Chapter 16
Data Decisions and Ethics: The Case of Stakeholder-Engaged Research

Melody S. Goodman, Kristyn A. Pierce, James M. DuBois, and Vetta Sanders Thompson

16.1 Introduction

The established standard frameworks of ethical review used by institutional review boards (IRBs), which are designed to maintain compliance with regulations and to protect individual research participants from harm, do not always adequately address the ethical challenges that arise within research relationships and projects. The shortcomings of the "one-size-fits-all" ethical frameworks used in traditional research are especially apparent regarding data-related decisions in stakeholder-engaged research (SER). Therefore, normative ethical guidelines are needed to ensure ethical conduct related to data ownership, use, publication, and sharing in SER (Wilson et al., 2017).

Kouper and colleagues conducted a study among individuals who work with data, the majority of whom were from social science disciplines, to understand how

M. S. Goodman (✉)
Department of Biostatistics, School of Global Public Health, New York University, New York, NY, USA
e-mail: melody.goodman@nyu.edu

K. A. Pierce
Department of Pediatrics, NYU Langone Health, New York, NY, USA
e-mail: kristyn.pierce@nyulangone.org

J. M. DuBois
Center for Clinical and Research Ethics, Division of General Medical Sciences, Washington University School of Medicine, Saint Louis, MO, USA
e-mail: duboisjm@wustl.edu

V. S. Thompson
Brown School, Washington University in St. Louis, Saint Louis, MO, USA
e-mail: vthompson22@wustl.edu

various stakeholders work together to make decisions regarding data sharing (Kouper et al., 2020). Almost a quarter of participants perceived funding agencies as being responsible for making key decisions regarding data. However, when asked who should take on this responsibility, 65% said individual researchers should be responsible for the governance of the data. The research team ultimately concluded that data governance should be shared by the "data ecosystem" (Kouper et al., 2020). The complexity of this ecosystem is defined by the following:

> 1) the varied and substantial number of organizations and institutions involved, 2) individual actors who act both on behalf of their employing organizations and the larger collective communities, 3) the presence of organizations with economic interests, 4) the belief that everyone is responsible for data governance, 5) and the absence of a shared vision and collective action organizations that balance the influences of the various actors. (Kouper et al., 2020, p. 133)

Overall, data governance in the United States, and globally, is a complex issue with serious implications for society. Here, we will focus on ethical challenges involving ownership, use, and publication of data collected during a stakeholder-engaged research study.

Several ethical challenges may arise during SER, but the potential benefits will outweigh the challenges if researchers plan accordingly. By establishing meaningful partnerships with community stakeholders, researchers can act in a sensitive, ethical manner when collecting, managing, sharing, and publishing data obtained in SER. Investigators should take the time to help the community understand funders' expectations for data ownership and availability. If there are requirements to deposit data, the community must know that the data may be available to be analyzed by parties with whom they have no agreement or relationship and when this might occur. Beyond issues of funding, important negotiation topics include budgets and barriers to community participation in research, data sharing, and dissemination. The discussion is not simply whether data are shared, but also the format and method of sharing. How data and findings are shared affects whether they are useful to the community and meet community needs and SER expectations. Without these discussions, unexpected dissemination of findings may undermine trust and violate informal assumptions and understandings based on SER principles (see Chap. 3, this volume).

16.1.1   Definition of Data

The National Institutes of Health (NIH) defines *data* as "recorded information, regardless of the form or media on which it may be recorded" (National Institutes of Health, 2021, par. 1). Based on this definition, data include "writings, films, sound recordings, pictorial reproductions, drawings, designs, or other graphic representations, procedural manuals, forms, diagrams, workflow charts, equipment descriptions, data files, data processing or computer programs (software), statistical records, and other research data" (National Institutes of Health, 2021, par. 1).

Policies and practices related to data ownership, publication, and sharing must consider all of these forms of data.

16.1.2 Ownership, Stewardship, Use, Publication, Sharing, Repository, and Open Access

Before addressing the myriad ethical challenges surrounding data in SER, we will broadly discuss the terms *ownership*, *steward*, *use*, *publication*, *sharing*, *repository*, and *Open Access*, and their applications to research data (Table 16.1).

First, the entity who has ownership, or the legal right to make decisions regarding the data, may be determined by institutional policies, contracts, case law, and how the project is funded (DuBois et al., 2018a; National Institutes of Health, 2021; Blum, 2015). The owner of research data is usually the institution where the research takes place (Blum, 2015). A data steward is the individual who is responsible for making data findable, accessible, interoperable, and reusable on a long-term basis (Jansen et al., 2019). The challenge with SER is that the data steward (e.g., Principal Investigator), not the data owner (e.g., institution), negotiates with nonacademic stakeholders on how the data will be used, stored, and shared (Glass & Kaufert, 2007; Flicker et al., 2007). Although the data steward controls data use, it is unclear who will make the ultimate decision about data storage and sharing: the data steward or the data owner (Blum, 2015).

Data use refers to the set of processes, conditions, and contexts in which the steward uses the data, such as in certain analyses or to answer various research questions (Coburn & Turner, 2011). In SER, it is important to make sure that all partners have access to the data in useful formats and have permission to use the data for their needs (e.g., program evaluation, advocacy, grant writing, publication). Every partnership should have clear guidelines on how the data will be used and should ensure that the guidelines are appropriate for and applied consistently across all members of the partnership (Minkler, 2004).

Table 16.1 Definitions of key terms related to data

Key term	Definition
Data ownership	Individual or entity that has the legal right to possession of data and is allowed to make decisions regarding data
Data Steward	Individual or entity responsible for activities regarding the long-term care of and access to data
Data use	Various ways in which the owner employs the data
Data publication	Making publicly available the results derived from data
Data sharing	Making data available for use by other individuals or entities
Data repository	A digital "warehouse" where data are deposited, stored, managed, and made available to other individuals or entities
Open access data	Data that are made available to be freely accessed and used without any restrictions

Publication occurs when the data steward has made the results derived from the research study publicly available via scientific journals or other public forums (National Institutes of Health, 2021). The NIH defines data sharing as the process of making data as "widely and freely available as possible while safeguarding the privacy of participants, and protecting confidential and proprietary data" (National Institutes of Health, 2003b, par. 3). A data repository, where data are often shared, is a central domain where datasets may be submitted and then stored, maintained, archived, and shared with other researchers (National Institutes of Health Office of Data Science Strategy, 2022). Finally, open access data are data that are available for free and are unrestricted in terms of how they may be used (United Nations Educational Scientific and Cultural Organization, 2022).

In SER, it is important to make a distinction between data sharing within the project team among the contributing stakeholders and making data publicly available. Sharing data between academic researchers and community partners is inherent in SER, as stakeholders contribute to the interpretation and validation of data (Minkler, 2005). However, broader data sharing—with those who have not been involved in the collection of data—should be discussed with community partners early in collaborations. It may cause some concerns, as community partners may lose control over secondary analyses, yet making the data available for such analyses may be mandated by funding agencies. Case 16.1, "Just Because You Can, Doesn't Mean You Should," illustrates how perspectives on secondary data use may differ between academic and community stakeholders.

Case 16.1: Just Because You Can, Doesn't Mean You Should
Dr. Joseph Gone, a community psychologist and a member of the Gros Ventre tribal community of the Fort Belknap Indian reservation in Montana (Gone, 2017), reflected on a time when during a presentation at a tribal college in Montana, hoping to critically discuss Indigenous research methodologies, he described a Crow Indian religious tradition and included a photo to supplement the description. Conference audience members who were American Indian did not receive his presentation well. Dr. Gone later learned from his audience that his presentation was viewed as unethical, due to Indigenous beliefs that sacred knowledge and sacred objects should not be publicized. In his reflection, he notes that all the historical data he shared, including images of various sacred items, were openly available from public sources, although one audience member suggested that seeing the photograph "felt like violence." In this situation, Dr. Gone grapples with his identity as a part of the American Indian community and his need to be sensitive to their sacred beliefs, while also maintaining his dedication to academic inquiry as a researcher. This account further underscores two key considerations for data sharing. First, the way data are interpreted and shared is meaningful to the community to which they belong. Second, researchers can never be certain of the effects of their research unless they consult with the individuals who created the knowledge in the first place.

