

Health IT Policy Committee – September 18, 2009

Models for Data Storage & Exchange, Aggregate Data, De-identification/ Re-identification

Claudia Williams

Director, Health Policy and Public Affairs

Markle Foundation

Thank you for the opportunity to speak to the Health IT Policy Committee today.

Health Information Technology (IT) has great potential to improve health and health care if we focus on two clear goals: (1) unleashing the potential of networked information to achieve measurable health improvements, reduce costs and engage patients; and (2) maintaining public trust by making privacy a critical enabler of information sharing and health IT adoption.

Today's panel addresses this second goal of privacy and public trust, and in particular the architecture choices to support trusted information sharing. My remarks will focus on three key points:

1. **Adopt a Framework-Based Approach** - The full array of foundational privacy principles, sound network design and strong governance and accountability are all needed and must work together to assure trusted information sharing.
http://connectingforhealth.org/resources/20080822_policy_brief.pdf¹
2. **Ensure that Policy Guides Technology** - Policy goals must shape technology choices, including standards and architecture, and not vice versa.
3. **Stimulate Innovative Models for Protecting and Sharing Information** - Public investments should support and encourage innovative models to achieve our health goals and protect information.

Better Use of Information is Needed to Improve Health and Reduce Costs

The potential of networked information to achieve measurable health improvement is enormous. Access to and use of critical information—recent lab values, home monitoring results, discharge summaries, medication fill histories—is the lifeblood of health improvement.

- But this critical information is often not available when and where it is needed. For instance, primary care physicians only have hospital care summaries one third of the time when they first see recently discharged patients.²
- Redundancy, inefficiency and unneeded administrative overhead result from information gaps. On average physicians spend three weeks per year simply interacting with health plans for a total national cost of \$31 billion annually.³
- The net effect of a lack of information for clinical decisions is that innovations that can improve care are disseminated and adopted painfully slowly. It takes 17 years to achieve wide-spread adoption of a new evidence-based practice in health.⁴
- There is inconsistent delivery of proven care. For instance, adults only receive 55 percent of recommended care.⁵

To address these issues, we will need a 21st century health information ecosystem characterized by:^{6,7}

- Trusted, distributed and dynamic access to information by authorized users that will support patients and providers in making the best decisions and improving care.
- Information management and architecture models that can achieve the privacy and security protections required while limiting complexity and cost.
- Leveraging distributed analysis across data sources when collection of identifiable information is not necessary.
- Requiring feedback loops in quality, research and public health to support rapid learning and high quality care.
- Enabling vital health information sharing among authorized users across organizational and geographic boundaries while protecting privacy.

Connecting for Health Common Framework

Core Privacy Principles

Openness and Transparency

- Communicate policies to participants and individuals
- Provide privacy notices to consumers
- Involve stakeholders in developing information sharing policies

Purpose Specification

- Specify the purpose of the data collection effort clearly and make it narrowly suited to the need

Collection Limitation and Minimization

- Assure that only data needed for specified purposes are being collected and shared

Use Limitation

- Establish processes to ensure that data are only used for the agreed upon and stated purposes
- Establish what data access is permitted for each user

Individual Participation and Control

- Allow individuals to find out what data have been collected and who has access, and exercise meaningful control over data sharing
- Give individuals access to information about them, and the ability to request corrections and see audit logs

Data Integrity and Quality

- Provide that data are relevant, accurate, complete and up-to-date

Security Safeguards and Controls

- Establish tools and mechanisms to provide that data are secured against breaches, loss or unauthorized access
- Establish tools and approaches for user authentication and access

Accountability and Oversight

- Establish who monitors compliance with policies and procedures for handling breach
- Produce and make available audit logs

Remedies

- Establish mechanisms for complaints
- Establish remedies for affected parties to compensate for harm caused by breach

Privacy is a Critical Enabler of Health IT. A Policy Framework is Needed

The success of the American Recovery and Reinvestment Act of 2009 (ARRA) will depend in no small part on whether the public and health industry participants trust that information will be protected.

In 2005, Markle Connecting for Health articulated a policy framework for enabling information sharing while protecting privacy.⁸ The framework

(http://www.connectingforhealth.org/commonframework/docs/P1_CFH_Architecture.pdf)

hinges on nine core privacy principles (see box) derived from fair information practice principles (FIPPS) that have guided information-sharing efforts worldwide since the 1970s.⁹ The principles require that limits be set on data collection and use, that patients have access to and reasonable control over their health information, and that security safeguards are adopted.¹⁰

(<http://www.connectingforhealth.org/commonframework/docs/Overview.pdf>).

As Bob Gellman points out and as we discuss in a recent policy brief authored with the Center for Democracy and Technology (<http://www.cdt.org/healthprivacy/20080221consentbrief.pdf>), no one mechanism, including patient consent, can on its own protect information.¹¹

Individual participation and control is certainly one of the elements of the framework, but it is most meaningful if buttressed by the other principles, policies and practices.

