended misuse	*exposed*	managed	avoided	avoided	avoided	managed
Stewardship Legal barriers	*exposed*	*exposed*	avoided	avoided	managed	avoided
Functionality Limited data access	avoided	avoided	*exposed*	*exposed*	managed	*exposed*
Functionality Limited algorithm access	*exposed*	avoided	managed	avoided	managed	*exposed*
Functionality Erroneous algorithm	avoided	managed	*exposed*	managed	avoided	managed
Barriers Tied to other tools	managed	avoided	managed	avoided	avoided	*exposed*
Barriers Prohibitive cost	managed	managed	managed	managed	managed	avoided
Barriers Complex explanations	avoided	managed	avoided	avoided	*exposed*	managed

Table IV. Data-sharing models and their potential risks

(4) The need to create a secure research platform for work which links health, social service, justice and other databases across the entire spectrum of determinants of health:

- options warranting discussion include creating a single IRB (rather than individual IRBs in each institution), creating an overarching certification or other means to ensure all IRBs operate in a consistent manner enabling acceptance of each other's determinations without further review, etc.

(5) Ensuring education in key aspects of Public Health law for all Public Health degree and other research graduate program students, developing a certification for those wanting to follow a career path toward becoming a chief privacy officer or privacy commissioner.

Standardized competencies and an international certification for the emerging profession of "Informatician" is also a critical development (Karras *et al.*, 2009), which should further help ensure basic knowledge of privacy principles, security standards, ethical frameworks and legal constructs in the collection, use and disclosure of personal information. Public Health's dynamic problem sphere centers around:

- surveillance that requires multiple data sources;
- information sharing that requires management of inter-organizational aspects of data;

- combining different types of data which presents new problems (e.g. the sum is bigger than the parts); and
- the need to share health data with non-medical organizations which can present barriers to timely response in an outbreak or crisis.

What will it take to move forward and achieve a sustainable data strategy in Public Health? The proliferation of geo-spacial data linked in GISs provides an opportunity in Public Health to realize a vision of real-time population surveillance at a community level. The benefits can be significant, enabling Public Health to enlist new strategies in managing disease outbreaks, environmental hazards, bio-terrorism threats and now potential cyber threats to critical environmental infrastructure. However, ethical issues at the heart of public trust are of paramount importance because they are intrinsic to the necessary framework needed for such future systems (Olvingson *et al.*, 2002; Eysenbach, 2009).

We suggest that this dialogue already has started in various countries. For example, at a national level there have been initiatives in the USA, Canada, the UK and Australia. In 2014 the National Committee on Vital and Health Statistics and the President's IT Advisory Committee supported the development of a National Health Information Infrastructure with a privacy and security framework requiring a more comprehensive level of protection for all Public Health data (www.ncvhs.hhs.gov/wp-content/uploads/2014/07/140616lt.pdf). Canada has developed a trust model of privacy and security for their Pan-Canadian e-Health Record, of which Public Health has a dedicated information system to manage immunizations and outbreak surveillance on a national level (https://sl.infoway-inforoute.ca/downloads/Panorama%20Overview.pdf). In 2012 the UK Medical Research Council established a consortium of researchers to strengthen the UK's capability in analyzing and linking data within the National Health System's database of 62 million records, establishing a supportive infrastructure for safe data sharing between 24 UK academic institutions (www.mrc.ac.uk/research/initiatives/health-and-biomedical-informatics/initiatives-in-informatics-research). Australia has developed a National e-Health Security and Access Framework overarching across the healthcare sector. Their Personal Controlled Electronic Health Record seeks to balance patient privacy with clinical information requirements (www.oaic.gov.au/privacy/privacy-act/e-health-records). At the state or provincial level, calls for public comment have been issued on agency-request legislation in Washington State, and on a recent special report by the Office of the Information and Privacy Commissioner for British Columbia ("A Prescription for Legislative Reform: Improving Privacy Protection in BC's Health Sector" available at www.oipc.bc.ca/report/special-reports).