16.2 Benefits of Sharing Data

The NIH posits that "data sharing is essential for expedited translation of research results into knowledge, products, and procedures to improve human health" (National Institutes of Health, 2003a, par. 2). The NIH's stance rests on several principles. First, by allowing public access to data, researchers demonstrate transparency as well as their accountability and dedication to sound scientific methods, encourage public oversight, and foster public trust (DuBois et al., 2018a; Langat et al., 2011). Second, when making data publicly available, researchers also pave the way for novel research to be conducted. A variety of researchers can diversify analyses and opinions on a particular set of data and, therefore, support the collaborative nature of contemporary science (DuBois et al., 2018a; Chauvette et al., 2019; Langat et al., 2011). Third, allowing previously collected data to be "re-purposed" in secondary analyses is also economically responsible for institutions, trainees, and those who funded the original research (e.g., taxpayers, health organizations) and demonstrates good stewardship of the data (DuBois et al., 2018a; Chauvette et al., 2019; Langat et al., 2011). Finally, sharing of data supports data democratization and, therefore, has the potential to broaden our ability to produce knowledge and promote social justice for communities and society as a whole (Reynolds & Sariola, 2020; Langat et al., 2011).

16.3 Responsible Use of Data in Stakeholder-Engaged Research

The main goal of SER is to make meaningful improvements for the community or group that is the focus of the research. Therefore, researchers must be mindful of how data are used. To ensure data are used in a way that positively affects the community, researchers must ensure that community partners are actively involved in the dialogue regarding the analysis, interpretation, and dissemination of the study's findings (Minkler, 2004). As the popularity of SER methods grows, so does the demand to share these data with government and funding agencies and to publish findings in academic journals (Fienberg et al., 1985; Office of Science and Technology Policy, 2013; Shelby, 2000). Before making decisions about publication or data sharing, however, research teams must determine and consider who owns the data, possible benefits and burdens of sharing data, and how to share data safely. These considerations may be difficult to discern, as there is no explicit policy in the United States that entails guidelines on how to share data (Langat et al., 2011).

16.3.1 Data Ownership

Before any governing decisions can be made regarding data collected in SER, it is important to identify who has the legal right to make such decisions. Several factors influence who owns data. First, most research grants are given to an institution rather than to an individual or group of researchers. Therefore, when a project is funded by a grant, the data are typically owned by the grantee—that is, the institution that is the primary recipient of funds (DuBois et al., 2018a). According to the NIH Grants Policy Statement, the grant recipient institution owns the rights to the data resulting from the project (National Institutes of Health, 2021). However, when research is funded by an internal grant within an institution, rules regarding ownership are typically dictated by institutional policy (DuBois et al., 2018a) but may be negotiated when third parties (e.g., community partners) are involved (Blum, 2015). Ownership ordinarily confers significant rights and responsibilities, and in SER, community partners usually want some control over how data are used and shared. Negotiating this control can pose challenges.

When project teams discuss data ownership, these conversations are had in good faith. However, in grant-funded research, the players involved in the discussion (e.g., academic researchers, community stakeholders) do not own the data collected; the institution owns the data. For projects not funded by grants but conducted by researchers at academic institutions with institutional resources, data ownership may fall within the university's intellectual property and copyright policies (Carroll, 2015). It is important that researchers are transparent with their stakeholders about such policies and that they continue to act in good faith. It should be clear to stakeholders that their ownership rights to data are equivalent to the ownership rights of the researchers on the study team. It should also be clear to all stakeholders that academic researchers may not have the final decision about ownership rights, as demonstrated in Case 16.2, "Data Ownership of Biological Samples."

Case 16.2: Data Ownership of Biological Samples
A thought-provoking outlook on the complexity of data ownership is offered in the case of Washington University vs. Catalona (Wash. Univ. v. Catalona, 437 F.Supp.2d 985, 1002 [E.D.Mo.2006], 2006). The case was opened to determine the lawful owner of biological samples that individuals had contributed to the Washington University (WU) Genito-urinary (GU) Biorepository for cancer research. The founder of the GU Biorepository, Dr. William J. Catalona, eventually left WU for a different institution and, unable to gain permission to take the samples with him, urged former research participants and patients to consent to the re-location of their samples. Ultimately, the court found that based on the original consent forms signed at WU, WU was the sole owner of the samples. Under the facts of the case, the "donors" of the biospecimens did not retain any ownership that allowed them to authorize the transfer of their samples to a third-party individual. Although the

Case 16.2 (continued)

biological samples were literally part of the participants at one point, the signed informed consent forms at WU ultimately were the basis for determining who owned the data.

Since this case, to avoid future lawsuits, most research institutions have instituted "template" language in their consent forms and contracts clarifying matters of data ownership and sharing. Researchers and community partners may wish to initiate early discussions with oversight offices when changes to template language are desired. When data are complex or qualitative, such as biologic specimens or interview recordings, participants may feel especially "connected" to these data and assume ownership over them. Therefore, in SER, it may be important to revise template language to clearly define who owns the research data and who has the right to grant permission for secondary data users. It may be useful to also attempt to outline the intended lifecycle of these data to ensure there is no confusion among researchers and stakeholders.

16.3.2 Data Decisions

Stakeholder engagement principles direct researchers toward equitable interaction with communities, but what constitutes an equitable interaction is not self-evident (Minkler, 2005; Wallerstein & Duran, 2006). Arguably, equitable interaction requires community participation in decision-making in addition to researcher commitment to share data and study findings with the community in ways that support social change (Ansley & Gaventa, 1997). Increasingly, funders require that grantees share data to the extent possible with the broader research community, so these expectations should then be conveyed to stakeholder partners and included in formal agreements. However, the conditions and requirements of data repositories may affect the ability of community organizations and other stakeholders to access data that they assisted in generating. Community organizations may not be able to demonstrate that they have adequate capacity to secure the data as required, the ability to obtain regulatory approvals, or the background and credentials to conduct analyses deemed of value. For example, it is possible that the community may not have the infrastructure to store and protect the data or an ethics board in place to ensure proper use of the data. Because of these constraints, it may be difficult for researchers to ensure community access to the data, given funders' repository access guidelines. However, if researchers provide support to community organizations to try to overcome these infrastructural and regulatory barriers, repositories may be helpful in promoting ongoing community access to the data. Repositories offer long-term data preservation and sharing, but even still, project teams lose control over who can access the data and what types of questions are explored with the data once the data are made available in a repository. As research environments change, such as the type of community (e.g., grouped by diagnosis or geographic location), type of data being collected, where the data are stored, or research question being explored, other issues and concerns may arise.

Minkler (2004) discusses that a community or community partner organization may desire to suppress findings—for instance, findings that are perceived as negative and/or damaging to the community. In addition to encountering community impressions of the implications of the findings for community well-being and image, researchers may encounter ethical dilemmas, such as community backlash or pressure from industry, if they have not considered the policy impact of their work (Minkler, 2004). On the other hand, suppression of results may be considered antithetical by researchers given the importance of "open science" and transparency in academic research. However, it is reasonable for communities to be concerned about the ways that research, interpretation of findings, and their dissemination may affect legislative and political activities. Yet researchers may be less aware of legislative and policy activities and, thus, less prepared to address community concerns about these issues.

In consideration of the SER call for researchers to commit to work for social change, community voice matters. However, Wang et al. (2016) note that in any partnership, there may be disagreement about the type or scale of the change sought from the project (also see Chap. 2, in this volume). Conflicts related to the nature of the sought change may exist within the community and can lead to differences among stakeholders in their beliefs about how the data should be used, analyzed, published, and shared. It is important for researchers to discern how differing perspectives may affect the work of the partnership and to have processes in place to resolve any conflicts that may arise. It is important that conflict resolution does not mute minority voices, especially related to ethical concerns. Case 16.3, "Balancing the Scales Between Advocate and Academic," outlines a situation when a researcher faced conflict regarding study outcomes as well as conflict regarding protection of study participants. The researcher sought small changes to improve quality of life for a community but was met with backlash when the results of the study catalyzed policy change that could have jeopardized the livelihoods of some community members.

