CAMBRIDGE

Cambridge University Press
978-1-107-02087-0 - Privacy, Confidentiality, and Health Research
William W. Lowrance
Excerpt
More information

1 Introduction

This book examines the ways society can best promote the use of personal information for health research, and *at the same time* protect informational privacy and confidentiality. Much is at stake.

Occasionally skeptics frame the overall issue as privacy versus research, casting research as mainly serving the intellectual curiosity and self-advancement of scientists. This is shortsighted. Like everyone else, scientists hope for fame and fortune, or at least a good reputation and reasonable income. And having an intense sense of curiosity is a requisite for being a scientist. But the work is awfully demanding, and true discovery moments are rare. Researchers work hard to solve health-related problems, as do the institutions that host and support them. To the extent that there is a balancing in properly established research, the issue is privacy versus the advancement of health through research.

Throughout, the book reflects the author's conviction that both health research and privacy protection are public-interest causes. The realizations that lead to such a conviction are familiar and the logic is commonsensical, but the reasoning deserves to be summarized at the outset.

Health research as a public-interest cause

From conception onward, everyone is exposed to myriad health risks. Many of the risks are either attracted, caused, intensified, or transmitted by human activity, and many can be mitigated or compensated for by human activity. Society at every scale of organization, whether village, city, provincial, national, or supranational, accords high priority to the reduction of health threats, the promotion of health, and the provision of health care. What good health means for any person at any stage of life is relative to his or her risks and resources. But despite the difficulty in defining it precisely, health is universally valued as a core human need and a condition for living life in dignity.

Nowadays we live thoroughly interconnected lives in which illness risks and costs are widely shared, although not equally or fairly. The burdens of

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-107-02087-0 - Privacy, Confidentiality, and Health Research
William W. Lowrance
Excerpt
More information

illness and disability suffered in resource-poor and unstable countries are burdens on the world in the toughest economic, social, and security senses, as well as in high moral senses. The poor of the world share vectors of contagion with the prosperous, and the prosperous share vectors of unhealthy lifestyles with the less prosperous. Pathogens can propagate at the speed of airline travel or food transport. Epidemics and natural disasters strike with no respect for political boundaries. We are vulnerable together.

None of us can know what diseases and disabilities we or our families or friends will fall victim to as our lives go on – and when those afflictions occur, we tend to hope fervently that a great many people's experiences have been studied in depth and have led to the development of proven effective diagnostic and curative, or at least palliative, techniques and products for coping with them. As we all stand to benefit from the knowledge commons, we should all contribute to the knowledge commons to help others, including people we will never know. Participation in research is a moral opportunity.

Relating to all this are costs. In all developed countries and increasingly in the developing countries, health care is an industrial-scale activity, the costs of which are paid for at least partly, and in many cases almost entirely, by the state. These costs, and cost-effective provision of prevention and care, are and always will be of crushing concern to states, and so states depend on research to understand causes, evaluate what works best and is most cost-effective, innovate, and improve. Although private sector healthcare providers and insurers have different financial considerations, they too must worry about costs and quality, and they too depend on research for understanding and improvement.

In many ways, it is the poor of the world who stand to gain the most from research, relative to their health and socioeconomic burdens – not only as regards such infectious scourges as tuberculosis, malaria, AIDS, schistosomiasis, leprosy, and river blindness, but also as regards such draining noncommunicable afflictions as infant malnutrition, diabetes, cardiovascular diseases, and cancer.

An encouraging thing is that the fruits of health research, whether basic knowledge, practice guidance, techniques, or products, tend to propagate extraordinarily efficiently and widely. There will always be budgetary and cultural limitations to the application of research results and new healthcare technologies, even in the wealthiest communities. But no other universal progressive endeavor comes close to the critical screening and efficiency with which scientific and medical ideas and information are generated, distributed through journals, the web, and conferences, evaluated for quality and relevance, and translated into practice. Knowledge

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-107-02087-0 - Privacy, Confidentiality, and Health Research
William W. Lowrance
Excerpt
More information

developed anywhere about mastitis, burns, glaucoma, migraine, organic solvent toxicity, wood dust allergies, hospital hygiene, and countless other matters can be put to use around the world.

