Achieving a Nationwide Learning Health System

Charles P. Friedman,* Adam K. Wong,   David Blumenthal

Published 10 November 2010; Volume 2 Issue 57 57cm29

We outline the fundamental properties of a highly participatory rapid learning system that can be developed in part from meaningful use of electronic health records (EHRs). Future widespread adoption of EHRs will make increasing amounts of medical information available in computable form. Secured and trusted use of these data, beyond their original purpose of supporting the health care of individual patients, can speed the progression of knowledge from the laboratory bench to the patient's bedside and provide a cornerstone for health care reform.

According to conventional wisdom, 17 years elapse before a new element of validated clinical knowledge finds its way into routine clinical practice in the United States (1). Although there is undoubtedly considerable variance around this estimate, the latency between biomedical discovery and care implementation is clearly too great. A more efficient, effective, and safe health care system requires a more rapid progression of knowledge from the lab bench to the bedside. Adoption of health information technology and trusted "meaningful use" (2) of patient data can help speed this process. In this Commentary, we present our vision of a nationwide biomedical learning system and describe the key contributory roles of meaningful use and additional components required to move the United States in its entirety toward this critical goal.

THE POTENTIAL: HEALTH INFORMATION TECHNOLOGY ADOPTION AND MEANINGFUL USE

The American Recovery and Reinvestment Act of 2009 introduced the concept of meaningful use of health information technology to improve health care and population health across the United States and authorized the payment of incentives to eligible health professionals and hospitals that achieve meaningful use. Meaningful use requires adoption of certified electronic health records (EHRs), secure mobility of health information, and reporting of quality measures (3). As the United States progresses toward President Obama's goal that every American will benefit from an EHR, massive amounts of clinical

information will be stored in electronic form (4). At the same time, achievement of meaningful use of these EHRs will enable this clinical information to flow securely from the site where it was collected to a different location where the information has an authorized use. In practice settings that achieve meaningful use, the clinical information will be represented by using precisely defined standards that have been adopted for use throughout the United States. Standardized representations ensure that the meaning of clinical information is preserved as the data move to new locations.

The accumulation through EHR adoption of these computable, liquid, standardized data creates an enormous potential for the U.S. health system to conduct clinical and translational research, assess and improve the quality of health care, and survey the health of the public at speeds approaching real time. These goals can be achieved by moving data, on an as-needed basis, from the panoply of locations where they are collected to one or more investigative centers where they are aggregated and analyzed for a specific purpose. Rapid data aggregation enables the creation of large, scientifically valid samples that can then be used to draw powerful inferences about populations. When this process can happen routinely, with mechanisms in place to establish and maintain public trust that the process is secure and private, the nation will have substantially progressed toward establishing a so-called rapid learning health system (5–7).

Adoption and meaningful use of EHRs are necessary to establish a nationwide learning health system and to create a foundation for its construction. Therefore, federal resources that directly promote the adoption and meaningful use of EHRs also move the nation toward a learning system (8). However, although necessary, EHR adoption and meaningful use are not sufficient to achieve this goal; additional components are required to achieve our vision of a highly participatory biomedical learning system in the United States (Fig. 1).
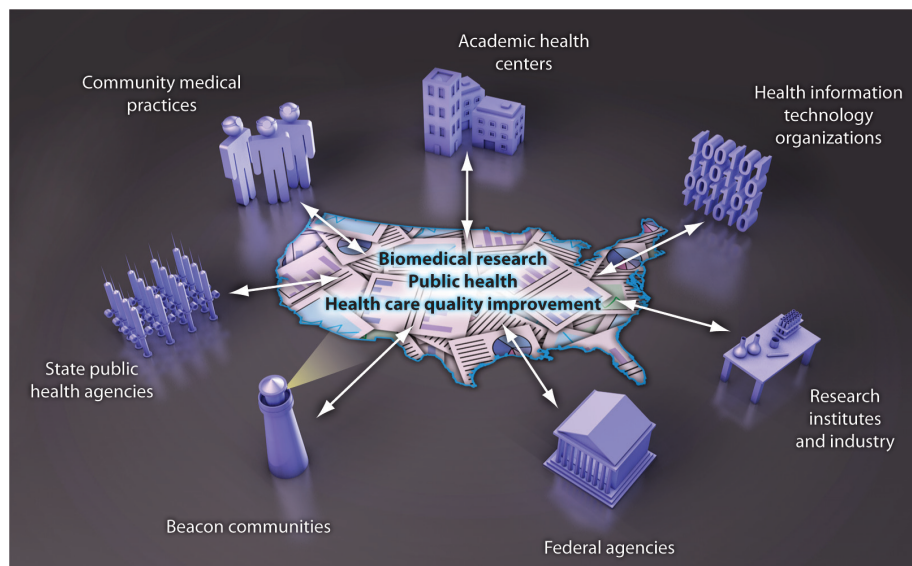