Case 16.3: Balancing the Scales Between Advocate and Academic
A perspective offered by Dr. Steve Wing, an epidemiologist at the University of North Carolina (UNC), sheds some light on the many ethical dilemmas researchers may face in regard to policy and social change (Wing, 2002). In the late 1990s, Dr. Wing and his team partnered with the Concerned Citizens of Tillery (CCT), a local grassroots organization in rural North Carolina, to study the effect of industrialized hog production on environmental health and quality of life for those in the community. The expansion of confined animal feeding operations (CAFOs) at the time in Halifax County caused concerns for residents about economic development, environmental pollution, and contamination of drinking water. With funding from the National Institute of Environmental Health Sciences (NIEHS), the SER project found evidence that CAFOs were more concentrated in low-income, predominately Black areas as well as in areas where residents relied on well water for drinking.

Case 16.3 (continued)

Furthermore, individuals in these areas reported increased negative health symptoms due to CAFO emissions.

With the support of his CCT partners, Dr. Wing presented these findings at a large academic meeting and shared them with the press. While the release of information was both exciting and empowering to his stakeholder partners, Dr. Wing faced severe backlash from pork industry leaders and state legislative bodies. The North Carolina Pork Council demanded copies of all of the study documents under the North Carolina Public Records Statute. Although Dr. Wing was determined to maintain the confidentiality of participants, lengthy legal action eventually led the UNC attorney to agree to release the de-identified study documents. Although the released documents did not contain any explicit identifying information, there remained a concern for loss of community trust and breach of confidentiality due to the sparse population in which the research was conducted.

Dr. Wing describes two complex ethical dilemmas he faced while working on this important environmental justice project. First, the results would differentially impact the community. The outcomes from the project were intended to improve the quality of life of individuals who were ill from the CAFO-produced environmental contaminants. However, the project also had the potential to jeopardize the livelihoods of community members who relied on the pork industry for income. Second, Dr. Wing grappled with the tensions between his obligation to protect the confidentiality and trust of participants and the essential scientific culture of data transparency and availability for re-analysis.

16.3.3   Data Publication Decisions

Project teams should establish dissemination guidelines that address who can publish project information, what information can be shared outside of the project team, when information can be shared, what contributions qualify for authorship on publications, the standard acknowledgements that are included for project publications, who makes decisions about what can be published, and how conflicts about publication will be resolved (Holkup et al., 2004; Minkler, 2005; Israel et al., 1998). It is recommended that partnerships discuss these intended outputs (e.g., publications or press releases) of the project in the early phase of the research project (Prehoda et al., 2019; Glass & Kaufert, 2007). These conversations should result in mutual agreement between researchers and community partners regarding how the data are analyzed, interpreted, and then disseminated. Moreover, community members must be included in the decisions regarding what is published after the project ends.

Both formal and informal ways exist to ensure that publication is mutually agreeable to both the researchers and stakeholder partners. For example, in their collaboration with a Native research team, Holkup and colleagues implemented a policy whereby all team members edit, review, and are listed as co-authors on every publication resulting from their research (Holkup et al., 2004). However, some journals

Table 16.2 Authorship criteria as outlined by ICMJE

Recommended guidelines for authorship (must meet all criteria):
Significant contribution to concept or design of the study; or collection, analysis, or discussion of the data for the manuscript
Critically drafting or editing the manuscript for key intellectual content
Approval of the final body of work to be submitted for publication
Acceptance of accountability for the published work and dedication to ensure any questions regarding the legitimacy or scholarship of the work are resolved appropriately

Table adapted from International Committee of Medical Journal Editors (2022)

require more than recruitment and data collection participation to qualify for authorship, which will exclude some partnership members depending on their specific contributions to the work (International Committee of Medical Journal Editors, 2022). This stipulation leaves the potential for inequity between collaborating parties; therefore, requirements for authorship (Table 16.2) should be conveyed to stakeholders as early as possible in research partnerships and all partners should be given the opportunity to participate in all activities necessary for authorship.

One option in cases where the contributions of individual partners may not warrant authorship per journal guidelines is to consider "corporate authorship" in which one or more byline authors write an article "on behalf of" a group, and group members are all listed within the article; alternately, many journals now publish the names of nonauthor contributors (International Committee of Medical Journal Editors, 2022). Contributors who do not qualify as authors are typically listed either individually, or in a group, under a single heading with their contributions specified (International Committee of Medical Journal Editors, 2022).

Nevertheless, policies must be in place for determining how to proceed when researchers and stakeholders cannot reach an understanding about publication of certain findings. It is necessary for project teams to develop conflict resolution procedures as part of the partnership development process (Minkler, 2004; Mikesell et al., 2013; Han et al., 2021; Wilkins, 2011). A dissemination plan should be created in order to ensure all partners are on the same page regarding the study outputs and so that information that may be harmful to the community is not published (Bodison et al., 2015; Glass & Kaufert, 2007). For example, communities who are already stigmatized, such as racial subgroups or individuals with certain medical diagnoses, would face even more stigma if study results depicted them in a bad light. Ethical SER should always maximize benefits and minimize risks to communities, so the implications of the results of the study should be heavily considered.

16.3.4 Data Agreements

Once the researchers and stakeholders negotiate the data use issues outlined previously, they should create a mandatory formal agreement, or memorandum of understanding (MOU), which outlines how the partnership will handle the data (Glass &

Kaufert, 2007; Flicker et al., 2007; Mikesell et al., 2013). It is recommended that memoranda of understanding detail information regarding data ownership, decision-making processes, roles of each involved party, interpretation and analysis of the data, control and use of the data, and publication and dissemination of the results (Flicker et al., 2007; Glass & Kaufert, 2007). To be consistent with the ever-evolving relationship between researchers and community stakeholders, these agreements should retain some flexibility and be regularly re-negotiated as necessary (McDavitt, 2019).

16.4   Barriers to Responsible Data Sharing

16.4.1   Lack of Infrastructure

Given the varying, sometimes conflicting, guidelines on data sharing, it is common that the individual who is responsible for a dataset, most commonly the Principal Investigator (PI), may end up being the sole steward of said data (Kouper et al., 2020). Data are typically shared via submission to a data repository, along with any contextual materials (e.g., codebooks, field notes, interview recordings, drawings by participants, data collection protocols) required for meaningful interpretation of data (McLeod & O'Connor, 2020; Meyer, 2018). Repositories in the United States and other regions have their own specific data-sharing guidelines, and processes and requirements may differ (DuBois et al., 2018a). Given limited funding for research in general, it may be difficult to allocate research funds—and also time—to properly and responsibly prepare the data for repository sharing (McCurdy & Ross, 2018). McLeod and O'Connor argue that much current guidance regarding data sharing does not acknowledge how unique and complex qualitative data and their accompanying ethical and methodological dilemmas can be; instead, more of the guidance applies to quantitative data (McLeod & O'Connor, 2020). Regulations from the Health Insurance Portability and Accountability Act (HIPAA) specifically list all of the identifiers to remove from the dataset, but these guidelines have varied local interpretations and, thus, leave data sharing, whether qualitative or quantitative, to be prone to risk (United States Department of Health and Human Services, 1996; Freymann et al., 2012). Due to the general lack of consistent guidance, much of the responsibility for data stewardship is placed into the hands of the researcher and, therefore, could leave room for variable, questionable, or even (unintentionally) harmful data-sharing practices (Kouper et al., 2020).

Regardless of funders' guidelines, researchers may be inclined to take individual ownership over the data and to be reluctant to share (Broom et al., 2009). Due to the "competitive reality of science" and the growing demands on researchers from their own institutions, they may also be reluctant to share what they consider to be their hard-earned data (Langat et al., 2011). The withholding of "intellectual property" may be particularly tempting regarding qualitative data collected during SER, due to the investments of time, resources, and effort required to establish meaningful

partnerships and to cultivate the trust required for knowledge sharing (Langat et al., 2011).