Health research, then, is of great public-interest importance, and this is underscored by the fact that much of it is financially supported or conducted by government bodies, by organizations to which governments grant nonprofit tax-exempt status, and by international organizations in which governments participate. At many points this book will refer to the public-interest justification of research policies and practices that serve the common human good, and to the contributions to the common good that members of the public make by volunteering to participate in research or allowing data about themselves to be studied.

Privacy protection as a public-interest cause

Simply and profoundly: Privacy should be respected because people should be respected. Despite its personal and contextual relativity and the difficulty in defining it, as will be discussed in Chapter 3, privacy is widely valued as a core human need and a condition for living life in dignity.

For the research enterprise – researchers and all the institutions that support, regulate, govern, and disseminate the results of research – attending assiduously to privacy and the relations of confidentiality that serve it is essential to earning the trust that encourages the public to become involved in research, be candid and generous in answering questions or allowing information about themselves to be used, be unselfish in providing biospecimens, and stay involved in projects as long as is scientifically useful.

If these matters are not carefully attended to, the people to whom data relate may be offended on principle by a violation of confidentiality promises. They may resent intrusions following wrongful disclosure, such as improper clinical trial recruitment approaches or unwanted disease-targeted marketing. They may be, or fear that they may be, exposed to embarrassment, defamation, stigmatization, harassment, extortion, identity theft, or financial fraud, or denial of access to health or life insurance, employment, job promotion, or loans. And they and their sympathizers may drop their goodwill toward the institutions or research programs. Offending researchers or institutions may suffer negative publicity, litigation, or financial losses. A breach of confidence can cost months or years of remediating effort by university administrators or clinical or corporate managers and their lawyers and public affairs

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-107-02087-0 - Privacy, Confidentiality, and Health Research
William W. Lowrance
Excerpt
More information

staff, and a sullied reputation can be a burden for a long time. A research team may be denied access to data or data-collection opportunities, and even a whole line of research may suffer by association.

Health data vulnerability

For calibration one might well wonder how many privacy intrusions have occurred and what harms people have incurred. Thousands of unwarranted disclosures, losses, and thefts of health *care* data, and some successful hacking attacks, have occurred in a number of countries; however, relatively few direct material harms to data-subjects have been documented and not many lawsuits have been pursued in the courts (although some incidents may have been pursued out of court, off the public record). Probably credit card abuse and identity theft have been the main harms. Most of the violations that have been confirmed have resulted from careless, incompetent, or illegal actions of the sorts that professional care and organizational safeguards should be able to prevent. But the scale and sensitivity of most healthcare data systems, and the fact that they carry patient payment details, will always make them a potential target of attack.[1]

Health *research* data have mostly been spared so far. This may be because of precaution, or luck, or because healthcare records are more tempting targets for intrusion in that they tend to be easier to pry into than most research databases, reveal more readily comprehensible health information, and carry more exploitable financial details. None of this justifies being lackadaisical: Any of the dread events recited above could be incurred by health research data. Several research databases have been lost, stolen, or hacked into in the last few years. And health research is changing in ways that are creating new vulnerabilities.

The challenges

Ethical, legislative, regulatory, technical, administrative, and day-to-day operational adjustments are being made everywhere to try to cope with the issues raised in this book. It is important to have the assortment of challenges in mind from the beginning, partly because most of the issues

[1] For documentation of incidents, see Privacy Rights Clearinghouse, "Chronology of data breaches, 2005–present," healthcare subset: www.privacyrights.org/data-breach. US Department of Health and Human Services, healthcare data security breach notifications: www.hhs.gov/ocr/privacy/hipaa/administrative/breachnotificationrule/breachtool.html.

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-107-02087-0 - Privacy, Confidentiality, and Health Research
William W. Lowrance
Excerpt
More information

manifest themselves as issue-clusters and have to be dealt with in concert. For concision they are set out here in bulleted lists, with a few examples.

