HHS Public Access

Author manuscript

AJOB Empir Bioeth. Author manuscript; available in PMC 2015 June 20.

Published in final edited form as:

AJOB Empir Bioeth. 2015 ; 6(2): 12–22. doi:10.1080/23294515.2014.949906.

Benefits and Risks in Secondary Use of Digitized Clinical Data: Views of Community Members Living in a Predominantly Ethnic Minority Urban Neighborhood

Robert J. Lucero,
Columbia University School of Nursing

Joan Kearney,
Columbia University School of Nursing

Yamnia Cortes,
Columbia University School of Nursing

Adriana Arcia,
Columbia University School of Nursing

Paul Appelbaum,
Columbia University College of Physicians and Surgeons

Roberto Lewis Fernández, and
Columbia University College of Physicians and Surgeons

Jose Luchsinger
Columbia University College of Physicians and Surgeons

Abstract

Background—There is potential to increase the speed of scientific discovery and implement personalized health care by using digitized clinical data collected on the patient care experience. The use of these data in research raises concerns about the privacy and confidentiality of personal health information. This study explored community members' views on the secondary use of digitized clinical data to (1) recruit participants for clinical studies; (2) recruit family members of

Address correspondence to Robert J. Lucero PhD, MPH, RN, Columbia University School of Nursing, 617 W. 168th Street, New York, NY 10032, USA. rl2451@columbia.edu.

AUTHOR CONTRIBUTIONS

R. J. Lucero supervised the study overall, participated in data collection, assisted in the analysis, and led the writing of the article. J. Kearney provided expertise and direct support in developing and carrying out the qualitative methods, and assisted in writing the article. Y. Cortes and A. Arcia were directly involved in conducting the focus-group analysis as the primary and secondary coders, respectively, and assisted in writing the article. P. Appelbaum and R. L. Fernández assisted in conceiving the study and assisted in writing the article. J. Luchsinger conceived of the study and design, participated in data collection, provided assistance with the study overall, and assisted in writing the article.

CONFLICT OF INTERESTS

Dr. Roberto Lewis Fernández has received research support from Eli Lilly and Company. The other authors report that they have no conflicts of interest to disclose.

ETHICAL APPROVAL

This study was approved by the institutional review board at the Columbia University Medical Center.

persons with an index condition for primary studies; and (3) conduct studies of information related to stored biospecimens.

Methods—A qualitative descriptive design was used to examine the bioethical issues outlined from the perspective of urban-dwelling community members. Focus groups were used for data collection, and emergent content analysis was employed to organize and interpret the data.

Results—Thirty community members attended one of four focus groups ranging in size from 4 to 11 participants. Five critical themes emerged from the focus-group material: (1) perceived motivators for research participation; (2) objective or "real-life" barriers to research participation; (3) a psychological component of uncertainty and mistrust; (4) preferred mechanisms for recruitment and participation; and (5) cultural characteristics that can impact understanding and willingness to engage in research.

Conclusions—The overriding concern of community members regarding research participation and/or secondary clinical and nonclinical use of digitized information was that their involvement would be safe and the outcome would be meaningful to them and to others. According to participants, biospecimens acquired during routine clinical visits or for research are no longer possessions of the participant. Although the loss of privacy was a concern for participants, they preferred that researchers access their personal health information using a digitized clinical file rather than through a paper-based medical record.

Keywords

patient data privacy; genetic privacy; community–institutional relations; minority health; data sharing; research ethics

---

The adoption of electronic health records has more than doubled since 2012 among eligible doctors' offices (>50%) and hospitals (>80%) in the United States (Department of Health and Human Services 2013). The result of this diffusion of health information technology in data-intensive environments is the collection and storage of vast amounts of patient data and the potential of data reuse. Clinical data collected through electronic health records and/or accessible in digitized formats have the potential to increase dramatically the speed of scientific discovery and implement personalized health care (Blumenthal and Tavenner 2010). The collection and retrieval of digitized clinical information can enable the creation of individual and networked databases and registries that facilitate large-scale studies (Sturmer et al. 2011).

The secondary use of digitized clinical data utilizes personal health information for uses outside of direct patient care, including clinical and systems research, quality and safety measurement, public health, and payment (Safran et al. 2007). Health care data available to researchers in digitized files can increase analytic efficiency and reduce costs associated with manually extracting clinical information from paper-based medical records (Weiner and Embi 2009). This potential wealth of information could provide the basis to examine very large numbers of individuals from different groups (e.g., racial/ethnic, diagnostic categories) across different clinical settings over time or compare the effectiveness of various treatments.

Equally important, the secondary use of stored digitized clinical data could help more fully account for under-represented groups and mitigate concerns about generalizability of clinical studies. The large number of individuals represented in digitized data warehouses has the potential to reduce dependency on oversampling to achieve adequate numbers for reliable estimates of findings across multiple groups. Additionally, secondary use of personal health information associated with digitized clinical files could provide researchers with information about family members who could be recruited to enhance representation of underrepresented groups in studies.