In today's world social media has become a partner in surveillance, where data is "big" and data linking more easily achieved through zip codes and geo-location. There is a real need for Public Health leaders to frame an on-going conversation in terms of what information is needed and how individual privacy is safeguarded while serving other needs of the public. Where to start? Engaging a wider group of stakeholders, including regulators, researchers, health service administrators and subject matter experts in technology and security who can contribute meaningful strategies for safe data sharing is key. There also is a need to consult citizens who use social media technologies during times of crisis, in terms of learning about their expectations and willingness to share information publically. For example, during the Japanese Tsunami crisis, citizens used social media for identifying safety issues, to locate individuals in crisis, to identify potential environmental hazards that may

affect Public Health and to provide information and support for those in the affected region (Peary *et al.*, 2012). Social media has been used by citizens and health professionals in a similar way for responding to disease outbreaks (Hines and Sibbald, 2015). The following issues for such conversations have been compiled from literature reflecting the realities of Public Health in an Informatics 2.0 world:

- meeting Public Health data requirements from within clinical care information systems by assuring flexible access to data during disease outbreaks;
- creating common national standards between Public Health and clinical entities that may be automated within normal business practices;
- educating other healthcare providers about changing roles in Public Health;
- funding Public Health capacity and infrastructure so it can play a vital role in value-added clinical information exchanges such as electronic medical or health record systems, health information exchange networks, etc.;
- identifying, developing and disseminating tools to summarize and transform complex data into meaningful information;
- developing nationally recognized trust models to enable appropriate data sharing;
- developing new/novel scientific methods (e.g. prediction and probability vs explanation); and
- viewing "the cloud" as a platform for data collection/linkages/analytics.

Conversations on these issues must keep in mind distinctions between medical and Public Health ethics, as well as ethical constructs of other partners (e.g. academic researchers, journalists, lawyers, information technology private companies, non-profit foundations, etc.) while continuing to maintain a balance between individual and public interests.

References

Eysenbach, G. (2009), "Infodemiology and infoveillance: framework for an emerging set of Public Health informatics methods to analyze search, communication and publication behaviour on the internet", *J Med Internet Res*, Vol. 11 No. 1, p. e11, available at: www.jmir.org/2009/1/e11/ (accessed April 14, 2015).

Hines, D. and Sibbald, S.L. (2015), "Citizen science: exploring its application as a tool for prodromic surveillance of vector-borne disease", *Canada Communicable Disease Report*, Vol. 41 No. 3, pp. 63-67, available at: www.phac-aspc.gc.ca/publicat/ccdr-rmtc/15vol41/dr-rm41-03/surv-4-eng.php (accessed April 14, 2015).

Karras, B., Davies, J., Koo, D., Richards, J., O'Carroll, P., Miner, K., Oberle, M., Corn, M., Detmer, D., Foldy, S., Hanrahan, L., Hare, G., LaVenture, M., Lynch, C., Roderer, N., Ross, D., Savel, T. and Sondik, E. (2009), *Competencies for Public Health Informaticians, 2009*, US Department of Health and Human Services, Centers for Disease Control and Prevention, Atlanta, GA, available at: www.cdc.gov/InformaticsCompetencies/ (accessed April 14, 2015).

Olvingson, C., Hallberg, J., Toomas, T. and Lindqvist, K. (2002), "Ethical issues in Public Health informatics: implications for system design when sharing geographic information", *J Biomed Informatics*, Vol. 35 No. 3, pp. 178-185.

Peary, B.D., Shaw, R. and Takeuchi, Y. (2012), "Utilization of social media in the East Japan earthquake and tsunami and its effectiveness", *Journal of Natural Disaster Science*, Vol. 34 No. 1, pp. 3-18.

Robertson, J. (2013), *States' Hospital Data for Sale Puts Privacy in Jeopardy*, Bloomberg, San Francisco, CA, available at: www.bloomberg.com/news/articles/2013-06-05/states-hospital-data-for-sale-puts-privacy-in-jeopardy (accessed April 14, 2015).

Further reading

FIPPA (1996), "Freedom of Information and Protection of Privacy Act [RSBC 1996] c.165", available at: www.bclaws.ca/civix/content/complete/statreg/1198514681/?xsl=/templates/browse.xsl

HIPPA (1996), "Health Insurance Portability and Accountability Act of 1996 [104-191]", available at: www.hhs.gov/ocr/privacy/hipaa/administrative/statute/hipaastatutepdf.pdf

PIPA (2003), "Personal Information Protection Act [SBC 2003] c.63", available at: www.bclaws.ca/civix/content/complete/statreg/1922970521/?xsl=/templates/browse.xsl

PIPEDA (2000), "Personal Information Protection and Electronic Documents Act [SC 2000] c.5", available at: /laws-lois.justice.gc.ca/eng/acts/P-8.6/index.html

Corresponding author

Paulette Lacroix can be contacted at: placroix@placroix.ca