Although data sharing is crucial to increasing the production of knowledge and improving public health, sharing data may not be appropriate in some situations. Given the lack of general guidance regarding sharing data with repositories and the concern for breach of confidentiality, researchers partnering with communities may be inclined to request a data-sharing exemption from the funder (Kouper et al., 2020; McLeod & O'Connor, 2020; United States Department of Health and Human Services, 1996; Freymann et al., 2012; Sturges et al., 2015). Investigators may request an exemption due to risk of identification of individuals in small communities or the potential for exploitive analyses and reporting by outside investigators (Holkup et al., 2004; Guishard, 2018). The NIH's Final Policy on Data Management and Sharing, which went into effect in 2023, states that researchers should consider control of access to certain data when re-identification of participants poses the risk of increased harm (National Institutes of Health, 2020). However, the policy lacks explicit guidance on how to properly apply the sharing exceptions or how to request full exemption from data sharing to safeguard sensitive data (U.S. Department of Health and Human Services, 2020). Given the lack of clear-cut guidance on responsible data sharing, it is recommended that investigators become familiar with laws and regulations regarding human subjects' data and that they exercise caution when drafting agreements for data sharing with their community partners. Relevant laws and regulations for research data are outlined in Table 16.3.

Table 16.3 Laws and policies governing the stewardship of human subjects' research data

Law/policy	Agency	Brief summary
Health Insurance Portability and Accountability Act (HIPAA)	U.S. Department of Health and Human Services	Federal law to protect patients' protected health information
Family Educational Rights and Privacy Act (FERPA)	U.S. Department of Education	Federal law to protect students' education records
Federal Policy for the Protection of Human Subjects ("Common Rule")	U.S. Department of Health and Human Services	Federal policy to protect individuals who volunteer to participate in research studies
NIH Data Sharing Policy (effective until January 25, 2023)	National Institutes of Health	Policy requiring all grantees to include a data-sharing plan in their grant proposal and share their final research data
NIH Policy for Data Management and Sharing (effective January 25, 2023)	National Institutes of Health	Update to policy (as above) requiring grantees to submit a Data Management and Sharing plan with their grant application
NSF Data Sharing Policy	National Science Foundation	Policy requiring all grantees to share their research data with other researchers

Sources: National Institutes of Health (2003a, 2020), U.S. Department of Health and Human Services (2016), U.S. Department of Education (2021), National Science Foundation (2022), and United States Department of Health and Human Services (1996)

16.4.2   Ethical Considerations

In addition to the structural challenges that arise regarding the sharing of data, various ethical hurdles must also be considered. Stewards of data collected during community-academic partnered projects have an obligation to ensure that participants' privacy is safeguarded when they make decisions regarding data sharing and dissemination. Concerns regarding re-identification of the data can vary from participant to participant, based on the contextual implications of being identified (Banks et al., 2013). Even if data are de-identified using the HIPAA requirements, recent studies have shown just how easy it can be to re-identify data by using obscure additional data such as genome sequence or credit card metadata (Gymrek et al., 2013; de Montjoye et al., 2015). Chauvette and colleagues posit that full anonymization can never actually be guaranteed (Chauvette et al., 2019).

On the other hand, some research participants may want to be identified with the information that they share with researchers and may even be displeased if their contributions are diminished or even invisible because of de-identification practices (DuBois et al., 2018b). However, a participant can self-identify with published data and even share this information in public domains, which may put other participants and collaborators who want their identities to remain confidential at risk for identification (McCurdy & Ross, 2018). For instance, an individual who discusses their participation in a study on their publicly available social media account may lead to other participants being easily identified. Breaches of confidential information, such as immigration or disease status, can put individuals at risk for legal problems, blackmail, stigmatization, and psychological harms, among others (DuBois et al., 2018a, b; McCurdy & Ross, 2018). For example, individuals participating in a study of criminal behavior could face risk of imprisonment or even the death penalty if they are connected with their data.

In the case of qualitative data collected via key informant interviews or focus groups in community-academic partnerships, the complexity and variability of these data pose additional challenges. First, qualitative data are dependent on the context in which they were collected and the relationships that existed between researchers and the community at the time of data collection (Chauvette et al., 2019). When research participants are also research partners, often many things are left out of the record (DuBois et al., 2018a; McCurdy & Ross, 2018). Although contextual evidence, such as field notes, are encouraged to be archived with qualitative datasets in data repositories, it may be hard for a researcher who was not involved with the project to understand the notes of another researcher. Further, the subjectivity of qualitative data leaves room for bias, which ultimately could lead to misinterpretation of the data and the possibility of negative outcomes for the community (DuBois et al., 2018a; McCurdy & Ross, 2018). Additionally, in secondary analysis, mechanisms are not in place to ensure quality and meaningfulness of the findings (Chauvette et al., 2019). For example, in a re-analysis of their own data after some time had passed, Mauthner et al. (1998) found that even though they were aware of the context in which the data were collected, deriving meaningful

answers to the questions posed in secondary analysis was difficult. There are no "checks and balances" in secondary analysis of data to ensure results are contextualized appropriately.

The duty to disseminate results must be balanced with duties to protect and respect community partners (Minkler, 2004). The risk of data misuse is an important consideration when preparing any type of data to share, especially qualitative data. Current data use policies do not guard against researchers using secondary sources to examine inherently racist hypotheses or research questions, essentially using data resources as "a modern tool of exploitation and scientific racism" (Guishard, 2018, p. 404). In cautious anticipation of such issues when sharing data, researchers should discuss these risks with community stakeholders in the early stages of the project (Minkler, 2004). Furthermore, it is necessary for repositories to consider the potential misuse of data and implement safeguards against this into their processes for accessing data.

16.5   Community Engagement to Strengthen Ethical Data Practices

Engaging with community members is important to ensure data are ethically handled. Because data-related decisions occur at many stages of the research project, there are several opportunities to seek input from community stakeholders to strengthen these decisions. We will focus on the benefits of including community advisors in discussions of informed consent, data collection methodology, and participant confidentiality, as well as the various issues to be considered in these conversations about research data.

16.5.1   Informed Consent

Informed consent documents have implications for data use, ownership, and sharing (DuBois et al., 2018a). When considering open access for research data, it is impossible to predict the future "life cycle" of the data (Chauvette et al., 2019). Therefore, it may not be easy for researchers to adequately inform participants about the specifics of how their data may be used down the line (DuBois et al., 2018a). Further, participants may wish to contribute their insights only to the specific research question outlined in the informed consent process (Chauvette et al., 2019). Case 16.4, "Context of Consent Matters," outlines a situation in which researcher-stakeholder discussions about data ownership could have informed the consent process to prevent unforeseen conflict later in the data lifecycle.

Case 16.4: Context of Consent Matters

Banks et al. (2013) describe the story of a project manager at a museum working on a stakeholder-engaged project with individuals to create their digital stories to be accessioned and displayed by the museum. During the consent process, the museum staff informed potential participants that if they agreed to participate, the museum would have a copyright over the digital material and the stories would be 'fixed,' or unchanged from the original submission. After the project ended and their stories had already been displayed, two participants requested either complete deletion, or amendment to their stories, which challenged the original agreement to the museum's definitive ownership and 'fixity' of the digital stories. It also challenged the manager's dedication to the important tenets of co-production of knowledge in research, forcing them to balance the terms of the project's regulatory agreements while also respecting participants' wishes regarding their data. Ultimately, the situation raised the question about whether the digital data of these participants should be treated as "regular museum objects." This situation raises an interesting dialogue about the ownership of unconventional forms of data. If the museum sought stakeholder input regarding ownership of the stories from study inception, they may have been more equipped to mitigate ethical conflicts down the line. Finally, this case illustrates differences between legal and ethical solutions; if the museum legally owned the stories, they thus had the legal right and arguably the ethical responsibility to consider making exceptions to their initial agreement.

16.5.2   Data Collection Methodology

Community voice (i.e., input from individuals within the community) can affect the research methodology used and ultimately the data collected. Academic researchers may be constrained by their usual ways of collecting data, whereas community members may recognize the need to use alternative strategies to gain information (Black et al., 2018). They may be aware of community-based resources to facilitate data collection (e.g., community theater, religious gatherings, recreational sporting events) as well as alternative methods (e.g., photovoice, digital narratives). Alternatively, community participation in research may be influenced by the research method (Adhikari et al., 2019). In other words, community members may be reluctant to contribute their data based on the protocol and the time and effort required to participate. Engagement in qualitative methods is common, but that does not mean that quantitative strategies cannot sustain community engagement. The measure selection and data collection processes complicate engagement because of time requirements for participation in the decision-making, as well as skill and capacity considerations. Therefore, research teams must engage in conversations with the community during study inception to ensure the protocol includes data collection methods that are reasonable to the community and allot resources to

facilitate this. For example, community members may be able to participate along with organizational members and research consultants as long as they are provided relevant training and adequate compensation for their participation (Israel et al., 1998; Komaie et al., 2017, 2018; Goodman et al., 2018; McGowan et al., 2015; Coats et al., 2015).