However, this secondary use of digitized clinical data in research raises concerns about the privacy and confidentiality of personal health information (Peddicord et al. 2010). Patients have cited privacy as a concern in connection with the adoption of electronic health records by health care providers and hospitals (Blumenthal 2009). According to a 2005 California HealthCare Foundation survey, 67% of Americans are concerned about the privacy of their health records (Bishop, Holmes, and Kelly 2005). A greater proportion of ethnic/racial minority respondents (73%) expressed worry over the privacy of their health records. In a 2007 Harris Interactive poll, a majority of respondents (51%) said that the use of electronic medical records makes it more difficult to ensure patients' privacy, but nearly two-thirds said that the benefits of electronic medical records outweigh the privacy risks (Krane 2007).

Although physicians and consumers may perceive benefits of using electronic health records (e.g., improved quality of care, increased efficiency), potential threats to patient confidentiality with secondary use of digitized clinical information cannot be discounted. Harmful linkages or disclosure of personal health information may occur at the hands of persons authorized to access patients' records, but who use the information in ways that violate a patient's request for, or expectations of, confidentiality. Digitized information may also be accessible to persons not authorized to have access, also known as secondary users. Commercial entities, such as insurance carriers, may want to combine clinical information with other data sources (e.g., administrative billing codes) and sell the aggregated data for surveillance purposes.

Theoretically, all of these potential misuses of clinical information can be circumvented with attention to an individual's autonomy (Beauchamp and Childress 2012). By respecting individuals' rights to act intentionally, creating conditions for self-directed decision making, and removing controlling influences from existing second and third parties (i.e., health care system and commercial entities, respectively), consumers can become instrumental in determining future uses of their health information.

The role of autonomy, however, is less clear in research based on the secondary use of digitized clinical data. Do traditional consent requirements actually protect patient interests, or would effort be better spent on measures that enhance data security after participants have consented to use of their clinical information? Does reliance on patient consent create unreasonable barriers to conducting research based solely on the secondary use of digitized clinical data (i.e., consent to retrieve data for every research study)? Should the full scope of regulatory protection be applicable to such research studies when individuals will not undergo any type of intervention, and risks—other than those associated with loss of

confidentiality—are minimal? Given the potential of secondary use of digitized clinical data for research, especially for studies that focus on hard-to-reach, vulnerable, and minority populations, the cost of inappropriate regulation may be high.

Although there is a strong argument for increased participation of underrepresented groups in biomedical research, and one potential way to address this is through the secondary use of digitized clinical data, there are few studies that have examined the thoughts and opinions of individuals from underrepresented groups regarding secondary clinical and nonclinical use of digitized clinical data. Thus, from the perspective of community members who self-identified primarily as Hispanic or Black and also those from non-ethnic/racial minority groups, this study explored the secondary use of digitized clinical data to (1) recruit participants for clinical studies, including observational and clinical trials; (2) recruit family members of persons with an index condition for primary research studies; and (3) conduct studies of information related to stored biospecimens.

By examining the perspectives of individuals whose personal health information is potentially available to researchers, this study aimed to investigate how those who are providing personal health information view the potential risks and benefits of secondary clinical and non-clinical uses of such information. The study focused on issues particularly affecting investigators conducting research on minority health and health disparities, family caregiver burden, and comparative effectiveness (i.e., comparing two or more interventions for the same disease/outcome to identify which yields the best results), for which clinical data in electronic health records can be a significant source of information (Berger et al. 2009; Burton, Anderson, and Kues 2004; Shields et al. 2007).

METHODS

Embedded within the community-based participatory research framework, a qualitative descriptive design was used to examine three bioethical issues outlined from the perspective of community members (Sandelowski 2000; 2010). Community-based participatory research is an approach that enables community members to participate actively in research with a goal of influencing change in health, systems, programs, or policies. Focus groups were used for data collection, and emergent content analysis (Hsieh and Shannon 2005) was employed to organize and ultimately interpret the data. Since this is exploratory work, this approach was chosen so as not to limit the amount or kind of information we received in this early phase of the work.

Participant Sampling and Recruitment

A purposive sampling frame was used to recruit lay adult community members from the Washington Heights/Inwood neighborhood surrounding Columbia University Medical Center (CUMC). Potential focus-group participants were recruited through community-based organizations that serve the needs of Hispanic community members, primary care clinics, and the Columbia Community Partnership for Health of the CUMC. The Columbia Community Partnership for Health is a commonly used venue for community meetings as well as research recruitment and activities. With the assistance of a well-established

community advisory board, individuals were identified for possible participation in focus groups.

Community members of different sex, race/ethnicity, age, and education levels were recruited to elicit a broad range of responses. Because shared life experiences result in more in-depth discussion, participants with similar demographic characteristics were grouped together (Bernard 1995). Maximum variation in the sample was sought to gather data from participants who (1) had already participated in past research projects; (2) had refused research participation in the past; and (3) had never considered participating in research.

Procedures

After approval by the CUMC institutional review board, participants were approached and their informed consent obtained; focus groups were held at the Columbia Community Partnership for Health, a community-based site about a half-mile from CUMC. Groups were held every other month in the evening for a duration of 2 hours each. Food and beverages were provided for each session, and each participant was compensated $25.