Policy Guides Technology

Over the years we have seen this framework translated into very specific practices within the health sector. This policy-driven approach means that when data are needed for public health, research, quality or some other authorized use, the purpose must be specified and only the data necessary for achieving that objective is shared. The other principles when taken together buttress each other. Data use and sharing are made transparent through immutable audit logs. Data stay as close as possible to where they are captured, and are shared according to specific needs and with specific purpose. In contrast, a technology driven approach often starts with technical requirements and is driven by technology decisions. While the latter approach operates without policy requirements or constraints, technology decisions inevitably result in policies made by default that can be misaligned with or inadaptable to policy goals that are established after the fact.

Many of the Markle Connecting for Health Common Framework privacy principles are directly addressed by new ARRA requirements, including for breach notification, accounting of disclosures, giving patients access to electronic information, and guidance on minimum necessary.

These principles should guide and shape clear policies and technology choices, including how information is discovered, exchanged, analyzed and stored as we share it across the health care system.

Innovative Models for Protecting and Sharing Information

I will walk through a few examples—from health information exchange, research, quality reporting and public health—of how these privacy principles can translate into operational decisions about how information is shared across the health care system (as opposed to within entities).

These examples are not meant to serve as uniform ready-made solutions, but rather as illustrations of how we can use technology and architecture to reach our goals: improving health and health care by sharing information while leveraging technical approaches that are privacy-protective.

Health Information Exchange

Applying the principles of purpose specification, transparency, collection limitation, data integrity and quality result in architectures in which data are locally controlled, and are shared as needed to fulfill specific purposes. Data remain distributed (they are not comingled in one database) while being

“discoverable” using directories and other technical tools that prevent the need to disclose all of the underlying data.

The Mid-South e-Health Alliance shares lab, imaging, diagnosis and discharge summary information with emergency rooms, hospitalists and primary care providers at eight hospitals and health systems in Memphis.¹² Local control, clear policies for information sharing, clear contractual and enforcement policies and a distributed model that does not require all participants to share their data in one logically centralized repository, have all contributed to trusted information sharing among “traditionally competing” entities.¹³

This example echoes some of the points Marc Overhage made. He discussed how privacy was “architected” into health information exchange through the **Indiana Health Information Exchange** (IHIE). Providers are custodians of their data. Patient information can be identified across the network using a directory and shared with specific rules guiding its use.

Research

Emerging approaches for research and analysis benefit from the computational power of distributed information sources without the costs, time lags, privacy risks and data quality issues that develop when creating new aggregated databases that must first collect, clean and centralize data before they can be used for analytic purposes.

The HMO research network is a consortium of 15 HMO organizations that has conducted collaborative multi-site studies on a wide range of clinical and health policy topics including medication safety, cancer care quality and cardiac disease management. Each site applies a common research question to its own local data. Results are reported and aggregated.¹⁴ It is one example of a distributed health data network allowing researchers to ask the same questions across multiple similarly structured databases housed in different organizations. Only appropriate levels of summary data or results are returned to the researcher; not all of the source data.¹⁵

Quality Reporting

The concept of collection limitation and minimization does not just mean stripping data fields that are not needed before we share whole data sets, but also sharing information in the least revealing form—

anonymized, hashed or as aggregated results and not underlying data—that allow users to answer critical questions without unnecessary exposure.

Many quality improvement efforts require participants to share personally identifiable health information in order to aggregate and analyze quality information. **New York City's Primary Care Information Project** has taken a different approach. Physicians' EHRs directly generate quality measures, and these summary results are reported to the Citywide Quality Reporting System with built-in mechanisms for audit. Only the needed results, not identifiable health information, are shared to support quality improvement.¹⁶

Public Health

The **DiSTRIBuTE** initiative¹⁷ takes a similar approach to flu surveillance. Hospitals and clinics report simple aggregate flu counts (in whatever manner they use to determine whether a patient has flu-like illness), not underlying identifiable health information or the atomized fields required to determine whether the patient has flu-like illness centrally. The system performs as well as traditional flu reporting in identifying emerging outbreaks. It is also proving to be highly accurate, timely, cost-effective and by definition creates fewer exposure risks than surveillance approaches that attempt to collect detailed underlying data fields. Like the other models we have discussed, it represents an innovative response to the question of how to achieve our public health objectives while protecting information by clearly specifying the purpose, collecting just the information that is necessary to truly accomplish the task and keeping the detailed data and some of the analysis as close to the source as possible.

To be sure, there is no one-size-fits-all technical approach. Every effort should start by defining why information is being shared and with whom and what the clear purpose is. Guided by this clear purpose and core privacy principles, only then can it determine what information should be shared and with what technical approach.

Conclusion

We must use public funding to encourage adoption of successful models that use technology and architecture to protect information as it is shared, guided by a clear set of information policies that create trust.

