Privacy, Trust, and the Public's Comfort with Sharing Health Data with Third-Party Commercial Companies

by

Marie Grace Trinidad

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Health Infrastructures and Learning Systems)
at the University of Michigan
2020

Doctoral Committee:

Professor Sharon Kardia
Assistant Professor Zach Landis-Lewis
Assistant Professor Jodyn E. Platt, Chair
Research Professor Matthew Reed

Marie G. Trinidad
mgracet@umich.edu
ORCID ID: 0000-0002-0615-767X

Dedication

This dissertation is dedicated to my grandparents. Because of your many sacrifices and incomparable grit, my brother and I could dream bigger dreams.

Acknowledgements

The process of producing a single Doctor of Philosophy requires an incredible amount of time, patience, and generosity from so many people. I will spend a lifetime in gratitude and service to you all.

None of this is possible without Jodyn Platt. In the last four years there have been many ups and downs for me, and she patiently guided me through all of them. Because of her, academia is a more collegial, generous, and inclusive place. It is a privilege to work with her and I am so grateful.

I will not be able to articulate my gratitude to Sharon Kardia for her guidance and mentorship, in part because her lessons are still unfolding in my life. It's with unparalleled wisdom and brilliant humor that she teaches and mentors her students. I have notebooks scribbled with her insight into not only statistics and writing, but on life, achieving balance, and how to make even the brutal parts of work and life maybe a tiny bit more enjoyable.

To Zach Landis-Lewis, thank you for your teaching and your kindness. I never thought I would say I love ontologies, but I love ontologies! That is because of your teaching and the passion that you share with your students. To Matt Reed, thank you for welcoming this student from Learning Health Science into Design Science and for making her feel like she was in her intellectual home. Thank you for asking me what I saw beyond this journey and academia. To you both, I'll never be able to repay the hours, energy, and attention you've devoted to nurturing this one student. Yours is a humbling gift.

To Chuck Friedman, thank you for your vision. I landed at the University of Michigan the same year you did. The day you presented the then brand-new Health Informatics Graduate Certificate at the School of Public Health I rearranged and planned my two-year program around

what you created. I did not know then that you were also planning a PhD program, but four years after that first Health Informatics presentation I was again one of the first to apply to yet another opportunity you created. “Thank you” isn’t enough.

To Anne Sales and Gretchen Piatt: I am laughing as I write this because I remember, and you remember, that first year PhD student. Thank you for your guidance that was gentle when necessary and tough when necessary. Thank you for your tireless commitment to your students.

Without the intervention of Allen Flynn, I don’t know if I would have had the confidence to apply to be a part of Learning Health Sciences. His belief in me when I had my doubts propelled me into a PhD program and into the most challenging role I have taken thus far. Thank you for your gentle nudge and for encouraging me to integrate my design and public health backgrounds with book recommendations, meetings and conversations both before and during my first year as a PhD student, and for being a generous friend.

Thank you to the best research team ever: Mina Raj, Paige Nong, Daniel Thiel, Tevah Platt, Philip Amara, Kerry Ryan, Kayte Spector-Bagdady, and Melissa Creary. I am looking forward to rejoining our Thursday meetings.

My time at the University of Michigan started in 2011 at the School of Public Health. I did not know then how moving from California to Michigan would completely change my life. Thank you to the following people who walked with me during this journey. Thank you to my first mentors at the School of Public Health: Brian Zikmund-Fisher, Andrew Maynard, Palmer Morrel-Samuels, and Scott Roberts. I was also gifted an amazing opportunity to spend time with the folks at the Taubman College of Architecture and Urban Planning. Somehow, they made me a designer. Thank you to Sean Vance, Joy Knoblauch, Robert Adams, Caitlin Cashner, Lizzy Baskerville, Samira Daneshvar, and Casey Carter.

Thank you to Jack Kufahl and Sol Berman, for providing me with advice and confidence that I needed, when I needed it. To the rest of my tribe at Michigan Medicine (in no particular order): Jennifer Wallis, Deb Haslam, Imani Williams, Alex Jolliet, Brett Miller, Molly McPhall,

Jennifer Pietras, Steve Gendler, John Herlocher, Buzz Nau, Jeremy Hallum, Ryan Echlin, George Brown, Jim Deneen, Bruce Taylor, Amy Yamasaki, Sue McDowell, Susan Niepoth, JD Jordan, Javan Thompson, Scott Bolak, Kenny Hill, Dave Cuevas, Joe Fodor, Stratos Kotzabassi, Nick Chrumka, Larry Bowling, Roger Burns, Chris Chapman, Marc Stephens, Ted Hanss, Cindy Leavitt, Michael Warden, Fusen Li, Drew Montag, Tim Nolan, Stephanie Dascola, Dr. Johmarx Patton, Ryan Heynard and everyone I'm probably forgetting to mention. Thank you for being my Michigan family, my mentors, my advisors, and my friends.

To Dan Stokols, David Meyer, and Joe DiMento at the University of California, Irvine — thank you for your teaching and your belief in my abilities. You patiently watched me climb each step and gave me the final boost I needed to launch into a new life at the University of Michigan. To all the teachers in my life that encouraged me: Thank you for demanding always more of me than I thought I was capable of and expecting my best. I've carried your words of encouragement with me for all these years.

To my Dad and Mom, and William – thank you for everything.

To Tracy and Eddie Baldus – you've supported me through thick and thin. Thank you.

To the greatest family of friends anyone can ask for – Brunch Club: Leif Zeng, Ellen Schmidt, Michele and Scott Goci, Carol Im, Renee Casey, Cameron Watkins, Fernando Yarza, Colleen McClain, Crista DaVia, Lanie and Tim Burns, Amanda Sullivan, Courtney, Brian, and Grayson Simko; the Villa Way Crew: Philip Morgan, Jun Yu; my LHS brothers and sisters: Oliver Gadabu, Nik Koscelniak, Mike Roth, Jeff Vlasic, Kyle Kerbawy, Emily Kobernik, Rama Mwenesi, Elliott Brannon, and all the HILS students that come after us; my daily PhD support buddy: Elizabeth Umberfield; and, finally, to Robert Avers. You've all made me laugh and listened while I cried. I am so grateful.

It took a village.

Plans fail for lack of counsel, but with many advisers they succeed (Proverbs 15:22 – NIV)

Preface

This dissertation was completed in the midst of the COVID-19 pandemic. Epidemiology, privacy concerns, and scrutiny of health behaviors of the public filled new headlines and everyone I knew suddenly became interested in fomites and virology. As I wrote, the world navigated isolation and fear as shelter-in-place orders blanketed the globe. School was cancelled. University students were sent home. The economy tanked. People lost their jobs. Many lost their lives and loved ones. The University of Michigan Rackham Graduate School elected to push back all deadlines, including my dissertation filing deadline, from May 2020 to August 2020. The dissertation defense process, a time-honored routine of formal presentation followed by questions and deliberation, moved to a video chats and screen sharing. Today, five months after the declaration of a state of emergency, restaurants that haven't shuttered their doors forever are re-opening their doors. Hair and nail salons are putting clients back on their schedules. People can visit their doctors and dentists for routine visits. The university we temporarily left behind will be forever changed by these events, and as I emerge from this dissertation writing hovel, we all step out into a new normal.

Table of Contents

List of Tables

List of Figures

Abstract

Healthcare partnerships with third-party commercial companies have been met with reservations from the public about the privacy of health information and concerns about how this data is used. While research points to the need to provide patients greater control over the use of their data, or notification of data use, it is not yet clear how to move forward with this effort while balancing the needs of researchers for quality data sets. To better understand and characterize the public's comfort with third-party commercial companies and perhaps manage and address the public's concern, in this dissertation I examine the relationships between the public's comfort with sharing health data with third-party commercial companies for patient and business purposes in relation to the public's comfort with demographic characteristics, perceived healthcare access, trust in the health system and trust in providers, privacy concerns, and altruism. I also explore the effect of a past data breach and concern about recent data breach events, comfort with researchers, quality analysts, commercial companies, and law enforcement, confidence in existing health data laws, and desire for greater control over health data or notification of data use on the public's comfort with sharing health data with third-party commercial companies.

In this dissertation I present the results of a survey of the US public (n = 1841) to assess comfort with sharing health data with third-party commercial companies for patient or business purposes. Weighted Ordinary Least Squares (OLS) Regression analysis was used to first estimate the relationship between comfort with third-party commercial companies for patient and business purposes (dependent variables) and the aforementioned independent variables, followed by stepwise regression modeling to estimate a full model of contributing factors to the public's comfort with sharing health data with third-party commercial companies.

In the final analysis, variables that were significantly ($p < 0.05$) related to comfort with data sharing for patient purposes included: comfort with researchers and commercial companies, comfort with law enforcement accessing genetic data, altruism, individuals between the ages of 45-59, and educational attainment. Statistically significant variables associated with data sharing

for business purposes included: comfort with researchers, quality analysts, and commercial companies, comfort with law enforcement accessing genetic data, concern about Memorial Sloan Kettering's startup company, Paige.AI, and employment status. The most salient factors associated with respondent's comfort with sharing health data with third-party commercial companies for both patient and business purposes were trust in the health system, confidence in existing laws and policies, and desire for notification.

The results of this study suggest that increasing trust in the health system may have a greater impact on the public's comfort than efforts to address privacy concerns alone. Desire for notification was also more important to the public's comfort with third-party commercial companies than the desire for control over health data. Patients may be better served by focusing on efforts to build trust in healthcare organizations and by providing notification of health data use instead of more granular control over health data use.

Chapter 1 Introduction

1.1 Background

In September 2018, the New York Times, in partnership with Propublica, published a story detailing the undisclosed conflict of interest entanglements of former Memorial Sloan Kettering Cancer Center's Chief Medical Officer, Dr. Jose Baselga. Dr. Baselga, a renowned oncologist who revolutionized treatments for breast cancer, failed to disclose to Memorial Sloan Kettering the extent of his corporate connections (Ornstein & Thomas, 2018a). In response to the negative public press, all corporate affiliations of Memorial Sloan Kettering were subsequently scrutinized. One corporate affiliate, Paige.AI, launched in February 2018 with $25 million in venture capital and a promise to "transform how cancer is diagnosed", was reviewed for its exclusive deal to use the Cancer Center's archive of 25 million patient tissue slides and the diagnostic work of hundreds of pathologists. A second article published by the New York Times about the Memorial Sloan Kettering and Paige.AI arrangement noted that Memorial Sloan Kettering pathologists have "strongly objected [to the deal], saying that it is unfair that the Paige.AI founders and medical partners received equity stakes in a company that relied on the pathologists' expertise and work amassed over 60 years" and "questioned the use of patient's data – even if anonymous – without their knowledge in a profit-driven venture". The article went on to note that patients were "nervous that their health data was being commercialized by the institution" (Ornstein & Thomas, 2018b). Former Memorial Sloan Kettering Cancer Center patient Steve Petrow wrote in an open letter to the hospital: "My sense of betrayal only deepened when ProPublica and the New York Times reported on an artificial intelligence startup called Paige.AI [...] Are the slides of my cancer among them? My mom's? My sister's? I'm uneasy wondering whether they are being commercialized without our consent, or even without our being notified. The hospital claims that the data are anonymous, but anonymous data these days has a habit of somehow becoming identifiable" (Petrow, 2018).

The Paige.AI project, however, had been approved by Memorial Sloan Kettering's institutional review board (IRB) for human subject research, and, despite the characterization in the New York Times, the actual tissue samples of patients were not going to be shared with the corporate startup venture, only de-identified images and notes. In response to the story printed in the New York Times, on September 23, 2018, Memorial Sloan Kettering put out a press release stating "there are several similar computational pathology efforts underway across the country that have emerged from academic medical centers and have faculty founders", and that "the research and sharing of images and data complies with Memorial Sloan Kettering rules and legal requirements" (Memorial Sloan Kettering, 2018). Despite this press release, however, and despite Memorial Sloan Kettering's regulatory and legal compliance, articles continued to be written about the possible ethics violation of the deal, and partner healthcare systems proceeded to distance themselves and their patients from Memorial Sloan Kettering's corporate start up effort, reassuring their communities that "the New York-based health system does not have access to [our] patient data, diagnostic information, or tissue images" (Huang, 2018).

I open with this story because of the complicated issues at work for Memorial Sloan Kettering. First, the unflattering story in the New York Times diminished national and institutional trust in Memorial Sloan Kettering's leadership. In response to this deterioration of trust, all corporate ties to the institution fell under increased scrutiny and criticism, despite their adherence to existing protocols (IRB) and despite the de-identified nature of the information shared with corporate partners such as Paige.AI, which follow the existing regulations outlined in the Health Insurance Portability and Accountability Act (HIPAA). The questions, then, are these: if meeting legal and regulatory requirements is insufficient to quell patient concerns and negative public press, what can Memorial Sloan Kettering and other healthcare institutions in similar circumstances do to anticipate and mitigate public fallout over the use of personal medical data and what do these stories tell us about future uses of healthcare data in other artificial intelligence and big data efforts? If patients indicate concerns about the commercialization of their data, what do those concerns imply for how partnerships are created? Approaching answers to these questions can help prevent violations of patient and consumer trust.

As patients behave as healthcare consumers and seek leading edge innovation and treatment, healthcare systems can feel pressured to pursue corporate partnerships that promise innovation such as Paige.AI without the input of patients and the public. While these partnerships may hold significant breakthroughs for patient care, few healthcare systems have obtained explicit consent for sharing, and, as of this writing, progress has not been forthcoming. Recent reports on the use of IBM Watson for Health has shown that the data product has done little to improve or impact care, and in some instances recommended "unsafe and incorrect treatment recommendations"(Ross & Swetlitz, 2018). Internal documents from IBM Watson blame these issues on the "small number of synthetic cases" provided to IBM Watson by Memorial Sloan Kettering. In response, IBM Watson and other companies focused on artificial intelligence will likely request larger and larger samples of raw patient data.

