

Health Data Stewardship: What, Why, Who, How

An NCVHS Primer

—EXECUTIVE SUMMARY—

What is health data stewardship?

Health data stewardship is a responsibility, guided by principles and practices, to ensure the knowledgeable and appropriate use of data derived from individuals' personal health information. These uses include (but are not limited to) data collection, viewing, storage, exchange, aggregation, and analysis. A central concept of data stewardship is accountability, which resides in a named data steward with formal responsibility for assuring appropriate use of health data, and with liability for inappropriate use. Health data stewardship supports the benefits to society of using individuals' personal health information to improve understanding of health and health care while at the same time respecting individuals' privacy and confidentiality.

Why is health data stewardship important?

Health data stewardship has taken on great practical urgency because of the increase in availability of electronic health data; growing recognition of the value of electronic data in improving health care and population health; the acceleration in the use of information and communication technology; and awareness of the potential risks associated with incorrect or inappropriate uses of health data.

Who should practice data stewardship?

Everyone who collects, views, stores, exchanges, aggregates, analyzes, and/or uses electronic health data should practice data stewardship. This includes health care organizations, clinicians, payers, information exchanges, vendors, the quality improvement community, health statistics agencies, researchers, and caregivers.

What specific practices are suggested for data stewardship?

Thought leaders have identified essential practices and principles for health data stewardship. They include transparency about use; identification of the purpose for data use; participation of individuals; security safeguards and controls; de-identification (when relevant); data quality, including integrity, accuracy, timeliness, and completeness; limits on use, disclosure, and retention; oversight of data uses; accountability; and enforcement and remedies.

Where can I find more information on health data stewardship?

The section on major themes (p. 5-6) and Appendix A summarize leading documents on health data stewardship and related topics and provide reference information.

Preface

Today's health information technology offers powerful opportunities for benefit as well as potential for harm because of the ease it affords for the aggregation, use, and reuse of personal health data. To realize the potential benefits associated with data use without increasing the risk of harm, it is critically important that all persons and organizations with access to health data in health care, health statistics, research, and policy adhere to a set of agreed-upon data stewardship principles and practices. The principles are designed both to protect the rights and privacy of the persons whose data are involved, and to assure the quality and integrity of the data and their uses.

The National Committee on Vital and Health Statistics (NCVHS) is a Federal Advisory Committee that has advised the Department of Health and Human Services on health information policy since 1949. Current members are listed in Appendix B. In recent years, NCVHS has not only developed its own recommendations on data stewardship but also engaged in broader dialog on the subject with many other leaders and health policy makers. NCVHS created this primer to synthesize current thinking about how health data should be managed and to make this information widely available, along with references to the rich analytic and policy literature on the subject. The primer does not represent an endorsement of any positions except those of NCVHS.

This primer is intended for all those who need to understand data stewardship and how to practice it, including those who collect, view, store, exchange, aggregate, analyze, and/or use personal health data. The intended audience encompasses a broad spectrum that extends from visionaries establishing new directions in the health and health care industry to students entering relevant fields. Patients and research subjects who want to know how their personal health data should be managed also may find the primer useful.

Harry Reynolds
Chairman, NCVHS
September 2009



National Committee on Vital and Health Statistics (NCVHS)
U.S. Department of Health and Human Services, Hyattsville, MD,
September 2009.

Authors: Susan Baird Kanaan and Justine M. Carr, M.D.

NCVHS is the public advisory body on health information policy to the Secretary of Health and Human Services. NCVHS members in 2009 are listed on the back page.

ncvhs.hhs.gov

Health Data Stewardship: What, Why, Who, How

An NCVHS Primer

Health information technology has the potential to improve the quality and affordability of health care, reduce medical errors, reduce health disparities, improve population health, increase prevention and coordination with community resources, and improve the continuity of care across health care settings. Yet the same technology that makes possible previously unimagined achievements also presents potential risks. These risks include not just outright misuses of personal information for private gain or to cause harm, but also the use or propagation of inaccurate or incomplete information or its transmission at the wrong time, to the wrong destination, or in the wrong transmission mode.

What can be done to ensure that personal health information is put to appropriate and beneficial uses and that misuses are prevented? *Data stewardship* is an important part of the answer. The fundamental tenet of data stewardship might be expressed as *Do unto the data of others as you would have others do unto yours*. While this may sound simple, how can it be accomplished?