Focus-group questions were open-ended and centered on the three main bioethical issues of interest (see Appendix). To ensure rigor, each session employed a focus-group team approach (Hennink and Diamond 2000), consisting of a moderator and observer/note-taker. Participants were identified only by first name for the purpose of reviewing individual contributions in debriefing sessions among the focus group team. Participants were informed in the written and verbal consent that their names would not be linked to their contributions, and data would be reported in the aggregate. Groups were led by a senior member of the research team (RJL and/or JL) and the study coordinator, who filled these roles. Both researchers had experience in focus-group methods, having conducted previous research in Washington Heights/Inwood.

Data Analysis

Focus-group transcripts were analyzed using an emergent content analysis approach, and organized with NVivo 9 software. The analysis was completed in a stepwise iterative process to ensure accuracy. The focus group recordings were transcribed and analyzed in the original language (e.g., Spanish). Following first-pass examination of transcripts by a designated primary coder, a secondary coder, and one of the study principal investigators (PIs) (RJL), a second phase of analysis ensued in which the lead coder developed detailed codes in consultation with the PI. These codes went through additional review in conference with the entire coding team and the qualitative consultant (JK), yielding a total of 25 final unit codes. Reliability testing was then conducted on transcripts of 25% from each group by the primary and secondary coders. Cohen’s kappa values ranged from a low of .69 on one code to a high of 1.0 on five codes. For all but one code, kappa coefficients were .77 or above.

After establishing reliability for codes, thematic analysis was conducted, with unitary codes integrated into final critical categories or themes. These categories reflected participants’ concerns and opinions regarding the bioethical issues of interest in this study. Again, a systematic protocol was followed whereby regular conferencing within the coding team was

supplemented by meetings with the qualitative consultant. All final categories were conferenced to 100% agreement by the team. Selected Spanish quotes used in the results were translated by the bilingual focus-group coders.

The scientific adequacy of this project was considered as it applied to the four principles of scientific rigor or trustworthiness (Guba 1981). Credibility was addressed by triangulating heterogeneous sources across focus groups and involving multiple investigators in data collection and analysis. Purposive sampling and verbatim transcription assured transferability. Dependability was emphasized through the auditing protocol and the clear articulation of analysis decisions. Finally, confirmability was addressed by the continual reflexivity inherent in the process of conducting the study and analyzing the data.

RESULTS

Thirty adult community residents attended one of four focus groups ranging in size from 4 to 11 participants. Two of the focus groups were conducted in Spanish and two in English. Characteristics of focus-group participants based on self-reported information are described in Table 1.

Five critical themes emerged from the focus-group material. These included (1) participants' perceived motivators for participating in research; (2) objective or "real-life" barriers to participation; (3) a psychological component of uncertainty and mistrust of the process; (4) preferred mechanisms for recruitment and participation; and (5) cultural characteristics that may impact understanding of and willingness to engage in research. These themes are briefly described here.

Perceived Motivators to Participate in Research

Factors that motivated informants to engage in research included relevance of the research topic and the personal, family, or community benefits derived from their participation in research as well as financial incentives. The idea of benefit extended to the utility of the scientific knowledge gained for themselves and their community. One informant highlighted the notion of scientific benefit or knowledge to be gained in this simple but imaginative exemplar: "It's like passing a ball. You learn something, then you tell your friend, 'Hey, listen. Do you know about this?'"

Another informant spoke about being motivated by financial need and the potential of gaining a direct benefit from an experimental intervention despite the challenges associated with participation:

> At that particular time I was extremely financially in a bind … I wanted to get the benefit, absolutely, so just to drop out I would've been there already 45 minutes and now I'm gonna stop 'cuz I'm uncomfortable?

For some informants, financial incentives were secondary motivators to participate in research. Other informants remained silent on this matter. However, it appears that informants in some instances might be motivated to risk or even give up more than they bargained for in exchange for direct personal financial gain.

Practical Barriers to Research Participation

Practical barriers to research participation were related to challenges in the environment, such as lack of transportation, access to information, and communication about ongoing research studies. These barriers are present in one informant's remark:

> A lot of times there's research things going on that's not publicly known, so how are we to go about finding out about certain research programs that's not advertised.

Not surprisingly, the lack of a "common language" to communicate research opportunities and respond to recruitment efforts was mentioned as a barrier. Notably, informants expressed hesitation to take part in research that involves a clinical intervention. Many informants reported that they would not volunteer for a study if it required them to consume an experimental drug or undergo a surgical procedure. An informant's comment that reflects this view was:

> When I think of research, I don't think of doctors. I think about information. I don't think about going to take a medicine, be a guinea pig. I'm talking about information, sit there like we doing now … I'm not gonna get medicine … I just find out what it's about.

A majority of the informants echoed the value of communicating lived experiences in focus groups as a mechanism for exchanging information regarding health and disease with each other and investigators. They viewed focus-group participation as preferable to "invasive" forms of research as a mechanism for gathering information and knowledge about health and disease that they could take back to the community.

Subjective Barriers to Research Participation (Psychological Component)

Informants provided compelling messages about fear, uncertainty, and lack of trust regarding volunteering for research. The interplay of these messages was expressed within a rubric of privacy violations based in revealing highly personal information and third-party access to information regardless of confidentiality clauses. As an informant conveyed, "If there is no permission, without one's consent, of course it is an invasion of privacy." This concern extended to the secondary use of digitized personal health information without their prior consent regardless of how investigators intended to use the information (e.g., primary research of a specific clinical condition or mining information to identify family members).