Fig. 1. A nationwide network. Meaningful use of EHRs, widespread participation by multiple diverse entities, and an appropriate technical architecture can spur the construction of a highly participatory rapid learning system that stretches from coast to coast. The resulting rapid learning system can be used, for example, to support biomedical research and augment public health data, with the ultimate goal of improving the quality of health care.

Office of the National Coordinator for Health Information Technology, U.S. Department of Health and Human Services, Washington, DC 20201, USA.
*Corresponding author. E-mail: charles.friedman@hhs.gov

THE VISION:
A RAPID LEARNING SYSTEM

We envision a so-called federated approach to a national learning system. In a federated system, data remain in place until they are needed elsewhere for a particular purpose. Predicated on a policy framework that ensures public trust in the process, organizations that are members of a learning system are eligible to place queries to all other members who would then provide relevant information to address the query. Following are some examples of how such a learning system might operate.

Example 1. An institution that is planning a clinical trial for a new drug to be tested in a specific class of patients wishes to know whether a sufficient number of such patients exists to support the trial as designed. This institution places a query to the learning system: "How many patients who meet these specific eligibility criteria does your institution have?" All members of the learning system would receive the query, and many would reply with an answer expressed as a numerator (the number of patients who fit the criteria) and possibly a denominator (the total number of patients evaluated) as well. This allows the institution that is planning the study to determine whether the proposed sample size is feasible and to develop an appropriately designed strategy for patient recruitment.

Example 2. An outbreak of an infectious disease occurs in a specific part of the country, and the disease begins to spread. Once it is apparent that an outbreak has occurred, the learning system is mobilized to track the disease's spread. As new cases are diagnosed, these data are stored in the EHRs at health care practice sites. In response to a daily or more frequent query, electronic case reports are moved from each practice site to aggregation points in the local, state, and national public health system, making possible real-time nationwide surveillance of the spread of the disease.

Example 3. A new drug is approved for routine use. The learning system is engaged to monitor the new drug's safety. As patients begin using the new drug, any side effects anticipated from the clinical trials are captured in the EHRs as part of the health care of these patients. In a manner that ensures individual privacy, these findings may be routinely transported in an automated manner from the EHRs in which they are collected to federal oversight agencies and to the company that is manufacturing the drug. In addition, researchers who suspect unanticipated adverse events could send a query to the learning system to ascertain the prevalence of such events in a national sample. In both scenarios, the reports supplied by participating members include not only the occurrence of the event but also contextual data that aid in the interpretation of adverse event information.

Example 4. In Example 3, the myriad clinical data obtained from large numbers of patients who are taking a new drug may reveal that patients who display particular physiological characteristics would benefit from a modified dosage of the drug. These findings can lead to the rapid development of a decision-support rule, compatible with almost all deployed EHRs, that is nationally disseminated and incorporated in the decision-support components of these EHRs. When the drug is prescribed, the rule will generate a suggestion to modify the drug dosage in only those patients for whom the change is indicated.

Each of these scenarios demonstrates how the time for disseminating new scientific achievements can be reduced from the current average of 17 years to 17 months, 17 weeks, or almost real-time through a nationally scaled and connected learning system. The system is currently conceived as a voluntary membership organization. The incentive to join rests on a principle of reciprocal benefit. Those who agree to make their data available to the system for response to questions from other members can place queries to the system themselves. The greater the size of the system, the greater the validity of the inferences drawn from the studies it enables. In the future, a global learning health system might be achievable through agreements among individual nations or engagement of multinational organizations such as the European Union, which has outlined such a system for its member nations (*9*).