16.5.3 Protection of Participants

Meaningful engagement with community co-researchers can increase protection of participants' anonymity. Since regulations direct IRBs to consider protection of individuals, potential risks to the community may not be obvious and, thus, not considered. As experts, stakeholders can inform researchers on certain vulnerabilities of individuals in the community and, therefore, can help identify ways to ensure that these individuals are included with the appropriate protections in place (Adhikari et al., 2019). If a study offers incentives or "singles out" a marginalized community in a way that is perceived to be excessive by outsiders, stigma toward that community may be compounded (Adhikari et al., 2019). In addition, community members can inform researchers about risks that may not be obvious and can help to ensure that the appropriate measures are taken to mitigate these risks (King et al., 2014). For example, stakeholder partners may be able to identify certain data, such as general location or personal narratives, that are not obvious to researchers that could implicate participants, and they could advise that these data are to be omitted in the data de-identification process.

16.5.4 Removing "Non-traditional" Identifiers from the Dataset Before Sharing

It is important to note that meeting HIPAA regulations is a minimum requirement and that these regulations do not represent the gold standard for data de-identification before submission to a data repository, especially regarding SER and qualitative research (United States Department of Health and Human Services, 1996; Freymann et al., 2012). Researchers have described several situations in which they made certain data available that seemed harmless to them, such as metadata in biomedical images, descriptions of experiences, genomic data, and even general location, only to later realize that this jeopardized the anonymity of all study participants (Banks et al., 2013; Chauvette et al., 2019; Wing, 2002; McCurdy & Ross, 2018; Meyer, 2018; Zimmer, 2010; Gymrek et al., 2013). For instance, anonymous users were able to triangulate publicly available, de-identified genomic data with surname and additional metadata, like age and state, to identify the owner of the genomic sequence (Gymrek et al., 2013). It is recommended that investigators share all

reports and outputs with community collaborators to check for identifying information (Holkup et al., 2004). Engaging with community partners during the data preparation phase, and when drafting the data management plan, can ensure that all nonobvious identifiers are omitted from the data.

16.6   The Road to Responsibly Shared Data: Roles and Responsibilities

16.6.1   Role and Responsibilities of the Principal Investigator

As the primary stewards of the data collected during community-academic partnered projects, PIs play a key role in preparing these data for sharing. To start, PIs should include a data-sharing plan in the initial grant proposal; oftentimes, the funder may require this. The NIH recommends that the data-sharing plan outline the "expected schedule for data sharing, the format of the final dataset, the documentation to be provided, whether or not any analytic tools also will be provided, whether or not a data-sharing agreement will be required…and the mode of data sharing" (National Institutes of Health, 2003b). A data-sharing agreement should be considered in all SER partnerships (National Institutes of Health, 2003b; Jarquín, 2012). The data-sharing agreement should be developed with clear communication among all members of the partnership and should contain information that distinctly outlines the intended data-sharing process with well-developed content (Jarquín, 2012). In candid conversations with stakeholders, the PI can gain a better understanding of any indirect identifiers in the data that may not be obvious to remove before sharing and also whether or not they need to request a data-sharing exemption in the case of sensitive content (National Institutes of Health, 2003b; Holkup et al., 2004). Finally, the PI should ensure that the de-identified data are shared in a timely manner. The PI should share the data with community partners as soon as the data are available and with those outside of the partnership when the main findings are accepted for publication (National Institutes of Health, 2003b).

Given the myriad barriers to data sharing, conversations among researchers and stakeholders about potential barriers can begin to diminish the reluctance to share. Researchers can discuss the benefits of data sharing in their respective sector and open a dialogue surrounding standards and best practices for sharing (National Institutes of Health, 2003b). It is important that each of the stakeholders in the partnership has access to the data in a format that is useful for them. Given the time and resources expended by both communities and researchers to facilitate successful SER projects, it is especially important to share SER data to maximize their potential benefits. As SER data are shared, the potential exists for an emergence of common data elements that may aid in data harmonization and may increase partnership synergy and research outputs for the engagement science field (Holt & Chambers, 2017).

Sharing data in a manner that is responsible and respectful can be time-consuming and expensive. Many funding agencies, including the NIH, allow investigators to budget for data-sharing costs. As new requirements go into effect, investigators will want to budget accordingly, both for data-sharing costs and for the time and effort required to obtain input from community partners about the data-sharing plan.

16.6.2   Role and Responsibilities of Funders

While some U.S. funding agencies (e.g., Centers for Disease Control and Prevention, Food and Drug Administration) require grant recipients to submit a data-sharing plan as part of the yearly renewal application, it is clear that this is not yet the norm (Tenopir et al., 2020). In their review of over 300 academic research papers, Terry and colleagues found that 65% of these papers did not include any information about how to locate or access the data (Terry et al., 2018). In other words, almost two-thirds of the papers reviewed did not provide access to the data on which they were written, conflicting with the emerging demand by funders to share data. Funding agencies must engage with researchers and other stakeholders to clearly define the benefit of data sharing, and journals must do their part in encouraging data sharing. It is recommended that funders take steps to support the development of good practice standards for data sharing and for a strong infrastructure to host data (Terry et al., 2018). Funders must also provide adequate funding for data preparation, sharing, and archiving (National Institutes of Health, 2003a, b). The funders who do require data-sharing plans should make them publicly accessible and should even standardize them in order to encourage and aid researchers in their preparation for data sharing (Terry et al., 2018).

Funders may want to consider whether "traditional" data-sharing policies must always apply to SER. Current data-sharing policies are vague and do not specify where or how data should be shared, only that the plan for sharing be included in the application for funding, which is typically before researcher-stakeholder partnerships have had some of these important discussions (Pearce & Smith, 2011; Langat et al., 2011). SER requires trust between researchers and nonacademic stakeholders about who will have access to the data and how the data will be used; ultimately, the timelines for building trust and developing the data-sharing plan are often in conflict. In submitting an official data-sharing plan before funding is granted, the researcher becomes the sole steward of the data, and the tenet of equality in SER may be violated before the partnership even begins (Pearce & Smith, 2011; Kouper et al., 2020). Therefore, funders should allow some flexibility in the timeline to submit the data-sharing plan for SER projects to allow project teams adequate time to discuss and decide on how the data will be shared. It is important that the funder's requirements for data sharing be aligned with stakeholder engagement principles. Funders should allow sufficient time for discussions surrounding data stewardship after the partnership is formed and should allow for flexibility to change and adapt the data-sharing plan as the project progresses (Israel et al., 1998). Researchers

work hard to establish partnerships with nonacademic stakeholders, and these partnerships have been beneficial to both the scientific process and scientific discovery (Balls-Berry & Acosta-Pérez, 2017). Therefore, it is important that data-sharing policies not cause partnership conflicts that can derail progress and damage relationships.

16.6.3   Roles and Responsibilities of Nonacademic Stakeholders

Community voice is important to the process of data sharing, but it must always be balanced against the norms of open science. Community partners and researchers should work together to develop protocols to protect the community from the harmful effects of sharing information inappropriately (Holkup et al., 2004). Community partners act as liaisons between the researchers and the community when negotiating the MOU, the data-sharing plan, and the dissemination plan (Glass & Kaufert, 2007; Flicker et al., 2007; Mikesell et al., 2013; McDavitt, 2019; Bodison et al., 2015). The scientific knowledge held by investigators and the cultural knowledge of stakeholder partners are both necessary to prevent unnecessary repercussions to community members by identifying what could happen if data are released prematurely, if they are released in an insensitive manner, or if the data contain any indirect identifiers (Flicker et al., 2007; National Institutes of Health, 2003b; Holkup et al., 2004). In addition, nonacademic stakeholders are crucial to the dissemination of data in ways that are useful and actionable to the community, key constituencies, and other project stakeholders (Israel et al., 1998; Boyer et al., 2007). Beyond dissemination of results, community members can provide input on interventions or policies that would be beneficial to their community. The nonacademic stakeholders' contribution to data interpretation is inherent in SER, and it is important for stakeholders to contribute to the development of contextual documents for data sharing, including the identification of limitations for secondary analysis.