New scientific opportunities. We are in an era of unprecedented scientific opportunity, and information technologies are providing reach, speed, memory capacity, and search flexibility as never before. For the health sciences, it's an exhilarating time. Among the many advances, research is benefiting from:

- generation of entirely new kinds of data (real-time digital images of brain activity, genomic and other -omic data, human microbiome data, i.e., detailed ecological characterization of the trillions of creatures that colonize everyone's body . . .);
- development of new modes of data capture, storage, and transmission (networked electronic health records, wearable biosensors, Internet-mediated gathering of input from dispersed research participants . . .);
- reclassification of many health problems based on deeper understanding of causal factors instead of, or in addition to, their clinical appearance;
- integration of social and behavioral research with biomedical and health services research (large life-course and multigenerational studies, behavioral toxicology, mapping of health and healthcare disparities across populations . . .);
- amassing of ever-growing mountains of data in healthcare administrative databases, payment databases, and disease and condition registries, most of which can be accessed, under various conditions, for research;
- construction of large-scale research platforms (health-data linkage systems, research biobanks, genome-wide association databases, clinical trial networks, social research data archives . . .).

Chronic policy problems becoming worse. Serious privacy and confidentiality impediments continue to hamper research, however, notably among them:

- uncertainties and disagreements around the formal construal of "personal data" or "personally identifiable information," and the notion of identifiability generally;
- debate about the ethical and legal appropriateness of consent to complicated research and the sharing of data via research platforms, at least as consent tends to be applied currently;
- dispute over the acceptability of broad consent to unspecified, perhaps indeed unspecifiable, future research;
- lack of clarity about how to deal with privacy and confidentiality implications for relatives of people involved in research;

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-107-02087-0 - Privacy, Confidentiality, and Health Research
William W. Lowrance
Excerpt
More information

- public – and researcher – apprehension about the legal power of researchers to resist forcible access to sensitive research data by the police, courts, banks, health or life insurers, or other external parties;
- inconsistencies and redundancies among the multitude of laws, regulations, and guidelines, many of them vague, duplicative, outdated, or just not relevant for contemporary research;
- onerous, inefficient, and costly procedural requirements for complying with all the laws, regulations, and guidelines.

Exacerbations of scale. Many challenges are growing in complexity as scale increases, as with:

- in many situations, reduction of direct control by patients over how data are used, and reduction of direct control by physicians and medical institutions that provide access to patient data and biospecimens for research;
- increased risks to privacy from the growth of data and biospecimen holdings, pooling and interlinking of data-sets, and the generally increasing geographic and institutional dispersing of data and biospecimens;
- decentralization and outsourcing of much data storage and analysis, potentially diffusing accountability.

Recently arising or intensifying concerns. Among the newer concerns are ones having to do with:

- the use of kinds of data and biospecimens that in many jurisdictions have been ethically or legally off-limits for research until fairly recently (residual newborn blood screening specimens, abortion registry data, stem cells . . .);
- the availability of some entirely new kinds of data that can be useful for research but that can carry clues to individuals' identities, habits, and connections (geospatial tracing of people's movements and exposures, broadcasting of personal details via online social networking . . .);
- privacy risks collateral to the rapid and extensive data sharing that is being widely promoted;
- the security of computerized data generally;
- genomic science's headlong assembly line decoding of personal origin-and-fate factors that society is not well prepared to interpret or make judicious use of;
- genotype data as a potentially identifying or tracing tag when linked with otherwise non-identified data;
- the possibility of aggressive demands for access to research data under freedom of information or anti-terrorism laws;
- threats to privacy as torrents of data are transmitted across borders as research continues to globalize, in a world that has hardly globalized privacy protections.

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-107-02087-0 - Privacy, Confidentiality, and Health Research
William W. Lowrance
Excerpt
More information

2 Data, biospecimens, and research

A few essential notions

Clarity of conception and vocabulary is essential when discussing the subjects of this book. Most of the following notions may seem commonplace and hardly in need of definition, but most of them become the subject of scholarly debate or close legal scrutiny from time to time. (Unless otherwise noted, the definitions are ones consistent with common usage but are the author's attempt to be precise and clear for the purposes of this book.)