1.2 Problem Statement

The confluence of data breaches, broken trust, and the creep of cookies and web trackers gathering and analyzing 2.5 quintillion bytes of data daily (Marr, 2018) has brought concerns about the use and power of personal data to the forefront of news. In the case of Memorial Sloan Kettering, it appears current regulations and approaches are insufficient to quell patient concerns about the use of their healthcare data, and that organizational and individual reputations can be severely damaged if these reservations about data use are not addressed. ***In this dissertation I explore the predictors for comfort with data sharing with commercial companies for patient purposes*** (i.e. improving the diagnosis and treatment of other patients, developing tailored care predictions) and for ***business purposes*** (i.e. information storage by third party companies, sale of de-identified information to pharmaceutical companies). Healthcare systems do indeed blur the use of patient and commercial use and benefit, however, there may be an underlying assumption that the patient is aware of how business uses of patient data improve patient care. In the case of Sloan Kettering, healthcare organizations may emphasize the innovative nature of the partnership and the benefits accrued to the health system instead of tangible improvements to patient care (CooperKatz, 2018). I dichotomize patient uses and business uses, even though these two concepts are intertwined, interrogating comfort when data is shared for patient purposes, i.e., to improve care, diagnosis, or treatment, versus comfort when data is shared for business purposes, i.e., the sale of de-identified data for artificial intelligence efforts, allowing for examination of

the effect of communicated purpose of use on comfort with sharing health data with third-party commercial companies. This division is consistent with existing models of consumer willingness to provide access to PHI. Previous research indicates patients desire more control if their health data will be used for profit-generating research (Willison et al., 2009), and are more willing to provide access to their health information if the potential health benefits to the public are clear (Anderson & Agarwal, 2011; Castell & Evans, 2016). The public has been found to be less accepting of data sharing partnerships not only when the public health benefits of the partnership are not made clear, but also if the data sharing relationship was determined to be of only private benefit (Castell & Evans, 2016). In a study of US veterans, participants expressed to the study team that research studies must have "high value with an 'overall impact on society' and not be 'an academic exercise' and should consider whether 'just a few hundred [people] or several thousands' would benefit" (Damschroder et al., 2007).

The goal of this dissertation is to examine public comfort with the sharing of patient health information with third-party commercial companies for patient and business purposes and examine the differences in comfort that arise from these two presentations of use. This research also explores the factors associated with comfort with both purposes. Based on a survey of the general public (see Appendix), my research addresses the following questions:

- ***Paper 1:***
    - ***Which demographic and health characteristics are associated with comfort with providing health data to third party commercial companies for patient purposes and business purposes?***
    - ***To what extent do privacy concerns impact this comfort?***

- ***Paper 2:***
    - ***To what extent are a) attitudes towards trust in the healthcare system, b) trust towards providers, c) altruism, d) experience with a past data breach, and e) concern about recent data breach and data-use-violation events associated with comfort with sharing health data with third party commercial companies for patient purposes and business purposes?***

- ***Paper 3:***
    - ***To what extent is comfort with researchers, quality analysts, commercial companies broadly, and law enforcement associated with comfort with sharing health data with third party commercial companies for patient and business purposes?***
    - ***What is the relationship between confidence in existing laws governing health data and comfort with sharing health data with third-party commercial companies?***
    - ***What is the relationship between a) desire for control over health data and b) desire for notification of data use and comfort with sharing health data with third-party commercial companies for patient and business purposes?***

In the following sections I provide definitions important to this research, followed by a review the policy and regulatory landscape that shape the rules for sharing health information as well as additional background on the emergence of new corporate entities in the context of digitized healthcare, and their impact on the public. I then consider relevant theories of privacy, trust, and transparency. I follow this brief review of the literature with the conceptual framework that informs my dissertation research and the key concepts hypothesized here to be associated with comfort with data sharing.

Defining third-party commercial companies

Throughout this dissertation I refer to "third-party commercial companies". Third-party commercial companies are those companies that fall outside of usual "covered entities", covered entities being healthcare systems or providers who transmit any health information, health care plans, and health care clearinghouses (billing services, community health information systems, etc.) (Office for Civil Rights (OCR), 2013). Traditionally, "third-party" referred to administrators or payors of healthcare expenses—intermediaries that network with other providers and systems to fairly price medical billing claims or facilitate in some other capacity tasks critical to the administration of a health system. More recently, however, "third-party" has also come to refer to companies that purchase de-identified health data from providers, payors,

and pharmacies in order to gain market strategy insights (Arndt, 2018). The business of third-party medical data trading has grown tremendously, with some companies leading this field making over $2 billion in revenue in one year alone (Tanner, 2016). Once a data transaction is completed between a covered entity and a third-party commercial company, that company is then able to re-sell this de-identified information on the secondary data market. This dissertation research focuses on third-party commercial companies that support the growing field of precision health, and precision oncology in particular. Precision oncology leverages genomic, electronic health record, and other large data sets for the treatment of cancer and other diseases. Because genomic sequencing and data stewardship is beyond the scope of many health systems, health providers engage private companies to collect, analyze, and interpret data. Costs of genomic sequencing are made more affordable by negotiating data sharing agreements in which commercial companies may keep the data for future use.

Defining comfort

In this dissertation I employ the word "comfort" to describe the public's "acceptance", "willingness", "or openness" to sharing health data. As an example, "comfort with sharing data" can be used in place of "openness to sharing data". This usage has been used similarly in previous research (Dhopeshwarkar et al., 2012; McKnight et al., 2002a; O'Brien et al., 2019).

1.3 Policy and regulation shape the rules for sharing health information: Electronic Health Records and Patient Privacy

1.3.1 Health IT Legislation and The Development of EHRs

Health information technology legislation in the United States has spurred the adoption of Electronic Health Records (EHRs) and, with these newly available data, the possibility of unprecedented insight into the quality of healthcare delivery across populations and organizations, as well as for the individual patient. Legislation key to this transformation include the Health Information Technology for Economic and Clinical Health (HITECH) Act enacted under Title XIII of the 2009 American Recovery and Reinvestment Act (ARRA), as well as Title III of the 21st Century Cures Act signed in 2016 (Burde, 2011). Under HITECH, the United States Department of Health and Human Services (HHS) was tasked with spending $25.9 billion

to expand adoption of health information technology. Incentives for adoption and definitions of "Meaningful Use", or the use of certified EHR technology that improves the quality of care, spurred healthcare systems across the country to direct enormous amounts of time and resources towards transitioning their paper-dependent organizations to certified EHR systems by January 2014—as well as financial penalties of a 1% reduction in Medicare reimbursement if the transition was not demonstrated by 2015 (Burde, 2011).

As previously mentioned, one requirement of "meaningful use" is the implementation of a certified EHR technology to facilitate health information exchange (HIE) to improve the quality of care. As this occurred, healthcare and businesses turned their attention toward a new frontier of possibilities digitized data could unlock for healthcare quality improvement, healthcare operations, and research. The 2016 Precision Medicine Initiative of the National Institutes of Health (NIH) and its paired research program, "All of Us", the concept of a Learning Health System defined by the Institute of Medicine (IOM) in 2015, and the release of HealthKit by Apple in 2014 (Farr, 2014) are just some of the efforts in electronic healthcare data that have gained momentum since HITECH's incentivization of electronic health records. The 2016 Cures Act expands on HITECH efforts and tasked the Office of the National Coordinator (ONC) by developing or supporting a "trusted exchange framework and common agreements to address policies and practices between health information networks" (Rucker, 2018) and setting electronic health system interoperability standards for "(1) vocabulary and terminology, (2) content and structure, (3) transport, (4) security, (5) services, and (6) querying and requesting information for access, exchange, and use" (Upton, 2015).

1.3.2 Health Information Exchanges and New Risks

Within the HITECH act is the provision that among the responsibilities of the HIT policy committee established by the HITECH act, the HIT policy committee may make recommendations in the additional areas of "I) the appropriate uses of a nationwide health information infrastructure for the purposes of (I) the collection of quality data and public reporting; (II) bio surveillance and public health; (III) medical and clinical research; and (IV) drug safety." The section goes on to include: telemedicine, technologies that facilitate the continuity of care among health settings, methods to facilitate secure access by an individual to

such individuals protected health information, and any other technology with the greatest potential to improve the quality and efficiency of healthcare (HITECH Act, 2009). Health information exchange (HIE) is the expansive effort to allow health care professionals and patients to appropriately access clinical information, and as indicated in the previous sentence, there are myriad uses of healthcare data and myriad problems that accompany those uses. As an example, healthcare leaders have voiced reservations about wholesale publishing of medical claims information, indicating that such information is difficult to compare – procedural costs differ from institution to institution due to physician expertise, types of cases seen, and varied operation costs (T. Curran, personal communication, October 29, 2019). Additionally, growth in value-based payments and risk sharing have accelerated efforts to share health information, but many smaller healthcare organizations still lack the technological infrastructure necessary to securely provide this information to partner institutions.

Increased data connectivity poses greater risks to the privacy and security of personal health information (PHI). Electronic information can be more easily located, accessed, and duplicated than paper records. Over two thousand healthcare data breaches have been reported to the Department of Health and Human Service Office for Civil Rights since 2009, resulting in the disclosure of 194,853,404 healthcare records as of 2018, the equivalent to 59.8% of the US population. The number of healthcare data breaches reported yearly are up 83% since 2010, with a 157% year-over-year increase in the number of individual compromised records (HIPAA Journal, 2019). The public, enduring a steady stream of data breaches and compromised passwords from all service sectors, have responded by indicating their desire for greater control over all data, including healthcare data (Caine & Hanania, 2013).

In 2016, the Obama administration noted that "*the success of the digital economy ultimately relies on individuals and organizations trusting computing technology and trusting the organizations that provide products and services that collect and retain data. That trust is less sturdy than it was several years ago because of incidents and successful breaches that have given rise to fears that corporate and personal data are being compromised and misused*" (Donilon, 2016). Deterioration of trust can compromise the quality of medical data because it may affect willingness to disclose necessary health information, with potentially life-threatening

consequences (Agaku et al., 2014; ONC, 2015). Patients have reported withholding information from their provider or avoiding treatment entirely because of privacy concerns (Agaku et al., 2014; The Privacy Advisor, 2012) or engaging in "defensive" practices to limit the amount of sensitive information in their health records (Rothstein, 2011).

Although detailed health information can be of benefit to public health, to improvements in the quality of healthcare, and in realizing the goals of precision medicine, the research and data model implied by the use of artificial intelligence in healthcare are incompatible with current policies governing the usage of patient data (Kuchinke et al., 2016) and requires leveraging clinical data for purposes beyond individual care and payment for services. If, however, trust in healthcare remains low or decreases, and consequently information important to the physician-patient encounter is withheld, "benefits may be realized incompletely or inadequately at best." (Francis & Francis, 2017).

1.3.3 How is privacy protected in healthcare?

Privacy and ethics in healthcare and research are currently guided by the following: The Health Insurance Portability and Accountability Act (HIPAA), the Federal Policy for the Protection of Human Subjects, also known as the "Common Rule", broadly by the Privacy Act of 1974, and more recently by the Health Information Technology for Economic and Clinical Health Act (HITECH) signed as part of the American Recovery and Reinvestment Act in 2009. Warning that "the net effect of computerization is that it is becoming easier for record keeping systems to affect people then for people to affect record keeping systems", the advisory committee on automated personal data systems recommended in 1972 that "although there is nothing inherently unfair in trading some measure of privacy for a benefit, both parties to the exchange should participate in setting the terms" (Medicare et al., 2013). That same committee proposed a framework of data use that included five principles of use: 1) forbid the use of secret databases, 2) require that people know what records are kept about them, 3) require a statement of clear purpose and consent as purposes shift, 4) provide the ability to correct or amend and 5) provide adequate security and reliability (Nissenbaum, 2009). The resulting of this committee was the Privacy Act of 1974, establishing these principles as the Code of Fair Information Practice, which govern the collection, maintenance, use, and dissemination of

personally identifiable information and prohibits disclosure of this information unless one of twelve exemptions can be applied. These exemptions are 1) "need to know", which authorizes intra-agency disclosure, 2) required disclosure under the Freedom of Information Act (FOIA), 3) "routine use" or compatibility with the purposes for which the data was collected, 4) the US Census, 5) statistical research where the record is not individually identifiable, 6) National Archives or historical information whose value warrants preservation by the US government, 7) law enforcement requests, 8) compelling circumstances affecting the health or safety of an individual, 9) congressional documents, 10) the duties of the General Accountability Office, 11) as part of a court order, or 12) as part of debt collection activities. The Privacy Act forms the general guidelines by which identifiable information collected by federal agencies is protected and provides a common protection backdrop for all other privacy protections.

The Health Insurance Portability and Accountability Act (HIPAA), signed by President Clinton in 1996, established the federal policies and procedures that protect health insurance coverage and the privacy and security of health information. HIPAA is comprised of two parts, Title I and Title II. Title I of HIPAA protects an individual's ability to procure insurance regardless of their pre-existing conditions and prohibits group plans from restricting benefits for a specific disease or treatment. Title II establishes the policies and procedures governing the privacy and security of individually identifiable health information. Protected health information includes names, birth dates, telephone number, and other identifying numbers. Under HIPAA, patients reserve the right to request a list of individuals and entities their protected, identifiable healthcare information has been disclosed to. HIPAA's shortcomings were known even at the point at which it became law and was derided for being "too narrow" by applying only to "covered entities" – clinicians, health care facilities, pharmacies, health plans, and health care clearinghouses – HIPAA does not apply to individuals or groups outside of these roles (Cohen & Mello, 2018).

The Common Rule was introduced in 1981 and revised in 1991, and provides our basic ethical principles in research involving human subjects. Origins for the Common Rule are found in the Belmont Report, a 1978 document that defined core principles for ethical human research in response to a number of abusive, unethical clinical studies that had been conducted

without oversight, focusing specifically on the Tuskegee Syphilis Experiment (cite the Belmont report). Outlined in the Common Rule are protections for "vulnerable populations", or any disadvantaged population, and guidelines for Institutional Review Boards (IRB), informed consent, and assurances of compliance. These protections now form the requirements for ethical enrollment and treatment of human subjects in clinical study. Studies that follow these guidelines are presumed to be acting ethically and in the best interests of their enrolled study population. Because HIPAA was signed in 1996, there was some mismatch between the compliance requirements of HIPAA and the Common Rule. A Common Rule revision, first announced in 2011, attempted to resolve some of these mismatch issues. In January 2017, the Revised Common Rule was published. Included in the Revised Common Rule were 1) new policies providing researchers with the option of broad consent for future research and 2) exemptions for secondary research involving identifiable information if the research is regulated by HIPAA.

