

TRIBAL COMMUNITY PERSPECTIVES ON GENOMICS RESEARCH AND DATA SHARING:
A MIXED-METHODS STUDY

By

Krystal Sara Tsosie

Dissertation

Submitted to the Faculty of the
Graduate School of Vanderbilt University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in

Interdisciplinary

May 31, 2022

Nashville, Tennessee

Approved:

David G. Schlundt, Ph.D.

Jana Lauderdale, Ph.D., R.N., F.A.A.N.

Jada Benn Torres, Ph.D.

Roger Chalkley, Ph.D.

Copyright © 2022 Krystal Sara Tsosie
All Rights Reserved

To my Ancestors, who made this possible
and
For my own F1 generation, Benjamin and William, for making this worthwhile
and finally
To my best friend and husband, Timothy, for your infinite support

ACKNOWLEDGMENTS

This research was supported by the National Institute on Minority Health and Health Disparities of the National Institutes of Health under award number 1U54MD010724, as awarded to the Stanford Precision Health for Ethnic and Racial Equity (SPHERE) Transdisciplinary Collaborative Center. This center award has been instrumental for the creation of the Native BioData Consortium, an Indigenous-led biological and data repository that now stands as its own 501c3 research nonprofit institute to advance precision medicine and genomics research for Tribal nations within the United States. I am so profoundly grateful to serve as one of the co-founders of this amazing Tribal resource alongside my Indigenous scientist peers. Truly, most of my thanks for this research goes to the Tribal community members of the Cheyenne River Sioux Tribe in South Dakota. Their insights and wisdom related to Indigenous data sovereignty and data equity is born from several lifespans of experience, and I hope that this work can adequately reflect even an iota of its due justice.

On this project, more specifically, I am grateful to Dr. Lisa Goldman Rosas for guiding me on approaches for qualitative analyses for this dissertation. I also thank Jan Vasquez for co-moderating the focus groups as part of the second aim. I am extremely grateful to Jonathan Kim, who stepped into the Native BioData Consortium as a community health interventionist and has assisted me in coding transcripts resulting from this study. I am also thankful to Dr. Jennifer Li-Pook-Than who helped coordinate the administrative aspects of the project, especially in terms of navigating multiple institutions and for serving as an advocate for Indigenous-centered research. I am also grateful to Dr. Mike Snyder who continually serves as an advocate for Tribes to exercise sovereignty over their own data. In addition, I am extraordinarily grateful to Joseph Yracheta, my great friend and colleague who kindly introduced me to the larger research consortium and has repeatedly served as a support system as a fellow junior Indigenous scholar.

I am also particularly grateful to my dissertation committee. Drs. David Schlundt, Jana Lauderdale, Jada Benn Torres, and Roger Chalkley have served as my mentors and advocates during the long and difficult process of completing dissertation benchmarks. I am incredibly fortunate that I found such excellent mentors, and I cannot thank them enough for their guidance and patience. It is solely because of their advocacy that I am able to complete this PhD.

TABLE OF CONTENTS

	Page
LIST OF TABLES	vii
LIST OF FIGURES	viii
LIST OF ABBREVIATIONS	ix
1 Introduction	1
1.1 Scope of Work	1
1.2 Overview	1
1.3 Brief Description of Aim 1	2
1.4 Brief Description of Aim 2	3
2 Conceptual Overview	4
2.1 Significance	4
2.1.1 Achieving Health Equity through Data Equity	4
2.1.2 Prior Qualitative Research Studies on Tribal Data Sharing and Precision Medicine	5
2.1.3 Towards Tribally Driven Research	6
2.2 Innovation	7
2.3 Approach	8
2.3.1 The Native BioData Consortium	8
2.3.1.1 Overview	8
2.3.1.2 Consortium History	9
2.3.1.3 Mixed Methods Study Design	9
2.4 Study Partners	10
2.5 Ethical Approval	10
3 Aim 1: Survey	11
3.1 Overview	11
3.2 Methods	11
3.2.1 Eligibility	11
3.2.2 Recruitment	11
3.2.3 Measures	12
3.2.3.1 Participant Demographics	12
3.2.3.2 Past Research Participation	12
3.2.3.3 Factors Related to Willingness to Participate in Research	13
3.2.3.4 Concerns About Genetic Research	13
3.2.3.5 General Genetics Knowledge	13
3.2.4 Definitions	13
3.2.5 Other Methods Considerations	14
3.2.5.1 Reading Level	14
3.2.5.2 Collecting Tribal District Data	14
3.3 Data Analysis	15
3.4 Results	15
3.4.1 General Characteristics	15

3.4.2	Influence of Residence by Tribal District	15
3.4.3	Prior Engagement in Health and Genetics Research	19
3.4.4	Factors and Concerns about Participating in Research	21
3.4.5	General Genetics Knowledge	25
3.5	Discussion	25
3.6	Limitations	26
4	Aim 2: In-Depth Focus Groups	27
4.1	Overview	27
4.2	Methods	27
4.2.1	Eligibility	27
4.2.2	Focus Group Personnel	27
4.2.3	Recruitment of Focus Group Cohorts	27
4.2.4	Moderated Discussions	29
4.2.4.1	Perspectives on Health and Genetics Research	30
4.2.4.2	Data Sharing and Data Deposition	30
4.2.4.3	Tribal Concerns on Data Sharing	31
4.2.4.4	Possible Advantages of Tribal Participate in Research	31
4.2.5	Audio Recordings and Transcription	32
4.2.6	Coding Strategy	32
4.3	Findings	33
4.3.1	Positive and Negative Motivators to Research Participation	33
4.3.2	General Thoughts about Genetics	34
4.3.3	Perspectives on dbGaP and Federal Data Deposition	35
4.3.4	Tribal Concerns on Data Sharing	36
4.3.5	Emerging Themes	36
4.3.5.1	Data Sharing	36
4.3.5.2	Trust	37
4.3.5.3	Research Benefits	39
4.3.5.4	Tribal Regulation	39
4.4	Discussion	40
4.4.1	Limitations and Study Errors	40
4.4.2	Synthesis	40
5	Further Considerations	44
5.1	Strengths	44
5.2	Limitations	44
5.3	Implications for Practice, Policy, and Research	45
A	Participant Survey	46
A.1	Survey Consent Form	46
A.2	Pre-Survey	47
A.3	Survey	48
B	Focus Group Materials	54
B.1	Recruitment Script	54
B.2	Focus Group Demographic Survey	55
B.3	Focus Group Moderator Guide	58
	REFERENCES	63

LIST OF TABLES

Table		Page
3.1	Demographics of Survey Participants	16
3.2	Analyses to Examine Differences, by Tribal District	18
3.3	Participation in Research, by Demographics	20
3.4	Examining Whether Perspectives Differ by Group	24
3.5	General Genetics Knowledge	25
4.1	Codebook	42
4.2	Demographics of Focus Group Participants	43

LIST OF FIGURES

Figure		Page
2.1	Indigenous Community Concerns Related to Genomics Research	5
2.2	Spectrum of Tribal Data Governance	7
2.3	Mixed Methods Study Design: Sequential Explanatory Approach	9
3.1	Map of CRST Tribal Voting Districts	15
3.2	Factors Considered when Deciding Whether to Participate in Health Research	22
3.3	Concerns About Participating in Genetics Research	23
4.1	Flow Chart: Recruitment and Sampling Strategy	29
4.2	Themes and Subthemes	41

LIST OF ABBREVIATIONS

AI/AN	American Indian/Alaska Native
CDC	Centers for Disease Control
CRST	Cheyenne River Sioux Tribe
dbGaP	Database of Genotypes and Phenotypes
GDS	Genomic Data Sharing
GWAS	genome-wide association studies
HGDP	Human Genome Diversity Project
IHS	Indian Health Service
MBIRI	Missouri Breaks Industries Research, Inc.
NBDC	Native BioData Consortium
NIH	National Institutes of Health
REDCap	Research Electronic Data Capture
US	United States

CHAPTER 1

Introduction

1.1 Scope of Work

The survey and focus groups that constitute Aims 1 and 2 as described below are part of a larger body of research from my time in this PhD program at Vanderbilt University. Since Tennessee does not have any state or federally recognized Tribes, my doctoral work consisted of a number of Tribal community-engaged research projects employing qualitative and quantitative approaches in two Tribal nations in the Dakotas region of the United States. Here, I present a project from one of those Tribal communities.

For the project described here, I designed all aspects of the mixed-methods approach and the accompanying study design details, including the aims, scope, methods, creation of qualitative research tools, data collection, and analyses. I wrote the survey and focus group questionnaires and Moderator Guide plus any accompanying forms for the participants. I also wrote all the consent documents and oversaw university and Tribal ethical approvals. For the survey, I cleaned the raw data and performed all statistical analyses, created study figures, and interpreted study results. In addition to designing and co-moderating the focus group cohorts, I oversaw transcription, took a lead on code development and coding, and also performed all qualitative and interpretations. Much of the background materials also reflects my additional work from my time at Vanderbilt University.

1.2 Overview

Precision medicine will likely fail to redress disproportionately high rates of chronic conditions in American Indian and Alaska Native (AI/AN) communities without equitable approaches for the collectivization and sharing of genomics and health data from Indigenous communities (Tsosie et al., 2021b,c). Despite over a couple of decades of large-scale genomics research projects aimed at increasing recruitment of diverse populations (1000 Genomes Project Consortium, 2012; All of Us Research Program Investigators et al., 2019; Cavalli-Sforza, 2005), studies have been unable to successfully engage AI/AN participants (Need and Goldstein, 2009; Bustamante et al., 2011; Popejoy and Fullerton, 2016) due to a variety of factors, including but not limited to a historic distrust of Western science (Faye, 2004; Dukepoo, 1998; Dodson and Williamson, 1999), unethical procurement of genomic data by researchers, cultural incongruity with research aims, and lack of pre-existing Tribal policies and infrastructure to assert Indigenous data sovereignty in the domain of health and genomics data (Tsosie et al., 2021a). Bolstering understanding of these complex issues is increasingly important as biomedical research moves towards broad data sharing policies. While the perspectives

of Indigenous leaders have been sought on research, governance, and data sharing regarding the collection and use of specimens and data, it is equally important to ascertain the perspectives and viewpoints of Indigenous community members directly since this demographic is more representative of Indigenous research participants who are recruited to genomic and health studies.

This study is also important for establishing data sharing and governance policies for the Native BioData Consortium (NBDC), which is the first Indigenous led 501c3 research nonprofit organization that serves as a biological and data repository to advance genomics and health research for Tribal nations in South Dakota. Though it is a new Tribal research entity, the NBDC uniquely leverages a longstanding research collaborative network spanning over 35 years that includes Stanford University and local Tribal research partners within the Cheyenne River Sioux Tribe (CRST). We specifically engaged Lakota enrolled members of the CRST who reside within and proximal to Eagle Butte, South Dakota, which, as of the 2000 United States Census, has the largest population density for this rural region of the State, at 675.4 people per square mile. Most of the residents in this Tribal city are 80.29% Native American, which will enable specific targeting of Indigenous community members.

The National Institutes of Health (NIH) has recognized the importance of respecting Tribal nations' sovereign authority over Tribal members' data. This study has the potential to guide Tribally-led data initiatives, which have not yet existed in the US since the NBDC was established in 2018. Insights gained from this study are likely to signal a new age of positive and equitable engagement of Tribal members in precision medicine research.

Hence, I have designed and conducted a mixed methods study that combines quantitative and qualitative approaches to characterize Tribal community members' thoughts and perspectives on genomics research and data sharing with the following Specific Aims:

1.3 Brief Description of Aim 1

Survey Tribal community members to explore positive and negative factors contributing to trust and participation in genomics and health research.

I administered a Research Electronic Data Capture (REDCap) survey to adult CRST enrolled members (N=100-150) who were conveniently sampled at Tribal health fairs throughout the six CRST Tribal districts. The survey comprised 40 close-ended questions to include: demographic questions, factors driving willingness to participate in research, data privacy and study design concerns, general genetic knowledge, and prior participation in research. Descriptive analyses were conducted using Excel and STATA, additional analyses (t-test, ANOVA, Fisher's exact test) to illustrate any significant subgroup differences and to identify factors statistically associated with willingness to engage in research, if any. *I originally hypothesized that some*

factors (e.g., Tribal districts with higher % non-Indigenous residents) will be associated with increased willingness to participate in research studies.

1.4 Brief Description of Aim 2

Convene focus groups to further explore Tribal community members' perspectives on (a) collection of de-identified information and (b) data sharing and deposition in federal and Tribal databases.

I purposively sampled from each stratum (Tribal district, age, gender, and maximum education attainment) of self-identified individuals from Aim 1 who were interested in participating in, ideally, 4 focus groups of 10 participants each (N=40). In addition to a demographic questionnaire and general genetic knowledge assessment, participants were asked in-depth questions via moderated, 1-hour audio-recorded sessions about sharing of Tribal research data among researchers and opinions on benefits of genomics research. Audio recordings were transcribed (Rev.com) and checked by an independent reader before being coded by reviewers for content analysis using Dedoose. *I hypothesized that Tribal community members would express distrust for data sharing in federal databases compared to Tribally-managed repositories.*

CHAPTER 2

Conceptual Overview

2.1 Significance

2.1.1 Achieving Health Equity through Data Equity

Tribal community members of the US fare worse health outcomes compared to the dominant population across multiple health measures (Jacobs-Wingo et al., 2016). Disparities in health among Tribal members are rooted in contemporary sociocultural and systemic factors, including structural and economic barriers to accessing preventative health resources, persistently underfunded Tribal health systems, and geographic rurality—which are rooted in colonialism and complex histories of Tribal-trust treaty relationships. Though apparent, the true extent of health disparities among Tribal communities is unknown. In addition to under-reporting, a lack of harmonized data capture among heterogeneous Tribal, state, and federal health systems makes it difficult to understand the true extent of health inequities in US Tribal communities. What is readily apparent, however, is the enormous paucity of genomic and health data available on Tribal members to even begin addressing gaps in precision medicine.

Despite efforts to increase diversity in genomics research studies, Indigenous people constitute the lowest represented ethnic minority group in genome-wide association studies (Bustamante et al., 2011; Need and Goldstein, 2009) which are overwhelmingly represented by individuals of European ancestry (Mills and Rahal, 2019). The lack of diversity among research participants might limit the generalizability of genomic-based prevention strategies due to underlying differences in genetic architecture across population groups. Scientists have problematized this inherent bias in precision health in terms of needing to increase “inclusion of diverse populations” via larger sample sizes of individuals recruited to genomic studies to ensure that the benefits of precision medicine are attainable by all (Bentley et al., 2017; Cooke Bailey et al., 2020). Hence, an implicit line has been drawn between increasing inclusion of diverse populations—including Indigenous people—and increasing genomic and health equity.

Merely increasing the availability of genomic data from Tribal communities, however, should not be considered sufficient for redressing concerns related to Tribal members’ lack of engagement in biomedical research studies, which are complex and (especially for US Tribal nations) intricately linked with federally recognized sovereign status (**Figure 1.1**). After all, many of the concerns (e.g., bio-commercialism, bio-prospecting, cultural incongruities, biological reification of negative stereotypes) (Dukepoo, 1998; Dodson and Williamson, 1999; Faye, 2004) that Indigenous people expressed about the Human Genome Diversity

Project of the 1990s to early 2000s resemble similar concerns related to the current, rapid generation and accumulation of sequencing data that is being collectivized in ever-increasingly large datasets. Therefore, tying genomic and health equity to increased recruitment without correcting research practices deemed extractive by Indigenous communities endangers continuation of that same concerns of misappropriation of Indigenous genomes and data. Additionally, telling Indigenous people that they will miss out on precision health benefits if they do not participate in genetics research without also changing underlying power imbalances might cause undue coercion and stricture (Tsosie et al., 2021c).