This primer summarizes the major principles, practices, and resources associated with health data stewardship. It also points to resources that further elucidate how to use data responsibly. This information is important for clinicians, researchers, and policymakers. It is also relevant for patients and research subjects, to inform them about the appropriate management of their personal health data.

Figure 1 illustrates examples of sources, users, and uses of personal health data today.

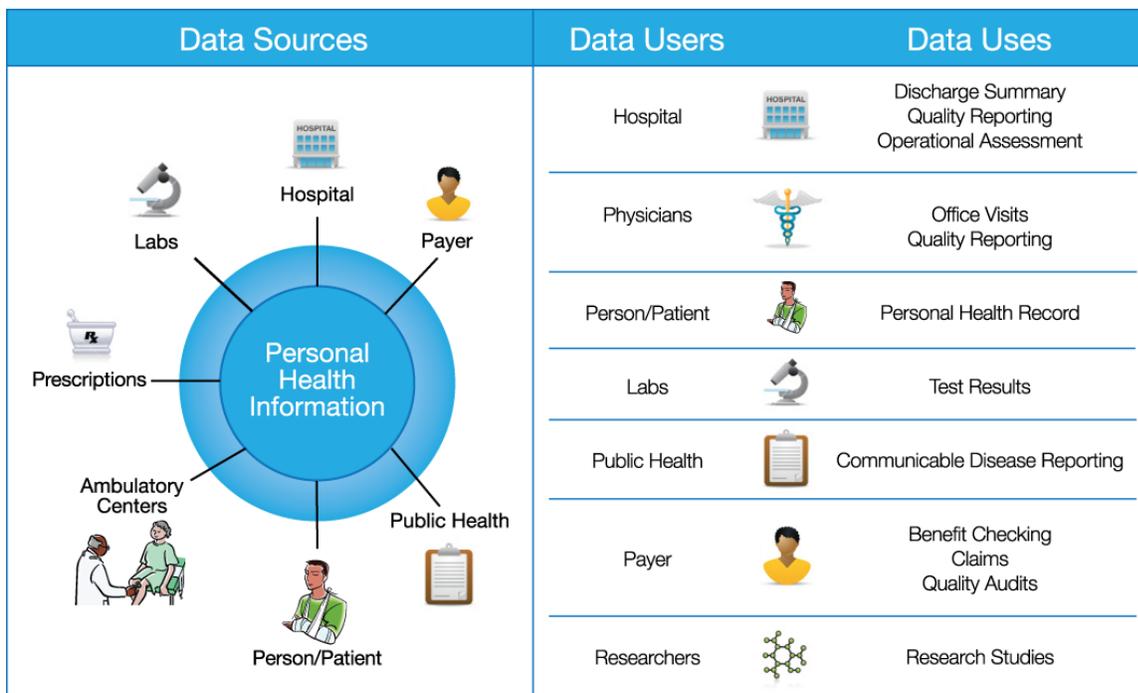


Figure 1. Data sources, users, and uses. (Modified from *Foundation for eHealth Initiative, November 2004.*)

What Is Health Data Stewardship, and Why Is It Important?

The term *health data stewardship* refers to the responsibility of ensuring the appropriate use of personal health data. The purpose of stewardship is to realize the greatest possible benefit from the effective and appropriate use of data while minimizing the risk of harm. This is accomplished by strengthening the chain of trust and accountability through principles and practices that apply to collecting, viewing, storing, exchanging, aggregating, analyzing, and using data from individuals' personal health information. Data stewardship approaches and principles apply to the uses of both identified and de-identified data. Specific issues, solutions, and data flow may differ from setting to setting. However, a consistent component should be the presence of a *data steward*—an overseer or guardian with final authority and accountability for appropriate use of health data. This may be a formal position or an assigned accountability.

Given the potential for both benefits and risks from the sharing and reuse of health data, consensus is building in the United States and beyond that everyone who touches individual health data for any purpose—health care, research, quality assessment, population health monitoring, payment, and more—must understand and practice data stewardship. Future assessments of health data usage are likely to look for stewardship mechanisms; and knowledgeable patients, families, and research subjects will expect these mechanisms to be in place and operating continuously.