16.7   Conclusion and Recommendations

It is clear that there is still a lot of work to be done in order to address the ethical concerns and lack of consensus regarding the sharing of data collected in SER. In order to respect community stakeholders while also allowing open access to data to support transparency in science, several recommendations for the ethical use of data in SER have been discussed, at many different levels in the data ecosystem. On a macro scale, regulatory agencies should establish data governance guidelines that are tailored more appropriately to nonquantitative data (especially data in unconventional forms). Institutions should work to reduce barriers for data sharing—for

instance, by increasing funding and resources allotted for data repository preparation (DuBois et al., 2018b). They should also create standard operating procedures to guide researchers in their data-related decisions (Chauvette et al., 2019). Journals and other platforms that disseminate scientific knowledge should provide researchers with adequate time to properly prepare the data for sharing. Additionally, these entities should implement policies to ensure proper credit is given to the original researchers when data are used in secondary analyses (Chauvette et al., 2019). Finally, researchers have an important obligation to engage in an open dialogue with diverse stakeholders in the community (Wilson et al., 2017). The dialogue between researchers and community stakeholders should ultimately result in a consensus regarding the lifecycle of the data collected from the project, which should be reflected in the MOU and other partnership agreements, as outlined below.

16.7.1 Ethically Guided Solutions: Memorandum of Understanding and Partnership Agreements

Formal (e.g., MOU, partnership agreement) and informal (e.g., partnership meeting and communication practices) methods exist for addressing the ethical considerations for research data that emerge in SER. For these methods to reduce ethical missteps, transparency and information sharing are important aspects of implementation. Implications for data use, sharing, and ownership must be considered during discussions regarding funding transparency for both the funder and institution, research implementation, and of course, data management and use. The requirements and details of the discussion will vary, depending on the individual project and setting (Adhikari et al., 2019). Table 16.4 summarizes topics for discussion regarding research data and provides examples of leading questions that should be considered when researchers and stakeholders begin cultivating their partnership and developing their formal agreements.

Although the principal investigator is usually the one to secure and manage funding, it is important to be transparent regarding allocation of funds and to share the responsibility and decision-making power with stakeholder partners (Israel et al., 1998). Budgetary discussions can help teams forecast the costs incurred at several stages of the project, such as compensation for team members and cost of data preparation for sharing (Adhikari et al., 2019; McCurdy & Ross, 2018). Once an understanding is established regarding budget usage, partners will then be able to consider aspects of the project that will affect implementation and conduct of the project such as community representation, social and human capital, and regulatory requirements (Glass & Kaufert, 2007; Holkup et al., 2004; Gilbert, 2006; Adhikari et al., 2019). Ultimately, researcher-stakeholder partnerships should discuss several data-related decisions such as intended outcomes, data ownership and usage, authorship opportunities, and dissemination to ensure that the community's interests are

Table 16.4 Memorandum of understanding and partnership agreements regarding data use, sharing, and publication: ethical considerations and questions to lead discussions between partners

	Topics of importance	Questions to consider
Funding transparency		
Funder	Funder's allowable and nonallowable expenses	How much time and effort will be required to prepare the data for submission to a repository? Is this time and effort accounted for in the budget?
	Funder's policies and implications for individual and organizational partners	Will the community partners have indefinite access to the data? What is the timeline for finalization and submission of the data-sharing plan?
Institution	Options to share indirect amounts	How will the indirect costs be used? Can some of these funds be used to support data preparation activities?
	Institutional policies and implications for individual and organizational partners	What are the institution's requirements for sharing data with outside groups?
Research implementation		
	Community partner participation	Who will represent the community? Who will serve as a liaison? Are community partners aware of the guidelines for data sharing?
	Community partner capacity	Do the community partners have the time and social capital to actively participate? Are the community partners appropriately compensated for their time and contributions to the work (e.g., data-sharing plan discussions, data preparation)?
	Community partner benefit	What are the intended outcomes of this project? Will project data be available in formats that are useful to community partners?
	Institutional review board (IRB) requirements	Is there a single IRB? Which institution(s) require IRB approval? What information is needed to obtain IRB approval? Was data sharing incorporated in the IRB approval?
Data management and use		
	Data ownership and access	At what point are data management and use addressed with community partners? What are the community partner's expectations regarding data use? Who owns the data? Who has access to the data? How is access to the data obtained? Who determines who has access to the data?
	Requirements for data deposits and access	Are the community partners aware/agreeable that the data will be made publicly available? Who is responsible for preparing the data for deposit? Do the community partners have a role or input into data preparation? Who approves the data deposit? Who has access to the deposited data?

Table 16.4 (continued)

	Topics of importance	Questions to consider
	What is disseminated and in what form	Are there any data that should not be shared? Are there any potential unwanted outcomes? How should data be disseminated (e.g., de-identified, aggregate form)? Who determines what data should be disseminated? How are community partner preferences for data dissemination determined? Is the research team responsible for analyses desired/required by community partners and other stakeholders?

respected and that trust is sustained (Banks et al., 2013; Glass & Kaufert, 2007; Holkup et al., 2004; Wang et al., 2016; Minkler, 2005; Prehoda et al., 2019).

Developing and executing MOUs can be time-consuming and cause delays to projects. To address this issue, some institutions are turning toward "accelerated research agreements," because starting from scratch and negotiating each detail leads to a lot of delays and inconsistencies (National Center for Advancing Translational Sciences, 2022). In principle, these agreements also educate community partners on things they might want to think about. Accelerated research agreements can provide efficiency and transparency, and they can protect the interests of partners (Clinical and Translational Science Awards Consortium, 2022; National Center for Advancing Translational Sciences, 2022). However, often, less time and flexibility exist to negotiate elements that individual community partners might want that are tailored to specific details or components of a project.

References

Adhikari, B., Pell, C., & Cheah, P. Y. (2019). Community engagement and ethical global health research, *31*(1), 1–12. https://doi.org/10.1080/11287462.2019.1703504

Ansley, F., & Gaventa, J. (1997). Researching for democracy & democratizing research, *29*(1), 46–53. https://doi.org/10.1080/00091389709603114

Balls-Berry, J. E., & Acosta-Pérez, E. (2017). The use of community engaged research principles to improve health: Community academic partnerships for research. *Puerto Rico Health Sciences Journal, 36*(2), 84.

Banks, S., Armstrong, A., Carter, K., Graham, H., Hayward, P., Henry, A., Holland, T., et al. (2013). Everyday ethics in community-based participatory research. *Contemporary Social Science, 8*(3), 263–277. https://doi.org/10.1080/21582041.2013.769618

Black, G. F., Davies, A., Iskander, D., & Chambers, M. (2018). Reflections on the ethics of participatory visual methods to engage communities in global health research. *Global Bioethics, 29*(1), 22. https://doi.org/10.1080/11287462.2017.1415722

Blum, C. (2015). Access to, sharing and retention of research data: Rights & responsibilities. Council on Governmental Relations, Washington, DC. https://www.cogr.edu/sites/default/files/access_to_sharing_and_retention_of_research_data-_rights_%26_responsibilities.pdf

Bodison, S. C., Sankaré, I., Anaya, H., Booker-Vaughns, J., Miller, A., Williams, P., & Norris, K. (2015). Engaging the community in the dissemination, implementation, and improvement of health-related research. *Clinical and Translational Science, 8*(6), 814–819. https://doi.org/10.1111/cts.12342

Boyer, B. B., Mohatt, G. V., Pasker, R. L., Drew, E. M., & McGlone, K. K. (2007). Sharing results from complex disease genetics studies: A community based participatory research approach. *International Journal of Circumpolar Health, 66*(1), 19–30. https://doi.org/10.3402/ijch.v66i1.18221