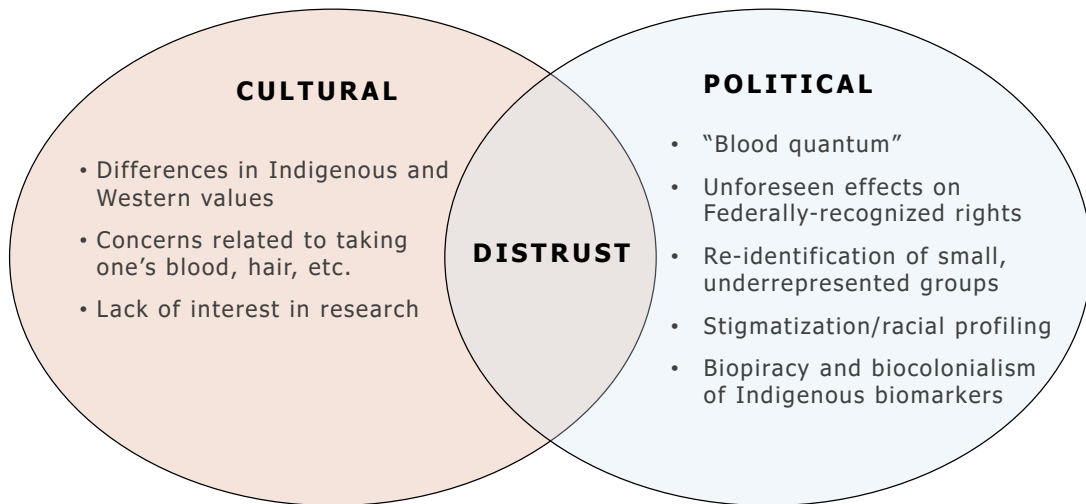


Figure 2.1: Indigenous Community Concerns Related to Genomics Research

Therefore, an important benefit of this study is the potential to foster pathways for empowering Indigenous genomic data sovereignty, defined here as the right of Indigenous nations and people to exercise autonomy to protect their interests related to genomic and health data.

2.1.2 Prior Qualitative Research Studies on Tribal Data Sharing and Precision Medicine

The application and interest of mixed methods studies has been relatively recent in biomedical and health groups, and proposals and training for qualitative methods has only been recently offered by the NIH within the past 8 years (Curry et al., 2013) compared to the long-standing use of these methods in other fields. The NIH initially established its Genomic Data Sharing Policy in 2014 to promote the sharing of large-scale human and non-human genomic data generated from NIH-funded research (Paltoo et al., 2014). It follows then that one of the inaugural qualitative research studies to explore Tribal perspectives on data sharing also occurred in this same relative time frame. A 2014 paper by James et. al. in *Genetics in Medicine* was one of the first studies to explore pathways to building trust in data sharing and control in the context of university-

Tribal research partnerships (James et al., 2014). As follows, **most studies aimed at elucidating Tribal perspectives on data sharing and precision medicine utilize Tribal leaders and Indigenous academics as key knowledge experts.**

Later in 2016, Hudson et. al. engaged Indigenous Advisory Panel members of Te Mata Ira, a three-year project to explore Māori perspectives on genomics research to inform culturally-appropriate guidelines for biobanking (Hudson et al., 2016). In 2019, a study in *Ethnicity and Disease* conducted semi-structured interviews with Tribal leaders, clinicians, researchers, policy makers, and Tribal research review board members about policy suggestions to house, share, and grant access to Tribal data (Garrison et al., 2019). Perspectives from AN leaders and researchers on translation of genomics research to health and data sharing were discussed at a workshop whose dialogue was summarized in 2020 (Hiratsuka et al., 2020). Most recently, a 2022 publication outlines the results of a democratic deliberative approach that assesses Tribal leaders' views on precision medicine research and Tribal oversight (Trinidad et al., 2022). While gauging the thoughts and perspectives of Tribal leaders and Indigenous researchers is an important endeavor, these studies are usually placed in the context of academic settings with previously established relationships with non-Indigenous researchers. Hence, there is an assumption that an equitable research partnership may already be in place that does not reflect the true dynamics of research ongoing in different contexts. **While useful, these studies utilized informed Tribal key experts and should not be considered wholly representative of Tribal community members who might not regularly engage in research discourse.**

The increased numbers of Indigenous researchers who are trained in qualitative approaches has enabled more recent studies that increasingly engage the perspectives of representative Tribal community members. For instance, a 2020 study conducted semi-structured interviews to include N=21 adult AN patients recruited from a Tribal health system to seek community perspectives on communicating about precision medicine (Woodbury et al., 2020). The most highly powered Tribal community-based study to seek insight on issues related to genomics research was more specifically geared towards collecting Navajo community members' insights on the potential changes to the Navajo Nation's moratorium on genetics research. Of N=690 Navajo survey respondents, 59 percent agreed that instituting Tribal data sharing protections was important before allowing genomics research within their community (Claw et al., 2021).

2.1.3 Towards Tribally Driven Research

A recent and solely existent systematic international review of health research literature reveals that 90 percent of publications related to Indigenous data governance were published in the last 8 years (2013-2021) (Griffiths et al., 2021), thus reflecting a heightened awareness by global Indigenous nations to recognize the importance of Indigenous self-determination and self-governance of health research data. Of the publications positing

data governance frameworks that define practices, processes, and roles by which Indigenous data should be governed, only two frameworks explicitly focused on Indigenous data, the CONSIDER statement and the CARE principles (Carroll et al., 2020; Huria et al., 2019). The CONSIDER (CONSolIDated critERtia for strengthening the reporting of health research involving Indigenous Peoples) statement provides a checklist to aid in reporting health research back to the community that engaged in the research and its suggestions are aligned with the general stages of research to include relationship-building, methodologies, analysis and findings, and dissemination. The CARE principles (collective benefits, authority to control, responsibility, and ethics) are largely guiding principles meant to co-exist alongside the FAIR principles (Wilkinson et al., 2016) that encourage open data sharing, but serve to also remind the importance of community-held data as a means of advancing data equity.

As important as it is to create frameworks that contextualize data generation within Indigenous communities, usually these conversations center non-Indigenous community partnerships to originate that data on behalf of Indigenous peoples. Indigenous communities should also be self-mobilizing towards models that govern Tribally-generated data at all phases of research (American Indian Law Center, 1999). While true models of Tribally Driven Research (TDR) do not yet exist, it is reassuring to see that community-based participatory research within Tribal community contexts is increasingly advancing along the spectrum (**Figure 1.2**) towards Tribal data ownership and stewardship.

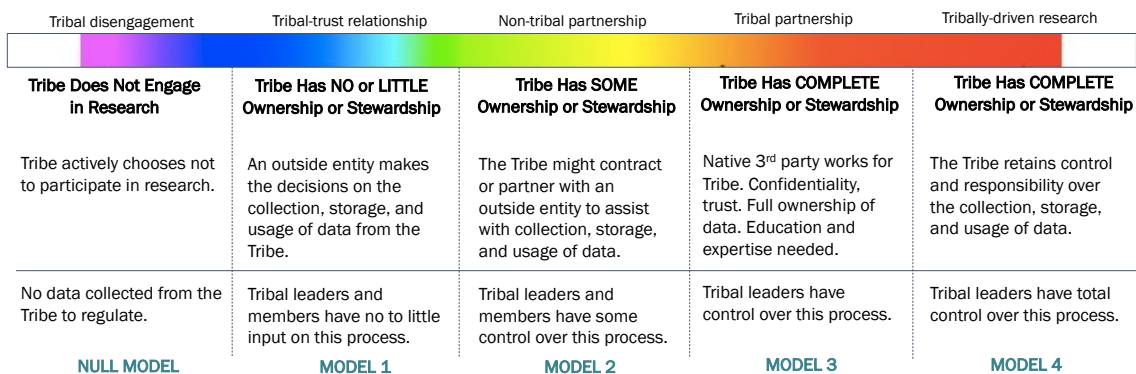


Figure 2.2: Spectrum of Tribal Data Governance

2.2 Innovation

Federal research agencies including the NIH and CDC have recognized the importance of consulting with Tribal partners as part of fostering nation-to-nation relations between the federal government and US Tribes. Effective consultation also entails bidirectional communication, mutual understanding, and truly informed consent and decision-making to respect Tribal rights to self-governance. The NIH, in particular, has recently

solicited public input on its Genomic Data Sharing (GDS) Policy (Office of the Director, 2021) (initially issued in 2014) to ensure that federally-supported biomedical research is responsive to broader policy issues concerning genomic information. As innovations in clinical, research-oriented, and consumer uses of genomic data introduce new changes for the collection, use, and consent of genomic information, the NIH recognizes its responsibility for ensuring that researchers properly balance data sharing and data privacy. However, the issues related to collection of genomic data from AI/ANs are unique and must be carefully considered. While the NIH has also issued a request for input from Tribal nation leaders on its NIH Policy for Data Management and Sharing (Office of the Director, 2022), there needs to also be a more rigorous approach for consulting and engaging Tribal community members on issues related to their data.

It is important to gauge the knowledge of Tribal community members about research uses and interest in their genomic data since Tribal community members, not just Tribal leaders, are largely the ones targeted for recruited into genomic research studies. Furthermore, it is important that any study that aims to gain perspectives on Tribal community members' perspectives on genomic research, data sharing, and data governance is conducted by Indigenous researchers.

2.3 Approach

2.3.1 The Native BioData Consortium

2.3.1.1 Overview

This study serves an important part of the process in developing an Indigenous-led biological and data repository for Tribes in South Dakota that is responsive to Tribal community members' needs. Efforts to build new long-term precision and genomics health resources in these settings must be driven by and overseen by local authorities, where ethical, sovereignty, and logistical issues can be clearly addressed and remedied. Therefore, it is important to ascertain from Tribal community members directly their views on genomics and health research and data sharing as part of the local capacity building for this initiative.

Through NIMHD grant 1U54MD010724-01, the Stanford Precision Health for Ethnic and Racial Equity (SPHERE) Transdisciplinary Collaborative Center was able to facilitate partnerships between the Cheyenne River Sioux Tribe (CRST) and Stanford University to form the Native BioData Consortium (NBDC), a locally-owned and -controlled 501c3 non-profit dedicated to the cultivation of local scientific knowledge and ownership of its products. The NBDC specializes in processing, storage, and archiving of Tribal samples and data. They also help create ethical, policy, and legal best practices for the Tribe and research participants. This has resulted in the creation of a biological lab resource in the Lakota community in South Dakota, which we propose to leverage here as a foundation upon which to build out an Indigenous health research hub.

2.3.1.2 Consortium History

The research relationships between the CRST and Stanford University are greatly facilitated by the previous work of Missouri Breaks Industries Research, Inc. (MBIRI), a private-owned business on the CRST with two satellite locations on the Oglala Sioux Reservation and the Spirit Lake Dakota Reservation, which has 35 years of experience in biomedical and epidemiological research. Therefore, with this long-established and ongoing history of local Tribal community member participation in research, this study is positioned to evaluate Tribal community members' awareness of research projects that have been present within their own Tribal nation.

2.3.1.3 Mixed Methods Study Design

Mixed methods studies, in which quantitative and qualitative research approaches are combined in a single study, can be valuable in biomedical research, where the complementary strengths of each approach can contribute to a greater global understanding on complex phenomena than either approach alone (Curry et al., 2013; Fetters et al., 2013). Here, I describe a sequential explanatory approach (**Figure 1.3**) in which a survey tool was carried out prior to the beginning of the focus group study period. The intention of the two-stage approach was defined a priori, with the use of a common sample (I drew a purposive sample from the focus group based on the quantitative element of the initial survey), and I aimed to integrate both findings into a global understanding of factors contributing to perspectives gained from both Aims of the study.

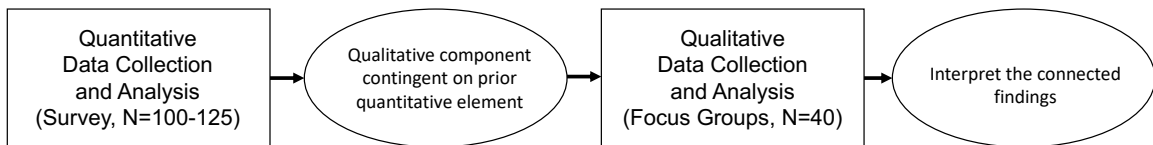


Figure 2.3: Mixed Methods Study Design: Sequential Explanatory Approach

The advantage of a survey study phase is that our research group had the resources to fund up to N=150 Tribal community members' time for a research modality that had a short engagement period of 15-20 minutes or less. Additionally, unlike previous studies (see "Prior Qualitative Research Studies"), it was important for us to determine the extent by which Tribal community members have similar representative views on factors contributing or deterring from genomics and health research participation.

The other advantage of starting with the survey is that it enabled us to more easily design focus group cohorts using demographic data collected from the quantitative survey phase. While focus groups certainly enable in-depth questioning of target topics, it can be difficult to compose focus groups that are representative (Powell and Single, 1996) of the general Tribal community without this type of demographic characteristics

available beforehand. Additionally, by allowing survey respondents to self-designate their willingness to participate in the second phase focus group study, we also alleviated the need to employ more logistically burdensome recruitment tactics such as cold-calling or use of snowball sampling which can result in unrepresentative focus group cohorts. Cold-calling techniques are also generally discouraged from use by Tribal research regulatory boards.

The goal of the focus groups is to facilitate in-depth discussions to further expand on some of the insights potentially gained from survey responses. Since garnering representative Tribal community perspectives is key, it was important that we utilize moderated sessions to ensure that there is some rigor with data collection between one focus group cohort and the next. For this, it was important that we predetermine a Moderator Guide to ensure that the data collection itself does not serve to bias any perspectives or insights gained from the focus groups.

2.4 Study Partners

This study is funded by Stanford University under the BRAICELET (Bio-Repository for American Indian Capacity, Education, Law, Economics and Technology) grant (NIMHD Grant Number 1U54MD010724-01) through the Stanford Precision Health for Ethnic and Racial Equity (SPHERE) center.

Mike Snyder, PhD, served as Principal Investigator (PI) and oversees resource management to the project. Lisa Goldman-Rosas, PhD, served as methods collaborator to guide coding and analyses. As a local Tribal-based research organization, the Missouri Breaks Research Industries Inc. (MBIRI) conducted local recruitment, administered survey materials, and assisted checking transcripts from audio. The Native BioData Consortium serves as the final data repository for the original data. Additionally, the NBDC was the primary site of data analyses. The NBDC conferred with local Tribal research regulatory processes to approve and oversee research conducted on CRST Tribal lands.

2.5 Ethical Approval

Multiple Institutional Review Boards (IRBs) approved this protocol. Stanford University, as the funding institution, approved this study. Upon approval by the CRST Tribal Health committee and full review by the Great Plains Area IRB, the protocol has also undergone approval via an exempt review process by Vanderbilt University.

CHAPTER 3

Aim 1: Survey

3.1 Overview

An important part of the community-engaged process is to seek understanding of the thoughts and perspectives members of the Lakota community have on genomics research and its implications on data sharing and policy development. To gather this information, a REDCap survey of 40 close-ended questions was designed with a series of questions to determine demographic factors (e.g., gender, age in years, CRST Tribal enrollment status, Tribal district residence, maximum level of education attainment, and home Internet accessibility), factors driving willingness to participate in research (such as who is funding the research or types of health-related research questions), data privacy and study design concerns (for instance, ability to access individual results after study completion, Tribal research approval), general genetic knowledge, and prior participation in research.

3.2 Methods

3.2.1 Eligibility

AI adults (over the age of 18) who are enrolled community members of the Cheyenne River Sioux Tribe were eligible. Since we have only research approval from CRST, members of neighboring Tribal nations (such as Pine Ridge or Standing Rock Sioux Nations) were excluded. We asked that employees or close relatives of employees of Missouri Breaks Research Industries, Inc. do not participate. Owing to the longstanding history of Tribal community recruitment into studies by MBIRI, we asked participants if they have ever participated in previous local research studies, however we did not exclude them based off this criterion alone. It was important to ask these questions because past research participants might be more likely to have positive views about engaging in research especially since they previously made the decision to participate in research.

3.2.2 Recruitment

September 2019 (prior to the COVID-19 pandemic) was the last season that permitted ongoing Tribal health fairs within the six CRST Tribal districts. Since the nature of the health fairs was bound to attract Tribal community members interested in Tribal health issues, we decided to utilize preexisting booth space to recruit eligible participants as a convenient strategy. Potentially eligible participants were handed a tablet device preloaded with an offline version of the REDCap survey by MBIRI study personnel. Since the Tribal health fairs

take place in outdoor, rural settings, Internet accessibility was not assured. Hence, syncing locally stored data from the tablet to WiFi after the health fair event ended greatly facilitated easy capture of participants' responses. Eligibility was determined by a pre-screening questionnaire of 5 qualifying questions to ensure participants met eligibility criteria.

3.2.3 Measures

A survey was designed to include: demographic questions, factors driving willingness to participate in research, data privacy and study design concerns, general genetic knowledge, and prior participation in research.

3.2.3.1 Participant Demographics

To keep the expected length for the survey tool as minimal as possible, only basic demographic questions were asked. For instance, general information such as maximum educational attainment (i.e., did not complete high school, some college, and degree level), age, or gender could influence viewpoints on willingness to participate in health and/or genetics research studies.