Key Principles and Practices of Health Data Stewardship

Several leading organizations and agencies, including the National Committee on Vital and Health Statistics (NCVHS), have been engaged in a robust public discussion over the last several years to refine the concept of data stewardship and develop the practical applications. Appendix A summarizes key documents and references that have emerged through that process. The principles and practices outlined in these documents can be grouped into four categories: individual rights; the responsibilities of the data steward; needed security safeguards and controls; and the accountability, enforcement, and remedies that should be in place.

1. Individual rights

- Access for an individual to his/her own health data
 - An individual has the right to know what is in his or her health records.*
- Opportunity to correct one's own data
 - Improving the accuracy of the data enhances health care and the utility of the data.*
- Transparency for the individual about the use(s) of his/her data
 - Making an individual aware of what information exists and how it will be used builds trust. The individual should be notified in advance of policies, procedures, and technology, including what information will be shared under what circumstances.*
- Individual participation and consent for the use of the data
 - Allowing an individual to make decisions about electronic exchange of his or her data also builds trust. The degree of choice may vary with the type of infor-*

mation, the purpose of the exchange, applicable law, population health needs, and other factors.

- Education

In order to participate meaningfully, consumers should be educated about the principles and practices governing appropriate use of their health data and the potential benefits (and risks) of health data use for the public.

- Other rights to privacy of personal health information as set forth in state and federal laws and regulations

2. Responsibilities of the health data steward

Data stewardship requires the presence of a data steward in every organization that handles health data. This is either a formal position or an assigned accountability with responsibility for the following areas:

- Adherence to an appropriately determined set of privacy and confidentiality principles and practices
- Appropriate use of information from the standpoint of good statistical practices (such as by not implying cause and effect when the data only point to correlation)
- Limits on use, disclosure and retention
- Identification of the purpose for a specific use of the data
- Application of “minimum necessary” principles
- Verification of receipt by the correct recipient, wherever possible
- Data de-identification (HIPAA-defined and beyond)
- Data quality, including integrity, accuracy, timeliness, and completeness

3. Security safeguards and controls

To ensure the confidentiality, integrity, and availability of health data, data stewardship requires the implementation of administrative, technical and physical safeguards to protect the information and minimize the risks of unauthorized or inappropriate access, use, or disclosure.

4. Accountability, enforcement, and remedies

Data stewardship requires policies that specify appropriate use and identify clear accountability. Mechanisms are needed for detection of failure to follow policy. When a failure occurs, consequences to the accountable party must be enforced, along with remediation for the individual whose data are involved.

Major Themes in an Evolving Discussion

Use of data and understanding of how to implement data stewardship principles are developing in parallel with information technology and its applications. The concept of data stewardship and

role of the data steward emerged initially in the information management arena, where the focus was on improving data quality for effective decision-making. In the health care arena, the quickening pace of information technology use has heightened awareness of the importance of data stewardship for personal health data. This section briefly reviews key themes and participants in the evolving discussion of health data stewardship.

The passage in 1996 of the *Health Insurance Portability and Accountability Act (HIPAA)* resulted in a substantial increase in the number of electronic health care claims. HIPAA led to the *Privacy Rule* of 2001, which required entities that submit electronic claims to safeguard patients' health information¹ by restricting access to it and seeking patient permission to disclose it in certain circumstances (outside treatment, payment or operations).

Starting in 2005, Connecting for Health, a public-private collaborative of the Markle Foundation, published *The Common Framework*, a collection of technical and policy documents to help health information networks share information while protecting privacy and allowing for local autonomy and innovation.^{2,3}

The quality improvement community has generated a significant literature through its thoughtful examination of ways to balance health care quality improvement efforts with protection of patients' privacy and other rights.⁴ The absence of consensus on a good strategy for de-identifying and aggregating data is a major issue for quality improvement, as it is for much public health work and research.