Nonetheless, informants preferred that investigators use electronic health records over paper-based records held in their health care provider's office when gathering personal health information. This preference is reflected in the following statement. "I feel the computer is better. Don't go to my doctor and pull my folder out … I'm letting it all on paper. What the doctor puts on the computer is … different." This comment and other similar ones seem to convey that subjective health and/or personal concerns are more likely to appear on paper-based records, whereas electronic health records only contain the provider's "objective" health assessment. Thus, informants may believe that they are protected from researchers or third parties accessing confidential communications.

Informants were also clear about what their expectations are of researchers when being recruited by telephone to participate in a study. These expectations were based on full disclosure about the qualifications of the caller and how the caller identified them as a potential research participant for the study. An exemplar of this expectation was conveyed in the following remark:

> Who is this that's calling me and why are they calling me? They don't wanna divulge the information that they looked at my records to find out I'm able to participate in this group. They have to tell me why they're calling.

Preferred Mechanisms for Research Participation

Informants highlighted the importance of partnerships between researchers and clinical providers and face-to-face contact with a member of the research team as factors associated with deciding to participate in research. Informants expressed a preference for being recruited for research through the auspices of their "doctor." An informant who had previously participated in research expressed the following:

> I would have felt more comfortable if it come from my doctor who I've established a relationship with, but this is a stranger and they're asking all these personal questions about my health or my history and then my family's history and then did you have this or did you have that or this and this and this.

Informants talked about the need to have a discussion with their doctor to get advice about safety and whether they were suited for the study. An informant mentioned that they would "talk to the doctor to see if it benefits me, 'cuz they have more knowledge of my medical condition and stuff and they would tell me yes or no." Moreover, face-to-face contact with a member of the research team was preferred over telephone contact as a mechanism for fostering confidence in the proposed research. A participant stated, "They can call me, but set up an appointment. Then, I can talk to him [researcher] directly."

Cultural Characteristics for Consideration

Concerns in this category included "ownership" of biospecimens once they had been collected by health care providers. Informants agreed overwhelmingly that once specimens are collected these samples are no longer perceived as belonging to them. An example of this is, "After they have drawn my blood … it is no longer mine, I have already given it to you." This sentiment was also expressed when informants considered personal health information they release to researchers. There seemed to be an assumption that once the information is collected it no longer belongs to the research participant.

A significant translational barrier was encountered among Spanish-speaking informants. There was notable misunderstanding between the concepts of scientific research and medical examination/procedures (i.e., translated in Spanish, scientific research and medical examination/procedures are expressed similarly using the term investigación). This misinterpretation was evident when one informant reported, "I had a clogged artery and they placed a small clamp here in the heart and performed the investigation." In addition, participants began to ask for clarification from each other regarding the definitions of scientific investigations and clinical exams. Those who had previously participated in

research provided examples of their experiences to highlight the differences between the two.

DISCUSSION

Researchers have examined barriers to recruitment of research participants from underrepresented populations (Connell et al. 2001; Dennis and Neese 2000; Neufeld et al. 2001; Preloran, Browner, and Lieber 2001). This remains a significant issue since adequate representation of under-represented groups often poses a scientific challenge to biomedical researchers.

However, less is known about the attitudes and opinions that individuals from underrepresented groups have toward secondary clinical and nonclinical use of personal health information and their willingness to be active participants in allowing use of such information. Based on a nationally representative U.S. sample, African Americans, compared with Whites, and those with lower educational levels were significantly less positive about the use of medical information for research (Brown and Moyer 2010). While the availability of digitized clinical data portends to provide a mechanism for increasing the number of under-represented individuals in biomedical research, concerns about privacy and confidentiality remain unexplored among the vast majority of individuals who leave behind personal health information.

Participants in this focus-group study of community members residing in a largely ethnic/racial minority neighborhood expressed a complex mix of attitudes toward use of electronic health information for research purposes. They were concerned that secondary use of their personal health information for research recruitment constituted a privacy violation. This sentiment reflected participants' fear, uncertainty, and lack of trust regarding research. These reactions are consistent with findings from previous studies that have found concerns that data would be used to portray communities as having health problems rather than to find solutions to these problems (Fouad et al. 2000; Gooden et al. 2005; Herring et al. 2004).

Nonetheless, if researchers were to have access to personal health information for recruitment and other research purposes, participants preferred that they review digitized clinical files rather than having access to paper-based medical records. Participants seemed to assume that confidential conversations (including non-health-related information) with their health care provider were not recorded in an electronic health record but discoverable only in a paper-based medical record.

With regard to participation in research involving digitized personal health information, participants' overriding concerns related to safety and benefit from the research. Participants noted that their enrollment would be predicated on the importance of the topic of the research to them, their family, or their community, and the likelihood of benefit from their involvement.

As a means of seeking reassurance about the safety of research participation and the trustworthiness of the researchers, these participants wanted face-to-face contact with members of the research team during the recruitment process and full disclosure about how

they were identified to be contacted for participation, including any access that the research team had to their personal health information. Expectations of direct interactions between investigators and potential researchers are typical in Hispanic populations. *Personalismo* or the valuing and building of inter-personal relationships is thought to encourage the development of warm and friendly relationships and deconstruction of overly formal connections (Santiago-Rivera, Arrendondo, and Callardo-Cooper 2002). Moreover, *confianza* (i.e., trust), another familiar concept among Hispanic populations, facilitates both access to potential research participants and *personalismo* with a "special quality of openness" (Lewis-Fernandez and Kleinman 1994, 69).