The federated approach to a learning system contrasts sharply with more centralized approaches—typically used within single organizations—that establish large, persistent repositories of clinical information. In a centralized approach, data are moved to a central repository in anticipation of future uses, before there is a specific need to do so. Large amounts of data reside in these repositories for extended periods of time. This approach is unlikely to be workable on a national scale. Organizations are understandably reluctant to move data beyond their own boundaries absent a clear and specific need to do so, and patients will be less likely to consent to allow this to happen. While the U.S. federal government does have the authority to require reporting of limited data concerning specific conditions that affect the public health (*10*), we believe that a voluntary system with reciprocity of benefits is more likely to gain widespread acceptance and support among patients, care providers, academic and industry researchers, health system administrators, and other key stakeholders.

Several organizations have built learning systems for specific purposes aligned with their missions. Examples in the private sector include Kaiser-Permanente and many academic medical centers, such as the Mayo Clinic, Intermountain Health, Duke University, and the Cleveland Clinic (*11*). Exemplary federal initiatives include the U.S. National Cancer Institute's Cancer Biomedical Informatics Grid (caBIG), a network that connects the cancer community, and the integrated health information systems of the U.S. Veterans Health Administration (*12, 13*). Collectively, these efforts represent an enormous base of experience on which a nationwide effort can draw. These various initiatives have also demonstrated their potential benefits—such as Kaiser-Permanente's early detection of the long-term side effects of Vioxx (*11*). However, none of these efforts can scale directly to serve the entire nation. In general, each organization has evolved its own approach to technology, standards, and policies, all of which drive each entity's learning system and are not easily separated from the institutions' particular patient care and business practices.

BUILDING ON MEANINGFUL USE

Taking the learning system from an idea to a working reality will require mutually reinforcing technologies, standards, and policies created in specific anticipation of nationwide implementation. The national program to achieve EHR meaningful use will contribute many but not all of these.

Technologies. In many respects, the purely technical resources required to move data on demand, securely and using the Internet as the pipeline, already exist. A technical infrastructure for health information exchange, resting on a maturing infrastructure for broadband communication, is being established to support meaningful use. This infrastructure can be extended to provide the technical support for an expanded set of information exchange scenarios required for the learning system. For example, new services beyond those needed for meaningful use will support the asking of a question and the returning of an answer. Other services would support the secure transmission of data about a selected group of persons (rather than an individual patient) along with the metadata that describe the group.

Standards. Many different kinds of standards are required for the development of a rapid learning system. An accumulating set of data and communication standards that support meaningful use can be inherited by the learning system to help ensure that data retrieved from different system members are represented compatibly, ensuring in turn that the data can be aggregated and analyzed. In addition, the learning system will require standards for describing a question in such a way that all recipients and respondents will understand it. Standards for expressing the intent and design of a study are also needed.

Those conducting research and other investigative studies must know not only the results of observations, but also a great deal about how the observations were made. Data collected at different sources, even if the results are represented compatibly, will be amalgamable for valid research if and only if the observations were made with sufficiently similar methods. This requires the learning system to standardize metadata that describe the how, what, when, and where of data collection. Through access to rich metadata, researchers will be able to determine whether the data from elsewhere in the learning system meet the criteria for inclusion in their own studies.

Policies. Although several components of the policy infrastructure required for meaningful use will be applicable to the learning system, many new policies will be required. The vision of a federated national learning system inherits all of the discussion, over the past decade and longer, regarding data reuse and data stewardship (*14*). Public trust in the system is essential. A functioning learning system that supports clinical and translational research, public health information, and comparative effectiveness studies requires resolution of data ownership, patient consent for data reuse, and other key issues, in a sufficiently consistent way to allow the system to function, even though it may not be necessary to require all system members to adopt identical policies. The policy structure will need to definitively address patient consent for use of data in the federated environment. Where data flows can be initiated automatically, policies must explicitly define which functions can happen automatically and which ones require approval. A conceptual basis for these policies will flow from the privacy and security framework being developed to support meaningful use (*15*).