Initially, when this survey was being developed, we did seek additional CRST Health Committee approval to administer this survey electronically via a website link in order to capture Tribal community participants from all six CRST Tribal districts. Fortunately, we were able to recruit all requisite N=100-150 participants as budgeted for the survey phase of the mixed-methods design from the Tribal health fairs. Hence, there is a question about Internet accessibility in case we needed to employ online recruitment. It is rather useful to ask whether participants' have at-home Internet access as a general indicator of socioeconomic status without asking directly about household income, which could be a sensitive topic for Tribal members who live at or below poverty levels. Inquiring about home Internet access also enables us to plan for future study designs that more heavily deploy Internet surveys as a tool.

3.2.3.2 Past Research Participation

Participants were asked whether they recall ever being asked to participate in a health research study and/or a genetics research study (the present study excluded). For one, it is potentially useful to know if only certain demographic groups are being continually solicited by research groups for participation in studies. This can have implications for the generalizability of past, ongoing, and future studies that feature Tribal members from the CRST Tribal nation.

Participants were further asked if they have agreed or declined participation in health and/or genetics research in the past. Without knowing the details of these studies, merely knowing a participants' underlying

inclination to participate in research can also provide additional insight into responses from other survey questions.

3.2.3.3 Factors Related to Willingness to Participate in Research

The survey also included a series of twelve statements in which respondents were asked to use a Likert-type scale to rate their willingness to participate in studies matching a variety of real-world conditions. Presented as a table of radio buttons, respondents were asked to select from the following choices: "More willing to participate", "Less willing to participate", "I would not participate". These statements described different potential study aims of health-related research, in case a Tribal community member is more or less likely to participate in research that matches their own viewpoints or values. Other statements assessed whether the involvement of certain types of entities (for instance, a university or pharmaceutical company) would affect a respondent's willingness to participate in health and/or genetics research.

3.2.3.4 Concerns About Genetic Research

Following a similar Likert-type rating, respondents were asked to select either "This is a concern", "Not a concern", or "I would not participate in a genetics study" in response to a series of questions about the research study process. This gave participants an opportunity to select whether topics such as data accessibility, data anonymity, Tribal research approval, research benefit, alignment with Lakota culture, or secondary data use were of concern to individuals.

3.2.3.5 General Genetics Knowledge

Having a general understanding of genetics and familiarity with genetics terms might impact participants' understanding and perspectives related to genetics research. Therefore, it is important to assess participants' baseline genetics knowledge in order to better interpret other survey responses in context. There have been a number of previous works that have assessed participants' baseline understanding of genetics via short questionnaires comprising simple "true"/"false" choices to basic genetic facts (Falcone et al., 2011). We included eight short statements that participants must decide are "true" or "false" based on their prior knowledge of genetics and inheritance. Each of these eight statements has a correct answer. The total number of correct answers was then used to indicate a participants' familiarity with genetic concepts. We chose to only include eight statements in order to keep the time length of the survey manageable.

3.2.4 Definitions

We differentiate "health-related research" or "health studies" from "genetic research studies". For purposes of these analyses, it is important to distinguish any studies that generally or broadly relate to human health versus

genetic studies which might be more specific to heritable conditions or diseases or traits. "Health-related research" can broadly encompass studies which may not collect participant specimens (such as saliva or blood samples) but which may collect ethnographic research metrics such as opinions or perspectives, cultural knowledge relating to traditional medicinal practices, patient medical/clinical histories, or routine clinical examination results (e.g., blood pressure measurements). It is important that we make this distinction since "genetic research" may have a negative stigma from particularly Tribal community members. If there are any biases in perspectives related to genetics research, it is important to differentiate any captured responses from health research, generally.

3.2.5 Other Methods Considerations

Participants' perspectives and opinions about health and genetics research were measured across 40 closed-ended survey questions (**Appendix A**). After completing the survey, participants were handed their compensation (\$10) and provided a paper copy of their consent. Participants were also given the option to self-identify their interest as potential focus group participants for Aim 2. The estimated length of the survey was 15-20 minutes or less. The REDCap results were synced to NBDC servers and a de-identified, comma-delimited file was provided for analysis.

3.2.5.1 Reading Level

The Flesch-Kincaid scale rates the survey tool at a seventh grade reading level. The SMOG Index rates the survey at an eighth grade reading level. Based on readability formulas, the text scores generally at the sixth grade level with standard/average readability for those aged between 10-11 years old.

3.2.5.2 Collecting Tribal District Data

An important feature of the demographic questionnaire is the collection of Tribal district residence data from each participant. The CRST is divided into 6 voting districts, with roughly Districts 1, 2, and 3 serving as the western half of CRST and Districts 4 and 6 comprising the eastern half of CRST. District 5 consists of Eagle Butte as the center of CRST and serves as a central demarcation or dividing line between the two halves.

There is a racial and economic divide in that a higher percentage of non-Indigenous, white farmers reside more proximally adjacent to the Missouri River, on the eastern portion of CRST (particularly Districts 4 and 6). The longest occupied regions by Indigenous CRST members have historically resided further away from the river (Districts 1, 2, and 3) due to historic and discriminatory reasons. In order to capture these intra-community factors, we have asked participants in which Tribal district they reside (**Figure 3.1**).

3.3 Data Analysis

The collected data were analyzed using STATA. The data were analyzed using descriptive statistics such as frequency counts, mean, range (minimum and maximum), and standard deviation. Additional tests such as Fisher's exact test and ANOVA were also used as tests of association. We used Fisher's exact tests when testing the association of two categorical variables when the sample sizes or numbers of observations for some stratified data were under N=5. For small sample sizes, it is possible to calculate the significance of deviation from a null hypothesis exactly. When comparing means of a quantitative dependent variable across categorically-defined groups, we used ANOVA.

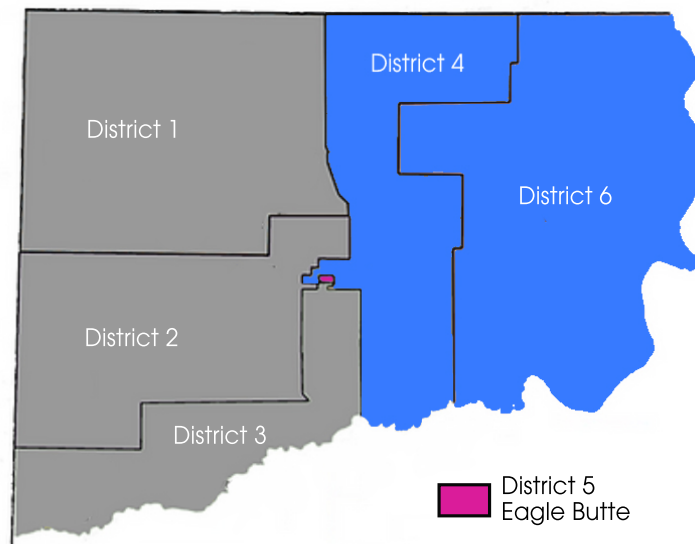


Figure 3.1: Map of CRST Tribal Voting Districts

3.4 Results

3.4.1 General Characteristics

A total of N=125 eligible participants completed the survey. As seen in **Table 3.1**, most of the participants self-identified as female (N=76; 60.8%). The mean age of participants was 42.7 years (range: 19-77). Over half of the participants reported a maximum educational attainment of high school or GED equivalent or lower (N=71; 56.8%) with only N=11 (8.8%) having graduated with a Bachelor's Degree. Only N=70 (56.0%) report having Internet access at home.

3.4.2 Influence of Residence by Tribal District

It was earlier hypothesized that differences by Tribal district would be reflected in the data owing to historic racial and socioeconomic divisions with CRST. More specifically, it was hypothesized that Districts 4 and 6

Characteristics	N	%
Eligible Respondents	125	100.0
Gender		
Male	45	36.0
Female	76	60.8
Other	2	1.6
Prefer Not to Answer	2	1.6
Age (years)		
N	122	
Mean \pm SD	42.7 \pm 13.7	
Median	40.0	
Range (Min - Max)	19 - 77	
Residence, by Tribal Voting District		
District 1	14	11.2
District 2	2	1.6
District 3	12	9.6
District 4	32	25.6
District 5	29	23.2
District 6	28	22.4
I Do Not Know	7	5.6
Prefer Not to Answer	1	0.8
Maximum Educational Attainment		
Did Not Complete High School	19	15.2
High School or GED Equivalent	52	41.6
Vocational or Technical Training	4	3.2
Some College, No Degree	24	19.2
Associate Degree	9	7.2
Bachelor's Degree	11	8.8
Master's Degree or Higher	0	0
Other	3	2.4
Prefer Not to Answer	3	2.4
Home Internet Access		
Yes	70	56.0
No	54	43.2
Prefer Not to Answer	1	0.8

Table 3.1: Demographics of Survey Participants

would present differences compared to Districts 1, 2, and 3.

Based on **Table 3.1** demographics, 48.0% of survey participants reside in the eastern part (Districts 4 and 6) of CRST compared to 22.4% who reside in the westernmost Districts 1, 2, and 3. Twenty-nine (23.2%) participants reside in District 5. District 5 comprises the city center of Eagle Butte, South Dakota, and is only approximately 1.5 square miles in area. District 5 also lies along the boundary between the eastern and western halves of CRST, so it difficult to further dichotomize Districts into a binary variable without further between-group analyses to assess whether stratification by Tribal District residence is meaningful.

Therefore, **Table 3.2** examines whether there are any demographic differences that are structured along District lines. We first conducted a frequency count of participants by gender, age category (years 18-33, 34-

49, 50-65, 66+), home Internet access, and maximum educational attainment. We re-categorized maximum educational attainment into a dichotomous variable ("No College" versus "Some College or Higher"). This allowed us to preserve numbers of observations within each cell and reduce the degrees of freedom to conduct appropriate tests of association, if applicable.

We also began assessing whether participants were more or less likely to be asked to participate in health or genetics research studies based on their Tribal District residence. If there were any associations observed, that could potentially relate to whether a participant's likeliness of being asked to participate in research is linked with geography.

Owing to the small sample size and the numbers of cells in **Table 3.2** with fewer than 5 observations, we utilized Fisher's exact test ($\alpha < 0.05$) to test for any nonrandom associations between Tribal District and the other categorical variables. There were no observed significant associations between Tribal District and gender, age category, or maximum education.

There does seem to be an association between Tribal District residence and being asked to participate in health research ($P=0.005$); likewise, Tribal District residence also appears to be associated with awareness of other ongoing health studies ($P=0.027$). In all other aspects, the influence of Tribal District residence appears to be minimal in this sample size of $N=125$ survey respondents.

	N (%)						Fisher's Exact Test P
	Total 117 (93.6)	District 1 14 (11.2)	District 2 2 (1.6)	District 3 12 (9.6)	District 4 32 (25.6)	District 5 29 (23.2)	
Gender							
Male	42 (35.9)	7	1	5	7	11	11
Female	71 (60.7)	6	1	7	23	17	17
Age (in years), by Category							
18-33	30 (25.6)	3	-	2	7	7	11
34-49	47 (40.2)	6	2	8	11	13	7
50-65	31 (26.5)	4	-	2	10	8	7
66+	9 (7.7)	1	-	-	4	1	3
Maximum Educational Attainment, Dichotomous							
No College	69 (59.0)	6	2	6	17	18	20
Some College or Higher	42 (35.9)	5	-	5	14	10	8
Home Internet Access							
Yes	65 (55.6)	8	1	7	17	15	17
No	51 (43.6)	6	1	5	14	14	11
Have Been Asked to Participate in Health Research Before							
Yes	48 (41.0)	4	1	1	20	9	12
No	47 (40.2)	9	1	7	17	15	17
Engaged in Health Research Before							
Yes	47 (40.2)	4	1	2	20	8	12
No	1 (0.9)	-	-	-	-	1	-
Aware of Other Ongoing Health Research in Community							
Yes	36 (30.8)	5	2	2	15	7	5
No	80 (63.4)	9	-	10	16	22	23
Have Been Asked to Participate in Genetics Research Before							
Yes	21 (17.9)	3	1	2	10	4	1
No	75 (64.1)	9	1	9	17	21	18
Engaged in Genetics Research Before							
Yes	20 (17.1)	3	1	2	10	3	1
No	1 (0.9)	-	-	-	-	1	-

Table 3.2: Analyses to Examine Differences, by Tribal District

3.4.3 Prior Engagement in Health and Genetics Research

Simply being asked to participate in research is different than making the decision to actually agree to participate in research. Whereas Tribal District residence might be associated with being sought for recruitment in a health research study (**Table 3.2**; $P=0.005$), we tested to see if there were any associations with agreeing to participate in a health or genetics study by gender, age category, or college attendance (**Table 3.3**).

In short, we could not observe whether gender, age category, or college attendance was associated with agreeing to participate in health or genetics research. We do observe, however, based on frequency that 47 out of 48 (98.0%) participants who were asked to be part of a study actually agreed to participate.

	Participated in Health Research			Participated in Genetics Research		
	Agreed	Decline	Fisher's Exact Test, P	Agreed	Declined	Fisher's Exact Test, P
Gender						
Male	14	1	0.523	7	-	0.433
Female	33	1		12	2	
Age (in Years), by Category						
18-33	15	2	0.403	6	-	0.787
34-49	13	-		8	1	
50-65	17	-		4	1	
66+	3	-		3	-	
Maximum Educational Attainment,						
Did Not Attend College	26	1	0.587	14	1	0.386
Attended Some College or Higher	19	-		3	1	

Table 3.3: Participation in Research, by Demographics

3.4.4 Factors and Concerns about Participating in Research

Participants were asked whether certain study benefits would contribute to a potential willingness to participate in health research. Perceived benefits of health research include but are not limited to: developing new disease treatments or therapies, advancing research into a condition affecting a loved one, improving existing health services, bringing new job and education opportunities, or fostering improved health services or outcomes to Tribal communities.

Study aims can also be determined by the entities performing the research. Additionally, who conducts the research is an important factor that relates to embedded trust or distrust in the research itself. Participants were also asked whether certain study partners would also affect their willingness to participate in research. Of interest, we asked whether a study led or partnered by a) the Indian Health Service, b) a charity, nonprofit, or patient-advocacy organization, c) university or academic institution, d) for-profit company that is not a drug company, e) federal institution other than the Indian Health Service, or f) drug company would have any bearing on willingness to participate. Since biomedical research corporations can operate under a for-profit or nonprofit structure that can be deterministic of study interests, we differentiated between these two corporate entity types. We also wanted to assess Tribal community members' trust in federal agencies, particularly the Indian Health Service, which provides healthcare to federally-recognized Tribal nations under Tribal-trust relations.

We ranked participant responses according to which factors contributed to the most willingness to participate (**Figure 3.2**). About 84.4% of participants stated that they would be more willing to participate in studies that bring new job and education opportunities in health, closely followed by studies that improve existing hospital or clinical services (82.4%). Studies that relates specifically to helping their Tribal community (81.6%) and will develop new ways to treat or prevent a disease (80.8%) were also ranked highly for participants' willingness to participate in health research. Interestingly, only 67.2% of participants stated that studies researching a disease affecting them or a friend or relative would contribute to a higher willingness to participate.

A larger variation in responses was seen among participants' opinions of study partners. For instance, over 70% of participants stated that they would be more willing to participate in studies led or partnered by the IHS (72.8%) or a charity (70.4%). Lowest ranked are studies led by a for-profit company (54.4%), federal institution other than IHS (52.8%), and drug companies (36.8%). In fact, 20.0% and 24.0% of participants stated that they would not participate in a study led by a for-profit company (that is not a drug company) or explicitly a drug company, respectively. This potentially points to the distrust that Tribal community respondents might have related to corporate interests in biomedical research.