On the whole, individuals are willing to participate in health-related research, but reasons for non-participation have not been well explored. Mistrust of researchers, aversion to donation of blood specimens, and concerns about potential misuse of information by a third party have been identified as factors for refusal to participate in health-related research (Bussey-Jones et al. 2010; Diaz et al. 2008; Green et al. 2006; Sanner and Frazier 2007). These causes for refusal to participate in research were also expressed in our study.

Researchers have reported high donation rates across ethnic/racial groups in studies that offered participants monetary incentives (Bussey-Jones et al. 2010; McQuillan, Pan, and Porter 2006; Mezuk, Eaton, and Zandi 2008). Our experience with Hispanic populations has shown that incentivizing participation in research can be viewed as demonstrating respect for their experience, expertise, and time.

On the other hand, there are reports of unwillingness to consent to specimen donation, underrepresentation of ethnic/minority/vulnerable populations in specimen collection efforts, or no representation whatsoever among various groups, including Blacks, Hispanics, women, and older adults (Diaz et al. 2008; McQuillan et al. 2006; Mezuk et al. 2008; Sanner and Frazier 2007).

Notably, participants seemed to draw different conclusions about the ownership of health information contained in medical records and biospecimens. Although they felt strongly that they retained ownership of their personal health information and were suspicious of unauthorized access, they generally believed that biospecimens—including samples of blood and tissue taken from their bodies—became the property of the hospital. This distinction appears to be based in a belief that ownership belongs to the holder of knowledge about the meaning of the material in possession (i.e., their firsthand knowledge of symptoms related to hypertension suggests a continued right to control that information, whereas a physician's diagnosis of cancer based on examination of a biopsy implies that the physician has the right to the tissue and to the information derived from it). An alternative explanation could be based in Parry's concept of the "pure gift," which is defined as "altruistic, moral and loaded with emotion" and free from a reciprocal expectation, and a way of solidifying relationships (Parry 1986, 466).

One limitation of this study was the process of sample selection. The participants we engaged included adult community residents of ethnic/minority groups confined to a small geographic area of Manhattan in New York City. Although it would have been valuable to

get a broader perspective by including more residents from other parts of New York City, this sample included predominantly Latino, Spanish-speaking participants. However, our study sample reflects closely the racial/ethnic representation, respectively, in most groups of Washington Heights/Inwood: White (13% vs. 17.6%), Black (17% vs. 7.3%), Asian (0% vs. 2.5%), Latino (67% vs. 71.0%), and Other (7% vs. 1.6%) (New York City Department of City Planning 2010).

A second limitation was that there might have been a self-selection bias of participants who agreed to participate in the study. There was an uneven representation of participants across the four focus groups, with two comprising 10 or more Latino Spanish-speaking participants, one with 4 of 5 White, non-Latino, English-speaking participants (one unreported), and one with 4 Black, English-speaking participants. A third study limitation was that there was a greater representation of older adult (i.e., 41–81+ years of age) compared to younger adult participants (i.e., 18–40 years of age). This age gap may be related to the sites that were targeted for recruitment, as well as the scheduling of focus groups. Future studies should sample from a broader geographic area to address the issue of representativeness in terms of race/ethnicity, cultural background, and age.

In conclusion, this study sought to elicit thoughts and opinions of community members from underrepresented groups that are contributing to the growing pools of digitized health data and are eligible for participation in bio-medical research focused on underserved populations. The thoughts and ideas provided by community members add new knowledge about current and future potential secondary clinical and non-clinical uses of digitized data.

Their perspectives also enrich the discussion of community engagement. Because of the privacy and confidentiality concerns raised by participants, it is important for research organizations and institutional review boards to build trust by developing outreach programs that engage communities in discussion about the use of their health information.

In that regard, the recommendations of the American Medical Informatics Association for a national framework for the secondary use of health data may be helpful (Safran et al. 2007). According to the authors:

> A more transparent dialogue with our citizens concerning the use of their health data is key to maintaining and strengthening the public trust, while enhancing the public's informed actions. (Safran et al. 2007, 7)

Their recommendations include the following: (1) develop transparent policies and practices; (2) focus on data access, use, and control—not on ownership; (3) discuss privacy policies and security; (4) increase public awareness of benefits and challenges; (5) create a taxonomy (i.e., a system of classification for possible non-clinical uses of personal health information); (6) address comprehensively the difficult, evolving questions; and (7) focus national and state attention on secondary use of health data.

By incorporating community members' views about the secondary use of digitized clinical data into data governance policies, organizations can determine the threshold at which data reuse is acceptable to community members. Community engagement can serve as a strategy in the development of policies aimed at preventing privacy and confidentiality violations.

Therefore, increased efforts at including lay community members on hospital boards and academic institutional review boards are needed. Additionally, there is a need for greater emphasis on educating health care providers and academic researchers about community concerns regarding data security. The risks that accompany misuse of personal health information and the fears that deter participation in research can be minimized with shared understanding about appropriate access and uses of stored digitized clinical information.