As detailed in earlier sections of this review, in 2009, the HITECH act was signed as part of the American Recovery and Reinvestment Act (ARRA). Intended to expand the use of electronic health record (EHR) system among providers, the HITECH act also widened the scope of privacy and security protections available under HIPAA with the addition of legal liability for HIPAA violations and the requirement for Notification of Breach, whereby patients must be notified of any unauthorized uses or disclosures of their health information. HITECH also expanded HIPAA's scope to also include "business associates" of covered entities, or any company or persons who conduct business with one of the previously mentioned covered entities. As healthcare information becomes increasingly digitized, the ability of these laws and regulations to meet the evolving needs of providers and patients will continue to face greater challenges. As healthcare research efforts such as precision health make use of increasingly large data sets, the line between research and practice becomes increasingly blurred.

1.3.4 How will big data threaten current legal conceptions of privacy?

Companies that collect large amounts of user data, such as Google and Facebook, profit from the insight gained from this data that is analyzed, packaged, and sold. If people are aware of data collection at all, it's likely at the initial point of contact, not as data are combined and recombined, sold and transferred to others, or put to uses that might never have been anticipated

at the time of the original collection (Ohm, 2010). Additionally, given the size and scope of data collection and use, approaches to controlling or protecting what is done with big data that rely on individual notice, monitoring, and choice are unlikely to be effective (Ohm, 2010).

HIPAA's general requirement for written patient authorization for disclosure of identifiable health information has led to routine practices for compliance within systems that conduct most research in universities and health care systems, but the health information landscape is growing to include entities who are not necessarily business associates of healthcare providers (i.e. Apple, Google, Fitbit) and are therefore not governed by HIPAA. These companies are taking increasing interest in medical records and in the purposeful or accidental generation of health information. Loyalty cards that track purchases of over-the-counter drugs can indicate purchasing patterns consistent with certain diseases or acute illnesses. Web search activity that is tracked and provided to a myriad number of businesses interested in more targeted marketing towards consumers can also provide information about healthcare concerns, age and health status, and overall lifestyle. One notorious example is Target's algorithmic determination that a teen was pregnant based on her search history, followed by the pre-emptive deployment of coupons for baby clothes and cribs to her home to the outrage of her father, who did not yet know his daughter was pregnant (Hill, 2012). The increasing availability of health information outside of the healthcare setting as well as the analytic capabilities conferred by big data methods threaten the assumption that data can actually be fully de-identified and protect an individual patient's privacy (Cohen & Mello, 2018).

1.3.5 New corporate entrants into the healthcare sector

Complicating these regulatory issues are the tech companies that "target shortcomings and legacy systems that are no longer efficient [in healthcare]" (Beaver, 2018). This interest is driven by the massive market transformation precipitated by the 2010 Affordable Care Act. In a 2016 report, the Department of Health and Human Services estimated a net reduction of 20 million uninsured adults since ACA's enactment (Uberoi et al., 2016). With 20 million new users, the U.S. healthcare system finds itself having to innovate in search of care optimization and cost savings. Within the existing healthcare industry, examples of these innovation attempts include Sloan Kettering, whose Paige.AI partnership was detailed in earlier sections, as well as

Mount Sinai, University of Maryland Medical Systems, Novant Health in North Carolina, and the Cleveland Clinic. Hardin Memorial Health in Kentucky and Health Quest Systems in New York have both announced partnerships with IBM Watson to improve clinical trial matching and extract unstructured health information from medical records (Park, 2019). Cedars-Sinai launched a health tech accelerator program to attract innovators with a range of healthcare disrupting ideas and the opportunity to test those ideas alongside physicians. Innovative technologies proposed include an interface for shared decision making, a wearable blood collection system, and content management platforms to streamline data collection for clinical trials (Sullivan, 2017).

Outside of the healthcare setting, marketplace interest in fitness trackers, health sensors, online vision tests, and clinical communications is outpacing other industries in attracting venture capitalist dollars. One example of a new healthcare entrant is Amazon's addition of "healthcare skills" to the Amazon Alexa voice-controlled speaker system. Working with Providence St. Joseph Health and pharmacy benefit manager Express Scripts, Amazon has made their online and voice platform HIPAA compliant to provide prescription tracking, assist parents in providing updates to their child's physician team to report on recovery after surgery, book appointments, and track blood sugar levels (Ramsey, 2019). Although now in partnership with healthcare systems, and therefore falling under rules HIPAA, other commercial healthcare skills hosted by Alexa were created before Amazon established HIPAA compliance in early 2019. These applications include timers for pregnancy contractions, weight loss assistants, sleep trackers, and food and calorie management voice assistants. Amazon skills and comparable applications on other platforms sit outside of traditional modes of healthcare delivery and therefore outside of the laws and policies that have thus far governed the use of healthcare data. While any one of these applications alone might have small risks for patient privacy, as with any large data source, in aggregate the data generated by multiple applications and "skills" can create a comprehensive profile of patient health.

In 2017, JOANY, a Los Angeles-based start-up designed as an "insurance concierge" service that assisted people in navigating their insurance company, find new physicians, and secure health services, recruited participants for a health insurance study by advertising online

and offering a $25 payment for participation. Qualifications for participation included purchasing health insurance for 2017 through Healthcare.gov and not receiving health insurance through an employer or university, among other requirements. Qualifying participants were then asked to text a photo of their health insurance card to the study team to complete their application to participate in the survey study (Kohler, 2017). Although the announcement indicated that JOANY was HIPAA compliant, it was not clear from the article promoting this effort what IRB approval the study received if any, and who participants could contact if they had additional questions. Interested participants may not have sought this information. As of this writing, however, JOANY health concierge has folded, a churn typical of start-up companies. It is unknown what was done with participants data or how that data is handled now that the company has ceased to exist. In their frustration, study participants and JOANY customers have taken to business review sites like Yelp.com to voice their concerns. One reviewer on Yelp.com writes: "I am deeply concerned because I had input my complete billing information on their site […] All I want is what they promised me: a less-hassled way to "renew" the health insurance I need for my family before this upcoming enrollment period ends. [I am] waiting on joany.com to let me know what they did with my financial and personal information (Anonymous, 2018)." An estimated 70-90% of all startup ventures "fail", or close within 10 years. Depending on the reporting source, the failure rate of digital health startups is higher than other industries at 98% (Beckers Hospital Review, 2016). With each sign-up and click to participate and subsequent start up failure, by degrees, individuals lose control over information and data they contribute.

1.3.6 The Public's Response

As of now, legal recourse for consumers who have entrusted health data to commercial companies that go out of business appears to be lacking. While frustrated customers wait for legal remedies, patients have already indicated their desire for more control over their data in established healthcare systems (Rothstein, 2011). Reservations about the use and sharing of health information predates the current rise of data breaches. In a 2003 literature review conducted before widespread use of current social media platforms, Sankar et al. reviewed 110 papers written between 1966 and 2001 on patient views of medical confidentiality and categorized papers into one of four types: 1) understanding and awareness, or the patient's understanding of the confidentiality of their healthcare information; 2) limits of access, or who

should be allowed access to their medical information; 3) effect on seeking care, or how patients' perceptions of the confidentiality of their data affects the decision to seek care; 4) effect on disclosure, or how perceptions of confidentiality influence what patients share with providers (Sankar et al., 2003). Their findings are consistent with health information survey research today: patients are 1) generally comfortable with the sharing of their health information with other providers; 2) patients generally accept that in exceptional circumstances (reason to suspect harm to the welfare of the patient) providers must share health information with law enforcement or other like individuals; and 3) patients reject or are uncomfortable with sharing their health information with employers, family, and third parties (Sankar et al., 2003). The research presented here focuses on this final point: If healthcare partnerships such as those detailed in earlier sections are only going to increase in number, under what circumstances are patients and the public more comfortable with sharing their health data with third-party commercial companies?

1.4 Trust and Privacy Concerns

1.4.1 Trust and Transparency

Without trust, our complex ecosystem of interdependencies would cease to function. I trust my bank to safeguard my money, a restaurant to provide me with food that won't make me sick, and my pharmacy to provide me the right medication in the right amounts. Patients trust that their physicians are well trained and able to treat their ailments. Patients also entrust providers with their most sensitive information with the expectation that data they generate in their healthcare encounter will be private and secure. Trust has been defined as "a willingness to impart authority and accept vulnerability to another in the fulfillment of a given set of tasks" (Platt et al., 2018), or "a bet about the future contingent actions of others" (Sztompka, 1999). We make these bets and accept vulnerability based on the information available to us about an entity's reputation, credentials, and applicable laws and regulations. The Health Information Portability and Accountability Act (HIPAA) is a regulatory control aimed at "reducing the risk of, and harm caused from, sharing and misuse of health data" (Platt et al., 2018), protecting the privacy of patient information and providing patients with the assurance of the law.

Research suggests that individuals are more likely to disclose personal information when higher levels of trust in the benevolence of the trustee and as well as trust in the integrity of the trustee exist (McKnight et al., 2002b). In marketing research, studies of trust economics show that buyers, or consumers of goods, are more likely to do business with "virtuous sellers" (Matouschek, n.d.). Although being virtuous may not mean the most profit in a single transaction, virtuous (or ethical) sellers maintain customer relationships and loyalty and over the long term and generate stronger brands than sellers who view every transaction with the greatest short-term profit in mind (Gensler, 2015; Sinek, 2009). Gaining and maintaining trust, and thus, business loyalty, is ever more important as healthcare tech companies compete for patient business and as the Centers for Medicare and Medicaid's Hospital Value-Based Purchasing program (VBP) shift organizational priorities toward the patient experience (CMS, 2019; Elliott et al., 2016). Whether patients are truly consumers of healthcare or whether this conception is helpful is debatable (Durrah, 2019; Gusmano et al., 2019), but healthcare systems will have to balance aggressive innovation and cutting edge research with the requirements of a strong trustee-trustor relationship.

In an article about building brand trust, Forbes magazine cited "transparency" as one of the 12 most important ways to build brand trust (Council, 2016) and in their 2018 Health Care Engagement Survey, Deloitte concluded that "to continue to foster consumer trust, organizations should focus on developing resources and tools that are centered around consumer needs, including privacy and clarity about how data will be used and shared" (Betts & Korenda, 2018). Approaches of patient data transparency are currently being promoted at the national level through the 2018 MyHealthEData Initiative, which aims to give "every American control of their medical data" and Medicare's Blue Button 2.0, "a secure way for Medicare beneficiaries to access and share their personal health data in a universal digital format, [enabling patients…] to connect their claims data to the secure applications, providers, and research programs they trust" and "foster increased competition among technology innovators to serve Medicare patients and their caregivers" (CMS Press Release, 2018)

1.4.2 Privacy

Addressing privacy concerns can improve trust in organizations that share health information, but characterizing the meaning and scope of "privacy concerns" is not a straightforward or simple undertaking. Privacy can be defined as "the right to be left alone" (Warren & Brandeis, 1890) or "custody of the facts of one's life" (McCreary, 2008). Privacy violations occur when system features do not align with social expectations or violate boundaries of personal preference (Nissenbaum, 2009). Boundaries of personal preference, however, are fluid and dynamic – information that would be considered of little value or importance in one context may be the very information valued and concealed in another. Communication privacy management (CPM) theory assumes people make choices regarding disclosure of personal information based on a "mental calculus" to determine whether to share or withhold information (Petronio, 2013), guiding their participation in shopper's cards, online and credit card transactions, and web browsing. To make these determinations, CPM posits two rules or guides that aid in decision-making: contextual factors and risk-benefit ratio criteria. Contextual factors include prior experiences and traumatic events. The risk-benefit ratio criteria include individual assessments of type and level of risk when deciding to disclose information—which include relational risks, stigma, reputation and security. In further research on Communication Privacy Management, privacy boundaries were found to be influenced by the "mission-relatedness of the request for information"—or the extent to which the information requested aids a fair, identifiable mission (Stanton & Stam, 2002). However, both of these approaches to privacy management, identification of fairness and risk-benefit analysis, require complete information to determine. Surveys show that despite concerns about the privacy of their personal information, individuals will continue to share their personal data in exchange for monetary benefit or convenience, also known as the "privacy paradox", named for the systematic inconsistencies in what people what people say their privacy preferences are and what they do (share information widely).

One explanation for the privacy paradox is the theory of incomplete information. The theory of incomplete information (TII) emerged from John Harsanyi's work on game theory, and assumes that all parties involved in a transaction are not equally informed, which thus impedes the ability of consumers to rationally calculate risks and benefits of engaging in a transaction

(Bandara et al., 2017). All users participating in online transactions are engaged in unavoidable information asymmetry as personal information that is provided in a transaction then multiplies in the course of information intake and processing, propagates as the information is analyzed and shared, and persists for an unpredictable span of time—placing the individual in a position of information asymmetry with respect to the party they are completing the transaction with. The negative utility that may come from future misuses of that information (credit card or personal information breaches, for example) is then impossible to calculate (Acquisti, 2004). In general, data subjects know less than the data holders about the magnitude of data collection and use of shared or collected personal data, or about the associated consequences. Because users lack critical information about true risks and benefits of information sharing and privacy, the immediate benefits of transacting overwhelm the privacy calculus, resulting in the privacy paradox (Bandara et al., 2017).

Six major categories of "unknowns" have been determined to exacerbate this information asymmetry: 1) a user often only has vague and limited knowledge of the actions she can take to protect her personal information, and she has limited knowledge of the actual or possible actions being taken by data gatherers to collect her information; 2) the decision to protect or trade information has unpredictable consequences; 3) the rapid rate of technological developments make true assessment unknowable; 4) desired actions may not be available—specifically, the user cannot regain control over information that has been shared with a third party; 5) users are in a daily negotiation of the effort necessary to evaluate everyday privacy decisions, and 6) privacy protection and invasion is often a by-product of other desired goods—the convenience of online shopping or cost-savings from participation in shoppers clubs interfere with privacy determinations, making it difficult to make true privacy valuations (Grossklags & Acquisti, 2007). However, in one of the few empirical studies on information giving and transaction behavior, Tsai et al. found that when the privacy information of a website is made easier to identify and understand, consumers were more willing to pay a "privacy premium" to conduct business with those websites that offered greater privacy protection (Tsai et al., 2011). The willingness to pay a premium indicates that resolving information asymmetry in transactions can resolve some of the paradoxical behaviors and concerns regarding privacy.

Examination of the role of trust and privacy concerns in comfort with sharing health data with third-party commercial companies allows us to begin disentangling these relationships in pursuit of actionable recommendations to assuage and mitigate patient concerns about the use of health data and close this asymmetrical relationship.