The system will also require a coherent but flexible organizational structure as well as policies governing membership and the ac-

tions of members. The policies must define general eligibility for membership in and the specific resources a member must bring to the system. Furthermore, these policies must distinguish between mandatory and optional behavior. For example: Under what circumstances would a member institution be required, rather than asked, to reply to a query posted to the system? Lastly, policies must clarify mechanisms for how compliance of members will be monitored and, if necessary, corrected. The experience of the National Information Governance Board of the United Kingdom provides an example of how such a governance mechanism could work on a national scale (*16*).

FINAL THOUGHTS: SLASHING THE 17 YEARS

The national aspiration for more effective, efficient, and safer health care requires the kind of rapid learning system we have described. A learning system can dramatically speed the creation and validation of new biomedical knowledge and translation of that knowledge into practice. Existing examples within specific organizations demonstrate the feasibility and signal the benefits of having a system that functions on a national scale. We have described what will be required to build many essential elements of a rapid learning system. Although meaningful use of EHRs provides an enormous boost to this effort, many challenges remain. The nation has only begun its progression to meaningful use. Those features of the rapid learning system that will not be direct byproducts of meaningful use will not build themselves.

Seen in this light, the nation's investments in EHR adoption and meaningful use constitute a "twofer." They will directly improve the care of individual patients and enhance some aspects of public health—and they will move the nation substantially toward the development of a rapid learning system. Carrying the nation the rest of the way to achieving a broadly participatory and functioning learning system will require coordination of effort, within and outside the federal government, of individual organizations that will inevitably be investing their own resources to advance their own capabilities as learning organizations. To the extent that these efforts align with progress toward a national system, they will advance a national agenda as much as each organization's unique mission.

REFERENCES AND NOTES

1. E. A. Balas, S. A. Boren, Managing clinical knowledge for health care improvement. In *Yearbook of Medical Informatics 2000: Patient-Centered Systems*, J. Bemmel and A. T. McCray, Eds. (Schattauer Verlagsgesellschaft mbH, Stuttgart, Germany, 2000), pp. 65–70.
2. Centers for Medicare and Medicaid Services, Meaningful use, www.cms.gov/EHRIncentivePrograms/35_Meaningful_Use.asp.
3. Health Information Technology for Economic and Clinical Health Act, Public Law 111-5, 42 U.S.C. 201, Sec. 4101(a) (2009); http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=111_cong_public_laws&docid=f:publ005.111.pdf.
4. J. J. Nadler, G. J. Downing, Liberating health data for clinical research applications. *Sci. Transl. Med.* **2**, 18cm6 (2010).
5. J. R. Slutsky, Moving closer to a rapid learning health system. *Health Aff.* **26**, w122–w124 (2007).
6. Institute of Medicine, Roundtable on value and science-driven health care. Learning Healthcare System series of publications (2010); http://www.iom.edu/~/media/Files/Activity%20Files/Quality/VSRT/RT%203%20pgr%20June2010.pdf.
7. A. P. Abernethy, L. M. Etheredge, P. A. Ganz, P. Wallace, R. R. German, C. Neti, P. B. Bach, S. B. Murphy, Rapid-learning system for cancer care. *J. Clin. Oncol.* **28**, 4268–4274 (2010).
8. D. Blumenthal, M. Tavenner, The "meaningful use" regulation for electronic health records. *N. Engl. J. Med.* **363**, 501–504 (2010).
9. M. Olive, H. Rahmouni, T. Solomonides, SHARE the journey, a European healthgrid roadmap. European Commission (2010); http://eu-share.org/roadmap/eHealth-SHARE-A4.pdf.
10. K. Benitez, B. Malin, Evaluating re-identification risks with respect to the HIPAA privacy rule. *J. Am. Med. Inform. Assoc.* **17**, 169–177 (2010).
11. High-Value Health Care Project, *How Registries Can Help Performance Measurement Improve Care* (The Brookings Institution, Washington, DC, 2010), pp. 10–11; www.rwjf.org/files/research/65448.pdf.
12. W. Wolfson, caBIG: Seeking cancer cures by bits and bytes. *Chem. Biol.* **15**, 521–522 (2008).
13. L. I. Backus, S. Gavrilov, T. P. Loomis, J. P. Halloran, B. R. Phillips, P. S. Belperio, L. A. Mole, Clinical Case Registries: Simultaneous local and national disease registries for population quality management. *J. Am. Med. Inform. Assoc.* **16**, 775–783 (2009).
14. M. Bloomrosen, D. E. Detmer, Informatics, evidence-based care, and research; implications for national policy: A report of an American Medical Informatics Association health policy conference. *J. Am. Med. Inform. Assoc.* **17**, 115–123 (2010).
15. D. Blumenthal, Launching HITECH. *N. Engl. J. Med.* **362**, 382–385 (2010).
16. H. Cayton, National Information Governance Board for Health and Social Care (U.K.). *Electronic Infrastructure for the Learning Healthcare System: The Road to Continuous Improvement in Health and Health Care.* An Institute of Medicine Workshop Series, Sponsored by the Office of the National Coordinator for Health Information Technology (2010); http://www.nationalehealth.org/IOM/4-HarryCayton.pdf.
17. **Acknowledgments:** The authors thank their colleagues D. Fridsma, F. Mostashari, R. Nelson, and W. Yu for their contributions to the development of these concepts. M. McGinnis and C. Grossmann of the Institute of Medicine also contributed useful insights that have been incorporated.