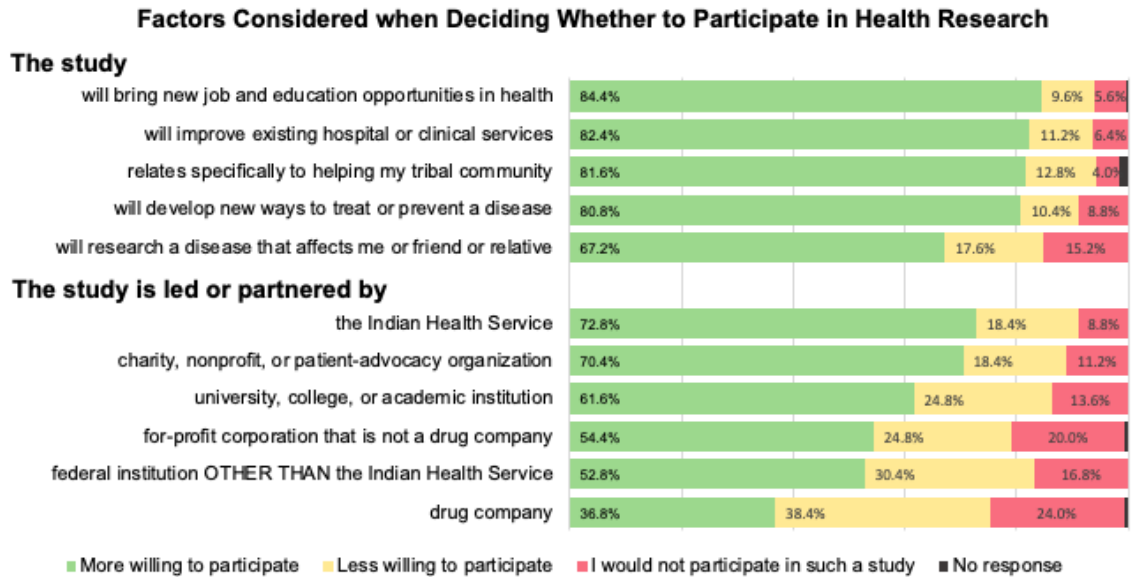


Figure 3.2: Factors Considered when Deciding Whether to Participate in Health Research

We also asked participants whether certain factors about genetics studies are of concern. We isolated a few general topics that have been discussed previously in relation to genetics studies, and we asked participants whether they considered these to be disconcerting or not. Topics included: unspecific broad consent, use of leftover biological materials from sampling, re-identification of anonymized data, who is able to access genetic data, Tribal community research benefit, public accessibility of research products including publications, and whether a study has undergone Tribal review (**Figure 3.3**).

Of highest concern was whether genetic data could be used in other studies without the knowledge of explicit consent of participants (63.2%). Participants also had concerns about what happens to leftover biological materials (61.6%), re-identification of anonymous data (61.6%), and whether study researchers understand Lakota culture (60.8%). Of least concern is whether participants are able to access their individual results after the study is completed (52.8%).

While **Figures 3.2 and 3.3** represent the views of the entire participant cohort, we wanted to determine if there were differences by gender, age category, college attendance, or Tribal District residence. Through a number of Fisher’s exact tests, we determined that, for the most part, there was no association between any of these demographic characteristics and response to factors or concerns (**Table 3.4**). Only the created “Attended College” variable had some association with a couple of factors and concerns.

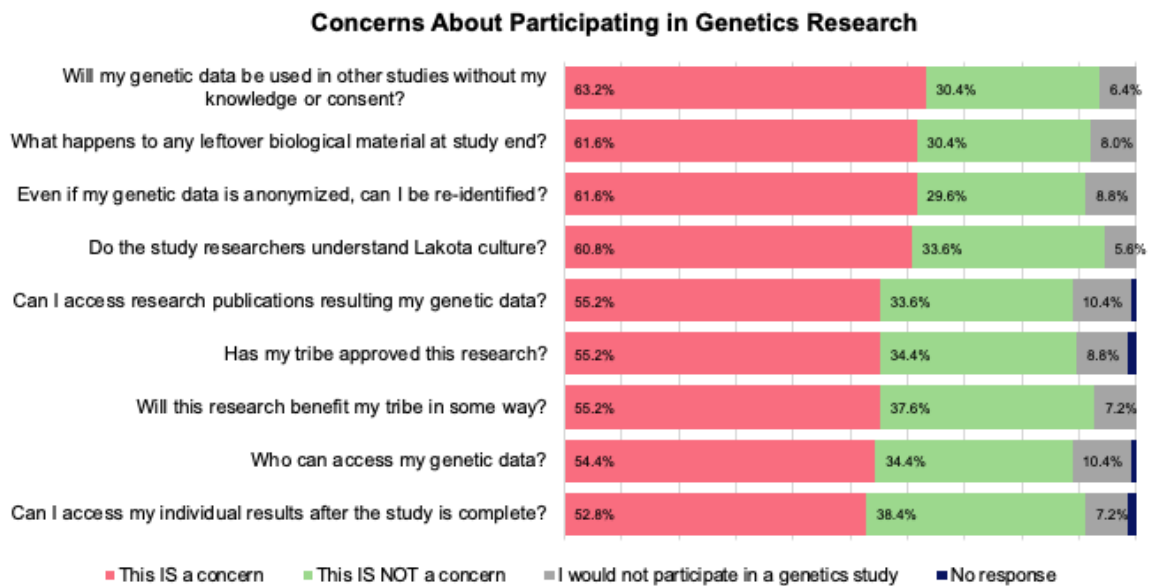


Figure 3.3: Concerns About Participating in Genetics Research

	Gender	Age (in Years), by Category	Attended College	District
I am more willing to participate in a study that:				
will bring new job and education opportunities in health	n.s.	n.s.	n.s.	n.s.
will improve existing hospital or clinical services	n.s.	n.s.	n.s.	n.s.
relates specifically to helping my Tribal community	n.s.	n.s.	n.s.	n.s.
will help develop new ways to treat or prevent a disease	n.s.	n.s.	n.s.	n.s.
will research a disease that affects me or friend or relative	n.s.	n.s.	0.038*	n.s.
is led or partnered by the Indian Health Service	n.s.	n.s.	n.s.	n.s.
is led or partnered by a charity, nonprofit, or patient-advocacy organization	n.s.	n.s.	n.s.	n.s.
is led or partnered by a university, college, or academic institution	n.s.	n.s.	n.s.	n.s.
is led or partnered by a for-profit corporation that is not a drug company	n.s.	n.s.	n.s.	n.s.
is led or partnered by a federal institution OTHER THAN the Indian Health Service	n.s.	0.050*	n.s.	n.s.
is led or partnered by a drug company	n.s.	n.s.	n.s.	n.s.
I agree with the following statement:				
Genetic research can help to improve the health of individuals	n.s.	n.s.	n.s.	n.s.
The following concerns me about participating in genetics research:				
Will my genetic data be used in other studies without my knowledge or consent?	n.s.	n.s.	n.s.	n.s.
What happens to any leftover biological material at study end?	n.s.	n.s.	0.010*	n.s.
Even if my genetic data is anonymized, can I be re-identified?	n.s.	n.s.	n.s.	n.s.
Do the study researchers understand Lakota culture?	n.s.	n.s.	n.s.	n.s.
Can I access research publications resulting my genetic data?	n.s.	n.s.	n.s.	n.s.
Has my tribe approved this research?	n.s.	n.s.	n.s.	n.s.
Will this research benefit my tribe in some way?	n.s.	n.s.	n.s.	n.s.
Who can access my genetic data?	n.s.	n.s.	0.023*	n.s.
Can I access my individual results after the study is complete?	n.s.	n.s.	n.s.	n.s.

Table 3.4: Examining Whether Perspectives Differ by Group

3.4.5 General Genetics Knowledge

We asked participants to determine whether basic statements about genetics were either "true" or "false". Then we assessed the percent correct for each participant and each statement (**Table 3.5**).

Of the 8 statements, N=117 (93.6%) of participants correctly indicated that "Genes are pieces of DNA" is a true statement. Participants also largely understood that healthy parents can have a child with an inherited diseases (84.8% correct), that half of one's genes are inherited from each biological parent (83.6%), and that genetic material is located inside of cells (81.6%). Participants most commonly erred on the statement "All health-related research involves genetics"; only 32.0% of participants indicated the correct choice.

The mean percent correct for all participants was 68.1% ± 16.8%.

True/False Questions (Total Respondents, N=125)	N (%)
1. All health-related research involves genetics. (False)	40 (32.0)
2. Healthy parents can have a child with an inherited disease. (True)	106 (84.8)
3. All serious diseases are inherited. (False)	77 (61.6)
4. A person with an altered (mutated) gene may be completely healthy. (True)	88 (70.4)
5. Genes are pieces of DNA. (True)	117 (93.6)
6. Half of your genes come from your mother and half from your father. (True)	104 (83.2)
7. Genes are inside cells. (True)	102 (81.6)
8. It is the father's chromosomes that decide if a baby is a boy or a girl. (True)	59 (47.2)

Table 3.5: General Genetics Knowledge

In order to determine if there were any group differences in the questionnaire scores, we performed ANOVA t-testing across demographic characteristics such as gender, age, and maximum education attainment (results not shown). Based on ANOVA testing, we show that performance on the general genetic knowledge assessment was associated with educational level ($p=0.0014$). This was that only characteristic associated with performance.

3.5 Discussion

Contrary to the original hypothesis, there were no observed differences between Tribal District residence and participant characteristics such as gender, age, or maximum education level. However, there was an association between Tribal district residence and being asked to participate in health research ($P=0.005$); likewise, Tribal district residence appeared to be associated with awareness of other ongoing health studies in the CRST community ($P=0.027$), assuming a level of significance, $\alpha < 0.05$. In all other aspects, the influence of Tribal district residence appeared minimal. We posit that Eagle Butte's proximity to District 4 likely has an effect on whether residents are likely to learn about ongoing research opportunities. Also, since MBIRI is located in Eagle Butte and has a longstanding history of recruitment in research studies, it is likely that community members already familiar with MBIRI were more likely to visit the Tribal health fairs and

agree to participate in this study.

We also determined that education level may impact participants' decisions to partake in research. We found that participants of higher educational attainment (some college or higher) were more likely to be willing to participate in a study that will research a disease that affects them or friend or relative ($P=0.038$) compared to those with no college experience. We also determined that participants with higher educational attainment were more likely to be concerned about who can access their genetic data ($P=0.023$) and what happens to their leftover biological material at the end of a study ($P=0.010$).

We also found that higher educational attainment was strongly associated with greater general knowledge about genetics ($P=0.0014$). Taken in conjunction with other findings from this study, we can demonstrate that those Tribal community members who were part of this study who have higher awareness about genetics are more likely to express concerns about data access and biospecimen management. This has huge implications for Tribal communities with lower college attainment levels in terms of how risks and benefits about health and genomics studies are disseminated to Tribal community members. It also has implications as Tribal communities increase their own numbers of members with higher education and training, particularly in fields related to science and biomedicine and data.

Trust in the type of entity conducting health research was an important factor for community members on whether to participate in health research. Participants were more likely to be willing to participate in studies led or partnered with trusted organizations such as the IHS or a charity or nonprofit. However, participants were less willing or even unwilling to participate in health research led or partnered with federal institutions other than IHS or drug companies. This is not an unsurprising result; US Indigenous peoples have had contentious relationships with federal agencies over the recent past several decades. It is important to note that Tribal community members are interested in the motives of entities in accessing their data. Hence, the observed wariness of participating in drug company sponsored research might relate to larger concerns of bio-commercialization and privatization of Indigenous data.

3.6 Limitations

The budget for this study limited the sample size to $N=100-150$. While this study size is sufficient for looking at broad perspectives and concerns related to health or genetic research, it limited further stratified analyses to examine the contribution of certain factors that might relate to differences in perspectives. The small frequency counts therefore restrict our ability to use this survey sample to make inferences on other CRST Tribal members without increasing the sample size. However, in terms of understanding motivators contributing to positive viewpoints of health and genetics research, this study was particularly useful.

CHAPTER 4

Aim 2: In-Depth Focus Groups

4.1 Overview

The goals of this aim included exploring Tribal community members' perspectives on the collection of de-identified information and gaining community insights into data sharing and deposition policies in federal and Tribal databases.

4.2 Methods

4.2.1 Eligibility

The eligibility criteria for the focus were the same as the previously defined by the survey study population. AI adults over the age of 18 years old who were enrolled members of the CRST were considered eligible to participate in the focus groups.

4.2.2 Focus Group Personnel

MBIRI personnel who were part of the survey and focus group recruitment were not involved with conducting the focus groups or analyses. Since KST designed the focus groups and the Moderator Script to guide the focus groups, KST decided that another person should serve as the moderator in order not to overly lead the focus group responses. A member of the SPHERE research team (JV) traveled from Stanford University to Eagle Butte, SD, in order to serve in the role as lead moderator. This person has experience in conducting focus group interviews among urban American Indians residing in the Bay Area. Additionally, she is a member of a Tribal nation in New Mexico. JV and KST conferred before the focus groups to ensure that both individuals were in agreement on conduct during the focus groups and to ensure that there was no confusion about the Moderator Guide (**Appendix B**). KST served as the note-taker and transcribed the focus groups live as a backup to the audio recordings of the focus group interviews. Her role as Moderator 2 was to ensure that all focus groups were asked the same questions for continuity.

4.2.3 Recruitment of Focus Group Cohorts

The logistics of recruiting focus group participants without a previously defined study cohort can be challenging. In order to build focus group cohorts that are similar in representation across age, education level, and gender, it was useful to have this kind of simple demographic information available beforehand from a subset of individuals who have expressed interest and availability in participating in more in-depth qualitative

approaches. Hence, we originally aimed to purposively sample focus group participants from the N=125 survey respondents (**Figure 4.1**).

At the end of the survey, we included questions in which participants were given a brief description of the focus group study aims and also days and times in which we planned to conduct the focus groups. If survey respondents were interested and available to be a part of the focus groups, then they could indicate which of the time slots worked for them. We also asked for interested respondents to leave their name and phone number at the end of the survey. These details were kept separate from their survey responses in order to ensure anonymity was maintained. In addition, since respondents encountered these questions through a self-guided REDCap survey on a tablet device, they should have felt no added pressure to express interest in participating in further research since these responses were not being observed by study personnel. A member of the MBIRI research team who was not part of the analyses then uploaded the survey responses from the tablets onto REDCap. They created two separate files: 1) a de-identified, comma delimited spreadsheet that contained the survey responses across participant IDs in which any participant identifiers were removed for data analyses for Aim 1, and 2) a de-identified spreadsheet that included only the demographic metrics (e.g., gender, age, maximum educational attainment, Tribal District residence) of those that indicated their interest and availability to participate in focus groups.

From these details, we aimed to select N=40 individuals into 4 focus groups (10 participants each) with similar representation by gender breakdown, age distribution, similar educational backgrounds, and with participants represented from all Tribal Districts. From the N=78 survey respondents that expressed interest in participating in focus groups, individuals that did not provide complete contact information (such as missing a phone number) were not included as a potential focus group participant. Since we had a majority of interested individuals who were women over the age of 40 years, we assigned these individuals a number and used a random number generator to include them into the focus groups.

A different MBIRI team member who is experienced in recruiting participants was then provided the list of participant IDs, which they matched to the contact information that survey respondents provided in order to schedule the focus groups.

We originally scheduled 4 focus groups comprising a total of N=38 participants over a single weekend in October 2019. These participants confirmed their commitment to attend prior to arrival. However, there was some confusion in the communication as to where the focus groups were to be held. The NBDC is located in a building in Eagle Butte, SD, called the Keya Building. Participants who were unaware of the Keya Building instead went to the similarly named Keya Café in town. Therefore, the first focus group on Friday night only had 5 out of 10 participants attend. After the confusion from the first focus group, the recruiter then recontacted the remaining focus group participants and clarified the study location. We had decided a priori that we

would consider 5 participants to be the minimum focus group size that we would consider convening; if it fell below that minimum number of participants, we would cancel the group interview. Upon recontacting participants, a couple of individuals switched their time slot from Sunday morning to one of the Saturday times. Unfortunately, only 3 individuals arrived for the fourth focus group time which was meant to occur on Sunday morning. We canceled the focus groups and paid participants the expected \$50 to compensate for their time and travel. Only N=20 participants were included in the focus group phase of the study.