Data are records of observations or actions, or, stated slightly more formally, patterns of symbols that stand for observed values or actions. They may be instrument readings, x-ray or scanner images, voice recordings, family lineage charts, interview responses, hospital billing records, or countless other results of looking, asking, listening, measuring, recording, or analyzing. In research, almost all data are now handled in digital form, even if this requires transcription or translation from nondigital formats. This greatly facilitates computerized analysis, of course. It also allows the distribution of data from site to site at close to the speed of light and at very low cost, which can be either wonderful or troublesome, depending on how the data are managed and used.

Information is data set within an interpretive context to generate meaning. Often information and data are taken to mean the same thing, but there is some advantage in using them differently in the context of research. Raw numbers, graphs, images, or long strings of digital bits – data – "mean" nothing until they are understood as representing hormone level, biopsy photomicrograph, quality-of-life score, model number of an implanted device, genetic sequence, cost of surgical episode, or whatever, and the sampling scheme, data collecting circumstances, observational system, descriptive scale, and framework of scientific or clinical understanding are taken into account. Data quality is always a concern, of course, as is the quality of the methodological and ethical documentation (*metadata*). Obviously, data can be incorrect, and information can be false.

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-107-02087-0 - Privacy, Confidentiality, and Health Research
William W. Lowrance
Excerpt
More information

Knowledge, as it is taken to mean when research is defined in regulations as "the pursuit of generalizable knowledge," can be thought of as widely accepted understanding, based on verified information, compatible with other knowledge, and perhaps proven useful in practical experience.

Data-subjects are the persons whom data are about, the people to whom data pertain. In the UK Data Protection Act, for example, "data subject means an individual who is the subject of personal data," with personal data defined immediately afterward. (Just to indicate how quickly such a definition can become contested: What about an easily contactable but noninvolved relative of a research volunteer about whom the study results hold implications? Has she or he become a data-subject de facto?)

Personal data, or individually or personally identifiable data, are data that are about real persons or that can be related to real persons by deduction from partial descriptors or linking with other data. How to formally distinguish data that are somehow person-related from those that are not, though, is a much more subtle matter than one might assume it to be. Identifiability and data-subject status will be discussed in Chapter 7.

(A note to readers: Whenever this book mentions either "data-subjects," "subjects," or "research participants," the alternatives should be understood as appropriate. Usually not much distinction is made, but "participants" suggests more aware and active involvement than "data-subject" does, and one can be a data-subject without being aware of it.)

Data handling, in ethical or legal senses relating to privacy, has to be construed comprehensively, in order to include any action that provides an opportunity to take the data into knowledge or affect them. Thus it must include such actions as collecting, receiving, holding, examining, analyzing, linking, altering, transferring, archiving, and destroying.

Data processing is the term used in the EU countries and many others for data handling. It connotes far more than the outmoded sense of rote data entry and analysis. The EU Data Protection Directive defines processing as "any operation or set of operations which is performed upon personal data, whether or not by automatic means, such as collection, recording, organization, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, blocking, erasure or destruction" (Article 2b).

Data disclosure, as it will be used in this book, means the divulging or transfer of, or provision of access to, data. It should not be confused with some statisticians' usage of disclosure to mean the revealing of

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-107-02087-0 - Privacy, Confidentiality, and Health Research
William W. Lowrance
Excerpt
More information

data-subjects' identities; the passing-on of data may or may not lead to the recipient's learning who the data are about. Whether a recipient actually looks at the data, does anything with them, or retains or destroys them is usually considered irrelevant as to whether disclosure has occurred. Any showing or passing along of data, whether authorized or unauthorized, whether careful, or careless, or malicious, whether orally, or via paper or electronic media, can amount to disclosure. (An exception might be if data are sent in error and the recipient proves, or at least reliably attests, that the data were not looked at or copied.) Often the act of disclosure in itself is regulated by ethics or law, regardless of what happens once the data are in the hands of recipients.