In recent years, these themes of privacy and security, quality improvement, and appropriate and effective data use converged in a series of inquiries into national policy and governance. One stimulus was a 2005 Institute of Medicine report on performance measurement that called for a National Quality Coordination Board.⁵ In that context, the AQA put forward a 2006 proposal for a "national health data stewardship entity" (NHDSE).⁶ Drawing on the AQA's work, the U.S. Agency for Healthcare Research and Quality (AHRQ) issued a Request for Information in mid-2007, asking for comments and inviting broad stakeholder discussion about the idea of establishing an NHDSE.⁷ AHRQ received 136 responses, which showed a relatively even split between those favoring and opposing the idea.⁸

The National Committee on Vital and Health Statistics also has helped advance the thinking about national health data stewardship policy. The Office of the National Coordinator (ONC) asked NCVHS to develop a stewardship framework for all uses of health data. The Committee submitted a report and recommendations to the Secretary in late 2007, based on input from 75 stakeholders and a large volume of written testimony.⁹ In the report, NCVHS stresses the need to strengthen the HIPAA Privacy Rule and Business Associate agreements. It also raises concerns about areas not regulated by HIPAA, including the definition and oversight of de-identified data and the uses of individually identifiable health data by non-HIPAA covered entities.

The American Medical Informatics Association (AMIA) also was working on data stewardship during the same period. AMIA convened a panel of experts in 2007 to consider a national framework for health data use, particularly aggregated data. The meeting led to publication of an AMIA white paper in late 2008.¹⁰ The white paper discusses the following stewardship principles for the management of health information: accountability, transparency, notice to patients, patient consent, permitted uses and disclosures, enforcement and remedies, plus several technical factors. The AMIA paper also provides useful references to 25 source documents.

In September 2008, the Center for Democracy and Technology (CDT) and the Health Privacy Project sponsored a full-day symposium with thought leaders on the topic of de-identified data. CDT subsequently released a white paper on the topic. (See Appendix A.)

The DHHS Office of the National Coordinator for Health Information Technology (ONC) released a Privacy and Security Framework for electronic health information exchange in December 2008.¹¹ Then in early 2009, the American Recovery and Reinvestment Act (ARRA) amended and extended HIPAA. The ARRA calls for the creation of a position of Chief Privacy Officer in ONC to advise the National Coordinator on issues of privacy, security, and health data stewardship. The Act also calls for a strengthening of the oversight and accountability of data use by the business associates of covered entities.

The matrix in Appendix A provides brief summaries of the analysis and recommendations of leading organizations concerning data stewardship and related topics. It is followed by notes and references on specific aspects of data stewardship, including several NCVHS reports and recommendations on related subjects.

Summary

There is wide agreement about the absolute necessity of making data stewardship a constant thread in all uses of individuals' health information for any and every purpose, including delivery of medical care. Data stewardship is the responsibility of everyone who collects, views, stores, exchanges, aggregates, analyzes, and/or uses electronic health data. The principles of data stewardship are now well established; the work of translating them into practice will continue to evolve as the urgency of data stewardship grows ever greater in the future.

Appendix A. Resources on Health Data Stewardship, 1996-2009

Organization	Title, date	Description
US Congress and DHHS	HIPAA Privacy Rule (2001) 2009 American Recovery and Re-investment Act (ARRA, 2009)	The 2001 HIPAA Privacy Rule provides an initial set of rules for protecting privacy in the context of using and exchanging health data. In the 2009 ARRA, Congress expands HIPAA's definition of covered entities and calls for ONC to appoint a chief privacy officer to advise on privacy, security, and data stewardship of electronic health information.
Connecting for Health	Common Framework (2005 & ff.)	Helps health information networks to share information among their members and nationwide while protecting privacy and allowing for local autonomy and innovation. Consists of a set of 17 mutually-reinforcing technical documents and specifications, testing interfaces, code, privacy and security policies, and model contract language. Developed by experts in information technology, health privacy law, and policy; has been tested since mid-2005.
Institute of Medicine	Report, <i>Performance Measurement: Accelerating Improvement</i> (12/05)	Acknowledges the progress in developing performance measures, but expresses concern about the duplication, gaps, and multiple independent reporting systems that are resulting from independent voluntary initiatives. Endorses a starter set of evidence-based measures acknowledged by major stakeholder groups. Also recommends formation of an independent National Quality Coordination Board within DHHS to set goals, designate or develop measures, ensure an adequate data repository system, identify a research agenda, and evaluate.
AQA Data Aggregation Workshop	Statement, National Health Data Stewardship Entity (4/06)	In response to the IOM report on performance measurement, recommends that "a public/private entity have primary responsibility of setting uniform operating rules and standards for sharing and aggregating quality and efficiency data." Proposes a mission, precepts, a scope of work, and next steps for such an entity. Also describes an AQA pilot project in six named sites "to explore approaches that measure individual physician, group, and system performance; aggregate data...; and generate reports...." Promises by the end of 2006 to make recommendations about what entities could fulfill the work described in this document.