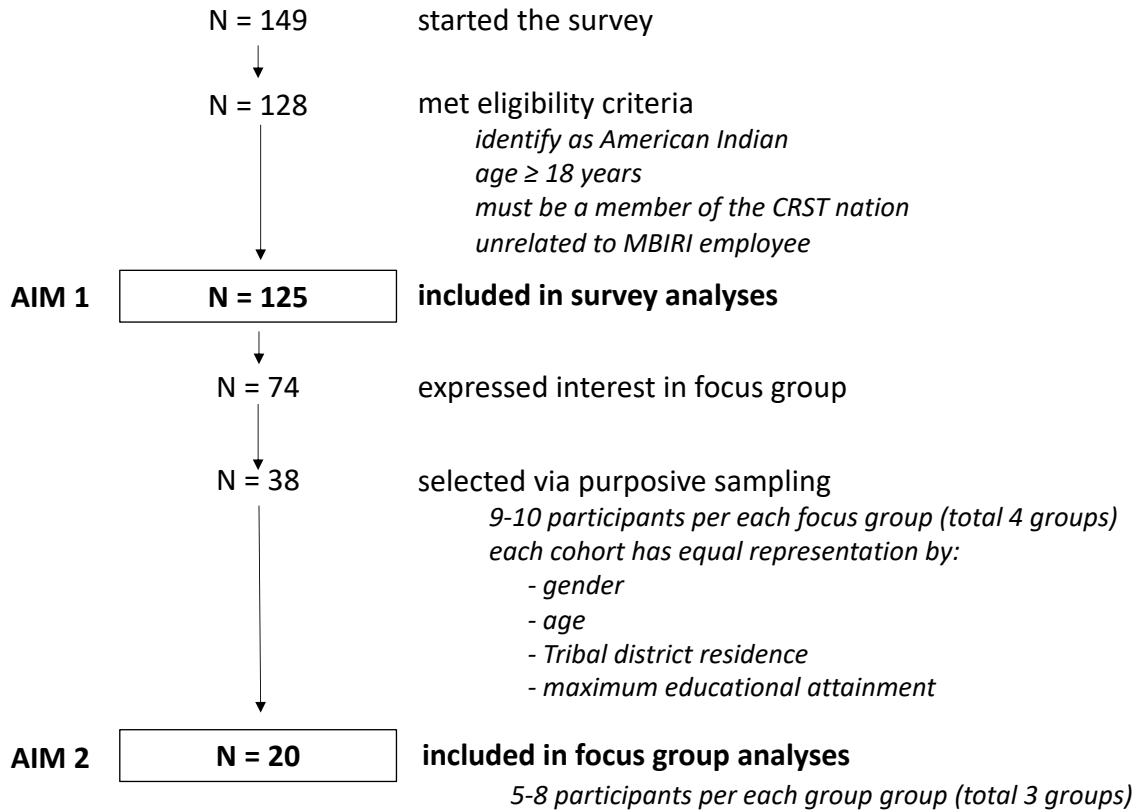


Figure 4.1: Flow Chart: Recruitment and Sampling Strategy

4.2.4 Moderated Discussions

As participants arrived, we asked them to fill out a paper questionnaire (**Appendix B**). This questionnaire included basic demographic questions such as gender, age (in years), Tribal enrollment status, Tribal district, maximum educational attainment, and home Internet access. We also asked participants what their one-way driving distance was on the day of the focus group. Since the information from the surveys were not linked with the focus groups, gathering these details again from participants was essential for understanding within-group and between-group characteristics.

Moderator 1 was advised to utilize the Moderator Guide (**Appendix B**) to lead the focus group discus-

sions. The Moderator Guide outlines a script to guide all aspects of the discussion: introduction of study personnel in the room, study objectives, "ground rules" to ensure everyone has been informed of their rights as research participant and for respectful conversation. The Guide also includes suggested conversation probes. Once participants were verbally consented (and provided an opportunity to exit the room if they wanted to decline participation), the audio recorders were turned on.

4.2.4.1 Perspectives on Health and Genetics Research

To begin the conversation, participants were asked if they have ever previously participated in a health-related research study. If so, participants were encouraged to describe motivators or driving factors that contributed to their decision to participate or not to participate in health research. Participants were also asked to recall any barriers that may have prevented them or other Tribal community members from participating in research. In recalling study details, participants were also asked whether there were particular research questions that interest them or would be of benefit to the community, for instance, in studying a disease or condition that affects their community or family members.

Following this, participants were asked to expand upon their perspectives and attitudes and opinions about genetics research. The moderators were specifically advised to probe for participants' understanding and knowledge about genetics. Of additional interest was whether participants had any concerns about genetics research being conducted in their community.

4.2.4.2 Data Sharing and Data Deposition

The next part of the discussion devoted a significant amount of time to assess perspectives related to data sharing and deposition of data in repositories. The Moderator Guide was particularly helpful in describing to participants, in a neutral and unbiased manner, the utility and interest in collecting health and genetic information from large samples of people across harmonized datasets to facilitate the study of factors contributing to disease. In addition, the Guide outlined the use of de-identified data (meaning no personal identifiers) to protect privacy and ensure confidentiality of private health information.

After hearing this initial description, participants were asked for their feelings and opinions about the process of collectivizing and harmonizing data. They were also asked if they would participate in a research study that collected health information that was de-identified. Of interest was to learn participants' thoughts about the sharing of data across researchers or collaborators. We wanted to understand if there were people or organizations in which participants would feel uncomfortable having their data shared.

The Moderators were guided on how to describe the interest of the NIH to institute data sharing policies that require the deposition of data of federally-funded research into a publicly accessible database such as

dbGaP (database of Genotypes and Phenotypes). The Guide also described the number of protections that are in place to ensure participants' information is secure, such as:

- All data stored must have personal identifiers removed
- Researcher access is limited to those who have been thoroughly checked by a governing committee
- Researchers can only use information in dbGaP for valid research questions
- Researchers agree to keep individuals' identities private
- Penalties are in place for violations to this agreement

After providing this description to focus group participants, they were asked: whether they would participate in a study knowing their data would be deposited into a public database, whether they thought the described protections were sufficient, whether they felt comfortable having researchers access their data, and if there were types of information they would not want to be deposited into a database.

4.2.4.3 Tribal Concerns on Data Sharing

Focus group participants were then asked a series of questions to compare and contrast their perspectives on data being stored and shared in a federal repository versus a Tribally-managed or -owned repository. To help illustrate the Tribal concerns related to the deposition of Tribal research data, focus group participants were provided a brief description of how different Tribes make the decision to engage or not engage in genetics research. For instance, they were informed that the Diné (Navajo Nation) instituted their own Tribal policies to prohibit genetics research in their communities, whereas other Tribes have permitted genetics research provided the research is overseen and approved by the Tribe. Additionally, participants were informed that some Tribes are mobilizing the create their own Tribal-specific database that is regulated by Tribal nations.

After this description, participants were asked whether they felt a Tribally-regulated database would serve as an adequate alternative to dbGaP. They were also asked if they would be more or less willing to engage in research if they knew their data would be deposited in a Tribally-regulated database as opposed to a federally-regulated one. Finally, participants were asked to elaborate on any concerns about the management of a Tribally-regulated database, including who should have access to it and whether the existence of such a resource would contribute to greater benefit to Indigenous communities.

4.2.4.4 Possible Advantages of Tribal Participate in Research

As a final topic of conversation, focus group participants were asked to reflect on the potential benefits of research for Tribal members. They were asked if there were any factors or determinants that would engage

beneficial genetic research for their communities.

4.2.5 Audio Recordings and Transcription

All participants verbally consented to participate in focus groups in the presence of audio recording devices before the group interviews commenced. Three voice recorders were placed around the tables which seated the moderator and the focus group participants. While a single voice recorder had the range to pick up audio clearly from any point in the room, as demonstrated during trial testing, all three devices served as backup for assurance.

Once the focus group interviews were completed, audio files were uploaded to Rev.com, an audio transcription service. We purchased transcription by freelancers in order to ensure accuracy. These transcripts only identified speakers by their role (for instance, "Moderator 1" or "Participant"). The transcripts were returned to us via text documents. KST compared the returned transcriptions to the notes taken live during the focus group sessions to make sure that there was agreement. For added measure, we contracted a local CRST Tribal member through MBIRI to listen to the audio recordings and double-check audio transcripts to ensure fidelity. If there was any confusion in the transcripts due to overlapping speaking turns, this individual was able to correct those details. KST then also listened to the audio and approved any tracked changes from the transcriptions.

The recorded portions of focus group sessions were 58-69 minutes in length. Once time for arrival and consent were factored in, participants were with us for about 90 minutes.

4.2.6 Coding Strategy

We utilized a conventional content analysis, in which codes were defined during data analysis (Hsieh and Shannon, 2005) but guided by topics from the Moderator Guide. Since we were interested in perspectives and viewpoints centered around positive and negative feelings about genetics and health research, data access, data management, data policies around data sharing, and Tribal data governance, a set of initial codes were generated from these related concepts and key words. KST then conducted multiple readings of each transcript to understand more fully the broad landscape of subjects and identified additional codes to comprise an initial codebook. With the assistance of JK, an NBDC employee with statistical training, we fine-tuned code definitions related to commercialization and privatization of data, and we added additional codes related to barriers/challenges that restrict research participation. Both KST and JK independently reviewed and coded the three focus group transcripts. As a team, we settled any differences by consensus. Findings were then integrated into a common narrative across participants, while re-contextualizing the conversations in relation to the themes mentioned in the Moderator Guide.

Our codebook is described in (**Table 4.1**). Coding occurred in Dedoose, a qualitative analytics software.

4.3 Findings

A total of N=20 adult CRST Tribal community members participated across 3 focus groups cohorts with between 5-8 participants each. As seen in **Table 4.2**, most of the participants self-identified as female (N=12; 60%). The mean age of participants was 46.1 years (range: 21-65). About half of the participants reported a maximum educational attainment of high school or GED equivalent or lower (N=9; 45.0%) with N=55 (55.0%) having at least some college experience. Most focus group participants resided in Tribal District 5 (50.0%) with a one-way driving distance to the focus group site of 9 miles or less (80.0%).

For the most part, these general characteristics align with the survey participants, who were overwhelmingly female, aged in their mid-40s, with a maximum educational attainment of some college or less.

4.3.1 Positive and Negative Motivators to Research Participation

When asked what factors or motivators contributed to their decision to participate in health research, there were more excerpts from the coded transcripts that listed positive motivators (N=17 excerpts) compared to negative motivators (N=7). Upon examination, focus group participants linked the intent of study and study aim with positive reasons why they or their family members have chosen to participate in research, thereby linking perceived benefit of research to themselves or family members (particularly dependents such as children or elders) or community with positive motivation to participate in research. Of research that participants viewed as beneficial, they listed studies looking at health factors contributing to cancer, cardiovascular conditions, autoimmune disorders including asthma, and HIV. However, one disease of study interest was predominantly mentioned by participants—diabetes. As encapsulated by one participant, "Diabetes is a big one. That's the big one . . . the dialysis unit here isn't large enough to accommodate the number of dialysis patients that we have on the reservation. So, I think studies like this would help to help motivate people that are capable of organizing those kinds of clinics so that we could treat our people here."

Additionally, participants listed logistical burdens linked with socioeconomic status as potential barriers to participating in research. Participants from all 3 focus groups listed the inability to find and pay for babysitters as a detractor to research participation. Since the focus group participants were predominantly women, one can assume that childcare would be huge factor for the cohorts. Also, many participants agreed that "getting into town" was a challenge for participation in research. Owing to the rurality of the CRST, it was mentioned that the cost of gas or for babysitters did not match the time and compensation for participating in research. It was also mentioned by older participants that they are physically limited in their ability to participate in research. Other factors such as "needing to have a computer" were also mentioned as logistical

challenges, particularly since N=8 (40.0%) of the focus group participants self-reported a lack of at-home Internet access.

4.3.2 General Thoughts about Genetics

We also assessed participants' perceptions about genetics. Since genetics research is increasingly involving associations drawn with health data, asking participants in an open-ended fashion about their thoughts about genetics and genetics research was particularly interesting. Without additional prompting, participants linked genetics with the longevity of traditional lifestyles sustained through "hundreds of generations" and pointed to changes in traditional diet and ways of living with changes in health status. There were enough statements from participants that pointed to current disease burden as result of changes to long-term ways of life.

I always think that genetics are about what comprises a person, a person's life within their community, and probably race and culture and cultural practice. Because those are developed over hundreds of generations. For instance, we live here, but this whole entire Northern Plains area, we live exclusively on what foods were available here. So, the genetics that are within us, I believe we get from the earth. And that genetics pairs us with a geographic area that we're accustomed to. And if we move outside of there, then we start from suffering different illnesses. And then the medicines and the foods that they would suggest to re-develop your health don't match your homegrown genetics. To me that's what genetics are.

I personally think that it would be stunted without some of the traditional history about our diet and our daily activities. Because that's what made us who we are and keeps us who we are. So, you see this whole dietary change and along with it comes this whole health change. So, you eliminate the traditional foods from the people that were here geographically prior to this whole food consumption change. The health is almost a hundred percent different and that's almost a hundred percent crippled compared to what the health status was 500 years ago back.

Closely tied to these sentiments were excerpts that were coded for racial purity (N=9). For instance, one male participant said, "Before Europeans come to this continent, the Native Americans were genetically pure. There was no colds, there was no flus, there was no sickness . . . Everyone was genetically sound. Through the introduction of European society, genetics made the human body more susceptible to diseases." While the notion of genetic purity is a common misconception, the participants do point out that certain diseases are relatively recent for Indigenous communities. Hence, certain diseases are associated with the influx of Europeans and, presumably, European genetic variation.

Once mentioned, participants then quickly linked these notions of racial purity with researchers' interest in studying Native Americans. As one participant stated, "Data gathered geographically is real key. Because it's your geographic birth area or birth region that I think really affects and adds to your DNA makeup." That individual continues to say that, owing to these geographical origins, data on Native Americans should only be examined with other Native Americans since they are, in their words, "unique study groups". Expanded further, this sentiment can have implications for multi-ethnic cohort studies which states that there is greater need for examining between-group variation to understand the contribution of genetic factors in health dis-

parities. Therefore, there needs to be greater communication about to benefit of studying Indigenous genetic variation for the direct benefit of Indigenous peoples to include Native Americans of the US.

4.3.3 Perspectives on dbGaP and Federal Data Deposition

The Moderator described the process of de-identification of data and deposition of data in datasets to facilitate the study of disease. Afterwards, participants were then asked whether they would agree to participate in a research study that also collected their health information, provided it was de-identified. The problem with this approach is that this was posed as a simple yes/no question, and the Moderator did not further probe for more details or elaboration. In response to whether participants would deposit their data into dbGaP, focus group A participants simply stated, "Yeah" or "I would." Participants in focus group C stated, "I'll let them do that [deposit data] with mine, so that I could find out what's all wrong with my side and my family and whatnot." Others in this cohort merely nodded in the affirmative. Focus group B had a slightly different reaction. One person stated they had no concerns about depositing their data into dbGaP "just so long as they don't chip us". Two other participants in the same focus group elaborated that "they" referred to "the government" and they heard a rumor on Facebook that people were being surveyed with implanted RFIDs.

To the discredit of the study team, we should have probed for further responses from, particularly, participants in focus groups A and C. In looking at the time stamps on speaking turns, it looks like the Moderator only spent 1-2 minutes on this question before moving onto the next part of the Moderator Guide which asked participants about their concerns, if any, with having their data accessed by researchers.

The Moderator script described the kinds of protections that are in place to ensure participants' information is protected, including sanctions to researchers who violate NIH Data Sharing Policies. After describing the process of scrubbing personal identifiers from data before depositing into dbGaP, participants were asked about their comfort level in participating in federally funded research.

Interestingly, focus group A independently stated that "as long as they [the government] aren't putting no kind of chip under my skin" that they are comfortable with the de-identification process. However, this individual stated the involvement of computers in the process raised concerns about data hacking.

A participant in focus group C expressed concern about invasion of privacy due to harmonization of individual data across different federal data systems. In their own words, they shared a personal anecdote expressing their distrust of federal housing of information:

Well, my experience with the government keeping information safe is, government-to-government will share their information. As a teenager, I got in trouble. They told me my records were sealed, I go to court again 15 years later for something else and they were bringing up my files from my juvenile record, which they told me were sealed. They used that against me. When you have experiences like that, it affects your freedom and your livelihood, it's kind of hard to trust them with any kind of information after something like that happens. So, I wouldn't.

Other participants expressed that they would be amenable to their data being shared as long as it served to benefit research. One participant stated, "I would just hope it [shared data] would be beneficial for other studies."

4.3.4 Tribal Concerns on Data Sharing

The moderated discussion then shifted to the topic of Tribal responses to the NIH Data Sharing Policy. In particular, participants were told that Tribes have varied responses to the NIH Data Sharing Policy; some have agreed to share Tribal members' data and some have instituted moratoriums on genetic research. Additionally, focus group participants were told that Tribes have the option of seeking an exemption to the Data Sharing Policy. In response to this, one participant stated that she "would feel better if Tribes were exempt, just for the fact that [she] wouldn't want all of our stuff out there that far out." Furthermore, she said that should "would like to know who's getting my information before I let it go."