Broom, A., Cheshire, L., & Emmison, M. (2009). Qualitative researchers' understandings of their practice and the implications for data archiving and sharing. *Sociology, 43*(6), 1163–1180. https://doi.org/10.1177/0038038509345704

Carroll, M. W. (2015). Sharing research data and intellectual property law: A primer. *PLoS Biology, 13*(8), e1002235. https://doi.org/10.1371/journal.pbio.1002235

Chauvette, A., Schick-Makaroff, K., & Molzahn, A. E. (2019). Open data in qualitative research*, 18*(January), 1–6. https://doi.org/10.1177/1609406918823863

Clinical and Translational Science Awards Consortium. (2022). *Accelerated Clinical Trial Agreement (ACTA)*. https://www.ctsacentral.org/articles/accelerated-clinical-trial-agreement-acta/

Coats, J. V., Stafford, J. D., Thompson, V. S., Javois, B. J., & Goodman, M. S. (2015). Increasing research literacy: The community research fellows training program. *Journal of Empirical Research on Human Research Ethics, 10*(1), 3. https://doi.org/10.1177/1556264614561959

Coburn, C. E., & Turner, E. O. (2011). Research on data use: A framework and analysis. *Measurement, 9*(4), 173–206. https://doi.org/10.1080/15366367.2011.626729

de Montjoye, Y. A., Radaelli, L., Singh, V. K., & Pentland, A. S. (2015). Unique in the shopping mall: On the reidentifiability of credit card metadata. *Science, 347*(6221), 536–539. https://doi.org/10.1126/science.1256297

DuBois, J. M., Strait, M., & Walsh, H. (2018a). Is it time to share qualitative research data? *Qualitative Psychology (Washington, D.C.), 5*(3), 380. https://doi.org/10.1037/qup0000076

DuBois, J. M., Walsh, H., & Strait, M. (2018b). It is time to share (some) qualitative data: Reply to Guishard (2018), McCurdy and Ross (2018), and Roller and Lavrakas (2018). *Qualitative Psychology, 5*(3), 412–415. https://doi.org/10.1037/qup0000092

Fienberg, S. E., Martin, M. E., & Straf, M. L. (Eds.). (1985). *Sharing research data*. National Academy Press.

Flicker, S., Travers, R., Guta, A., McDonald, S., & Meagher, A. (2007). Ethical dilemmas in community-based participatory research: Recommendations for institutional review boards. *Journal of Urban Health: Bulletin of the New York Academy of Medicine, 84*(4), 478–493. https://doi.org/10.1007/S11524-007-9165-7

Freymann, J. B., Kirby, J. S., Perry, J. H., Clunie, D. A., & Carl Jaffe, C. (2012). Image data sharing for biomedical research – Meeting HIPAA requirements for de-identification. *Journal of Digital Imaging, 25*(1), 14–24. https://doi.org/10.1007/s10278-011-9422-x

Gilbert, S. G. (2006). Supplementing the traditional institutional review board with an environmental health and community review board. *Environmental Health Perspectives, 114*(10), 1626. https://doi.org/10.1289/ehp.9005

Glass, K. C., & Kaufert, J. (2007). Research ethics review and aboriginal community values: Can the two be reconciled? *Journal of Empirical Research on Human Research Ethics, 2*(2), 25–40. https://doi.org/10.1525/jer.2007.2.2.25

Gone, J. P. (2017). "It Felt Like Violence": Indigenous Knowledge Traditions and the Postcolonial Ethics of Academic Inquiry and Community Engagement. *American Journal of Community Psychology, 60*(3–4), 353–360. https://doi.org/10.1002/ajcp.12183

Goodman, M. S., Gbaje, E., Yassin, S. M., Dias, J. J., Gilbert, K., & Thompson, V. (2018). Adaptation, implementation, and evaluation of a public health research methods training for youth. *Health Equity, 2*(1), 349. https://doi.org/10.1089/heq.2018.0077

Guishard, M. A. (2018). Now's not the time! Qualitative data repositories on tricky ground: Comment on DuBois et al. (2018). *Qualitative Psychology, 5*(3), 402–408. https://doi.org/10.1037/qup0000085

Gymrek, M., McGuire, A. L., Golan, D., Halperin, E., & Erlich, Y. (2013). Identifying Personal Genomes by Surname Inference. *Science (New York, N.Y.), 339*(6117), 321–324. https://doi.org/10.1126/science.1229566

Han, H. R., Xu, A., Mendez, K. J. W., Okoye, S., Cudjoe, J., Bahouth, M., Reese, M., Lee, B., & Dennison-Himmelfarb, C. (2021). Exploring community engaged research experiences and preferences: A multi-level qualitative investigation. *Research Involvement and Engagement, 7*(1). https://doi.org/10.1186/s40900-021-00261-6

Holkup, P. A., Tripp-Reimer, T., Salois, E. M., & Weinert, C. (2004). Community-based participatory research: An approach to intervention research with a Native American community. *Advances in Nursing Science, 27*(3), 162.

Holt, C. L., & Chambers, D. A. (2017). Opportunities and challenges in conducting community-engaged dissemination/implementation research. *Translational Behavioral Medicine, 7*(3), 389–392. https://doi.org/10.1007/s13142-017-0520-2

International Committee of Medical Journal Editors. (2022). *Defining the role of authors and contributors*. http://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html

Israel, B., Schulz, A., Parker, E., & Becker, A. (1998). Review of community-based research: Assessing partnership approaches to improve public health. *Annual Review of Public Health, 19*, 173–202. https://doi.org/10.1146/annurev.publhealth.19.1.173

Jansen, P., van den Berg, L., van Overveld, P., & Boiten, J.-W. (2019). Research data stewardship for healthcare professionals. In P. Kubben, M. Dumontier, & A. Dekker (Eds.), *Fundamentals of clinical data science* (pp. 37–53) https://library.oapen.org/bitstream/handle/20.500.12657/22918/1007243.pdf?sequence=1#page=42

Jarquín, P. B. (2012). Data sharing: Creating agreements in support of community-academic partnerships. Colorado Clinical and Translational Sciences Institute & Rocky Mountain Prevention Research Center, Aurora, CO. http://trailhead.institute/wpcontent/uploads/2017/04/tips_for_creating_data_sharing_agreements_for_partnerships.pdf

King, K. F., Kolopack, P., Merritt, M. W., & Lavery, J. V. (2014). Community engagement and the human infrastructure of global health research. *BMC Medical Ethics, 15*(1). https://doi.org/10.1186/1472-6939-15-84

Komaie, G., Ekenga, C. C., Sanders Thompson, V. L., & Goodman, M. S. (2017). Increasing community research capacity to address health disparities: A qualitative program evaluation of the community research fellows training program. *Journal of Empirical Research on Human Research Ethics, 12*(1), 55. https://doi.org/10.1177/1556264616687639

Komaie, G., Goodman, M., McCall, A., McGill, G., Patterson, C., Hayes, C., & Thompson, V. S. (2018). Training community members in public health research: Development and implementation of a community participatory research pilot project. *Health Equity, 2*(1), 282. https://doi.org/10.1089/heq.2018.0043

Kouper, I., Raymond, A. H., & Giroux, S. (2020). An exploratory study of research data governance in the U.S. *Open Information Science, 4*(1), 122–142. https://doi.org/10.1515/opis-2020-0010

Langat, P., Pisartchik, D., Silva, D., Bernard, C., Olsen, K., Smith, M., Sahni, S., & Upshur, R. (2011). Is there a duty to share? Ethics of sharing research data in the context of public health emergencies. *Public Health Ethics, 4*(1), 4–11. https://doi.org/10.1093/phe/phr005

Mauthner, N. S., Parry, O., & Backett-Milburn, K. (1998). The data are out there, or are they? Implications for archiving and revisiting qualitative data. *Sociology, 32*(4), 733–745. https://doi.org/10.1177/0038038598032004006

McCurdy, S. A., & Ross, M. W. (2018). Qualitative data are not just quantitative data with text but data with context: On the dangers of sharing some qualitative data: Comment on DuBois et al. (2018). *Qualitative Psychology, 5*(3), 409–411. https://doi.org/10.1037/qup0000088