Caught in the Web: Informed Consent for Online Health Research

Effy Vayena,[1]* Anna Mastroianni,[2] Jeffrey Kahn[3]

A context-specific approach to informed consent for Web-based health research can facilitate a dynamic research enterprise and maintain the public trust.

There has been an exponential growth in personal health data supplied by users of mobile devices, health apps, and the social Web (social networking sites, online disease support groups, and health-related information sites). These sources and data from tracking of consumers' online behavior coupled with advanced bioinformatics tools offer opportunities for use in health research [for examples, see (*1*, *2*)]. Tweets about disease outbreaks have been correlated with official public health surveillance data and have proven to be an important source of early outbreak detection (*3*).

A central ethical question is whether individuals who have provided personal information online in nonresearch contexts have consented to research uses. Here, we explore the issue of informed consent for health research performed using information collected from the Web, discuss some limitations of current practices, and offer recommendations for improving consent practices through a more tailored, context-sensitive approach that makes use of the dynamism of the Web-based context. Our proposals are rooted in the ethical imperative of protecting individual rights and respecting autonomy while enabling a dynamic research environment for the advancement of clinical medicine and public health.

A CHANGING CONTEXT FOR INFORMED CONSENT

Health-related research proposals with humans typically undergo prospective review by research ethics committees that ensure that the study is designed and conducted in an ethical manner that protects the privacy and autonomy of individuals through the informed consent process, balances risks and benefits, and ensures that subject selection is equitable. Informed consent requires that potential participants are provided with adequate information to make an informed and voluntary decision about research participation. Despite the standard practices of obtaining consent, there is a prevalent notion that the process is broken (*4*). How can an adequate consent process be achieved in health research that involves data collected in the Web environment?

Currently there is limited ethics guidance specifically for research with data collected on the Web (www.aoir.org/reports/ethics2.pdf). Reflected in the dearth of ethics guidelines is either a lack of acknowledgment of this growing area of research or perhaps a sense that this research should not be treated differently from other conventional areas of research. Further, although Web-based research is inherently interdisciplinary because of the range of research areas and diverse sources of data, limited dialogue exists among the represented disciplines in terms of a common set of research principles. New sources of online data and innovative health research applications and the increasingly disparate sources of data challenge traditional approaches to informed consent.

Conventional informed consent models are ill suited because they were not conceived in the context of the evolving applications and functionalities of social media that enable innovative research designs. In addition, traditional approaches to research ethics and informed consent include an ethical distinction between public and private information: The use of publicly available information typically does not require in-formed consent of the individual, whereas the use of "private" information may require consent depending on whether the information allows an individual to be identified. In an online world, the public-private distinction is increasingly blurred. Should explicit consent be required if a researcher collects and analyzes deidentified Facebook posts that reveal health status or health behaviors? Can health status information shared on social networking sites for patient communities (for example, PatientsLikeMe.com) be used for research without individual informed consent? Is such information properly characterized as public or private?