Organization	Title, date	Description
Agency for Healthcare Research and Quality (AHRQ)	Request for Information on a National Health Data Stewardship Entity (6/4/07)	Acknowledges the growing demand for health care data and the question of responsibility for safeguarding the data, and proposes formation of “a public-private national health care data stewardship organization (NHDSE). Using the AQA’s descriptions, asks for stakeholder responses to 25 questions about the need for such an entity; roles, responsibilities, relationships; challenges and risks; key stakeholders; governance models; funding; priorities; and more. AHRQ received 136 responses, 24 from organizations and 112 from individuals. They were about equally divided in favor of and against the idea of a national health data stewardship entity. The responses in their entirety and a summary of responses are posted on the AHRQ Website.
National Committee on Vital and Health Statistics (NCVHS)	Letter and Report, <i>Enhanced Protections for Uses of Health Data: A Stewardship Framework for “Secondary Uses” of Electronically Collected and Transmitted Health Data</i> (12/07) <i>Summary for Policy Makers</i> (April 2008)	Developed at the request of ONC, based on testimony and comments from 75 stakeholders; focused on national policy. Offers guiding principles and “considerations” for data stewardship; summarizes the major themes in the testimony, including HIPAA applications and limitations; and makes recommendations in 9 areas including chain of trust within HIPAA, transparency, individual participation and control, de-identification, security, data quality and integrity, oversight for specific uses, transition to the NHIN, and the need for additional privacy protections. (Also see other relevant NCVHS reports, listed below. ¹²)
American Medical Informatics Association (AMIA)	White paper, <i>Advancing the Framework: Use of Health Data—A Report of a Working Conference</i> , by M. Bloomrosen and D. Detmer (Nov/Dec 08)	Focuses on the need for a national framework for health data use to guide national policy and practice. Summarizes the discussions of expert panels in 2006 and 2007 on data stewardship as “a key building block in national framework.” Urges stakeholders “to deepen their understanding of data stewardship principles” and identifies 7 principles for managing health information. Discusses terminology issues and the genesis of AMIA’s emphasis on stewardship, with particular attention to issues related to the reuse of data for multiple purposes. Reviews the contributions of key policymaking groups and stakeholders in the U.S. and Europe. Proposes a taxonomy on dimensions of data use and a framework tool to assess the status of data uses.

<p>American Health Information Management Association (AHIMA)</p>	<p>Statement on Data Stewardship November 2008</p>	<p>Sets forth four major objectives and action steps. Objectives include: 1. Identify and authorize a partnership of stakeholders that will establish national coordination in a transparent process with other entities to set uniform rules and the requirements for principles of data stewardship to achieve uniformity and consistency of data. 2. Increase transparency in the operation and management of data access, use, and control. Individuals should have the opportunity to be informed of all potential uses of their health data. 3. Establish coordinated objectives at the federal, state, and local levels to improve data collection and use efficiencies and reduce the burden of cost. 4. Enable an effective method for the standardized release of data to approved agencies and organizations as permitted by law.</p>
<p>Office of the National Coordinator for Health Information Technology (ONC)</p>	<p>White Paper, <i>Nationwide Privacy and Security Framework for Electronic Exchange of Individually Identifiable Health Information</i> (December 2008)</p>	<p>Principles for a “single consistent approach to address privacy and security challenges related to electronic health information exchange through a network for all persons, regardless of the legal framework that may apply to a particular organization.” Elaborates eight principles addressing: individual access; correction; openness and transparency; individual choice; collection, use and disclosure limitation; data quality and integrity; safeguards; and accountability.</p>
<p>Center for Democracy and Technology (CDT)</p>	<p>White Paper, <i>Encouraging the Use of, and Rethinking Protections for De-Identified (and “Anonymized”) Health Data</i>. (June 2009)</p>	<p>Addresses the limitations of the HIPAA Privacy Rule in the face of technological innovations. Argues in favor of “strengthening the current de-identification standard, setting different levels of anonymization for different uses of data, requiring greater accountability for re-identification, and enforcing existing policies that are designed to place limits on the amount of data that can be collected and retained.”</p>