McDavitt, B. (2019). Dissemination as dialogue: Building trust and sharing research findings through community engagement. *Preventing Chronic Disease, 13*(3). https://doi.org/10.5888/pcd13.150473

McGowan, L. D., Stafford, J. D., Thompson, V. L., Johnson-Javois, B., & Goodman, M. S. (2015). Quantitative evaluation of the community research fellows training program. *Frontiers in Public Health, 3*(July), 179. https://doi.org/10.3389/fpubh.2015.00179

McLeod, J., & O'Connor, K. (2020). Ethics, archives and data sharing in qualitative research, *53*(5), 523–535. https://doi.org/10.1080/00131857.2020.1805310

Meyer, M. N. (2018). Practical tips for ethical data sharing, *1*(1), 131–144. https://doi.org/10.1177/2515245917747656

Mikesell, L., Bromley, E., & Khodyakov, D. (2013). Ethical community-engaged research: A literature review, *103*(12). https://doi.org/10.2105/ajph.2013.301605

Minkler, M. (2004). Ethical challenges for the 'outside' researcher in community-based participatory research. *Health Education & Behavior: The Official Publication of the Society for Public Health Education, 31*(6), 684–697. https://doi.org/10.1177/1090198104269566

Minkler, M. (2005). Community-based research partnerships: Challenges and opportunities. *Journal of Urban Health: Bulletin of the New York Academy of Medicine, 82*(Suppl 2), ii3. https://doi.org/10.1093/jurban/jti034

National Center for Advancing Translational Sciences. (2022). *Accelerated research agreements*. https://www.ara4us.org/

National Institutes of Health. (2003a). *Final NIH statement on sharing research data*. https://grants.nih.gov/grants/guide/notice-files/NOT-OD-03-032.html

National Institutes of Health. (2003b). *NIH data sharing policy and implementation guidance*. https://grants.nih.gov/grants/policy/data_sharing/data_sharing_guidance.htm

National Institutes of Health. (2020). *NOT-OD-21-013: Final NIH policy for data management and sharing*. https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-013.html

National Institutes of Health. (2021). *NIH grants policy statement: 8.2.1 rights in data (publication and copyrighting)*. https://grants.nih.gov/grants/policy/nihgps/html5/section_8/8.2.1_rights_in_data__publication_and_copyrighting_.htm

National Institutes of Health Office of Data Science Strategy. (2022). *Biomedical data repositories and knowledgebases*. https://datascience.nih.gov/biomedical-data-repositories-and-knowledgebases#repositories

National Science Foundation. (2022). *Dissemination and sharing of research results*. https://www.nsf.gov/bfa/dias/policy/dmp.jsp

Office of Science and Technology Policy. (2013). *Increasing access to the results of federally funded scientific research*. https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf

Pearce, N., & Smith, A. H. (2011). Data sharing: Not as simple as it seems. *Environmental Health, 10*(1), 1–7. https://doi.org/10.1186/1476-069X-10-107

Prehoda, E., Winkler, R., & Schelly, C. (2019). Putting research to action: Integrating collaborative governance and community-engaged research for community solar. *Social Sciences, 8*(1), 11. https://doi.org/10.3390/socsci8010011

Reynolds, L., & Sariola, S. (2020, June). The ethics and politics of community engagement in global health research. In *The ethics and politics of community engagement in global health research* (pp. 1–12). https://doi.org/10.1201/9781003011187-1

Shelby, R. (2000). Accountability and transparency: Public access to federally funded research data. *Harvard Journal on Legislation, 37*(2), 369–389.

Sturges, P., Bamkin, M., Anders, J. H. S., Hubbard, B., Hussain, A., & Heeley, M. (2015). Research data sharing: Developing a stakeholder-driven model for journal policies. *Journal of the Association for Information Science and Technology, 66*(12), 2445–2455. https://doi.org/10.1002/asi.23336

Tenopir, C., Rice, N. M., Allard, S., Baird, L., Borycz, J., Christian, L., Grant, B., Olendorf, R., & Sandusky, R. J. (2020). Data sharing, management, use, and reuse: Practices and perceptions of scientists worldwide. *PLoS One, 15*(3), e0229003. https://doi.org/10.1371/journal.pone.0229003

Terry, R. F., Littler, K., & Olliaro, P. L. (2018). Sharing health research data – The role of funders in improving the impact. *F1000Research, 7*. https://doi.org/10.12688/f1000research.16523.2

U.S. Department of Education. (2021). *Family Educational Rights and Privacy Act (FERPA)*. https://www2.ed.gov/policy/gen/guid/fpco/ferpa/index.html

U.S. Department of Health and Human Services. (2016). *Federal policy for the protection of human subjects ('common rule')*. https://www.hhs.gov/ohrp/regulations-and-policy/regulations/common-rule/index.html.

U.S. Department of Health and Human Services. (2020). *Attachment A – NIH data sharing policy: Implications of the NIH draft policy for data management and sharing on data derived from human participants*. https://www.hhs.gov/ohrp/sachrp-committee/recommendations/august-12-2020-attachment-a-nih-data-sharing-policy/index.html#_ftnref9

United Nations Educational Scientific and Cultural Organization. (2022). *What is open access?* https://en.unesco.org/open-access/what-open-access

United States Department of Health and Human Services. (1996). *Summary of the HIPAA privacy rule*. http://www.hhs.gov/ocr/hipaa

Wallerstein, N. B., & Duran, B. (2006). Using community-based participatory research to address health disparities. *Health Promotion Practice, 7*(3), 312–323. https://doi.org/10.1177/1524839906289376

Wang, C. C., Cash, J. L., & Powers, L. S. (2016). Who knows the streets as well as the homeless? Promoting personal and community action through photovoice, *1*(1), 81–89. https://doi.org/10.1177/152483990000100113

Wash. Univ. v. Catalona, 437 F.Supp.2d 985, 1002 [E.D.Mo.2006]. (2006).

Wilkins, C. H. (2011). Communicating results of community-based participatory research. *Virtual Mentor, 13*(2), 81–85. https://doi.org/10.1001/virtualmentor.2011.13.2.ccas1-1102

Wilson, E., Kenny, A., & Dickson-Swift, V. (2017). Ethical challenges in community-based participatory research: A scoping review, *28*(2), 189–199. https://doi.org/10.1177/1049732317690721

Wing, S. (2002). Social responsibility and research ethics in community-driven studies of industrialized hog production. *Environmental Health Perspectives, 110*(5), 437. https://doi.org/10.1289/ehp.02110437

Zimmer, M. (2010). 'But the data is already public': On the ethics of research in Facebook. *Ethics and Information Technology, 12*(4), 313–325. https://doi.org/10.1007/s10676-010-9227-5

Melody S. Goodman is the Vice Dean for Research and Professor of Biostatistics. She is a biostatistician and research methodologist. Her research interest is on identifying origins of health inequities and developing, as necessary, evidence-based primary prevention strategies to reduce these health inequities. She is a Fellow of the American Statistical Association (2021) and the inaugural recipient of the societal impact award from the Caucus for Women in Statistics (2021).

Kristyn A. Pierce is a Research Scientist and a biostatistician. She is interested in the application of statistical machine learning methodologies to understand and address health disparities for individuals with substance use disorder and other mental health disorders.

James M. DuBois is the Steven J. Bander Professor of Medical Ethics and Professionalism, Professor of Medicine and Psychology, and Director of the Center for Clinical and Research Ethics. He is an expert in bioethics and is focused on ethics across mental health research and organ transplantation, professional integrity, and moral development and education. He is the director of the NIH-funded Professionalism and Integrity in Research Program and a founding Editor of *Narrative Inquiry in Bioethics: A Journal of Qualitative Research*.

Vetta Sanders Thompson is the E. Desmond Lee Professor of Racial and Ethnic Diversity and a licensed clinical psychologist, educator, and health service provider. Her research is focused on racial identity, health services access, and socio-cultural determinants of health in diverse communities. She is the inaugural Associate Director of Diversity, Equity, and Inclusion in the Siteman Cancer Center at Washington University School of Medicine and recipient of The St. Louis American's 2022 Lifetime Achiever in Health Care award.