ACKNOWLEDGMENTS

We are grateful to the community residents of Washington Heights/Inwood for sharing their knowledge, expertise, and experiences with us in this study.

FUNDING

Preparation of the article was supported by the Columbia Center for Health of Urban Minorities (P60MD000206; P60MD000206-08S1; 3P60MD000206-09S2, Jose Luch-singer, PI).

APPENDIX: FOCUS-GROUP PROTOCOL

Research Information-Seeking Behavior

1. How do you find out about research studies going on at the Columbia University Medical Center?

   Probe: flyers, radio/TV/newspaper ads, word-of-mouth, contacted by research study personnel, recruited by health professional, other …

   a. Where do you learn about research studies taking place at the Columbia University Medical Center?

2. Now I have some questions about the decision-making process …

   a. What do you take into consideration when deciding whether or not to join a research study?

   Probe for: health relevance, commitment/extent of involvement/invasiveness of study, potential risk, incentive, perceived benefits, study personnel, etc.

   b. Who, if anyone, do you talk to when you are deciding whether or not to join a research study?

   Probe: relatives, spouses, nurses, pharmacists, friends, physicians/health professionals, research study personnel, etc.

   c. Why do you talk to that person in particular?

3. Now I have some questions about the process of joining a research study …

   a. How have you/do you go about joining a research study?

   Probe: do they call a number/hotline, sign up on the spot, other?

   b. What has been your experience with joining a research study?

   Probe: interaction with research personnel, screening/study criteria, consent form review, etc.

c. Who is your main contact on a research study?

Attitudes Toward Research Participation

4. What do you think are the benefits for community members of participating in medical research studies at the Columbia University Medical Center?

5. What do you think are the barriers for the community of participating in medical research studies at the Columbia University Medical Center?

Medical Records and Research Participation

Columbia University Medical Center stores all patient medical information in a centralized electronic system. By keeping medical information for all patients in a centralized electronic system, all doctors treating a patient at New York Presbyterian Hospital and Clinics have access to the patient's medical history and can be aware of the patient's medical conditions. Researchers, on the other hand, can use this electronic system of medical records to identify those patients who could qualify for their research studies.

6. What would be your reaction if—without your prior knowledge—someone from Columbia University Medical Center contacted you to invite you to participate in a research study based on information from your medical records?

Probe: feelings, thoughts …

a. Does it make a difference if the person who contacts you knows you personally, for example someone you know? How would it make a difference?

b. What do you think about this situation as it relates to the privacy of your information?

Probe: Would you find this to be an invasion of your privacy? Why or why not?

c. Would this be acceptable to you? Under what circumstances, if any?

7. What information about you, if any, could the hospital share with researchers who want to invite you to participate in their studies without getting your prior consent (approval)?

8. If you were the caregiver of a relative who is ill (i.e., has dementia or another illness that impairs their decision making abilities) and researchers have a study that focuses on the caregivers like you, what are ways in which a researcher should contact you about the study?

a. What, if anything, should the researcher know about you before contacting you?

b. Does it make a difference if the person who contacts you knows you personally, for example, your doctor or a health professional you know? How would it make a difference?

c. What do you think about this situation as it relates to the privacy of your information and the information of your family member?

Probe: Would you find this to be an invasion of your privacy? Why or why not?

d. Would this be acceptable to you? Under what circumstances, if any?

Use and Storage of Biospecimens

In many studies researchers are saving biological materials or biospecimens like urine, blood, tissues, or DNA for future research. This way, they will be able to run tests in the future that have not been invented yet and understand more about different conditions and treatments. The specimens are kept in what is called a "biospecimen repository."

9. What do you think about researchers holding on to your biospecimen for future research?

Probe: comfort level, good or bad perceptions, lack of knowledge about purpose of biospecimens, etc.

10. Is there anything else that you would like to share with us about any of the topics we discussed today?

Clinical Data Warehouse for Secondary Use of Personal Health Information

Data warehouse

There is a department at Columbia University that specializes in computerized information. They can create reports from the electronic medical records system using key words. For example, Mr. John, an investigator at Columbia University, needs to find people who have been to the hospital for chest pain and who live in ZIP code 10032. He can ask this group to create a list of all the individuals in the electronic system whose medical records indicate have had chest pain and live in ZIP code 10032. Mr. John can now use the information on this list to contact them and invite them to participate in a research study.

Probe: feelings, thoughts …

a. What do you think about this?

b. Now Mr. John (who is a trained investigator and will maintain the privacy of your records) goes to your doctor's office, obtains your medical record folder and sits down to review the documents, test results, etc. in that folder.

i. What do you think about someone reviewing your medical records *manually*?

ii. What do you think about someone reviewing your medical records *manually* first to see if you have a diagnosis that would make you eligible for the particular study before they contact you to invite you to participate in a study?

c. Now Mr. John (who is a trained investigator and will maintain the privacy of your records) sits at a computer to review your medical records, test results, etc.

i. What do you think about someone reviewing your medical records *electronically,* on a computer?

ii. What do you think about someone reviewing your medical records *electronically* first to see if you have a diagnosis that would make you eligible for the particular study before they contact you to invite you to participate in a study?