In response to a Tribally-regulated repository, one male participant stated that this would be "putting the cart before the house" because Tribes do not yet have the capacity to conduct their own research. However, another individual in a different focus group stated that data going to a Tribally-managed database "sounds more sensible" than going to dbGaP in order to shift ownership away from non-Tribal entities. On the topic of building research capacity in order to create a Tribally-regulated repository, one participant stated:

I think there's the professional degrees and abilities that go into that kind of a job. But I think there's a lot of our young people that are educated and moving elsewhere that I think could probably fill a role like that, but they're not very familiar with the home reservation and the home population.

Another participant in the same focus group later elaborated that "if we had a data bank like that, that could be utilized to educate individuals before they get into relationships."

4.3.5 Emerging Themes

While examining the coded excerpts to draw relationships to narrate the data, there were a few themes and subthemes that repeatedly emerged.

4.3.5.1 Data Sharing

On the topic of data sharing, several participants pointed to concerns related to **data security** and confidentiality. For instance, the threat of data being hacked seemed to be raised across 3 different excerpts. Some participants even raised the concern that anonymizing or de-identifying data may not provide enough security against the risks of hacking, re-identification, and even for providing fodder for facilitating racist assumptions about Native Americans. As exemplified, one participant said, "You can't put stuff out there that you don't want other people to know, even if it is confidential." Some participants, however, thought that

anonymizing data mitigated these potential risks. In focus group B, for instance, several participants nodded when one of the other participants stated, "It [genetic data] couldn't be used against anybody because it's anonymous, so it can only be used for something positive."

Coupled with the concern about data breaches is the concern of unconsented **data access**. One person stated, "I think you really need to prioritize who is actually seeing this information. I would want to know. I would want to know who's looking at my DNA and our genes." This statement relates to a general concern that researchers may be able to access Tribal community members' data for studies to which the Tribe or Tribal members did not consent. This is an intriguing concept. Just as researchers want more accessibility to Tribal participant data through more transparent data sharing policies, this person has touched upon a potential notion of facilitating transparency in a different way that enables participants to know who is accessing their data. Exploring pathways to increasing data access transparency can also relate to another group of themes related to trust.

Having more knowledge and oversight over data access and data decisions also entails an increased familiarity with **data agency** and **data stewardship**. For instance, one participant stated that "we should be given the validity in those kinds of [data] decisions, as a Tribe and as individuals." In terms of implications to data justice, another person stated that since researchers are "getting into our bodies, and the genetics, and that kind of data, we should at least be able to reserve the right of who is going to use it and what they're going to use it on."

In fact, many participants brought up examples related to historic and systemic injustices related to research harms perpetuated onto Indigenous peoples. Without re-enumerating these harms in great detail, of significance is the felt sentiment by this community member: "we almost literally don't own anything anymore. Even our statistics."

4.3.5.2 Trust

Besides coded excerpts relating to data access, the highest number of excerpts related to the topic of **trust**. In short, one participant succinctly stated, "I don't want my DNA in somebody's hands I don't trust."

Interestingly, participants expressed a number of positive statements related to a willingness to share data with non-Indigenous researchers. As one person said, "It's hard for us [Native Americans] to trust people. But sooner or later we're going to have to put our faith somewhere if we want to live healthier." Thus, the ability to trust researchers to access Tribal data is intrinsically linked to the derived utility of that data back to Tribal members. This is an important statement because it relates to the significance of returning benefits of health research back to the Tribal communities that are providing the data. As another participant stated, "I would share it with people who are going to make use of it."

Ensuring that health research directly benefits Tribes is especially important. There is a felt sentiment among participants that researchers might be unidirectionally benefiting from Tribal data. In particular, there was a felt notion that researchers procure funding for research but "stick it in their pockets and walk away." Further, participants shared a desire that they "see what's happening [with their data] and be involved in it."

Upon hearing the description of the NIH Data Sharing Policy and requirement for data deposition in a federal repository, all three focus group cohorts spent a considerable amount of time talking about distrust of the US government and concerns about federal oversight of their Tribal data. At least two individuals used the word "scary" to describe their fears related to government oversight of their data. As one participant stated:

It's a scary topic. It always has been for our people. Because they're... they mandate everything on our reservation. They control pretty much everything. How we live. What kind of funding were going to get. If we're going to get funding.

Another participant elaborated, "When you say 'government', we're all going to run. Nobody wants to deal with them." Of N=27 excerpts that were coded for mentions about the US government, none of them were positive. Many of the statements expressed concern about data harmonization and surveillance that potentially infringes on individual and community-level privacy. Participants asked, "What do they do with that information and where does it go?" As a nod to past relations between Tribal nations and the US government, one participant said, "When it comes to the government, us being where are, it never looks good for us. We always end up at the bottom."

While a number of participants pointed to trust in Tribal entities for acting on behalf of the welfare of their peoples, it was extraordinarily interesting to hear participants express distrust over their own Tribal members having too much access for their individual data. Statements of distrust over fellow Tribal members were found in transcripts from all three focus groups. To exemplify, one participant stated the following:

Everybody knows everybody . . . It's a small community. Even the professionals know the ones that are the patients and stuff, and it's a lot of personal relation and blood relations. And then we have nothing of value really here on the reservation except for information on somebody else. So that's how people a lot of people use it [access to sensitive data] to their advantage.

This distrust is not specifically identified due to being a Tribal member, it's important to note. As one participant said, "It doesn't matter if you're Native or non-Native, it's just the same." In other words, the distrust of having fellow Tribal members being able to access their data is more due to being part of a small, rural community rather than a characteristic of being Indigenous.

In response to these perspectives, the Moderator then asked focus group cohorts if they would be comfortable with members from other Tribal nations having access to their data. In response, participants affirmed that they would be comfortable with outside Indigenous peoples accessing their data.

4.3.5.3 Research Benefits

Much of the conversation related to trust, or rather distrust, was tied to a history of injustices perpetrated onto this Tribal community as a whole. As one participant said, "The Sioux are people. We're humans. We've been hurt for a long time." But a sustained sentiment throughout the focus group cohorts was that remediating these historic injustices could be achieved if Tribes derived direct or tangible benefits from research. As outlined earlier, some participants' distrusted outside researchers who derived perceived financial benefits from accessing Tribal research data. That perspective also extends to entities seeking to commercialize or privatize Tribal data. On Tribal data being potentially privatized for commercial gain, one participant advocated for "being strict with who gets to use it [Tribal data] and why they're using it. Because if they're going to use it, it should benefit that particular group." In terms of seeing results, participants wanted some assurances that those accessing their data were using it solely for health research to benefit their peoples.

4.3.5.4 Tribal Regulation

In terms of moving forward with health research, 2 out of 3 of the focus group cohorts spent some time discussing how Tribes can assert their sovereign abilities to self-regulate research and extend data sharing protections. Essentially, participants felt that CRST should have data sharing protocols already in place even before any Tribal member data is collected, and those policies should be binding under Tribal law. Additionally, any outside entities who want to access data from CRST members would have to get a referendum or motion from the Tribal council to approve their research request.

Participants also expanded this topic to postulate what resources would be required to create a Tribally-managed data repository. They felt that they would derive more use out of a Tribal repository than a non-Tribal repository. Since there was a shared concern about over-surveillance by fellow Tribal community members, one participant proffered, "I would be comfortable if it was not locally owned here . . . My suggestion would be to have it [Tribal data] sent to a Native American owned corporation for not only our Tribe, but for the United States, all the other Tribes." Another participant stated, "There are enough non-Indigenous people that are doctors and stuff. There aren't enough Indigenous people that are doctors and like that. So, we do need that outside help." Hence, in addition to passing laws and referendums on the Tribal side to ensure Tribal research interests are protected, some participants expressed more comfort with non-CRST Tribal members with professional expertise to establish a Tribally-regulated data repository within the bounds of CRST.

4.4 Discussion

4.4.1 Limitations and Study Errors

If afforded the opportunity to reconvene focus groups, we would conduct the focus groups slightly differently. The nested strategy of recruiting focus group participants from survey participants was particularly useful in designing focus group cohorts because it truly facilitated the ease of purposively designing the focus group cohorts. However, owing to physical distance in communicating with the recruitment team (KST was in North Dakota, JV was in California, and the MBIRI team was in South Dakota), there were errors in conveying important details to participants that resulted in a 50% reduction from the planned N=40 focus group participants to N=20. For one, we should have made it clear to individuals that the Keya Building is separate from the Keya Café. In anticipating some degree of loss of participants due to time and other extenuating factors, we should have ideally planned for a fifth focus group. However, since compensation to focus group participants comprised a majority of the budget, we would have been left in a lamentable position if we had an amazing retention rate.

We also should have been more intentional in trialing our Moderator Guide. While all study partners reviewed the Moderator Guide for content, participant comprehension, and timing, we did not anticipate having as many problems with ensuring that we were on the same page in regards to additional probing. For instance, in reviewing the transcripts for focus group B, the first half of the group interview comprised of a single individual dominating the conversation. Unfortunately this participant was prone to repeating the same points over and over. Under more controlled and gentle guidance, we could have perhaps engaged additional responses from other participants.

Perhaps owing to the small focus group cohorts (ranged from N=5-8 participants), the length of group interviews was at or under an hour when participants were originally asked to reserve up to 90 minutes. If we had more participants in each group, we probably could have elicited more varied responses.

4.4.2 Synthesis

We identified a series of themes and subthemes that perhaps can be encapsulated into a singular model (**Figure 4.2**). Here we show that, in starting with considerations about data (to include data sharing, data access, data security, and data stewardship), we are perhaps also relating these topics to larger ongoing issues related to trust and distrust. In turn, factors relating to building trust also feed into the larger aim of facilitating direct community benefits for research. Similarly, in turn, these benefits build into larger and long-term considerations for Tribes to self-direct their own data decisions. Hence, adequately fostering Tribal decision-making agency and authority has the potential to dismantle power imbalances or exacerbate them.

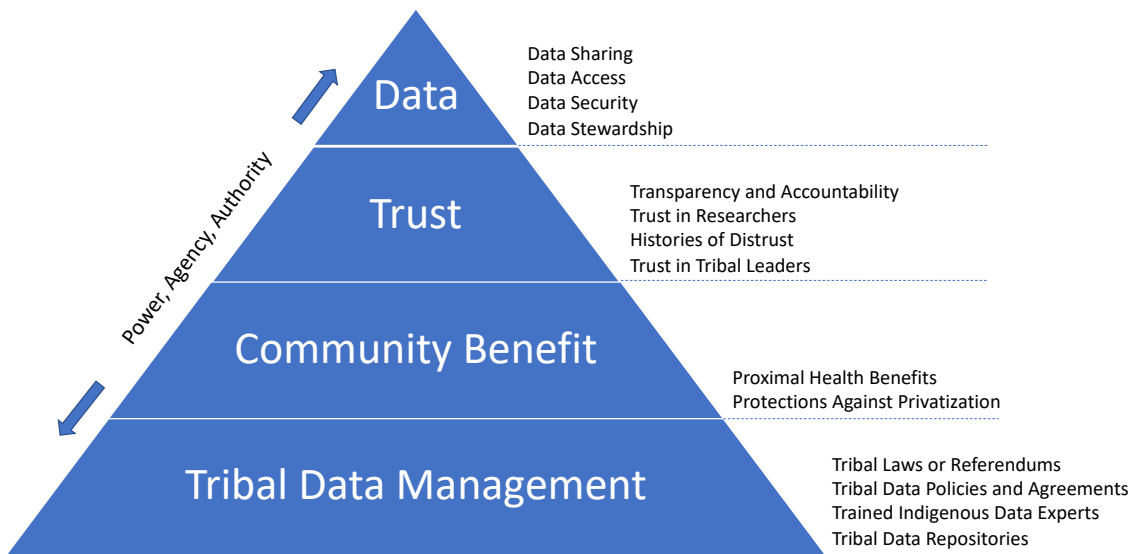


Figure 4.2: Themes and Subthemes

Codes	Definitions
1 motivators to participate	factors that encourage or discourage research participation
1.1 positive	factors that encourage research participation
1.2 negative	factors that discourage research participation
2 barriers to participation	social determinants preventing research participation
2.1 poverty	financial or economic barriers to research participation
2.2 childcare	inability to secure childcare to participate in research study
2.3 transportation	inability to find transportation to research study
2.4 other	other reasons that prevent research participation
3 feelings about research, generally	emotional perception or attitude relating to research
3.1 positive	
3.2 negative	
4 interest in research	reasons to decide to participate in research
5 concerns about research	something causing worry about participating in research
5.1 community concerns	concerns at the community-level
5.2 individual concerns	concerns at the private individual-level
6 distrust	reasons to regard an entity with suspicion
6.1 government distrust	distrust of federal government
6.2 distrust of community members	distrust of community members
6.3 distrust of non-Indigenous researchers	
7 injustice	lack of fairness or justice
7.1 historical injustice	longstanding lack of fairness or justice
7.1 systematic injustice	ingrained or embedded lack of fairness or justice
8 data access	ability to retrieve data stored in a database or repository
8.1 data privacy	privacy as a right to individual or community
8.2 data hacking	data leaks due to malicious intent
8.3 confidentiality	ensuring information is kept secure, as an ethical duty
8.4 data security	practical details related to data care and stewardship
9 data de-identification	removing personal identifiable information (PII) from data
9.1 general thoughts	
9.2 issues with sharing	
10 data privatization	data is privately owned, by individual or corporate entity
11 commercialization	data becomes co-opted by a corporate entity
12 non-genetic data	data that is not heritable
12.1 cultural data	
12.2 environmental data	
13 racial profiling	using race or ethnicity to discriminate or stereotype
14 racial purity	characterizing a racial or ethnic group as a % of a pure entity
15 research benefit	potential advantages for research participation
16 research risk	potential disadvantages for research participation
17 tribal regulation	regulation by tribal nation

Table 4.1: Codebook

Characteristics	Total (N=20)	Focus Groups		
		A (N=5)	B (N=7)	C (N=8)
Gender				
Male	6	1	1	4
Female	12	4	5	3
Other	-	-	-	-
Prefer Not to Answer	2	-	1	1
Age (years)				
N	20	5	7	8
Mean \pm SD	46.1 \pm 12.4	50.0 \pm 17.5	47.1 \pm 10.1	42.6 \pm 11.3
Median	47.0	54.0	45.0	42.5
Range (Min - Max)	21 - 65	21 - 65	36 - 63	23 - 54
Residence, by Tribal Voting District				
District 1	2	-	1	1
District 2	-	-	-	-
District 3	-	-	-	-
District 4	4	2	1	1
District 5	10	1	4	5
District 6	1	1	-	-
I Do Not Know	3	1	1	1
Prefer Not to Answer	-	-	-	-
One-Way Driving Distance				
0-9 miles	16	3	6	7
10-19 miles	1	1	-	-
20-49 miles	3	1	1	1
50-99 miles	-	-	-	-
100+ miles	-	-	-	-
Maximum Educational Attainment				
Did Not Complete High School	3	1	1	1
High School or GED Equivalent	6	3	1	2
Vocational or Technical Training	-	-	-	-
Some College, No Degree	7	-	2	5
Associate Degree	1	-	1	-
Bachelor's Degree	3	1	2	-
Master's Degree or Higher	-	-	-	-
Other	-	-	-	-
Prefer Not to Answer	-	-	-	-
Home Internet Access				
Yes	12	2	4	6
No	8	3	3	2
Prefer Not to Answer	-	-	-	-

Table 4.2: Demographics of Focus Group Participants

CHAPTER 5

Further Considerations

5.1 Strengths

Whereas previous studies on Indigenous peoples' perspectives on health and genetics studies predominantly utilize Tribal leaders or Indigenous researchers, this study represents another much-needed perspective that centers Tribal community members. It is important to solicit the perspectives and feedback of Tribal community members because they are less likely to regularly engage in discourse related to genetic and health data sharing compared to more knowledgeable Tribal leaders and partners. Yet, Tribal community members are more likely to be targeted for recruitment in genetic and health studies. Thus, understanding their attitudes and opinions about these topics is important to guide policies that ultimately affect them as research participants.

Garnering perspectives from Indigenous researchers also perhaps is poorly reflective of Tribal community members. Indigenous researchers are more likely to be younger in age and have higher educational attainment. As seen in the survey and focus groups, CRST Tribal members who participated in these two study arms were likely to be in the mid-40s with some college experience or less.

Without a previously defined pool of potential study participants from which to recruit, recruitment for focus group participants could have been logistically burdensome. Embedding the recruitment as part of the survey phase was particularly useful in alleviating this burden while also facilitating the capture of participant demographics to design the focus group cohorts.