1.5 Conceptual Model

The conceptual model for this dissertation builds on Anderson and Agarwal's Model of Consumer Digitized PHI Concerns (Anderson & Agarwal, 2011). Anderson's model, based on the Communication Privacy Management (CPM) theory previously discussed, explores under what circumstances individuals are willing to disclose identified health information, and what effect information type, requestor, purpose, trust, altruism have on willingness to disclose. Although Anderson and Agarwal hypothesized that the type of information provided and trust in the electronic medium would influence willingness to provide access to personal health information (PHI), in their analysis they concluded that 1) intended purpose, 2) requesting stakeholder, 3) the individuals health status emotion, and 4) altruism or empathy were the actual, final factors associated with willingness to provide access to personal health information.

In my dissertation I expand the Anderson and Agarwal model by exploring the public's comfort with sharing health data with third-party commercial companies for patient purposes versus business purposes (**Figure 1.1**), again, interrogating the public's comfort with sharing healthcare data with third-party companies when data sharing is expressed in terms of patient purposes, i.e., to improve care, diagnosis, or treatment, versus comfort when data is expressed in terms of business purposes, i.e., the sale of de-identified data for artificial intelligence efforts, with the goal of better understanding how presentation of third-party commercial partnerships affect comfort with sharing health data.. In the first analytical paper of this dissertation (Chapter 2) I examine how comfort with sharing health data with third-party commercial companies is related to demographic characteristics, self-reported health status, perceived healthcare access, and privacy concerns (Blue Arrows in Figure 1). In the second analytical paper of this dissertation (Chapter 3) I examine how comfort with third-party commercial companies is related to trust in the health system, trust in healthcare providers, altruism, experience of a past data breach, and concern about recent data breaches and data misuse events (Red Arrows in Figure 1).

In the final paper of this dissertation (Chapter 4) I provide a full model that includes the variables examined in the first and second analytical paper, and add in this third paper an examination of how comfort with third-party commercial companies is related to 1) comfort with researchers, quality analysts, commercial companies, and law enforcement; 2) the public's confidence in existing health data laws and policies; and, finally, 3) the desire for either more control over health data sharing or the desire for notification of data use (Purple Arrows in Figure 1).

In the following sections I provide a short review of each of these concepts as they relate to comfort with data sharing.

Conceptual Model

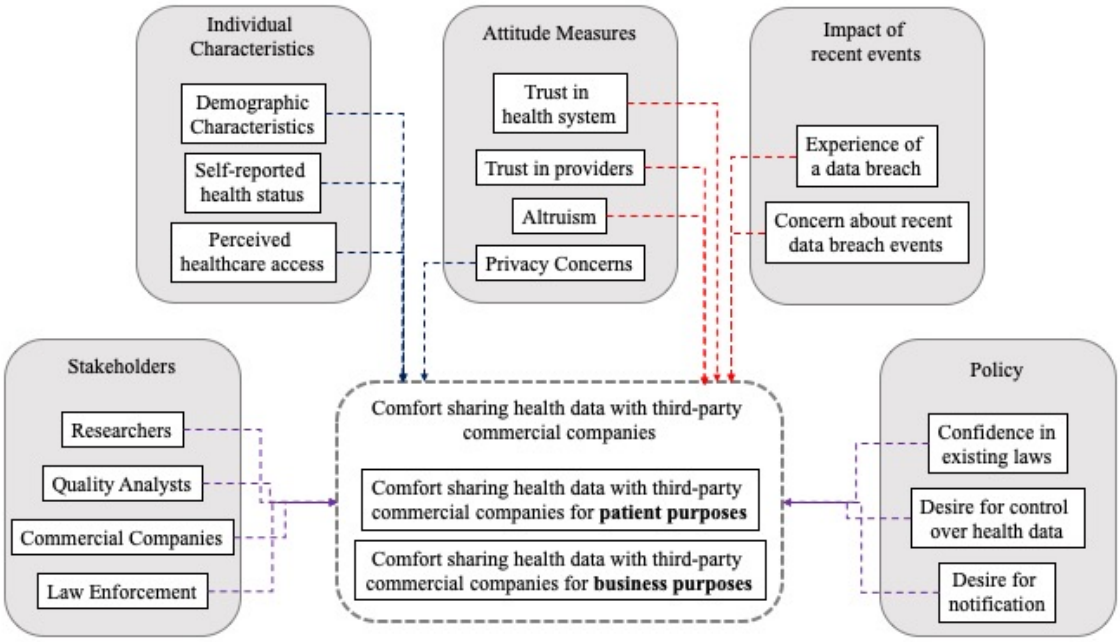

Figure 1-1 Conceptual Model of Dissertation Research

Table 1-1 Independent and dependent variables by chapter

	Ch 2	*Ch 3*	*Ch 4*
Dependent Variables			
Comfort with sharing health data with third-party commercial companies for patient and business purposes	X	X	X
Independent Variables			
Demographic characteristics	X	X	X

Self-reported health status	X	X	X
Perceived healthcare access	X	X	X
Privacy Concerns	X	X	X
Trust in Health System		X	X
Trust in Providers		X	X
Altruism		X	X
Experience of a Past Data Breach		X	X
Concern about recent data breach events		X	X
Comfort with data stakeholders			X
Confidence in existing laws			X
Desire for control over health data			X
Desire for notification of health data use			X

Sharing data for patient versus business purposes

In this dissertation I examine the public's comfort sharing health data with third-party commercial companies, distinguishing comfort when data is shared for patient purposes, i.e., to improve care, diagnosis, or treatment for themselves or for others, versus comfort when data is shared for business purposes such as selling de-identified data (outcome variables). This division allows for the examination of the effect of stated purpose on comfort with sharing health data with third-party commercial companies and is consistent with existing models of consumer willingness to provide access to PHI.

Previous research indicates patients desire more control if their health data will be used for profit-generating activities (Willison et al., 2009), and are more willing to or comfortable with providing access to their health information if the potential health benefits to patients are clear (Anderson & Agarwal, 2011). In a study of US veterans, participants expressed to the study team that research studies must have "high value with an 'overall impact on society' and not be 'an academic exercise' and should consider whether 'just a few hundred [people] or several thousands' would benefit" (Damschroder et al., 2007). This separation of patient-focused purposes from more general business purposes that employed in this dissertation allows for the examination of the impact of stated use on comfort with third-party commercial companies. Although business purposes and patient purposes are closely related, this distinction may not be readily made by patients or by the public. Based on previous research, one would expect greater comfort with third-party commercial use of health information when used for purposes that

clearly state benefits for patients and for healthcare provision, and less comfort when used for business purposes or those purposes that may be more abstract from improvements to patient health. Given that health information is currently used for both patient and business purposes, it is important to understand how factors such as concerns about privacy, healthcare access, and demographic characteristics, are associated with greater or lesser comfort.

In the following section I discuss the component independent variables of the conceptual model and revisit some of the attitudinal variables previously discussed (trust and privacy).

Demographic Characteristics

Although individual privacy attitudes are the result of individual experience and motivations (Petronio, 2013), differences in privacy concerns and willingness to share personal information have been found between men and women, with women evidencing greater privacy concerns and decreased comfort sharing personal information (Sheehan, 1999; Tifferet, 2019). Differences in privacy concerns have also been examined among age groups (Zhou & Salvendy, 2017), with studies showing privacy concerns decrease with increased internet experience (Lohse et al., 2000; Zhou & Salvendy, 2017). Younger adults spend more time online compared to their older counterparts, and 90% of US emerging adults aged 18-29 reporting using social media every day (Scott et al., 2017), leading to the hypothesis that younger people are more comfortable sharing their health information. Differences in willingness to share personal information has also been found according to educational attainment (Blank et al., 2014; J. Kim et al., 2019; Sheehan, 1999) and income (H. Lee et al., 2016; O'Neil, 2001), which find that as educational attainment and income increases, privacy concern also increases while willingness to share information decreases.

Perceived Healthcare Access and Health Status

Anderson and Agarwal introduce a conceptual model examining the role of emotions on willingness to disclose personal health information. In their research they found that highly negative emotions about health status (personal experience with a past or present cancer diagnosis) was associated with increased willingness to share PHI with pharmaceutical companies for clinical trial research. They hypothesize that this effect is due to the participant's

ability to recall what their cancer diagnosis felt like and are subsequently better able to envision the research benefits of sharing their health information as compared to individuals who have never received a cancer diagnosis (Anderson & Agarwal, 2011). Studies on the effect of health status broadly on participant willingness to share health information, however, have been contradictory—in one study, patients with self-rated fair or poor health were less willing to share their health information (Weitzman et al., 2010). In a study involving HIV patients, perceived healthcare access, or the patient's satisfaction with their ability to access necessary healthcare, was associated with increased willingness to share personal health information (Teixeira et al., 2011).

Privacy Concerns

Surveys of the public have routinely found that people are concerned about their personal privacy, but have also found that if individuals value the benefits of information disclosure enough, privacy concerns are outweighed. In the consumer sector, loyalty points, shopper discount cards, and customer accounts are reliant on this suspension of privacy concerns in exchange for monetary benefits or for convenience. Patients also appear willing to share health information in the research context if the benefit to other patients is clear (Damschroder et al., 2007). However, overall willingness to share information is modified by privacy concerns--individuals with greater privacy concerns express greater reluctance to share data even for patient purposes than those with less privacy concerns (Anderson & Agarwal, 2011). Empirical examination of the strength of the specific relationship between privacy concerns and comfort sharing health data with third-party commercial companies has not been provided in previous literature.

Trust in the Health System and Trust in Providers

Trust, or the willingness to accept vulnerability to the actions of others (Richards & Hartzog, 2015), has been shown to increase the likelihood patients will participate in research (K. K. Kim et al., 2017) and is a strong predictor of attitudes and behaviors in other areas such as online shopping (Wang & Tseng, 2011), customer loyalty (Kantsperger & Kunz, 2010), and self-disclosure on social media (Fogel & Nehmad, 2009; Taddei & Contena, 2013). While there are a number of approaches to defining and examining trust, four high level categories have been most

frequently used to capture the dimensions of trust: benevolence, integrity, competency, and predictability or fidelity (McKnight & Chervany, 2001). Benevolence means "caring and being motivated to act in one's interest rather than acting opportunistically"; integrity means "making good faith agreements, telling the truth, and fulfilling promises"; competence means "having the ability or power to do for one what needs to be done"; and predictability means "trustee actions that are consistent enough to be forecasted" (McKnight & Chervany, 2001). Existing research on trust, privacy, and control of information have differed in their results on the mediating or moderating relationship between each. Some have found that trust is the commanding variable, reducing privacy concerns and increasing willingness to share personal information overall (Fogel & Nehmad, 2009), while others have found that privacy is actually the commanding variable, increasing trust and leading to greater disclosure of personal information (Krasnova et al., 2010).

Altruism

Altruism is the prioritization of the needs of others even though no direct benefit may be conferred onto the individual. Individuals who consider themselves altruistic place high value on how their efforts may contribute to the well-being of others. As was stated earlier, studies on patient willingness to participate in healthcare research find that individuals are more willing to participate in research studies if their participation would help a friend or relative or has the potential to benefit society (Doukas & Hardwig, 2014; Reynolds & Nelson, 2007; Shavers et al., 2001).

Impact of Data Breaches and Concern about Recent Events

In a study of healthcare data breaches, healthcare systems lost a reported average of $408 per record due to detection efforts, notification, legal expenditures and fines, and lost business (Ponemon Institute, 2018). This same study found that healthcare organizations worldwide lost customers as a result of data breaches. Studies of data breaches in other sectors have found similar effects. Research on online shopping behavior found that in the wake of a data breach, customers engage in protective behaviors that include avoiding the online store entirely or doubling-down on protective monitoring efforts to mitigate any issues that might arise due to compromised personal information (M. Lee & Lee, 2012). Similar patterns of decreased

customer retention were also identified in a 2002 examination of hotel data breaches – a data breach event resulted in changes to the offline behavior of customers who reported being less willing to revisit and recommend that hotel (Belanger et al., 2002). Yet another study on the effect of data breaches found a negative and statistically significant impact on the market value of the breached companies (Acquisto et al., 2006).

Comfort with Researchers, Quality Analysts, Commercial Companies

Multiple studies have shown that on the whole, patients are willing to share their health data with researchers (Damschroder et al., 2007; Karampela et al., 2019; Seltzer et al., 2019; Spencer et al., 2016; Teixeira et al., 2011). Despite dissatisfaction about how the results of research studies were communicated back to participants, patients were also willing to share data beyond their personal health information for the purposes of research, including tax records and credit card histories (Seltzer et al., 2019). This willingness to share health information is, however, largely confined to only to efforts characterized as "research". As an example of willingness to share health data with third-party commercial companies, a 2019 survey conducted by the Chicago Booth/Kellogg School Financial Trust found that 93% of survey participants were unwilling to share their health data with Facebook (Promarket, 2020).

Comfort with Law Enforcement

Third-party access to health data also includes police. A 2019 survey conducted by the Pew Research Center found that 48% of Americans consider it acceptable for DNA testing companies (23andMe, Ancestry.com) to share their customers' data with law enforcement, while one-third (34%) of respondents said sharing with law enforcement was unacceptable, and 18% were unsure whether the practice was acceptable or unacceptable (Perrin, 2020). As third-party companies seek greater access to health data as well as data that can be considered health data (web histories, fitness applications), law enforcement also gains an increasingly robust data set with which to conduct investigations. In 2019, a Florida judge granted a warrant that allowed the police to search the complete genetic database of GEDMatch. The terms of the warrant included all customers of GEDMatch, including those who didn't opt-in to any data sharing agreement. Following this event, GEDMatch updated their terms of agreement and added a form asking users to consent to future searches. Notably, as of November 2019, only 185,000 of the

company's total population of 1.3 million users, a mere 7% of the total user population, have provided this consent (Tiller, 2019).

Confidence in existing laws and policies

As healthcare breaches increase year after year, patient confidence in the ability of the health system to control and prevent unauthorized access to their personal health information is decreasing (HIPAA Journal, 2017). In a 2016 survey, 89% of healthcare consumers reported withholding health information during their visit because of privacy and security concerns (Black Book Market Research, 2017). Research shows patients desire clear and consistent consequences for anyone who violates patient privacy, and for researchers to be held accountable for maintaining confidentiality (Damschroder et al., 2007), but changes to regulations governing information sharing have not yet occurred. In 2007, only 25% of participants were aware that researchers could use their medical records without explicit permission from the patient (Damschroder et al., 2007). In 2016, 81% of respondents to a Black Book Market Research survey (Black Book Market Research, 2017) reported concern that their chronic condition data was being shared with retailers, employers, or the government without their knowledge (Gooch, 2017).