Current discussions about the public-private dichotomy in the online world include a newer, richer concept that views privacy within "contextual integrity" (*5*). This approach argues for understanding the importance of the context in which information is located, and determinations of acceptable use are informed by expectations for the use of information within the context in question. This approach places heavier emphasis on the intent of individuals regarding access to personal information rather than on the traditional approaches that demand that researchers protect privacy as a condition of research.

Equally challenging to traditional concepts of informed consent is the control of

Mission control. Data sharing for health research requires an environment of trust.

[1]Institute of Biomedical Ethics, University of Zurich, Zurich 8032, Switzerland. [2]School of Law and Institute for Public Health Genetics, University of Washington, Seattle, WA 98195, USA. [3]Berman Institute of Bioethics, Johns Hopkins University, Baltimore, MD 21205, USA.
*Corresponding author. E-mail: vayena@ethik.uzh.ch

CREDIT: N. HOOGSLAG/IKON IMAGES/ALAMY

personal information on the Web. When subjects are traditionally asked to consent to the research use of their data, the limits of that use are spelled out in detail. Personal data posted on or collected by Web sites, however, can be sold or shared and subsequently used in research; thus, it is nearly impossible for users to maintain control of their data, its diffusion, and subsequent uses. As such, the notion of consenting to research use of data loses meaning when the use can involve many unknown researchers and uses in perpetuity. Such open-ended use of data renders the well-established right to withdraw consent to collection and use of personal data for research meaningless. A recent controversial European Union proposal attempts to address the "genie out of the bottle" problem by revising data privacy standards to include a digital "right to be forgotten"—users must be granted an option to delete personal data from the Web permanently (http://ec.europa.eu/justice/data-protection/document/review2012/com_2012_11_en.pdf).

DO WEB USERS CONSENT?

Health research using the Web is gaining momentum regardless of available guidance. How informed consent is treated and what it means varies substantially by site and project.

There are two broad types of data gathered through the Web that can be used in health-related research, either by the Web site owner or by third parties: (i) information actively supplied by the user (medical histories, genomic data, and Web posts), and (ii) personal information collected by the Web site while the user is interacting with the site (IP and e-mail addresses, searches, and location data). Both data types may be required by a Web site to enable it to provide the promised services to the user. Many Web services are provided to the user free of charge, while the content that users generate and personal data they provide become trading capital for the companies that provide the services. Web sites often authorize third parties to access their data sets for commercial and research purposes.

The disclosure to users of the potential uses of personal data vary dramatically from site to site. Further, no publicly available studies have yet documented whether users understand or are even aware of the potential uses of their data when they access a site. In reviewing a range of Web sites (6) that collect or contain health data for research, we have identified three general approaches

to consent: (i) research participation as a condition of use of the site, (ii) opt-in to research, and (iii) opt-out of research.

In what we term "condition of use" research participation, Web sites state in their terms of use, terms of service, or privacy statements that they maintain the right to use the data they collect for research, among other uses. By virtue of using the site, the user agrees to research participation. This is equivalent to a so-called browsewrap agreement, whereby the user agrees to the terms of use without any affirmative conduct, such as clicking an "I agree" button (www.eff.org/wp/clicks-bind-ways-users-agree-online-terms-service). It is most likely a carryover from the consumer-oriented sites that use browsewrap approaches as a basic disclosure of policy but without real expectation of careful review and affirmative consent.

The condition of use approach raises three potential concerns. First, the user provides a general consent to a range of uses, including research, rather than consent for a specific research project or research use, and is unable to access the site without giving broad general consent. Second, possible research use is often (but not always) listed among many other uses within the boilerplate language of disclosures and indemnifications, making it questionable that the reader will take notice. Third, the condition of use approach was crafted for consumer agreements to Web site use rather than to accommodate the requirements for informed consent in research. Thus, the condition of use approach stands in stark contrast to the conventional approach to informed consent in health research: a process carefully constructed so that (i) individuals are adequately informed about the project, (ii) the meaning of research participation is clear, (iii) the potential participant makes a voluntary agreement to participate, and (iv) the potential participant is offered the option to withdraw from the research. On this analysis, condition of use research participation does not meet the standards of informed consent except in the most limited and legalistic understanding of consent as agreement evidenced by accepted terms of use.