5.2 Limitations

Due to the study's short duration and limited budget, the sample is small. However, the goals of qualitative research are informed by an improved, in-depth understanding of complex phenomena, to which these findings do contribute on the topics of data sharing and data deposition. It would have also been beneficial to gather narratives from participants residing in Districts 1, 2, and 3. Additionally, while tasks such as focus group moderation and recruitment involved multiple individuals to minimize the potential for bias, more time should have been devoted to training the study team to ensure everyone was in alignment on study goals and practices.

5.3 Implications for Practice, Policy, and Research

This study emphasizes the importance of acknowledging a wide range of perspectives among Tribal community members and representatives.

We learned that Tribal community members highly value direct research benefits that are readily apparent. Downstream promises that, someday, research may contribute to health benefits are not sufficient for facilitating transparency and trust. Also, Tribal members are also calling for increased investment and research capacity building within their communities. As seen in the survey, participants rated new job and education opportunities higher than contributing to knowledge about a disease or condition affecting them or relatives. Additionally, Tribal community members are attuned to the motives of research entities interested in collectivizing and gaining access to their data. In fact, Tribal community members are more likely to trust Indigenous driven health research entities without a profit motive than to entrust the US government or for-profit companies with their data. This latter point should inform the NIH since they are increasingly partnering with industry in biomedical research. If the NIH wants to engender trust in Tribal communities, it should consider becoming more transparent about public-private partnerships in research.

We also learned from the survey that Tribal members are not as concerned about accessing their own data after a study ends, but rather they are more concerned about the broad sharing of their data with entities in which they are not explicitly informed. This is an important point to consider. Often researchers will point to the ability of participants to draw down their own data as a benefit of research. Yet, Tribal community members stated that this is less important compared to other design characteristics.

In terms of encouraging Tribal participation in studies, focus group participants indicated that researchers should consider cost of child caretakers and cost of fuel as part of their compensation for time and effort to participate in studies. Also, the relative lack of at-home Internet access points to study design considerations, especially if a researcher wants to deploy online surveys in Tribal communities.

In terms of federal requirements to deposit data, participants by-and-large point to potential future venues by which Tribes can utilize their sovereign authority to establish their data sharing policies and policies via Tribal laws or referendums. Of particular interest is the positive views of Tribal members for creating Tribally-regulated data repositories for their own research goals.

Appendix A

Participant Survey

A.1 Survey Consent Form

Purpose of Study: We are asking American Indian members of the Cheyenne River Sioux Tribe about their thoughts and views on health research in their community. This study will help us learn about health issues that are important to tribal members. It will also help us to develop a project to promote better health, improve health knowledge, and address concerns about genetic research in tribal communities like yours. We value your input and would like to hear from you.

Procedures: We ask that you complete the survey as best as you can. Please be honest. We will only ask you questions related to the study. This survey should take between 15-20 minutes to complete.

Study Compensation: To thank you for taking the time to fill out the survey, you will receive a \$10 gift card.

Eligibility: We are looking for American Indian adults (over the age of 18) who consider themselves to be community members of the Cheyenne River Sioux Tribe. This can mean that you are an enrolled member of CRST and/or you are an American Indian who lives within a 100-mile radius of Eagle Butte, SD. We ask that employees or close relatives of employees of Missouri Breaks Research Industries, Inc. do not participate.

Potential Risks: We will take every safeguard to protect your privacy. All of your answers will be kept private and will not be linked to your name in any way. Only your contact information may be linked to your name in order to mail your compensation (or if you want to help us with a future study).

Anticipated Benefits: Your participation will help to promote ethical research that can improve American Indian health. You will likely not receive any direct benefits by taking part in this study.

Voluntary Participation: Taking part in this study is voluntary. You have the right to choose not to take part in this study. If you do not wish to answer a question, you may choose to not answer or end the survey at any time. Your data will not be used if you do not complete the survey.

Data Privacy: We make every effort to keep your survey answers private. By law, some people from the National Institutes of Health and/or the Institutional Review Boards (which oversee this project) may need to see study materials, including this consent form, which could be linked to your name. We replace your name with a number to further protect your identity. All data we collect will be stored in locked files; only project staff will be able to see them.

In Event of Harm: None of the questions are intended to cause distress. In the unlikely scenario that a question triggers an emotional response, we can refer you to the Indian Health Services (IHS) or to your health care provider. You will not be paid for any loss due to lost wages or distress.

Funding Source: All funding for this study is provided by the National Institutes of Health.

Study Partners: This study is through the Bio-Repository for American Indian Capacity, Education, Law, Economics and Technology (BRAICELET). Partners include the Stanford Precision Health for Ethnic and Racial Equity (SPHERE) center at Stanford University in Palo Alto, CA, and Missouri Breaks Research Industries, Inc. (MBIRI) in Eagle Butte, SD.

If You Have Questions: If you would like to learn more about this study, please contact the local study administrator Marcia O'Leary at (605) 964-1260.

Your Rights as a Participant: You may contact Dr. Dewey Ertz of the Aberdeen Area IHS Institutional Review Board at this address: Great Plains Indian Health Service, Federal Building, 115 Fourth Ave SE, Aberdeen, SD, 57401. Or you may contact Dr. Ertz toll-free at (866) 331-5794.

Consent to Participate in this Study

I have read, or had read to me, this Survey Consent Form. I understand what it says and I know I can withdraw at any time. By continuing, I confirm my consent to take part in this study.

- I CONSENT TO PARTICIPATE → [GO TO QUALIFYING PRE-SURVEY]
- I DO NOT CONSENT → [END]

A.2 Pre-Survey

Thank you so much for agreeing to be a part of this study. Please answer the next five (5) questions to help us determine your eligibility for the remainder of the survey.

1. *Are you eighteen (18) years of age or older?
 - YES
 - NO → [SEND TO END MESSAGE]
2. *Do you identify as being American Indian or Native American?
 - YES
 - NO → [SEND TO END MESSAGE]
3. *Do you consider yourself a community member of the Cheyenne River Sioux Tribe (this can mean that you are an enrolled member of CRST and/or have a primary residence within a 100-mile radius of Eagle Butte, SD)?

- YES
- NO → [SEND TO END MESSAGE]
- I DO NOT KNOW

4. *Are you an employee of Missouri Breaks Research Industries, Inc. (MBIRI)?

- YES
- NO → [SEND TO END MESSAGE]

5. *Are you an immediate family member (i.e., spouse/partner, child) of AND are economically dependent on an employee of Missouri Breaks Research Industries, Inc. (MBIRI)?

- YES
- NO → [SEND TO END MESSAGE]
- I DO NOT KNOW

END MESSAGE

Thank you so much for your time. Unfortunately, your responses indicate that you are not eligible to take part in this survey. None of your responses have been recorded. We appreciate your willingness to take part in this study.

CONTINUE MESSAGE

Thank you so much for your time. Please answer the remaining questions as honestly as you can. If you do not feel comfortable answering a question, you may indicate the “Prefer not to answer” or “I do not know” options, if available. Some responses have an asterisk (*) next to them. Please answer them. Remember, you can choose to exit the survey at any time and your responses will not be recorded.

A.3 Survey

1. *With which gender category do you most closely identify?

- Male
- Female
- Other
- Prefer not to answer

2. *What is your age (in years)? → [open numerical box]

3. *What is your tribal enrollment status?

- I am an enrolled member of the Cheyenne River Sioux Tribe
- I am an enrolled member of a different tribal nation
- I am American Indian, but I am not an enrolled member of any tribal nation
- I am not American Indian
- Prefer not to answer

4. *In which tribal district do you live?

- District 1
- District 2
- District 3
- District 4
- District 5
- District 6
- None of these
- I do not know
- Prefer Not to Answer

5. *What is the highest level of formal education you have completed?

- Did not complete high school
- Completed high school or GED equivalent
- Vocational or technical training
- Some college, no degree
- Associate degree
- Bachelor degree
- Master degree or higher
- Other→ [open-text field box]
- Prefer Not to Answer

6. *Do you have access to the Internet at your home? (Some people might be taking this survey in other places.)
- Yes
 - No
 - I do not know
 - Prefer not to answer

From this point on, you will see the term “health-related research study” or “health study”.

This can include any study that has asked you questions about health or asked you to provide a biological sample (such as blood or saliva) for research purposes. These studies generally are not part of a treatment plan. You would have been presented a consent form just like at the beginning of this survey.

7. *Besides today, have you ever been asked to participate in a health-related research study before?
- I have been asked to participate in a health research study, and I agreed to take part
 - I have been asked to participate in a health research study, but I refused to take part
 - I have never been asked to participate in a health research study before
 - I do not remember or I do not know
8. *Besides this study, are you aware of any other health-related research that is currently being done in the Cheyenne River Sioux community?
- No, this is the only health-related study I have heard of in this community [GO TO Q10]
 - Yes, I know or have heard of other health-related study that has taken place in this community. [GO TO Q9]

9. What other health-related research studies have you learned of that are ongoing in the Cheyenne River Sioux community? (For instance, do you know who was conducting the study or why?) → [open-text field box]

The following are things you might consider when deciding to take part in a health-related research study. Which of the following would make you more or less willing to participate?

[PRESENTED AS A TABLE OF RADIO BUTTONS: “Less willing to participate”, “More willing to participate”, “Not something that I would consider”, “I would not participate in such a study”]

10. The study will develop new ways to treat or prevent a disease
11. The study will improve existing hospital or clinical services
12. The study will bring new job and education opportunities in fields related to health
13. I or someone close to me (friend or relative) is affected by the disease being studied
14. The study relates specifically to helping my tribal community
15. The study is partnered or led by a charity, nonprofit, or patient-advocacy organization
16. The study is partnered or led by a university, college, or academic institution
17. The study is partnered or led by a federal institution other than the Indian Health Service
18. The study is partnered or led by the Indian Health Institution
19. The study is partnered or led by a drug company
20. The study is partnered or led by a for-profit corporation that is not a drug company

More research studies are looking at how a person's genes might increase their chance in developing a disease or illness.

21. *Do you think that genetic research can help to improve the health of individuals?
 - Yes
 - No
 - I do not know
 - Prefer not to answer
22. *Have you ever participated in a genetic research study?
 - I have been asked to participate in a genetic research study, and I agreed to take part
 - I have been asked to participate in a genetic research study, but I refused to take part
 - I have never been asked to participate in a genetic research study before
 - I do not remember or I do not know

“Genetic data” is any information that can be learned from your genes. It is considered to be sensitive information because researchers can learn many things about you that may or may not

relate to your health. Since family members share similar genetics to you, your genetic data can also be used to learn about your family and ancestors, as well.

[PRESENTED AS A TABLE OF RADIO BUTTONS: “This is a concern”, “Not a concern”, “I would not participate in a genetics study”]

23. Who can access my genetic data?
24. Can I access my individual results after the study is complete?
25. Can I access any research publications that result from using my genetic data?
26. Even if my genetic data is anonymous (my name is removed), can someone link my identity later on?
27. Has my tribe approved this research?
28. Will my genetic data be used in other studies without my knowledge or consent?
29. Will this research benefit my tribe in some way?
30. When the study ends, what will happen with any leftover biological material (such as my saliva, blood, etc.)?
31. How much do the study researchers understand about Lakota culture?

Finally, the following are statements about genes and genetic diseases. To the best of your ability, please determine whether the statements are “true” or “false”.

[PRESENTED AS A TABLE OF RADIO BUTTONS: “True”, “False”]

32. All health-related research involves genetics.
33. Healthy parents can have a child with an inherited disease.
34. All serious diseases are inherited.
35. A person with an altered (mutated) gene may be completely healthy.
36. Genes are pieces of DNA.
37. Half of your genes come from your mother and half from your father.
38. Genes are inside of cells.
39. It is the father’s chromosomes that decide if a baby is a boy or a girl.

40. If you have any other thoughts, feelings, or opinions that you would like to share with us, we would love to hear from you. Please write in this box. → [open-text field box]

MESSAGE

That concludes the study portion of this survey. Thank you so much for your willingness in helping us learn about your community's thoughts and perspectives on health research.

We would like to ask you to fill in your name and mailing address in the following box so that we can mail you your compensation. This contact information will be stored separately from the survey you just took.

END SURVEY MESSAGE

Thank you so much for taking the time to answer all of these questions. Your suggestions and responses are truly insightful, and we will do our best to address as many of your concerns in the future program as it develops.

Appendix B

Focus Group Materials

B.1 Recruitment Script

[Read script parts in **bold** to potential participants]:

Good Morning/Afternoon/Evening. I am looking for (potential participant name).

[If person other than (potential participant name) answers the phone, ask if/when

(potential participant name) would be available. Just state that you are from “Missouri Breaks” and that you would like to ask if she’d like to be a part of a study.]

Hello, I’m name calling from Missouri Breaks Industries. Are you potential participant name?

I am calling because referee recommended your name as a possible participant in an upcoming study.

[If they do not know the name of the referee, you can still proceed if they are still open to being recruited.]

Missouri Breaks Industries has been working in Eagle Butte for the past 30 years to improve health outcomes for community members in a variety of different areas related to health.

We are now conducting a study to ask American Indian members of the Cheyenne River Sioux Tribe about their thoughts and views on health research in their community.

This is a focus group, so you will be in a room with 8-10 other community members and informally asked for your thoughts and feedback on some questions that we will present to you. It’s okay if you don’t know anything about health research.

We will not be collecting any samples from you. If you have some time, about two hours, we would be willing to compensate you for your time. If you are interested, we would provide a \$50 gift card to thank you for your time.

[If the person says that they are not interested, thank the person for her time and politely end the call.]

[Describe briefly the study.]

To give you a brief overview, we would ask that you and a few other community members come to Missouri Breaks, located on the west side of Eagle Butte, on a Saturday or Sunday that is convenient for all. As a group, we would ask you a few questions about what you think about some health and data questions that we would ask you.

We will not ask sensitive questions that require you disclose any private information, and you are free to only answer the questions that feel comfortable to you.

The responses we will be recorded, but we will not link your comments to your name.

We will, of course, leave some contact information should you want to remove yourself from the study in the future.

[Disclose that there are no benefits to the person for agreeing to be a part of the study. Any potential risks will be minimized by de-identifying their name from the recordings.]

Do you have any questions? Do you think you would like to take part in this focus group study?

[Schedule a time and location. Ensure that the number used to call them is the best contact number.]

If you have any questions, please feel free to call at (605) 964-1260.

[Thank them and end the call]

B.2 Focus Group Demographic Survey

Please answer the following questions as best as you can. These responses are important for us to be able to describe who participated in the focus group today.

1. With which gender category do you most closely identify?
 - Male
 - Female
 - Other
 - Prefer not to answer
2. What is your age (in years)? → [open numerical box]
3. What is your tribal enrollment status?
 - I am an enrolled member of the Cheyenne River Sioux Tribe
 - I am an enrolled member of a different tribal nation
 - I am American Indian, but I am not an enrolled member of any tribal nation
 - I am not American Indian
 - Prefer not to answer
4. In which tribal district do you live?
 - District 1
 - District 2
 - District 3

- District 4
- District 5
- District 6
- None of these
- I do not know
- Prefer Not to Answer

5. My *one-way* driving distance today is:

- 0-9 miles
- 10-19 miles
- 20-49 miles
- 50-99 miles
- 100+ miles

6. What is the highest level of formal education you have completed?

- Did not complete high school
- Completed high school or GED equivalent
- Vocational or technical training
- Some college, no degree
- Associate degree
- Bachelor degree
- Master degree or higher
- Other→ [open-text field box]
- Prefer Not to Answer

7. Do you have access to the Internet at your home?

- Yes
- No
- I do not know
- Prefer not to answer

8. Besides today, have you ever *been asked to participate* in a health-related research study before?

- I have been asked to participate in a health research study, and I *agreed* to take part
- I have been asked to participate in a health research study, but I *refused* to take part
- I have never been asked to participate in a health research study before
- I do not remember or I do not know

9. Do you think that genetic research can help to improve the health of individuals?

- Yes
- No
- Maybe
- I do not know

10. Have you ever participated in a genetic research study?

- I have been asked to participate in a **genetic** research study, and I *agreed* to take part
- I have been asked to participate in a **genetic** research study, but I *refused* to take part
- I have never been asked to participate in a **genetic** research study before
- I do not remember or I do not know

Please circle “true” or “false” to answer the following statements.