Desire for control and notification

Research on willingness to allow for personal health data to be used for research have found that in general, 96% of patients were willing to provide their data for research, yet 78% also indicated their desire for more control over how their information was used (Damschroder et al., 2007). Patients have consistently reported wanting to know how their health data may have contributed to helping others, and who was using their medical records for what purpose (Damschroder et al., 2007; J. Kim et al., 2019; Weitzman et al., 2010). Desire for notification persists even when there is high institutional trust (Damschroder et al., 2007).

1.6 Conclusion

Health systems must navigate patient privacy and innovation carefully as big data efforts in healthcare proliferate and the healthcare marketplace is inundated with new commercial entrants. Examination of the public's comfort with sharing health data with third-party

commercial companies is critical to mitigate public backlash and protect public and patient trust. In this dissertation, I consider a number of factors that may be associated with this comfort to inform future studies, health systems, and designers who can help bridge the gap between patients and the systems that serve them to ensure both privacy and trust.

1.7 References

Acquisti, A. (2004). Privacy and Security of Personal Information. In L. J. Camp & S. Lewis (Eds.), *Economics of Information Security* (Vol. 12, pp. 179–186). Kluwer Academic Publishers. https://doi.org/10.1007/1-4020-8090-5_14

Acquisto, A., Friedman, A., & Telang, R. (2006). Is There a Cost to Privacy Breaches? An Event Study. *Twenty-Seventh International Conference on Information Systems, Milwaukee*, 19.

Agaku, I. T., Adisa, A. O., Ayo-Yusuf, O. A., & Connolly, G. N. (2014). Concern about security and privacy, and perceived control over collection and use of health information are related to withholding of health information from healthcare providers. *Journal of the American Medical Informatics Association*, *21*(2), 374–378. https://doi.org/10.1136/amiajnl-2013-002079

Anderson, C., & Agarwal, R. (2011). The Digitization of Healthcare: Boundary Risks, Emotion, and Consumer Willingness to Disclose Personal Health Information. *Information Systems Research*, *22*(3), 469–490. https://doi.org/10.1287/isre.1100.0335

Anonymous. (2018, November 24). *JOANY - Downtown—Los Angeles, CA*. Yelp. https://www.yelp.com/biz/joany-los-angeles?hrid=N2YrwuHx30f9FczxeKbqng&utm_campaign=www_review_share_popup&utm_medium=copy_link&utm_source=(direct)

Arndt, R. Z. (2018, April 7). *How third parties harvest health data from providers, payers and pharmacies*. Modern Healthcare.

https://www.modernhealthcare.com/article/20180407/NEWS/180409938/how-third-parties-harvest-health-data-from-providers-payers-and-pharmacies

Bandara, R., Fernando, M., & Akter, S. (2017). The Privacy Paradox in the Data-Driven Marketplace: The Role of Knowledge Deficiency and Psychological Distance. *Procedia Computer Science*, *121*, 562–567. https://doi.org/10.1016/j.procs.2017.11.074

Beaver, L. (2018, July 24). *The top 5 startups disrupting healthcare within AI, digital therapeutics, health insurance, and genomics*. Business Insider. https://www.businessinsider.com/7-24-2018-digital-health-startups-to-watch-2018-7

Beckers Hospital Review. (2016, May 18). *98% of digital health startups fail—Here's why*. Becker's Health IT. https://www.beckershospitalreview.com/healthcare-information-technology/98-of-digital-health-startups-fail-here-s-why.html

Belanger, F., Hiller, J. S., & Smith, W. J. (2002). Trustworthiness in electronic commerce: The role of privacy, security, and site attributes. *The Journal of Strategic Information Systems*, *11*(3), 245–270. https://doi.org/10.1016/S0963-8687(02)00018-5

Betts, D., & Korenda, L. (2018, September 25). *Patient Engagement findings—2018 Health Care Consumer Survey | Deloitte Insights*. Deloitte Insights. https://www2.deloitte.com/us/en/insights/industry/health-care/patient-engagement-health-care-consumer-survey.html

Black Book Market Research. (2017, January 3). *Healthcare's Digital Divide Widens, Black Book Consumer Survey*. Black Book Market Research. https://blackbookmarketresearch.newswire.com/news/healthcares-digital-divide-widens-black-book-consumer-survey-18432252

Blank, G., Bolsover, G., & Dubois, E. (2014). A New Privacy Paradox: Young People and Privacy on Social Network Sites. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2479938

Burde, H. (2011). The HITECH Act: An Overview. *AMA Journal of Ethics*, *13*(3), 172–175. https://doi.org/10.1001/virtualmentor.2011.13.3.hlaw1-1103.

Caine, K., & Hanania, R. (2013). Patients want granular privacy control over health information in electronic medical records. *Journal of the American Medical Informatics Association*, *20*(1), 7–15. https://doi.org/10.1136/amiajnl-2012-001023

Castell, S., & Evans, H. (2016). The One-Way Mirror: Public attitudes to commercial access to health data. *Ipsos MORI Social Research Institute*, 161.

CMS. (2019, October 29). *Hospital Value-Based Purchasing*. CMS.Gov. https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/HVBP/Hospital-Value-Based-Purchasing.html

CMS Press Release. (2018, March 6). *Trump Administration Announces MyHealthEData Initiative to Put Patients at the Center of the US Healthcare System | CMS*. CMS.Gov. https://www.cms.gov/newsroom/press-releases/trump-administration-announces-myhealthedata-initiative-put-patients-center-us-healthcare-system

Cohen, I. G., & Mello, M. M. (2018). HIPAA and Protecting Health Information in the 21st Century. *JAMA*, *320*(3), 231–232. https://doi.org/10.1001/jama.2018.5630

CooperKatz. (2018, February 5). *Paige.AI Created to Transform Cancer Diagnosis and Treatment by Applying Artificial Intelligence to Pathology*. Business Wire. https://www.businesswire.com/news/home/20180205005557/en/Paige.AI-Created-Transform-Cancer-Diagnosis-Treatment-Applying

Council, F. A. (2016, September 19). *12 Ways New Companies Can Build Brand Trust*. Forbes. https://www.forbes.com/sites/forbesagencycouncil/2016/09/19/12-ways-new-companies-can-build-brand-trust/

Curran, T. (2019, October 29). *Data exchange at UPMC* [Personal communication].

Damschroder, L. J., Pritts, J. L., Neblo, M. A., Kalarickal, R. J., Creswell, J. W., & Hayward, R. A. (2007). Patients, privacy and trust: Patients' willingness to allow researchers to access their medical records. *Social Science & Medicine*, *64*(1), 223–235. https://doi.org/10.1016/j.socscimed.2006.08.045

Dhopeshwarkar, R. V., Kern, L. M., O'Donnell, H. C., Edwards, A. M., & Kaushal, R. (2012). Health Care Consumers' Preferences Around Health Information Exchange. *The Annals of Family Medicine*, *10*(5), 428–434. https://doi.org/10.1370/afm.1396

Donilon, T. E. (2016). Report on securing and growing the digital economy. *Commission on Enhancing National Cybersecurity*.

Doukas, D. J., & Hardwig, J. (2014). Patient Informed Choice for Altruism. *Cambridge Quarterly of Healthcare Ethics*, *23*(4), 397–402. https://doi.org/10.1017/S0963180114000073

Durrah, H. (2019). My Child Is Sick; Don't Call Her A 'Consumer.' *Health Affairs*, *38*(3), 502–505. https://doi.org/10.1377/hlthaff.2018.05012

Elliott, M. N., Beckett, M. K., Lehrman, W. G., Cleary, P., Cohea, C. W., Giordano, L. A., Goldstein, E. H., & Damberg, C. L. (2016). Understanding The Role Played By Medicare's Patient Experience Points System In Hospital Reimbursement. *Health Affairs*, *35*(9), 1673–1680. https://doi.org/10.1377/hlthaff.2015.0691

Farr, C. (2014, June 2). *Apple unwraps "Healthkit" alongside Mac, iPhone features | Reuters*. Reuters. https://www.reuters.com/article/us-apple-developers/apple-unwraps-healthkit-alongside-mac-iphone-features-idUSKBN0ED1V820140602

Fogel, J., & Nehmad, E. (2009). Internet social network communities: Risk taking, trust, and privacy concerns. *Computers in Human Behavior*, *25*(1), 153–160. https://doi.org/10.1016/j.chb.2008.08.006

Francis, L. P., & Francis, J. G. (2017). *Privacy: What Everyone Needs to Know*. Oxford University Press.

Gensler, A. (2015, July 28). *Trust is the most powerful currency in business*. Fortune. https://fortune.com/2015/07/28/trust-business-leadership/

Gooch, K. (2017, January 3). *Privacy issues drive health IT consumer skepticism: 10 Black Book survey findings*. Becker's Health IT. https://www.beckershospitalreview.com/healthcare-information-technology/privacy-issues-drive-health-it-consumer-skepticism-10-black-book-survey-findings.html

Grossklags, J., & Acquisti, A. (2007). What Can Behavioral Economics Teach Us about Privacy? In S. D. C. di Vimercati, S. Gritzalis, C. Lambrinoudakis, & A. Acquisti (Eds.), *Digital Privacy* (pp. 363–377). Auerbach Publications. https://doi.org/10.1201/9781420052183.ch18

Gusmano, M. K., Maschke, K. J., & Solomon, M. Z. (2019). Patient-Centered Care, Yes; Patients As Consumers, No. *Health Affairs*, *38*(3), 368–373. https://doi.org/10.1377/hlthaff.2018.05019

Hill, K. (2012, February 6). How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did. *Forbes*. https://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/#2e7213656668

HIPAA Journal. (2017, January 5). Patients Holding Back Health Information Over Data Privacy Fears. *HIPAA Journal*. https://www.hipaajournal.com/patients-holding-back-health-information-over-fears-of-data-privacy-8634/

HIPAA Journal. (2019, January 28). Analysis of 2018 Healthcare Data Breaches. *HIPAA Journal*. https://www.hipaajournal.com/analysis-of-healthcare-data-breaches/

HITECH Act, Pub. L. No. 111–5, 123 Stat. 115 (2009). https://www.hhs.gov/sites/default/files/ocr/privacy/hipaa/understanding/coveredentities/hitechact.pdf

Huang, B. (2018, September 25). *LVHN patient data not shared with for-profit company in Sloan Kettering trials*. Mcall.Com. https://www.mcall.com/health/mc-nws-lvhn-msk-paigeai-20180924-story.html

Kantsperger, R., & Kunz, W. H. (2010). Consumer trust in service companies: A multiple mediating analysis. *Managing Service Quality: An International Journal*, *20*(1), 4–25. https://doi.org/10.1108/09604521011011603

Karampela, M., Ouhbi, S., & Isomursu, M. (2019). Connected Health User Willingness to Share Personal Health Data: Questionnaire Study. *Journal of Medical Internet Research*, *21*(11). https://doi.org/10.2196/14537

Kim, J., Kim, H., Bell, E., Bath, T., Paul, P., Pham, A., Jiang, X., Zheng, K., & Ohno-Machado, L. (2019). Patient Perspectives About Decisions to Share Medical Data and

Biospecimens for Research. *JAMA Network Open*, *2*(8), e199550–e199550. https://doi.org/10.1001/jamanetworkopen.2019.9550

Kim, K. K., Sankar, P., Wilson, M. D., & Haynes, S. C. (2017). Factors affecting willingness to share electronic health data among California consumers. *BMC Medical Ethics*, *18*(1), 25. https://doi.org/10.1186/s12910-017-0185-x

Kohler, C. (2017, October 11). *This Health Insurance Survey Will Pay You $25 for 10 Minutes of Your Time* [Text]. The Penny Hoarder. http://www.thepennyhoarder.com/make-money/joany-paid-health-insurance-survey/

Krasnova, H., Spiekermann, S., Koroleva, K., & Hildebrand, T. (2010). Online Social Networks: Why We Disclose. *Journal of Information Technology*, *25*(2), 109–125. https://doi.org/10.1057/jit.2010.6

Kuchinke, W., Ohmann, C., Verheij, R. A., Veen, E.-B. van, & Delaney, B. C. (2016). Development Towards a Learning Health System—Experiences with the Privacy Protection Model of the TRANSFoRm Project. In *Data Protection on the Move* (pp. 101–134). Springer, Dordrecht. https://doi.org/10.1007/978-94-017-7376-8_5

Lee, H., Wong, S. F., & Chang, Y. (2016). Confirming the Effect of Demographic Characteristics on Information Privacy Concerns. *PACIS 2016 Proceedings*, 8.

Lee, M., & Lee, J. (2012). The impact of information security failure on customer behaviors: A study on a large-scale hacking incident on the internet. *Information Systems Frontiers*, *14*(2), 375–393. https://doi.org/10.1007/s10796-010-9253-1

Lohse, G. L., Bellman, S., & Johnson, E. J. (2000). Consumer buying behavior on the Internet: Findings from panel data. *Journal of Interactive Marketing*, *14*(1), 15.

Marr, B. (2018, May 21). *How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read*. Forbes. https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/

Matouschek, N. (n.d.). *Trust Economics: An Economist's Perspective*. Retrieved November 1, 2019, from https://www.kellogg.northwestern.edu/trust-project/videos/matouschek-ep-1.aspx

McCreary, L. (2008, October 1). *What Was Privacy?* Harvard Business Review. https://hbr.org/2008/10/what-was-privacy

McKnight, & Chervany. (2001). Trust and Distrust Definitions: One Bite at a Time. In R. Falcone, M. Singh, & Y.-H. Tan (Eds.), *Trust in Cyber-societies* (pp. 27–54). Springer. https://doi.org/10.1007/3-540-45547-7_3

McKnight, D. H., Choudhury, V., & Kacmar, C. (2002a). Developing and Validating Trust Measures for e-Commerce: An Integrative Typology. *Information Systems Research*, *13*(3), 334–359. https://doi.org/10.1287/isre.13.3.334.81

McKnight, D. H., Choudhury, V., & Kacmar, C. (2002b). The impact of initial consumer trust on intentions to transact with a web site: A trust building model. *The Journal of Strategic Information Systems*, *11*(3–4), 297–323. https://doi.org/10.1016/S0963-8687(02)00020-3

Medicare, C. for, Baltimore, M. S. 7500 S. B., & Usa, M. (2013, June 3). *Privacy Act of 1974*. CMS.Gov. https://www.cms.gov/Research-Statistics-Data-and-Systems/Computer-Data-and-Systems/Privacy/PrivacyActof1974.html

Memorial Sloan Kettering. (2018, September 23). *Memorial Sloan Kettering and Paige.AI*. Memorial Sloan Kettering Cancer Center. https://www.mskcc.org/press-releases/msk-and-paige-ai

Nissenbaum, H. (2009). *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford University Press.