Our policy strategy for networked health information should:

1. **Adopt a Framework-Based Approach**, requiring that information sharing efforts funded by public dollars address the three basic components of trusted information sharing: the set of core privacy principles, sound network design and strong governance and accountability. These policy requirements can be implemented through grants procurement under ARRA, recognizing that government has a role in implementing a privacy framework for health IT, particularly in efforts supported by public funds.¹⁸
2. **Ensure that Policy Guides Technology** by using the basic tenets of the fair information principles such as purpose specification and collection minimization in the design of quality, comparative effectiveness, information exchange and public health efforts.
3. **Stimulate Innovative Models for Protecting and Sharing Information** by investing in methodologies and approaches to address the analytic challenges of distributed data networks and analysis for quality and public health and developing approaches to share and use information that reduce unnecessary exposure through privacy-protective architecture.

¹ See Markle Connecting for Health. 2008. We Need a 21st Century Privacy Approach Allowing Americans to Protect and Share Health Information to Improve Quality, Policy Brief. http://connectingforhealth.org/resources/20080822_policy_brief.pdf (accessed September 14, 2009).

² A review of studies found that in the first visit after discharge, summaries of hospital care were only available to primary care physicians 12-34 percent of the time. See Kripalani, S., F. LeFevre, et al. 2007. Deficits in Communication and Information Transfer Between Hospital-Based and Primary Care Physicians. *JAMA* 297:831-841.

³ Casalino, L., S. Nicholson, D. Gans, et al. 2009. What Does It Cost Physician Practices to Interact with Health Insurance Plans? *Health Affairs* 28(4): w533-w543.

⁴ Balas, E., S. Boren. 2000. Managing clinical knowledge for healthcare improvement. In: Yearbook of Medical Informatics. Bethesda, Md. National Library of Medicine: 65-70.

⁵ McGlynn, E., S. Asch, J. Adams, J. Keeseey, J. Hicks, A. DeCristofaro, and E. Kerr. 2003. The Quality of Health Care Delivered to Adults in the United States. *The New England Journal of Medicine* 348 (26):2635-2645.

⁶ Diamond, C. 2008. Learning What Works: Presentation to IOM Roundtable on Evidence-Based Medicine, July 30, 2008.

⁷ See Diamond, C., F. Mostashari, C. Shirky. 2009. Collecting and Sharing Data for Population Health: A New Paradigm. *Health Affairs* 28 (2) 454-466.

⁸ See <http://www.connectingforhealth.org/commonframework/index.html> (accessed 9/16/09).

⁹ The FIPPS have shaped many US laws (including the 1974 Privacy Act, the HIPAA Privacy Rule and the Fair Credit Reporting Act) and privacy frameworks of Federal Agencies. See, for instance, DHS adoption of FIPPS as the basis for their privacy policy. http://www.dhs.gov/xlibrary/assets/privacy/privacy_policyguide_2008-01.pdf (accessed September 13, 2009).

¹⁰ See Markle Connecting for Health: The Connecting for Health Common Framework: Overview and Principles. Available at www.connectingforhealth.org (accessed 9/16/09).

¹¹ See Markle Connecting for Health. 2008. Beyond Consumer Consent: Why We Need a Comprehensive Approach to Privacy in a Networked World: Policy Brief. <http://www.cdt.org/healthprivacy/20080221consentbrief.pdf> (accessed September 14, 2009).

¹² As of April, 2009. See summary prepared for AMIA by Mark Frisse, April 2009. <http://www.markfrisse.com/docs/2009-april-mseha-talking-points.pdf> (accessed September 13, 2009).

¹³ Frisse, M., J.K. King, et al. 2008. A Regional Health Information Exchange: Architecture and Implementation. *AMIA Annual Symposium Proceedings*.

¹⁴ See <http://www.hmoresearchnetwork.org> (accessed September 16, 2009).

¹⁵ See http://www.hmoresearchnetwork.org/resources/toolkit/HMORN_VDWAnswers.pdf (accessed September 16, 2009).

¹⁶ See Diamond, et. al 2009, op cit, and Mostashari, F., M. Tripathi, M. Kendall. 2009. A Tale of Two Large Community Electronic Health Record Extension Projects. *Health Affairs* 28(2) 345-356.

¹⁷ See <http://isds.cirg.washington.edu/distribute/index.php> (accessed September 16, 2009).

¹⁸ See “assessment criteria” on pages 12-13 in: Carol Diamond. Testimony on Private Health Records: Privacy Implications of the Federal Government’s Health Information Technology Initiative before the Senate Committee on Homeland Security and Governmental Affairs, Subcommittee on Oversight of Government Management, the Federal Workforce, and the District of Columbia. Date: 2/1/2007. http://hsgac.senate.gov/public/index.cfm?FuseAction=Files.View&FileStore_id=06a0802a-dcda-4643-a888-46f589db93ba (Accessed September 16, 2009).