11. All health-related research involves genetics.

12. Healthy parents can have a child with an inherited disease.

13. All serious diseases are inherited.

14. A person with an altered (mutated) gene may be completely healthy.

15. Genes are pieces of DNA.

16. Half of your genes come from your mother and half from your father.

17. Genes are inside of cells.

18. It is the father’s chromosomes that decide if a baby is a boy or a girl.

When you have completed this, please hand to the group leader or any of the study personnel.

B.3 Focus Group Moderator Guide

Introduction:

Good Morning/Afternoon/Evening. My name is Moderator 1 and I'll be your group leader today for this very important discussion on health data and privacy. My role is direct the content and flow of our talk and to make sure that we cover the main topics.

I would like to introduce Moderator 2 who will be observing and assisting in this discussion.

I would like to introduce Transcriber/Note Taker who will be taking notes during this discussion.

Objectives and Agenda:

Currently, Stanford University and Missouri Breaks Research Industries, Inc. (or MBIRI) is conducting a study to explore community perceptions of health research within the Cheyenne River Sioux community.

Ground Rules:

If you haven't already done so, please read through the consent letter. Print your name and sign the consent document at then end if you agree to participate. If you do not wish to take part in this interview, you are free to leave now. Please also fill out the short questionnaire you have in front of you. There is a place for your name at the top.

Before we begin, I would like to go over a few basic ground rules for our discussion.

- First, whatever is discussed in this space is confidential, and should not be shared outside of this setting. It's important that we respect one another's privacy, and so I ask you to please not reveal to others outside of this group the identities of those who are present, or disclose to others what was discussed during our session.
- If you inadvertently refer to anyone by their full name, please be assured that any such information will be omitted from the focus group transcript.
- This session is being audiotaped. This allows me, as the interviewer, to focus on you, instead of trying to jot down specific details about the discussion. Please speak in a voice as loud as mine, so the microphone can pick it up.
- My team will write a report using the tapes. Our report will not mention your name. By assuring your anonymity, I hope that you will speak openly and honestly about today's topic.
- While we are asking your opinion and experiences with research and health in general, you do not need to share specifics about any personal health issues or diagnoses if you do not want to.
- So that we may have your full attention, it would be great if you could please turn off your cell phones

or smartphones or place them on silent. If you need to leave to run to step out to take a call, that's perfectly fine.

- And remember, there aren't any right or wrong answers, so you can give both positive and negative viewpoints. We are not looking for a consensus; we don't expect you to agree with each other. We want to hear about the range of your thoughts and opinions.

Any questions so far?

[Place the audio recorder between you and the participants. Ensure there is not a lot of background noise that would interfere with hearing the responses of the interview. Ensure that the note taker if using a laptop is not too close to the recording device.]

Before moving on to our group discussion, we would like each of you to briefly state your first name.

Past Research Participation Behavior

- On the questionnaire, we asked if you have ever participated in a health-related research study before.

For those of you who have been a part of a health study before:

- Can you briefly describe the health-related research you have participated in?

Probe: what were some motivations or driving factors that contributed to their decision?

- Were these studies being conducted within the Cheyenne River community?
- How did you find out about this research?

- Have you ever considered not participating in a research study?

Probe for: personal experiences or reasons why they might not want to participate in a research study

- What might be some barriers that might prevent either you or other community members in participating in health-related studies being conducted in the Cheyenne River Sioux community?

- What kind(s) of research questions would you feel would most benefit your community?

Probe: disease-related questions that might interest them, or if they know of a disease-related question that is prevalent in the community, or some environmental or other factors that might affect health outcomes

Attitudes Towards Genomics-Related Research

Not all health-related research involves asking people about their genetic backgrounds, but increasingly more and more researchers are interested in how a person's genetics might increase someone's risk for developing disease.

- What is your understanding of genetics research?

Probe for: participant understanding and knowledge of genetics

- Do you have any concerns about genetics research being conducted in your community?

Questions about Data Sharing and dbGaP

In genetics research, it is hard to study diseases that are complex in nature (such as diabetes or heart disease, which have many possible contributing factors) with only a few individuals. So this kind of research typically requires very large datasets that recruits as many participants as possible. Additionally, these studies also typically collect information from patient health records and combine it with genetic data. This enables researchers to see all the potential contributing reasons that might cause a disease to occur, and to see if there are any common features between all of the individuals who have a particular disease or illness.

Health information usually comes from a person's clinical record data provided by that person's care provider, with the consent of the participant. Then all of that information is de-identified, meaning that no personal identifiers – such as your name, date of birth, or SSN – would be attached to your records. This is done to protect the confidentiality of private health information.

- What do you think about this process?
- Would you agree to participate in a research study that also collected your health information, provided that it was de-identified?

What would you think if the researchers wanted to share your de-identified information:

- With other researchers working or collaborating together on the same study?
- Would there be people or organizations that you would feel uncomfortable with sharing your health information?

In the United States, most health research is funded by the federal government through research grants with an organization called the National Institutes of Health, or the NIH for short. In particular, the NIH has a data-sharing policy that requires health information collected during studies paid by the agency to be deposited into a publicly-available database called dbGaP, which stands for the database of Genotypes and Phenotypes. A “genotype” refers to genetic information derived from DNA, and a “phenotype” can refer to information about a person's health.

Researchers who are doing Federally-funded research are asked to deposit study data in de-identified form to dbGaP so that it can serve as a resource for other researchers around the world. To protect individuals'

data, there are a number of protections in place to ensure study participants' information is protected, such as:

- All data stored must have personal identifiers removed
- Researcher access is limited to those who have been thoroughly checked by a governing committee
- Researchers can only use the information in dbGaP for valid research questions
- Researchers agree to keep individuals' identities private
- Penalties are in place for violations to this agreement

In terms of what was just described:

- Would you want to participate in a research study knowing that your de-identified health and (possibly) genetic information would be deposited in this public database?
- Do you think that these protections are sufficient for protecting your health and genetic data if you were a study participant?
 - What else would you want to know?
- Would you feel comfortable having other researchers have access to your private, de-identified data?
- What kind(s) of information would you not want to be deposited into such a database?

Tribal Concerns on Data Sharing Policy

Because of concerns related to data privacy, many tribal nations have instituted their own policies to govern genetics research within their own communities. Some tribes, like the Diné or Navajo Nation, have even prohibited genetics research from occurring on their tribal lands. Other tribes have still permitted genetics research to be conducted in their communities, but with oversight by tribal boards that review and approve research practices before the study can begin.

The NIH data sharing requirement, however, still remains a point of contention for many tribes. Many tribal leaders would feel more comfortable allowing genetics research to occur in their communities if there were other options available. For instance, some tribal leaders are asking for the creation of a Tribal-specific database to store genetic and health information that would be regulated by tribal nations.

- Do you feel that having a tribally-regulated database for genetic and health information would be a good alternative to dbGaP?

- Would you want to participate in a genetics study if you knew your information would be stored in the tribally-regulated database?
- What would be your concerns about the management of a tribally-regulated database?
- Who do you think should have access to a tribally-regulated database?
 - Should non-Indigenous researchers have access to the tribally-regulated database?
- Do you think having a tribally-regulated database might encourage more research studies that benefit Indigenous communities?

Possible Advantages of Tribal Participation in Genetic Studies

Part of the reason we are asking these questions is because American Indians have the lowest participation rate in genetics studies. Much of this is due to historical trauma that American Indians have faced in the recent and not-so-recent past and distrust of non-Indigenous researchers entering tribal communities.

While it is great that tribes have been able to utilize their sovereign status to direct their own policies related to research that occurs on tribal lands, American Indians generally have poorer health outcomes than other minority groups in the United States. As medicine continues to advance with genetic technologies to aid in understanding disease, American Indians might fall further behind because they have not participated in genetic studies.

It is possible that some drugs might have different effects in American Indians than non-Natives, for instance. It is possible that learning more about how genetic factors that are specific to American Indians lead to disease in our communities might allow us have better health outcomes.

- Consider these potential benefits and everything we have discussed so far. Would you want to encourage genetic studies in your community? Why or why not?

END MESSAGE

We would like to thank you very much for your participation in this focus group discussion. Before we adjourn, is there anything else that you would like to share with us or feel that we should know about?

[TURN OFF AUDIO RECORDER]

REFERENCES

- 1000 Genomes Project Consortium (2012). An integrated map of genetic variation from 1,092 genomes. *Nature*, 491(7422):56–65.
- All of Us Research Program Investigators, Denny, J. C., Rutter, J. L., Goldstein, D. B., Philippakis, A., Smoller, J. W., Jenkins, G., and Dishman, E. (2019). The "All of Us" research program. *N Engl J Med*, 381(7):668–676.
- American Indian Law Center (1999). *Model tribal research code: with materials for tribal regulation for research and checklist for Indian health boards*. American Indian Law Center, Albuquerque, NM, 3rd edition.
- Bentley, A. R., Callier, S., and Rotimi, C. N. (2017). Diversity and inclusion in genomic research: why the uneven progress? *J Community Genet*, 8(4):255–266.
- Bustamante, C. D., Burchard, E. G., and De La Vega, F. M. (2011). Genomics for the world. *Nature*, 475(7355):163–5.
- Carroll, S., Garba, I., Figueroa-Rodríguez, O., Holbrook, J., Lovett, R., Materechera, S., Parsons, M., Raseroka, K., Rodriguez-Lonebear, D., Rowe, R., Sara, R., Walker, J., Anderson, J., and Hudson, M. (2020). The CARE Principles for Indigenous Data Governance. *Data Science Journal*, 19(1):43.
- Cavalli-Sforza, L. L. (2005). The Human Genome Diversity Project: past, present and future. *Nat Rev Genet*, 6(4):333–40.
- Claw, K. G., Dundas, N., Parrish, M. S., Begay, R. L., Teller, T. L., Garrison, N. A., and Sage, F. (2021). Perspectives on genetic research: Results from a survey of Navajo community members. *Front Genet*, 12:734529.
- Cooke Bailey, J. N., Bush, W. S., and Crawford, D. C. (2020). Editorial: The importance of diversity in precision medicine research. *Front Genet*, 11:875.
- Curry, L. A., Krumholz, H. M., O’Cathain, A., Plano Clark, V. L., Cherlin, E., and Bradley, E. H. (2013). Mixed methods in biomedical and health services research. *Circ Cardiovasc Qual Outcomes*, 6(1):119–23.
- Dodson, M. and Williamson, R. (1999). Indigenous peoples and the morality of the Human Genome Diversity Project. *J Med Ethics*, 25(2):204–8.
- Dukepoo, F. C. (1998). The trouble with the Human Genome Diversity Project. *Mol Med Today*, 4(6):242–3.
- Falcone, D. C., Wood, E. M., Xie, S., Siderowf, A., and Van Deerlin, V. (2011). Genetic testing and parkinson disease: assessment of patient knowledge, attitudes, and interest. *J Genet Couns*, 20(4):385–95.
- Faye, D. J. (2004). Bioprospecting, genetic patenting and Indigenous populations: challenges under a re-structured information commons. *J World Intellect Prop*, 7(3):401–28.
- Fetchers, M. D., Curry, L. A., and Creswell, J. W. (2013). Achieving integration in mixed methods designs-principles and practices. *Health services research*, 48(6 Pt 2):2134–2156.
- Garrison, N. A., Barton, K. S., Porter, K. M., Mai, T., Burke, W., and Carroll, S. R. (2019). Access and management: Indigenous perspectives on genomic data sharing. *Ethn Dis*, 29(Suppl 3):659–668.
- Griffiths, K. E., Blain, J., Vajdic, C. M., and Jorm, L. (2021). Indigenous and tribal peoples data governance in health research: A systematic review. *Int J Environ Res Public Health*, 18(19).

- Hiratsuka, V. Y., Hahn, M. J., Woodbury, R. B., Hull, S. C., Wilson, D. R., Bonham, V. L., Dillard, D. A., Avey, J. P., Beckel-Mitchener, A. C., Blome, J., Claw, K., Ferucci, E. D., Gachupin, F. C., Ghazarian, A., Hindorff, L., Jooma, S., Trinidad, S. B., Troyer, J., and Walajahi, H. (2020). Alaska Native genomic research: perspectives from Alaska Native leaders, federal staff, and biomedical researchers. *Genet Med*, 22(12):1935–1943.
- Hsieh, H. and Shannon, S. (2005). Three approaches to qualitative content analysis. *Qual Health Res*, 15:1277.
- Hudson, M., Southey, K., Uerata, L., Beaton, A., Milne, M., Russell, K., Smith, B., Wilcox, P., Toki, V., and Cheung, M. (2016). Key informant views on biobanking and genomic research with Maori. *N Z Med J*, 129(1447):29–42.
- Huria, T., Palmer, S. C., Pitama, S., Beckert, L., Lacey, C., Ewen, S., and Smith, L. T. (2019). Consolidated criteria for strengthening reporting of health research involving Indigenous peoples: the CONSIDER statement. *BMC Med Res Methodol*, 19(1):173.
- Jacobs-Wingo, J. L., Espey, D. K., Groom, A. V., Phillips, L. E., Haverkamp, D. S., and Stanley, S. L. (2016). Causes and disparities in death rates among urban American Indian and Alaska Native populations, 1999–2009. *Am J Public Health*, 106(5):906–14.
- James, R., Tsosie, R., Sahota, P., Parker, M., Dillard, D., Sylvester, I., Lewis, J., Klejka, J., Muzquiz, L., Olsen, P., Whitener, R., and Burke, W. (2014). Exploring pathways to trust: a tribal perspective on data sharing. *Genet Med*, 16(11):820–6.
- Mills, M. C. and Rahal, C. (2019). A scientometric review of genome-wide association studies. *Commun Biol*, 2:9.
- Need, A. C. and Goldstein, D. B. (2009). Next generation disparities in human genomics: concerns and remedies. *Trends Genet*, 25(11):489–94.
- Office of the Director (2021). Request for information on proposed updates and long-term considerations for the NIH Genomic Data Sharing Policy. NOT-OD-22-029, National Institutes of Health.
- Office of the Director (2022). Request for public comments on draft supplemental information to the NIH Policy for Data Management and Sharing: Responsible management and sharing of American Indian/Alaska Native Participant data. NOT-OD-22-064, National Institutes of Health.
- Paltoo, D. N., Rodriguez, L. L., Feolo, M., Gillanders, E., Ramos, E. M., Rutter, J. L., Sherry, S., Wang, V. O., Bailey, A., Baker, R., Caulder, M., Harris, E. L., Langlais, K., Leeds, H., Luetkemeier, E., Paine, T., Roomian, T., Tryka, K., Patterson, A., Green, E. D., and the National Institutes of Health Genomic Data Sharing Governance, C. (2014). Data use under the NIH GWAS Data Sharing Policy and future directions. *Nature Genetics*, 46(9):934–938.
- Popejoy, A. B. and Fullerton, S. M. (2016). Genomics is failing on diversity. *Nature*, 538(7624):161–164.
- Powell, R. A. and Single, H. M. (1996). Focus groups. *Int J Qual Health Care*, 8(5):499–504.
- Trinidad, S. B., Blacksher, E., Woodbury, R. B., Hopkins, S. E., Burke, W., Woodahl, E. L., Boyer, B. B., and Hiratsuka, V. Y. (2022). Precision medicine research with American Indian and Alaska Native communities: Results of a deliberative engagement with tribal leaders. *Genet Med*, 24(3):622–630.
- Tsosie, K. S., Claw, K. G., and Garrison, N. A. (2021a). Considering “respect for sovereignty” beyond the Belmont Report and the Common Rule: Ethical and legal implications for American Indian and Alaska Native peoples. *The American Journal of Bioethics*, 21(10):27–30.
- Tsosie, K. S., Yracheta, J. M., Kolopenuk, J., and Smith, R. W. A. (2021b). Indigenous data sovereignties and data sharing in biological anthropology. *Am J Phys Anthropol*, 174(2):183–186.

- Tsosie, K. S., Yracheta, J. M., Kolopenuk, J. A., and Geary, J. (2021c). We have "gifted" enough: Indigenous genomic data sovereignty in precision medicine. *Am J Bioeth*, 21(4):72–75.
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., Gonzalez-Beltran, A., Gray, A. J. G., Groth, P., Goble, C., Grethe, J. S., Heringa, J., 't Hoen, P. A. C., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S. J., Martone, M. E., Mons, A., Packer, A. L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M. A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., and Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3(1):160018.
- Woodbury, R. B., Beans, J. A., Wark, K. A., Spicer, P., and Hiratsuka, V. Y. (2020). Community perspectives on communicating about precision medicine in an Alaska Native tribal health care system. *Front Commun (Lausanne)*, 5.