O'Brien, E. C., Rodriguez, A. M., Kum, H.-C., Schanberg, L. E., Fitz-Randolph, M., O'Brien, S. M., & Setoguchi, S. (2019). Patient perspectives on the linkage of health data for research: Insights from an online patient community questionnaire. *International Journal of Medical Informatics*, *127*, 9–17. https://doi.org/10.1016/j.ijmedinf.2019.04.003

Office for Civil Rights (OCR). (2013). *190-Who must comply with HIPAA privacy standards* [Text]. HHS.Gov. https://www.hhs.gov/hipaa/for-professionals/faq/190/who-must-comply-with-hipaa-privacy-standards/index.html

Ohm, P. (2010). Broken Promises of Privacy: Responding to the surprising failure of anonymization. *UCLA LAW REVIEW*, *57*, 77.

ONC. (2015). Guide to Privacy and Security of Electronic Health Information. *Office of the National Coordinator for Health Information Technology*, 62.

O'Neil, D. (2001). Analysis of Internet Users' Level of Online Privacy Concerns. *Social Science Computer Review*, *19*(1), 17–31. https://doi.org/10.1177/089443930101900103

Ornstein, C., & Thomas, K. (2018a, September 8). *Top Cancer Researcher José Baselga Fails to Disclose Corporate Financial Ties in Major Research Journals*. ProPublica. https://www.propublica.org/article/doctor-jose-baselga-cancer-researcher-corporate-financial-ties

Ornstein, & Thomas. (2018b, September 20). *Sloan Kettering's Cozy Deal With Start-Up Ignites a New Uproar*. ProPublica. https://www.propublica.org/article/sloan-kettering-cozy-deal-with-start-up-paige-ai-ignites-new-uproar

Park, A. (2019, July 2). *12 AI initiatives launched by hospitals, health systems in 2019*. Becker's Hospital Review. https://www.beckershospitalreview.com/artificial-intelligence/12-ai-initiatives-launched-by-hospitals-health-systems-in-2019.html

Perrin, A. (2020, February 4). About half of Americans are OK with DNA testing companies sharing user data with law enforcement. *Pew Research Center*. https://www.pewresearch.org/fact-tank/2020/02/04/about-half-of-americans-are-ok-with-dna-testing-companies-sharing-user-data-with-law-enforcement/

Petronio, S. (2013). Brief Status Report on Communication Privacy Management Theory. *Journal of Family Communication*, *13*(1), 6–14. https://doi.org/10.1080/15267431.2013.743426

Petrow, S. (2018, October 3). Memorial Sloan Kettering, you've betrayed my trust. *STAT*. https://www.statnews.com/2018/10/03/memorial-sloan-kettering-betrayed-my-trust/

Platt, J. E., Jacobson, P. D., & Kardia, S. L. R. (2018). Public Trust in Health Information Sharing: A Measure of System Trust. *Health Services Research*, *53*(2), 824–845. https://doi.org/10.1111/1475-6773.12654

Ponemon Institute. (2018). *2018 Cost of Data Breach Study: Global Overview*. IBM Security. https://www.ibm.com/downloads/cas/861MNWN2

Promarket. (2020, February 7). The Real Price of Health Data: Americans Don't Want to Share Their Records for Free. *Pro Market*. https://promarket.org/2020/02/07/the-real-price-of-health-data-americans-dont-want-to-share-their-records-for-free/

Ramsey, L. (2019, April 4). *Amazon's Alexa can now schedule doctor's appointments and give you updates on your prescription drug shipments*. Business Insider. https://www.businessinsider.com/amazons-alexa-adds-healthcare-skills-2019-4

Reynolds, W. W., & Nelson, R. M. (2007). Risk perception and decision processes underlying informed consent to research participation. *Social Science & Medicine*, *65*(10), 2105–2115. https://doi.org/10.1016/j.socscimed.2007.06.021

Richards, N. M., & Hartzog, W. (2015). Taking Trust Seriously in Privacy Law. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2655719

Ross, C., & Swetlitz, I. (2018, July 25). IBM's Watson recommended "unsafe and incorrect" cancer treatments. *STAT*. https://www.statnews.com/2018/07/25/ibm-watson-recommended-unsafe-incorrect-treatments/

Rothstein, M. A. (2011, February 17). *Debate Over Patient Privacy Controls in Electronic Health Records*. The Hastings Center. https://www.thehastingscenter.org/debate-over-patient-privacy-controls-in-electronic-health-records/

Sankar, P., Mora, S., Merz, J. F., & Jones, N. L. (2003). Patient Perspectives of Medical Confidentiality. *Journal of General Internal Medicine*, *18*(8), 659–669. https://doi.org/10.1046/j.1525-1497.2003.20823.x

Scott, C. F., Bay-Cheng, L. Y., Prince, M. A., Nochajski, T. H., & Collins, R. L. (2017). Time spent online: Latent profile analyses of emerging adults' social media use. *Computers in Human Behavior*, *75*, 311–319. https://doi.org/10.1016/j.chb.2017.05.026

Seltzer, E., Goldshear, J., Guntuku, S. C., Grande, D., Asch, D. A., Klinger, E. V., & Merchant, R. M. (2019). Patients' willingness to share digital health and non-health data for

research: A cross-sectional study. *BMC Medical Informatics and Decision Making*, *19*(1), 157. https://doi.org/10.1186/s12911-019-0886-9

Shavers, V. L., Lynch, C. F., & Burmeister, L. F. (2001). Factors that influence African-Americans' willingness to participate in medical research studies. *Cancer*, *91*(S1), 233–236. https://doi.org/10.1002/1097-0142(20010101)91:1+<233::AID-CNCR10>3.0.CO;2-8

Sheehan, K. B. (1999). An investigation of gender differences in on-line privacy concerns and resultant behaviors. *Journal of Interactive Marketing*, *13*(4), 15.

Sinek, S. (2009). *Start With Why*. Penguin Group. https://simonsinek.com/product/start-with-why/

Spencer, K., Sanders, C., Whitley, E. A., Lund, D., Kaye, J., & Dixon, W. G. (2016). Patient Perspectives on Sharing Anonymized Personal Health Data Using a Digital System for Dynamic Consent and Research Feedback: A Qualitative Study. *Journal of Medical Internet Research*, *18*(4), e66. https://doi.org/10.2196/jmir.5011

Stanton, J. M., & Stam, K. R. (2002). Information Technology, Privacy, and Power within Organizations: A view from Boundary Theory and Social Exchange perspectives. *Surveillance & Society*, *1*(2), 152–190. https://doi.org/10.24908/ss.v1i2.3351

Sullivan, T. (2017, September 20). *Cedars-Sinai kicks off new health tech accelerator class*. Healthcare IT News. https://www.healthcareitnews.com/news/cedars-sinai-kicks-new-health-tech-accelerator-class

Sztompka, P. (1999). *Trust: A Sociological Theory*. Cambridge University Press.

Taddei, S., & Contena, B. (2013). Privacy, trust and control: Which relationships with online self-disclosure? *Computers in Human Behavior*, *29*(3), 821–826. https://doi.org/10.1016/j.chb.2012.11.022

Tanner, A. (2016, February 1). *How Data Brokers Make Money Off Your Medical Records*. Scientific American. https://doi.org/10.1038/scientificamerican0216-26

Teixeira, P. A., Gordon, P., Camhi, E., & Bakken, S. (2011). HIV patients' willingness to share personal health information electronically. *Patient Education and Counseling*, *84*(2), e9–e12. https://doi.org/10.1016/j.pec.2010.07.013

The Privacy Advisor. (2012, November 5). *Electronic Health Records vs. Patient Privacy: Who Will Win?* ID Experts. https://www.idexpertscorp.com/knowledge-center/single/electronic-health-records-vs.-patient-privacy-who-will-win

Tifferet, S. (2019). Gender differences in privacy tendencies on social network sites: A meta-analysis. *Computers in Human Behavior*, *93*, 1–12. https://doi.org/10.1016/j.chb.2018.11.046

Tiller, J. (2019, November 12). *If you've given your DNA to a DNA database, US police may now have access to it*. The Conversation. http://theconversation.com/if-youve-given-your-dna-to-a-dna-database-us-police-may-now-have-access-to-it-126680

Tsai, J. Y., Egelman, S., Cranor, L., & Acquisti, A. (2011). The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study. *Information Systems Research*, *22*(2), 254–268. https://doi.org/10.1287/isre.1090.0260

Uberoi, N., Finegold, K., & Gee, E. (2016). *Health Insurance Coverage and The Affordable Care Act, 2010-2016* [Issue Brief]. Department of Health and Human Services. https://aspe.hhs.gov/system/files/pdf/187551/ACA2010-2016.pdf

Upton, F. (2015, July 13). *H.R.6 - 114th Congress (2015-2016): 21st Century Cures Act* [Webpage]. https://www.congress.gov/bill/114th-congress/house-bill/6

Wang, T.-L., & Tseng, Y. F. (2011). A Study of the Effect on Trust and Attitude with Online Shopping. *International Journal for Digital Society*, *2*(2), 433–440. https://doi.org/10.20533/ijds.2040.2570.2011.0052

Warren, & Brandeis. (1890). The Right to Privacy. *Harv. L. Rev.*, *IV*(5). http://groups.csail.mit.edu/mac/classes/6.805/articles/privacy/Privacy_brand_warr2.html

Weitzman, E. R., Kaci, L., & Mandl, K. D. (2010). Sharing Medical Data for Health Research: The Early Personal Health Record Experience. *Journal of Medical Internet Research*, *12*(2). https://doi.org/10.2196/jmir.1356

Willison, D. J., Steeves, V., Charles, C., Schwartz, L., Ranford, J., Agarwal, G., Cheng, J., & Thabane, L. (2009). Consent for use of personal information for health research: Do people with potentially stigmatizing health conditions and the general public differ in their opinions? *BMC Medical Ethics*, *10*(1), 10. https://doi.org/10.1186/1472-6939-10-10

Zhou, J., & Salvendy, G. (Eds.). (2017). *Human Aspects of IT for the Aged Population. Applications, Services and Contexts: Third International Conference, ITAP 2017, Held as Part of HCI International 2017, Vancouver, BC, Canada, July 9-14, 2017, Proceedings, Part II* (Vol. 10298). Springer International Publishing. https://doi.org/10.1007/978-3-319-58536-9

Chapter 2 The Public's Comfort with Sharing Health Data with Third-Party Commercial Companies

2.1 Abstract

Background: Healthcare systems are using big data-driven methods to realize the vision of learning health systems and improve care quality. In so doing, many are partnering with third-party commercial companies to provide novel data processing and analysis capabilities while also providing personal health information to a for-profit industry that may store and sell data.

Objective: To describe the public's comfort with sharing health data with third-party commercial companies for patient and business purposes and how this comfort is associated with demographic factors (sex, age, race/ethnicity, education, employment, income, insurance status, and self-reported health status), perceived healthcare access, and concerns about privacy.

Methods: We surveyed the US public (n = 1841) to assess comfort with sharing health data with third-party commercial companies for patient or business purposes. We examined whether there was a difference between comfort with data sharing for patient or business purposes and then used univariate and stepwise regression modeling to estimate the relationship between comfort with third-party commercial companies for patient and business purposes (outcomes) and demographic factors, self-reported health status, perceived healthcare access, and privacy concerns.

Findings: The public is more comfortable sharing health data with third party commercial companies for patient purposes as compared to business purposes (paired t = 39.84, p < 0.001). Higher education was associated with greater comfort with sharing health data for patient purposes (β = 0.205, p< 0.001) and decreased comfort with sharing health data for business purposes (β = -0.145, p = 0.079). There was an inverse relationship between privacy concerns

and comfort with sharing health data for both patient (β = -0.223, $p < 0.001$) and business purposes (β = -0.246, $p < 0.001$). Participants ages 45-59 were less comfortable sharing health data with third party commercial companies for patient purposes (β = -0.154, $p = 0.0012$) than other age groups.

Implications: Proactive acknowledgment of privacy concerns and better communication of the steps being taken to protect the privacy of health data can increase patient comfort. Healthcare systems may be able to increase public and patient comfort with sharing health data with third-party commercial companies by emphasizing the patient-centered benefits of these partnerships.

2.2 Introduction

2.2.1 Background

In the fall of 2019, Google and Ascension announced a data partnership called "Nightingale". As part of the effort, Ascension, the largest non-profit healthcare system in the United States, moved identifiable patient records onto Google's cloud servers to begin data analysis on a subset of Ascension's patient population of 50 million people (Copeland & Needleman, 2019). News coverage of the partnership included language such as "secretly gathering personal health records (Griggs, 2019)" and "Google: You can trust us with the medical data you didn't know we already had (Brodkin, 2019)". What likely began as an exciting data-discovery partnership has since devolved into a full investigation by the Office for Civil Rights in the Department of Health and Human Services (Brodkin, 2019). This response by the public, however, was not unprecedented. At the time of the announcement in November 2019, Google and the University of Chicago were being sued for the use of identifiable patient records without consent (Wakabayashi, 2019), and in Fall 2018, news coverage revealing the details of Sloan Kettering's external startup venture known as Paige.AI resulted in an internal, system-wide review of all third-party commercial company data sharing agreements and a breakdown in community trust (Singer & Wakabayashi, 2019; Vincent, 2019). However poorly received they may have been, partnerships and electronic personal health information (ePHI) data sharing agreements like these are key to realizing the potential of big data efforts in healthcare: personalized medicine, better understanding of rare diseases, and reduction of prescription errors,

among other efforts. Third-party commercial companies have been routinely used by healthcare systems to extend the often-limited in-house capabilities of diagnostic testing and image analysis. But as healthcare systems increasingly participate in various data and technology ventures and third-party partnerships, the ethics of these partnerships and the responsibility of the healthcare system to the patient are being questioned.