By contrast, the opt-in approach enables users to agree to participate in a specific research project. Web sites that use opt-in may include a statement with information about the project followed by a link that leads to the project or a requirement to click an "I agree" button to allow research use of personal data. In contrast to condition of use, opt-in par-

ticipation requires an affirmative decision by the user before participating in research. This is equivalent to the common approach to software or other licensed goods in the online world, in which a user must agree to a licensing agreement (also known as a clickwrap agreement) before accessing the product or site. Yet, in that consumer context, research shows that users spend almost no time browsing the text of the agreement before clicking the box, making it unlikely that users opting in to research using this model will carefully read the agreement texts (7). Thus, allowing a research participant the opportunity to review information about the potential research use and confirming participation with an affirmative act comes closer to satisfying the conventional criteria of informed consent, but it is far from clear that the opt-in model achieves the goal of informed voluntary research participation.

With the opt-out approach, Web site users agree to research uses of their data unless they take action to exclude themselves from participation. Users thereby control their data, provided they are aware that the opt-out option exists. For example, some search engines offer users the option of opting out of tracking so that personal data are not collected or stored, although users typically must be sufficiently web savvy to locate the opt-out option. The opt-out approach to informed consent exists in conventional biomedical research when the study poses low risks or when obtaining opt-in consent is impractical and could undermine study design (for example, large-scale epidemiological studies and genome-wide association studies). In these limited contexts, the use of opt-out approaches has been thoroughly debated in the literature and by research ethics committees before being put into practice. Researchers in traditional settings have advocated for wider acceptability of this opt-out approach to informed consent, arguing that it facilitates research while still safeguarding autonomy (www.ieaweb.org/index.php?option=com_content&view=category&id=22&Itemid=54). Whether an opt-out approach to consent in health-related Web-based research satisfies conventional criteria of informed consent requires analysis by consumers, scientists, legal scholars, and research ethics committees.

RECOMMENDATIONS

The health research enterprise continues to evolve along with new possibilities for harnessing online personal health data from

the social Web. Yet, such research is still governed by rules set to address the issues faced in traditional clinical research; this creates a mismatch between the policy requirements for research protections and the types of issues faced in health research using online data. It is important and timely to develop guidance that specifically addresses this new research context. The values that underpin conventional health-research rules remain important. Respect for individual autonomy, balanced and equitable distribution of risks and benefits, and the right to privacy have not lost normative weight because the social Web challenges the approaches we have created to respect them. On the contrary, the challenge presents an opportunity to explore more nuanced and innovative ways of interpreting these values.

It is hard to imagine a useful "one size fits all" consent model for research in the Web environment. Appropriate consent models will depend on the mission of the site, sensitivity and identifiability of the data collected, purpose of the research, and risks and benefits of participation. An interactive process is better suited to meeting the criteria of informed consent. At a minimum, transparent disclosure of the research uses of online personal data are required.

Two types of consent models are well suited for health-related research on the Web. The first is based on the relatively new concept of the Portable Legal Consent (PLC), a legal framework for research consent developed by the Consent to Research project (http://weconsent.us; www.inspire-2live.org; www.sagebase.org). It allows participants who are willing to relinquish control of their personal information to attach a one-time research consent to their health and genetic data, which they upload themselves onto the Web site. The data can then be used for research purposes by any researcher who agrees to specific criteria: publication of research results in an open-access forum, no reidentification of participants, and no redistribution of data unless the data recipient agrees to the PLC conditions. The overarching goal of the PLC is to minimize barriers to data sharing and make research data more widely available. Transaction costs related to contacting participants for consent to individual studies are eliminated, privacy concerns are minimized as the data are deidentified to the extent possible, and the participants are informed of potential privacy risks before consenting. Participants may withdraw their data from the database at any time, but are clearly advised that once data are uploaded, it may not be possible to remove it from all sources (for example, from researchers who have already downloaded, shared, or used the data). The Self Contributed Cohort for Common Genomics Research Study (SCC-CGR) (http://weconsent.us/about-us) is the first to implement the PLC.