Appendix B. NCVHS Members and Executive Staff*

Harry Reynolds, Chair
Blue Cross Blue Shield of North Carolina

Jeffrey S. Blair, MBA
Lovelace Clinic Foundation

Justine M. Carr, MD
Caritas Christi Healthcare

Leslie Pickering Francis, JD, PhD
University of Utah

Larry A. Green, MD
University of Colorado

Mark C. Hornbrook, PhD
Kaiser Permanente Northwest

John P. Houston, JD
University of Pittsburgh School of Medicine

Garland Land, MPH
Natl. Assn. for Public Health Statistics & Information Systems

Carol J. McCall, FSA, MAAA
Humana

Blackford Middleton, MD, MPH, MSc
Partners Healthcare

J. Marc Overhage, MD, PhD
Regenstrief Institute, Inc.

Sallie Milam, JD
West Virginia State Government

William J. Scanlon, PhD
Health Policy Research & Development

Donald M. Steinwachs, PhD
The Johns Hopkins University

Walter G. Suarez, MD, MPH
Institute for HIPAA/HIT Education & Research

Paul C. Tang, MD
Palo Alto Medical Foundation

Judith Warren, PhD, RN
University of Kansas

Executive Staff

James Scanlon
DHHS Office of the Asst. Secretary for Planning & Evaluation

Marjorie S. Greenberg
National Center for Health Statistics/CDC

**As of July 31, 2009*

NOTES

¹ Also referred to as protected health information or PHI.

² <http://www.connectingforhealth.org/commonframework/>

³ See also the eHealth Code of Ethics, a framework developed by the Internet Healthcare Coalition for an e-Health ethics summit held in Washington, DC, Jan 31 - Feb 2, 2000. *Journal of Medical Internet Research*, 2000. <http://www.jmir.org/2000/2/e9/>

⁴ Notable recent examples:

Health Care Quality Improvement: Ethical and Regulatory Issues, Jennings B, Baily MA, Bottrell M, and Lynn J, eds. The Hastings Center, Garrison NY, 2007. <www.thehastingscenter.org>

On the Ethics of Using QI Methods to Improve Health Care Quality and Safety, Jennings B, Baily MA, Bottrell M, and Lynn J, eds. The Hastings Center, Garrison NY, 2007.

5 IOM Performance Measurement report brief:

http://www.iom.edu/Object.File/Master/35/324/PerformanceMeasurement_Brief.pdf

Full report available at <http://www.iom.edu/report.asp?id=31310>

6 www.aqaalliance.org/files/HealthDataSteward-July06.doc

The AQA, a public-private body, uses only the initials as its name. It was originally the Ambulatory Care Quality Alliance.

7 Federal Register Vol. 72, No. 106 (Monday, June 4, 2007). p. 30803-30805

8 The responses are posted and summarized on its Website. See for example the responses of the American Health Information Management Association (AHIMA) and the National Association of Health Data Organizations (NAHDO):

AHIMA (accessed 4/17/09):

http://www.ahima.org/dc/positions/documents/AHIMA_PositionStatementDataStewardship_001.pdf

NAHDO: (posted on AHRQ Website)

9 NCVHS Report to the Secretary of HHS on Enhanced Protections for Uses of Health Data:

<http://www.ncvhs.hhs.gov/071221lt.pdf>

10 J Am Med Inform Assoc. 2008;15:715–722. DOI 10.1197/jamia.M2905.

11 Office of the National Coordinator for Health Information Technology, USDHHS, *Nationwide Privacy and Security Framework for Electronic Exchange of Individually Identifiable Health Information*, December 15, 2008.

Also of interest: The International Security Trust and Privacy Alliance (ISTPA) has compiled the privacy frameworks of several countries. <http://www.istpa.org/>

12 Related NCVHS reports (posted on NCVHS Website, <http://www.ncvhs.hhs.gov/>):

- *Recommendations on Initial Functional Requirements for an NHIN*, October 30, 2006
- *Recommendations to the Secretary of HHS Regarding Privacy and Confidentiality in the NHIN*, June 22, 2006
- *Report to the Secretary of HHS on Personal Health Record (PHR) Systems*, September 9, 2005
- *Recommendations on Privacy and Confidentiality, 2006-2008*, May 2009.