Healthcare systems in partnership with third-party commercial companies, Ascension, University of Chicago, and Sloan Kettering included, are doing so in compliance with existing HIPAA regulations and approval from Institutional Review Boards as required (Ornstein & Thomas, 2018). The accompanying media coverage, however, strongly suggests that compliance with existing laws is insufficient for patients and for the public (Landi, 2020). Previous research on patient willingness or comfort with sharing healthcare data indicate reservations about the use of healthcare data outside of those services needed to provide direct care (J. Kim et al., 2019) and concern about the motivations behind the use of health data (Stockdale et al., 2019). Despite these reservations, patients are largely supportive of research efforts and generally look forward to the potential healthcare insights offered by large patient data sets (Doukas & Hardwig, 2014; Reynolds & Nelson, 2007; Shavers et al., 2001). The challenge for healthcare systems and their third-party commercial partnerships is to reconcile patient privacy concerns with the patient's desire to contribute to the potential of big data to improve care outcomes, quality, and healthcare efficiency as healthcare systems grow increasingly reliant on third-party commercial companies to realize big data goals. Examination of the public's comfort with sharing healthcare data with third-party commercial companies and their privacy concerns is needed to guide the manner in which future healthcare system partnerships with third-party commercial companies are planned and communicated, so that healthcare organizations can mitigate the possibility of public surprise and backlash.

2.2.2 Conceptual Model

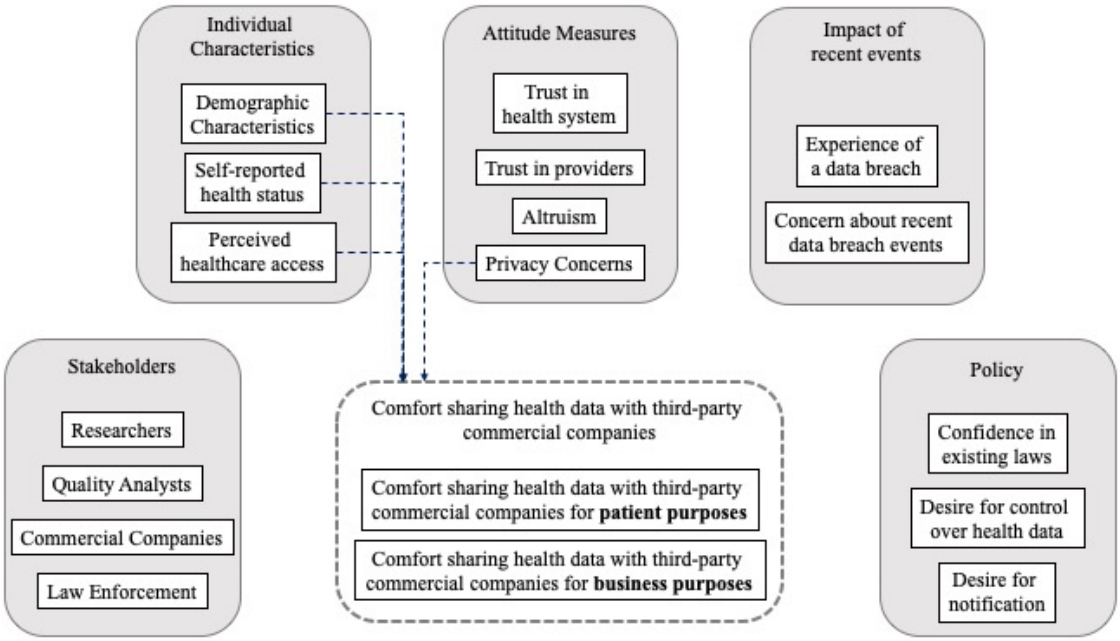

Figure 2-1 Conceptual Model of Dissertation Research – this analysis is focused on individual characteristics, perceived healthcare access and health status, and privacy concerns

Difference between business v. patient purposes

Healthcare systems do indeed blur the use of patient and commercial use and benefit, however, there may be an underlying assumption that the patient is aware of how business uses of patient data improve patient care. In the case of Sloan Kettering, healthcare organizations may emphasize the innovative nature of the partnership and the benefits accrued to the health system instead of tangible improvements to patient care (CooperKatz, 2018). We dichotomize patient uses and business uses, even though these two concepts are intertwined, interrogating comfort when data is shared for patient purposes, i.e., to improve care, diagnosis, or treatment, versus comfort when data is shared for business purposes, i.e., the sale of de-identified data for artificial intelligence efforts, allows for examination of the effect of communicated purpose of use on comfort with sharing health data with third-party commercial companies. Comfort can also be understood as "acceptance", "openness", or "willingness" to share health data.

The division we employ is consistent with existing models of consumer willingness to provide access to PHI. Previous research indicates patients desire more control if their health

data will be used for profit-generating research (Willison et al., 2009), and are more willing to provide access to their health information if the potential health benefits to the public are clear (Anderson & Agarwal, 2011; Castell & Evans, 2016). The public has been found to be less accepting of data sharing partnerships not only when the public health benefits of the partnership are not made clear, but also if the data sharing relationship was determined to be of only private benefit (Castell & Evans, 2016). In a study of US veterans, participants expressed to the study team that research studies must have "high value with an 'overall impact on society' and not be 'an academic exercise' and should consider whether 'just a few hundred [people] or several thousands' would benefit" (Damschroder et al., 2007). Given that health information is currently used for both patient and business purposes, it is worthwhile to examine how stated intent impacts comfort with third-party commercial companies and the relationship of this comfort with concerns about privacy, healthcare access, and demographic factors.

To further characterize comfort with sharing health data with third-party commercial companies, we explore the impact of privacy concerns, perceived healthcare access, and demographic characteristics on comfort with sharing health data with third-party commercial companies.

Privacy Concerns

Overall willingness to share information is modified by privacy concerns-- individuals with greater privacy concerns express greater reluctance to share data even for patient purposes than those with less privacy concerns (Anderson & Agarwal, 2011). However, empirical examination of the strength of the relationship between privacy concerns and comfort with sharing health data with third-party commercial companies has not been provided in previous literature.

Perceived Healthcare Access and Health Status

Highly negative emotions about health status (personal experience with a past or present cancer diagnosis) is associated with an increased willingness to share personal health information (PHI) with pharmaceutical companies for clinical trial research (Anderson & Agarwal, 2011). However, studies on the effect of health status broadly on participant willingness to share health

information have been contradictory—in one study, patients with self-rated fair or poor health were less willing to share their health information (Weitzman et al., 2010). In a study involving HIV patients, perceived healthcare access, or the patient's satisfaction with their ability to access necessary healthcare, was associated with increased willingness to share personal health information (Teixeira et al., 2011).

Demographics

Differences in willingness to share personal information has also been found according to educational attainment (Blank et al., 2014; J. Kim et al., 2019; Sheehan, 1999) and income (Lee et al., 2016; O'Neil, 2001). These studies have found that as educational attainment and income increases, willingness to share information decreases.

2.2.3 Study Objective

The aim of this study is to interrogate the public's comfort with sharing healthcare data with third-party companies when data sharing is expressed in terms of patient purposes, i.e., to improve care, diagnosis, or treatment, versus comfort when data is expressed in terms of business purposes, i.e., the sale of de-identified data for artificial intelligence efforts, with the goal of better understanding how presentation of third-party commercial partnerships affect comfort with sharing health data. We specifically examine how privacy concerns, perceived healthcare access, and demographic factors including self-reported health status are associated with both types of comfort with data sharing in order to provide insight on how healthcare systems and policy makers may navigate future data partnerships.

2.3 Methods

Respondent comfort with sharing health data with third-party commercial companies was captured using a 20-minute online survey of US adults. In the following section we explain how the concepts described above are operationalized in this survey, followed by an explanation of the statistical methods used to analyze this data.

2.3.1 Participants

Respondents were surveyed using the National Opinion Research Center's (NORC) probability-based, nationally representative sample of US adults, based on 2010 Census Information. NORC's national sample frame employs a two-stage probability sample design to select a representative sample of households in the United States, oversampling African American, Hispanic populations, as well as households 200% below the federal poverty level. Survey recruitment and deployment was done in May 2019. Data collection was completed by June 2019. Eligible participants (at least 21 years old and able to read and write in English) were contacted via email to participate in the online survey, resulting in a total of 2,157 participants (66% response rate). The first component of the survey was a short (90 seconds) animated video describing how health data of a fictional patient is shared through the duration of care—to insurers, billers, and analysts learning from the outcomes of treatment. Definitions of important terms such as "healthcare system", "healthcare providers", "electronic health record", "de-identified health information [or biospecimens]", and "commercial companies" were provided to survey participants wherever those terms appeared. "Commercial companies" was defined for respondents to this survey as "third-party companies that are not part of a hospital. For example, a third-party commercial company may conduct genetic tests and analyze information for a hospital or healthcare provider for a fee when a hospital is not able to conduct the test on their own." "De-identified [health information or biospecimens]" was defined for respondents in the following manner: "de-identified means that "identifying information" about you is removed from your health information. Identifying information includes things like your name, address, date of birth, etc."

NORC calculated post-stratification weights according to US Census demographic benchmarks for age, sex, household income, education, as well as race and ethnicity to reduce sampling bias. For the purposes of this paper, records with missing responses to one or more of the questions used in this analysis were not included, resulting in a final analyzed sample of 1,841 responses. This study protocol was approved by the University of Michigan Health Sciences Institutional Review Board.

2.3.2 Survey Design

Variables used in this study were derived from a 20-minute, 164-item survey created to examine knowledge, attitudes, and beliefs about data sharing. Privacy measures were adapted from Anderson's work on consumer willingness to disclose personal health data and the California Health Foundation's 2005 National Consumer Health Privacy survey (Anderson & Agarwal, 2011; Bishop et al., 2005). Privacy measures also include questions about deception and medical mistrust (Boulware et al., 2003; LaVeist et al., 2009) and have been used in previous studies (Platt et al., 2018).

2.3.3 Measurements used in this study

Public Comfort with Sharing Health Data with Third-Party Commercial Companies for Patient and Business Purposes

To explore public comfort with sharing health data with third-party commercial companies for patient purposes, respondents answered questions about "how comfortable" they were with three statements regarding data sharing with third-party commercial companies, each along a 4-point Likert scale. Participants were asked "How comfortable are you with a third-party commercial company using your DNA and health information to improve the diagnosis and treatment of cancer in other patients?" and "How comfortable are you with a third-party commercial company developing predictions about how you will respond to a particular cancer treatment?: "not at all comfortable" (1), "somewhat comfortable" (2), "fairly comfortable" (3), and "very comfortable" (4). Participants were also asked "how true" it was that "The organizations that have my health information and share it can use large amounts of data to improve patient care": "not true" (1), "somewhat true" (2), "fairly true" (3), and "very true" (4).

To examine participant comfort with sharing health data with third-party commercial companies for business purposes, participants were asked "How comfortable are you with a third-party commercial company storing your DNA and health information?"; "How comfortable are you with a third-party commercial company sharing predictions about how you will respond to cancer treatment with insurance companies?"; and "How comfortable are you with a third-party commercial company selling de-identified health information to a pharmaceutical company?". "Business purpose" in this research is understood as storage of health data beyond

the purposes of clinical care and sharing information with third-party commercial companies to improve their own business processes without explicitly stated direct benefit to patients. Respondents were provided with the options "not at all comfortable" (1), "somewhat comfortable" (2), "fairly comfortable" (3), and "very comfortable" (4). Indices for data use for patient purposes and business purposes were then calculated as the sum of participant responses to the three questions in each index divided by the number of questions.

The Cronbach's alpha for these questions was 0.766 for comfort with sharing health data with third-party commercial companies for patient purposes and 0.786 for comfort with sharing health data for business purposes.

Demographics

Demographic factors reported in this study include sex, age, race and ethnicity, education, income, and employment. The survey fielded by NORC provided with only two options for sex, male and female. Age was divided into four groups: 18-29, 30-44, 45-59, and 60+. Categories for race and ethnicity include "white, non-Hispanic", "black, non-Hispanic", "other, non-Hispanic", "Hispanic", "multiracial, non-Hispanic", and "Asian, non-Hispanic". Education was divided into four groups: less than high school, high school graduate, some college, or bachelor's degree or above. Employment was grouped into four categories: employed, not-employed, retired, or not working due to disability or other reasons.

Health variables

In addition to demographic information, we examine the effect of respondent's insurance status ("Are you now covered by any form of health insurance or health plan? Yes (1) or No (2)"), self-reported health status ("Would you say that in general your health is… "poor" (1), "fair" (2), "good" (3), "very good" (4), "excellent" (5)), and examine participant's perceived ability to access healthcare services at a satisfactory level via the perceived healthcare access index. The index is based on various aspects of the healthcare experience and is evaluated here using a five-item index, asking "how true" ("not true", "somewhat true", "fairly true", or "very true") the following statements were for participants: 1) "The healthcare system in this country is easy to use"; 2) "I can get the healthcare I need when I need it"; 3) "I get all the information I

need about my health from my healthcare provider"; 4) "I could access my electronic health record if I wanted to"; 5) "In general, I am satisfied with the treatment I receive from my healthcare provider". The perceived healthcare access index was then calculated as the sum of participant responses to these five items and then divided by the number of questions. The Cronbach's alpha was 0.820 for the perceived healthcare access index.

Privacy Concerns

To measure individual privacy concerns, respondent privacy attitudes were evaluated using a 4-item index, assessing their belief in the privacy protections of their healthcare system and whether they have concern information about themselves is being misused or could be used in a way that is harmful to the respondent. The component questions for the privacy index are: "1) My healthcare system respects my privacy; 2) I worry that private information about my health could be used against me; 3) I worry my health information is available to people who have no business seeing it; 4) There are some things I would not tell my healthcare providers because I can't trust them with the information". Each item asks respondents to rate "how true" each was for themselves on a Likert-type scale ranging from 1 (not true) to 4 (very true). The final privacy index score reflects the average of each participant's response to these four questions. The first component question of the privacy index, "my healthcare system respects my privacy", has been reversed-scored for inclusion in this index. The Cronbach's alpha was 0.771 for the questions used in this index.

2.3.4 Data Analysis

Descriptive statistics were estimated on all variables and are used to describe the demographic characteristics, perceived healthcare access, and privacy concerns of participants. A paired t-test examining the difference between comfort with sharing health data with commercial companies for patient purposes and comfort with sharing health data with commercial companies for business purposes was conducted to determine whether the difference between the two means is statistically significant.