11. *Discovery process*: Researchers can use the reports or lists mentioned in the example above to contact potential participants as long as the communication does not include personal health information such as any diagnoses. For example, Mr. John, an investigator at Columbia University, can contact all the people who have been to the hospital for chest pain and who live in ZIP code 10032. He can either mail them a letter or call them. However, in the letter he cannot say "you are receiving this letter because you have been diagnosed with chest pain and you live in ZIP code 10032 so you may qualify for our study" and if he calls, he cannot say "I am calling you today because you have been diagnosed with chest pain and you live in ZIP code 10032 so you may qualify for our study."

Probe: feelings, thoughts …

d. What do you think about this?

e. If you received a letter in the mail addressed to "Your name & Family" & when you opened it, it was a letter inviting you to participate in a study for an illness that it just so happens you had, how would you feel?

12. *Mode of contact*: There are various ways that people can be contacted for research.

a. Does it make a difference to you if the person contacting you is a physician you know, a doctor you don't know, or an investigator (may not be a doctor)?

b. Do you mind if a member of a research team, that is not a doctor, contacts you to invite you to participate in research?

c. How would you prefer that someone contact you? Mail, phone call, e-mail?

d. If you were sent a letter inviting you to participate in a study and informing you that we will be contacting you by phone unless you say otherwise within 10 days, how would you feel? What do you think about this?

e. What do you think about someone only accessing your contact information in your medical record to call you and invite you to participate in a research study? In other words, they have not reviewed your medical history yet.

REFERENCES

Beauchamp, TL.; Childress, JF. Principles of biomedical ethics. 7th. Oxford University Press; New York, NY: 2012.

Berger ML, Mamdani M, Atkins D, Johnson ML. Good research practices for comparative effectiveness research: Defining, reporting and interpreting nonrandomized studies of treatment effects using secondary data sources: the ISPOR good research practices for retrospective database analysis task force report—Part I. Value Health. 2009; 12(8):1044–1052. [PubMed: 19793072]

Bernard, RH. Research methods in anthropology: Qualitative and quantitative aproaches. AltaMira; Walnut Creek, CA: 1995.

Bishop, L.; Holmes, BJ.; Kelly, CM. National consumer health privacy survey 2005. 2005. Available at: http://www.chcf.org/~/media/MEDIA%20LIBRARY%20Files/PDF/C/PDF%20ConsumerPrivacy2005ExecSum.pdf

Blumenthal D. Stimulating the adoption of health information technology. New England Journal of Medicine. 2009; 360(15):1477–1479. [PubMed: 19321856]

Blumenthal D, Tavenner M. The "meaningful use" regulation for electronic health records. New England Journal of Medicine. 2010; 363(6):501–504. [PubMed: 20647183]

Brown M, Moyer A. Predictors of awareness of clinical trials and feelings about the use of medical information for research in a nationally representative US sample. Ethinicity and Health. 2010; 15(3):223–236. A.

Burton LC, Anderson GF, Kues IW. Using electronic health records to help coordinate care. Milbank Quarterly. 2004; 82(3):457–481. I. W. [PubMed: 15330973]

Bussey-Jones J, Garrett J, Henderson G, Moloney M, Blumenthal C, Corbie-Smith G. The role of race and trust in tissue/blood donation for genetic research. Genetics in Medicine. 2010; 12(2):116–121. [PubMed: 20098329]

Connell CM, Shaw BA, Holmes SB, Foster NL. Caregivers' attitudes toward their family members' participation in Alzheimer disease research: Implications for recruitment and retention. Alzheimer Disease and Associated Disorders. 2001; 15(3):137–145. [PubMed: 11522931]

Dennis BP, Neese JB. Recruitment and retention of African American elders into community-based research: lessons learned. Archives of Psychiatric Nursing. 2000; 14(1):3–11. [PubMed: 10692801]

Department of Health and Human Services. Doctors and hospitals' use of health IT more than doubles since 2012. 2013. Available at: http://www.hhs.gov/news/press/2013pres/05/20130522a.html

Diaz VA, Mainous AG, McCall AA, Geesey ME. Factors affecting research participation in African American college students. Family Medicine. 2008; 40(1):46–51. [PubMed: 18172798]

Fouad MN, Partridge E, Green BL, et al. Minority recruitment in clinical trials: A conference at Tuskegee, researchers and the community. Annals of Epidemiology. 2000; 10(8 Suppl):S35–S40. [PubMed: 11189091]

Gooden KM, Carter-Edwards L, Hoyo C, et al. Perceptions of participation in an observational epidemiologic study of cancer among African Americans. Ethnicity and Disease. 2005; 15(1):68–75. [PubMed: 15720051]

Green D, Cushman M, Dermond N, et al. Obtaining informed consent for genetic studies: the multiethnic study of atherosclerosis. American Journal of Epidemiology. 2006; 164(9):845–851. [PubMed: 16928727]

Guba E. Criteria for assessing the trustworthiness of naturalistic inquiries. Educational Communication and Technology Journal. 1981; 29(2):75–91.