Data sharing means the deliberate provision of access to data for use by others. This is a very important activity in research, as Chapter 10 will discuss.

Secondary use of data has long meant the use of existing data for a purpose different from the originally declared purposes. Although the observational techniques, choice of variables, and data quality can't be controlled the way they can in prospective studies, because studies of existing data can analyze (messy) real experience, they can help in understanding and improving (messy) real experience. Much constructive research depends on using existing data, such as clinical data. If the purposes, users, or auspices are different from the original ones, the implications for the data-subjects' privacy must be taken into account and compensating actions taken as necessary.

Differentiating secondary from primary uses can require judgment. This may hinge, for example, on how different the use is from the initial use, or on whether research is viewed under law or policy as being integral to the work of a healthcare system that is the source of the data. Some organizations, wanting to emphasize the integrated nature of their activities and make it clear that research is part of their mission, prefer to speak of "additional" instead of "secondary" use or to make little distinction, a stance that surely will become more prevalent, at least within large healthcare systems, as time goes on. In many situations, "secondary" can still be useful to connote that a formal threshold – such as approval by an ethics or other governance body, or de-identification of the data – has to be surmounted before the data can be used beyond the specified primary purposes or used by researchers outside the custodial circle. What is important is not the rubric but clarity about the nature of the use, the possible risks associated with the use, and requirements relating to the use.

Three important derivative uses of data-sets that not everyone thinks of right away are statistical case-controlling, searching for potential

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-107-02087-0 - Privacy, Confidentiality, and Health Research
William W. Lowrance
Excerpt
More information

candidates for research, and development and testing of analytic or privacy-protection methods.

The e-health revolution

e-health – the use of informatics in collecting and managing health-related data – takes many forms and can perform many functions, from medical recordkeeping, to practice management and procedure scheduling, to online prescription ordering (e-pharmacy), to automatic bedside or home collection and transmission of patient data, to the provision of clinical decision guidance based on accumulated medical experience, to the provision of care at a distance (telemedicine), to billing and reimbursing, to amassing research databases and facilitating the kinds of database research discussed throughout this book. This set of transformations in health care and research is now fully underway, even if some elements are developing faster than others.[1]

Electronic health records (EHRs) are comprehensive digital patient records carrying, or linked with, other health-related data. They are intended to be much more than just digital versions of conventional paper charts. The vision for EHRs includes their accumulating information from a person's health and healthcare experience over a long term, ideally throughout life, and being linkable with pharmacy dispensing data, biospecimens, medical images, family health and reproductive history, genomic data, disease and other registries, and other data. The challenge is to develop them so that they can manage the enormous variety of data that are handled in health care and payment, function efficiently as networked or at least intercommunicating systems among diverse settings, make information accessible at authorized point of need for a great many points and needs, and do all of this securely. Whether and how EHR systems will accommodate patient input and choices is being explored; they will remain at base formal accounts of medical encounters.

[1] A European overview is empirica GmbH for the European Commission, Karl A. Stroetmann, Jörg Artmann, Veli N. Stroetmann, *et al.*, *European Countries on their Journey towards National eHealth Infrastructures* (2011): www.ehealth-strategies.eu/report/eHealth_Strategies_Final_Report_Web.pdf. A Canadian overview is Don Willison, Elaine Gibson, and Kim McGrail, "A roadmap to research uses of electronic health information," in Colleen M. Flood (ed.), *Data Data Everywhere: Access and Accountability?* (Montreal, Quebec and Kingston, Ontario: McGill-Queen's University Press, 2011), pp. 233–251. In the US, much e-health activity is being shaped and supported under the US Health Information Technology Economic and Clinical Health (HITECH) Act and can be followed via: http://healthit.hhs.gov/portal/server.pt/community/healthit_hhs_gov_home/1204.

© in this web service Cambridge University Press www.cambridge.org