The PLC improves on existing consent processes in two important ways: (i) the conditions, rules, and restrictions that apply to all research and to researchers who access the data are transparent, and (ii) the risks to participants of sharing their data are clearly articulated. However, in contrast to traditional consent models, the PLC approach, in its proposed state, has some apparent shortcomings: (i) PLC participants must be willing to give up control over their personal health data, including a limited right to withdraw and the choice to decline specific research projects; (ii) deidentifying of data limits its usefulness for some research projects; (iii) PLC relies on a self-selected, well-informed population of computer-savvy users who are unlikely to be representative of the population at large; and (iv) PLC cannot be used for Web data collected for nonresearch purposes.

The shortcomings of the traditional consent process and PLC argue for an approach that is sensitive to the unique aspects of Web-based health research and that harnesses the dynamic aspects of the Web environment. Collaborative and context-specific consent employ the communicative and real-time features of the Web to facilitate a more dynamic approach to informed consent (8). Instead of the traditional approach of a one-time agreement that includes boilerplate text, a user could receive tailored information on research participation with specific choices of options relevant to his or her situation. Collaborative consent might provide a way to address a problem that has challenged even the traditional consent process: how to design effective ways of communicating information to prospective participants. Moreover, transparency requires commitment to clarity and the provision of accurate and appropriate information by researchers. Sites engaged in health research or that allow third parties to use their data for research must modernize communication by making use of the multimedia capabilities afforded by the Web.

PLC will give individuals a way to share (or cede control of) their health-related data. But what about prospective participants who desire greater control? Empirical evidence shows that people care about the way their data are used (9). Giving the user control will also contribute toward building trustworthy relationships and likely increase user participation in research. The "health-information altruist" (10) is willing to contribute to research for the common good. The claim behind the health data rights movement is that individuals should be able to control the sharing of their information (www.HealthDataRights.org). These trends can result in valuable contributions to research only if encouraged by an environment that is conducive to trust.

REFERENCES AND NOTES

1. B. H. Wagenaar, P. S. Sullivan, R. Stephenson, HIV knowledge and associated factors among internet-using men who have sex with men (MSM) in South Africa and the United States. *PLoS ONE* **7**, e32915 (2012).
2. J. Frost, S. Okun, T. Vaughan, J. Heywood, P. Wicks, Patient-reported outcomes as a source of evidence in off-label prescribing: Analysis of data from PatientsLikeMe. *J. Med. Internet Res.* **13**, e6 (2011).
3. R. Chunara, J. R. Andrews, J. S. Brownstein, Social and news media enable estimation of epidemiological patterns early in the 2010 Haitian cholera outbreak. *Am. J. Trop. Med. Hyg.* **86**, 39–45 (2012).
4. E. C. Hayden, Informed consent: A broken contract. *Nature* **486**, 312–314 (2012).
5. H. Nissenbaum, A contextual approach to privacy online. *Daedalus* **140**, 32–48 (2011).
6. E. Vayena, A. Mastroianni, J. Kahn, Ethical issues in health research with novel online sources. *Am. J. Public Health* **102**, 2225–2230 (2012).
7. R. Boehme, S. Köpsell, Trained to accept? A field experiment on consent dialogs. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (ACM, New York, 2010) pp. 2403-2406; http://dl.acm.org/citation.cfm?doid=1753326.1753689
8. P. Bernal, Collaborative consent: Harnessing the strengths of the internet for consent in the online environment. *Int. Rev. Law Comput. Technol.* **24**, 287–297 (2010).
9. J. Turow, J. King, C. J. Hoofnagle, A. Bleakley, M. Hennessy, Americans reject tailored advertising and three activities that enable it (University of Pennsylvania, Annenberg, 2009); http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1478214
10. I. S. Kohane, R. B. Altman, Health-information altruists— A potentially critical resource. *N. Engl. J. Med.* **353**, 2074–2077 (2005).