Hennink, M.; Diamond, I. Using focus groups in social resarch. In: Memon, A.; Bull, R., editors. Handbook of the psychology of interviewing. Wiley; Chichester, UK: 2000. p. 113-144.I

Herring P, Montgomery S, Yancey AK, Williams D, Fraser G. Understanding the challenges in recruiting Blacks to a longitudinal cohort study: The Adventist health study. Ethnicity and Disease. 2004; 14(3):423–430. [PubMed: 15328945]

Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qualitative Health Research. 2005; 15(9):1277–1288. [PubMed: 16204405]

Krane, D. Many U.S. adults are satisfied with use of their personal health information. 2007. Available at: http://www.harrisinteractive.com/vault/Harris-Interactive-Poll-Research-Health-Privacy-2007-03.pdf

Lewis-Fernandez R, Kleinman A. Culture, personality, and psychopathology. Journal of Abnormal Psychology. 1994; 103(1):67–71. [PubMed: 8040483]

McQuillan GM, Pan Q, Porter KS. Consent for genetic research in a general population: An update on the National Health and Nutrition Examination Survey experience. Genetics in Medicine. 2006; 8(6):354–360. [PubMed: 16778597]

Mezuk B, Eaton WW, Zandi P. Participant characteristics that influence consent for genetic research in a population-based survey: The Baltimore epidemiologic catchment area follow-up. Community Genetics. 2008; 11(3):171–178. [PubMed: 18376114]

Neufeld A, Harrison MJ, Hughes KD, Spitzer D, Stewart MJ. Participation of immigrant women family care-givers in qualitative research. Western Journal of Nursing Research. 2001; 23(6):575–591. [PubMed: 11569331]

New York City Department of City Planning. Table PL-P2 CD: total population, under 18 and 18 years and over by mutually exclusive race and Hispanic origin and total housing units New York City community districts, 1990 to 2010. 2010. Available at: http://www.nyc.gov/html/dcp/html/census/demo_tables_2010.shtml

Parry J. The gift, the Indian gift and the “Indian gift.”. Man. 1986; 21(3):453–473.

Peddicord D, Waldo AB, Boutin M, Grande T, Gutierrez L Jr. A proposal to protect privacy of health information while accelerating comparative effectiveness research. Health Affairs. 2010; 29(11): 2082–2090. [PubMed: 21041752]

Preloran HM, Browner CH, Lieber E. Strategies for motivating Latino couples’ participation in qualitative health research and their effects on sample construction. American Journal of Public Health. 2001; 91(11):1832–1841. [PubMed: 11684612]

Safran C, Bloomrosen M, Hammond WE, et al. Toward a national framework for the secondary use of health data: An American Medical Informatics Association white paper. Journal of the American Medical Association. 2007; 14(1):1–9.

Sandelowski M. Whatever happened to qualitative description? Research in Nursing and Health. 2000; 23(4):334–340. [PubMed: 10940958]

Sandelowski M. What’s in a name? Qualitative description revisited. Research in Nursing and Health. 2010; 33(1):77–84. [PubMed: 20014004]

Sanner JE, Frazier L. Factors that influence characteristics of genetic biobanks. Journal of Nursing Scholarship. 2007; 39(1):25–29. [PubMed: 17393962]

Santiago-Rivera, AL.; Arrendondo, P.; Callardo-Cooper, M. Counseling Latinos and La Familia: A practical guide. Sage; Thousands Oaks, CA: 2002.

Shields AE, Shin P, Leu MG, et al. Adoption of health information technology in community health centers: Results of a national survey. Health Affairs. 2007; 26(5):1373–1383. [PubMed: 17848448]

Sturmer T, Jonsson Funk M, Poole C, Brookhart MA. Nonexperimental comparative effectiveness research using linked healthcare databases. Epidemiology. 2011; 22(3):298–301. [PubMed: 21464649]

Weiner MG, Embi PJ. Toward reuse of clinical data for research and quality improve ment: The end of the beginning? Annals of Internal Medicine. 2009; 151(5):359–360. [PubMed: 19638404]

TABLE I

Inflation-Adjusted Mean and Median Workers' Compensation Costs by Claimant Age and Cost Type

		Age group					
	Total	**18–24 (n = 21,733)**	**25–34 (n = 36,018)**	**35–44 (n = 27,092)**	**45–54 (n = 16,360)**	**55–64 (n = 5,259)**	**65+ (n = 603)**
Type of cost							
Total ($)							
Mean	8,432	4,899	7,439	10,320	12,176	13,194	14,253
(SD)	(37,637)	(31,935)	(34,063)	(39,287)	(48,943)	(44,404)	(37,170)
Median	563	474	544	642	706	775	861
IQR	280–2,022	254–1,143	285–1,671	296–3,059	305–4,707	308–5,464	295–7,056
Medical ($)							
Mean	3,709	2,424	3,284	4,207	5,551	5,632	5,275
(SD)	(20,672)	(14,026)	(16,665)	(17,387)	(35,944)	(25,971)	(14,291)
Median	521	450	507	582	631	674	718
IQR	261–1,450	240–963	267–1,275	274–1,897	278–2,630	279–2,837	268–3,054
Indemnity ($)							
Mean	4,306	2,168	3,661	5,402	5,819	6,762	8,142
(SD)	(21,676)	(20,295)	(20,710)	(24,075)	(19,851)	(24,386)	(25,809)
Median	0	0	0	0	0	0	0
IQR	0–0	0–0	0–0	0–157	0–690	0–1,004	0–2,380

SD, standard deviation; IQR, inter-quartile range.

Costs ($) adjusted for inflation to2010 dollars.n = number of claims.