Weighted Ordinary Least Squares (OLS) Regression analysis was used to estimate the linear relationship between comfort with third-party commercial companies for patient and

business purposes and each demographic and health variable separately. We then estimated a multivariable model with all demographic and health variables and conducted a stepwise regression model to identify a parsimonious set of variables that explained the greatest amount of variability in the two outcomes – comfort with sharing data with commercial companies for business or patient purposes. For the stepwise regression model, we set statistical significance at α=0.05 ($p<0.002$) for inclusion and α=0.01 for exclusion, applying a Bonferroni correction to minimize Type I error. To enable comparison of effect sizes, regression coefficients were normalized (mean = 0, SD = 1).

2.4 Results

2.4.1 Sample Demographics

The resulting weighted sample of 1,841 participants shows a near even split between male and female participants (49% male). Approximately 12% of participants were under the age of 29, and 31% of participants were over the age of 60. Nearly 60% of participants identified as white non-Hispanic, 15% as black, non-Hispanic, 19% as Hispanic, 3% as Asian, non-Hispanic, 2% of participants identified race and ethnicity as "other", and 3% identified as multiethnic, consistent with 2016 data from the US Census Bureau (12% of the US population identifies as black or African-American, non-Hispanic). Nearly half of participants completed some college (46%), and 33% of participants have a bachelor's degree. While the proportion of participants with a bachelor's degree is consistent with national percentages (30%, 2016 census data), the proportion of participants with some college, no degree is much higher in this study than national percentages (21%, 2016 census data). Just over half of participants (59%) made an income less than $60,000, consistent with the median household income for 2018 (Guzman, 2019). Over half of participants (60%) had employment. Of the health questions included in this analysis, 89% of study participants reported having health insurance of some type, which is slightly lower than reported national percentages - 92% of the US population according to the 2018 US Census (US Census Bureau, 2019). The mean self-reported health score of participants was 3.08, suggesting that on average, the respondents were of "good" health.

Table 2-1 Demographic descriptive statistics

Table 2.1 Demographic descriptive statistics (N = 1841)			
		N	Frequency (weighted)
Sex			
	Male	903	49.05%
	Female	938	50.95%
Age			
	18-29	227	12.33%
	30-44	554	30.09%
	45-59	483	26.24%
	60+	577	31.34%
Race/Ethnicity			
	White	1086	58.99%
	Black, NH	273	14.83%
	Other, NH	30	1.63%
	Hispanic	358	19.45%
	Multiracial, NH	47	2.55%
	Asian, NH	47	2.55%
Education			
	Less than High School	73	3.97%
	High School	317	17.22%
	Some college	841	45.68%
	BA or above	610	33.13%
Income			
	Less than $60,000	1082	58.77%
	$60,000 or greater	759	41.23%
Employment			
	Employed	1112	60.40%
	Not employed	87	4.73%
	Retired	373	20.26%
	Disabled/Other	269	14.61
Insured			
	Is insured	1638	88.97%
	Is not insured	203	11.03%
Self-reported health			
	Range: 1 (Poor) to 5 (Excellent)		Mean: 3.08 (SD=0.92)

2.4.2 Public Comfort with Sharing Healthcare Data with Third-party Commercial Companies

Public comfort with sharing health data with third-party commercial companies was evaluated using two three-item indices: 1) comfort with sharing health data with third-party commercial companies for patient purposes (for themselves and for others), and 2) comfort with sharing health data with third-party commercial companies for business purposes (Table 1.2). The resulting mean of comfort with sharing data with third-party commercial companies for patient purposes was 2.54 (SD = 0.81) or between "somewhat comfortable" and "fairly comfortable". Roughly half of participants indicated that they were either fairly for very comfortable sharing data with third-party commercial companies for patient purposes (53.39% are comfortable with a third-party commercial company using their DNA and health information to improve the diagnosis and treatment of cancer in other patients, 49.16% are comfortable with third-party commercial companies developing predictions about how they will respond to a particular cancer treatment, and 47.80% believe that the organizations that have their health information and share it can use large amounts of data to improve patient care). Comfort with sharing health data with third-party commercial companies for business purposes had a resulting mean of 1.93 (SD = 0.85) or "somewhat comfortable". One quarter to one third of participants indicated they were either fairly or very comfortable with each of the component questions in comfort with sharing health data third-party commercial companies for business purposes (29.90% are comfortable with a third-party commercial company storing their DNA and health information, 31.02% are comfortable with a third-party commercial company sharing predictions about how they will respond to cancer treatment with insurance companies, and 24.39% are comfortable with a third-party commercial company selling de-identified health information to a pharmaceutical company). Figure 1 shows the distributions of the two indices.

A paired t-test was conducted on both comfort indices, the results of which show that there is a statistically significant difference between comfort with sharing health data with third-party commercial companies for patient purposes and comfort with sharing health data with third-party commercial companies for business purposes only, paired $t = 39.84, p < 0.001$.

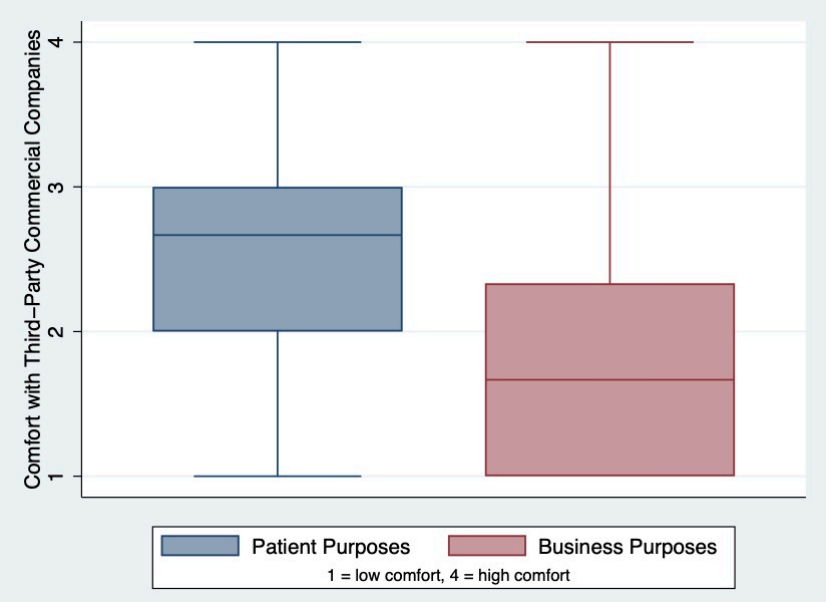

Figure 2-2 boxplot of public comfort sharing health data with third-party commercial companies for patient purposes (blue) and for business purposes (red)

Table 2-2 Descriptive statistics for survey questions used in indices measuring comfort with sharing health data with third-party commercial companies for patient purposes and business purposes

Table 2.2 Descriptive statistics for survey questions used in indices measuring comfort with sharing health data with third-party commercial companies for patient purposes and comfort with sharing health data with third-party commercial companies for business purposes. (N = 1841)		
	Frequency (% fairly or very comfortable/ fairly or very true)	Mean (SD)
Comfort with Sharing Health Data with Third-Party Commercial Companies for **Patient Purposes**		
How comfortable are you with a third-party commercial company using your DNA and health information to improve the diagnosis and treatment of cancer in other patients?	53.39%	2.58 (1.05)
How comfortable are you with a third-party commercial company developing predictions about how you will respond to a particular cancer treatment?	49.16%	2.48 (1.02)
The organizations that have my health information and share it can use large amounts of data to improve patient care	47.80%	2.56 (0.86)
Comfort with sharing health data with third-party commercial companies for patient purposes index (Cronbach's α=0.769)	Median: 2.67	2.54 (0.81)
Comfort with Sharing Health Data with Third-Party Commercial Companies for **Business Purposes**		
How comfortable are you with a third-party commercial company storing your DNA and health information?	28.90%	1.98 (1.01)
How comfortable are you with a third-party commercial company sharing predictions about how you will respond to cancer treatment with insurance companies?	31.02%	2.00 (1.04)

How comfortable are you with a third-party commercial company selling de-identified health information to a pharmaceutical company?	24.39%	1.81 (1.01)
Comfort with sharing health data with third-party commercial companies for business purposes index (Cronbach's α=0.786)	Median: 1.67	1.93 (0.85)

** Range of indices: 1 = not comfortable sharing health data with third-party commercial companies; 2 = somewhat comfortable sharing health data with third-party commercial companies; 3 = fairly comfortable sharing health data with third-party commercial companies; 4 = very comfortable sharing data with third-party commercial companies*

2.4.3 Perceived Healthcare Access

The resulting mean index score for the perceived healthcare access index was 2.82 (SD=0.75), which corresponds to "fairly true" for these questions, indicating fairly high confidence in participants' ability to access healthcare services at a satisfactory level. One-third (37.48%) of participants responded that it was fairly or very true that the healthcare system in the United States is easy to use, 70.23% responded that it was fairly or very true that they could get the healthcare they needed when they needed it, 62.42% responded that it was fairly or very true that they get all the information they needed about their health from their healthcare provider, 67.19% of participants responded that it was fairly or very true that they could access their electronic health record if they wanted to, and 73.01% of participants responded that it was either fairly or very true that they were satisfied with the treatment they received from their healthcare provider.

Table 2-3 Descriptive statistics for survey questions used in indices measuring Perceived Healthcare Access

Table 2.3 Descriptive statistics for survey questions used in indices measuring Perceived Healthcare Access (N=1841)		
	Frequency (% fairly or very true)	Mean (SD)
Perceived Healthcare Access Index		
The healthcare system in this country is easy to use	37.48%	2.22 (0.98)
I can get the healthcare I need when I need it	70.23%	3.02 (0.95)
I get all the information I need about my health from my healthcare provider	62.42%	2.82 (0.97)
I could access my electronic health record if I wanted to	67.19%	2.99 (1.06)
In general, I am satisfied with the treatment I receive from my healthcare provider	73.01%	3.04 (0.92)
Healthcare Access index (Cronbach's α=0.820)	Median: 2.8	2.82 (0.75)

* Range: 1 = "not true"; 2 = "somewhat true"; 3 = "fairly true"; 4 = "very true"

2.4.4 Privacy Concerns

Participant attitudes toward privacy were assessed using a four-item index (Table 1.3) examining various facets of privacy in healthcare. Just over half of participants (52.69%) responded that it was fairly or very true that their healthcare system respected their privacy, 35.58% responded that it was fairly or very true that they were worried health information could be used against them, 40.96% of participants indicated that it was fairly or very true that they worried their health information is being inappropriately accessed, and 24.12% responded that it was fairly or very true that they would withhold certain types of information from their care providers because of a lack of trust. One item in the index, "my healthcare system respects my privacy" was reversed so that higher Privacy Index scores consistently indicated greater privacy concerns. The resulting mean privacy attitudes index score was 2.22 (SD=0.78), or a privacy confidence of "somewhat true".

Table 2-4 Descriptive statistics for survey questions measuring privacy concerns

Table 2.4 Descriptive statistics for survey questions measuring privacy concerns (N=1841)		
	Frequency (% fairly or very true)	Mean (SD)
Privacy Index*		
My healthcare system respects my privacy**	52.69%	2.63 (0.91)
I worry that private information about my health could be used against me	35.58%	2.22 (1.07)
I worry my health information is available to people who have no business seeing it	40.96%	2.38 (1.05)
There are some things I would not tell my healthcare providers because I can't trust them with the information	24.12%	1.89 (1.00)
Privacy index (Cronbach's α=0.771)	Median: 2.25	2.22 (0.78)

* Range: 1 = "not true"; 2 = "somewhat true"; 3 = "fairly true"; 4 = "very true"
** This question has been reversed-scored for inclusion in this index

2.4.5 Univariate Model

Examination of comfort with sharing health data with commercial companies for patient and business purposes by demographic variables and privacy attitudes display slight increases and statistically significant differences in comfort with sharing health data with commercial

companies according to age, with participants between the ages of 45-59 indicating decreased comfort with sharing health data with third-party commercial companies for patient purposes compared to other age groups ($b^* = -0.102, p = 0.032$). Education displayed a small trend, with comfort with sharing health data with third-party commercial companies for patient purposes increasing as education increased (possession of a bachelor's degree: $b^* = 0.197, p = 0.002$). Examination of privacy attitudes and comfort with sharing health data with third-party commercial companies reveals that as privacy concerns increase, comfort with sharing health data with third-party commercial companies for both patient ($b^* = -0.260, p = 1.9 * 10^{-14}$) and business purposes ($b^* = -0.264, p = 5.7 * 10^{-14}$) decreases.

Table 2-5 Univariate associations for demographic factors, perceived healthcare access, and privacy concerns with comfort with sharing health data with third-party commercial companies for patient and business purposes

Table 2.5 Univariate associations for demographic factors, perceived healthcare access, and privacy concerns with comfort with sharing health data with third-party commercial companies for patient purposes and business purposes (N=1841)

		Patient Purposes (univariate)			**Business Purposes (univariate)**		
		b*	p-value	R^2	b*	p-value	R^2
Demographics							
Sex							
	Male	ref			ref		
	Female	-0.037	0.25	0.001	-0.042	0.20	0.002
Age							
	18-29	ref			ref		
	30-44	-0.078	0.085	0.007	-0.035	0.49	0.005
	45-59	**-0.102**	**0.032**		-0.091	0.087	
	60+	-0.029	0.53		-0.027	0.61	
Race/Ethnicity							
	White	ref			ref		
	Black, NH	-0.028	0.37	0.005	0.034	0.31	0.005
	Other, NH	-0.029	0.38		-0.014	0.64	
	Hispanic	-0.062	0.067		0.021	0.55	
	Multiracial, NH	-0.031	0.29		-0.039	0.15	
	Asian, NH	0.004	0.90		0.039	0.28	
Education							
	Less than High School	ref			ref		
	High School	0.098	0.14	0.014	0.001	0.99	0.011
	Some college	**0.126**	**0.034**		-0.045	0.55	
	BA or above	**0.197**	**0.002**		-0.117	0.14	
Income							
	Less than $60,000	ref			ref		
	$60,000 or greater	0.059	0.069	0.003	-0.028	0.41	0.001
Employment							

	Employed	ref			ref		
	Not employed	0.032	0.32	0.003	0.087	0.13	0.009
	Retired	0.022	0.44		0.014	0.63	
	Disabled/Other	-0.037	0.29		-0.034	0.28	
Insured							
	Has insurance	ref			ref		
	Does not have insurance	-0.060	0.057	0.004	0.021	0.46	0.001
Self-reported health							
	Poor	ref			ref		
	Fair	0.013	0.84	0.012	0.010	0.88	0.003
	Good	0.043	0.57		0.062	0.38	
	Very Good	0.073	0.29		0.010	0.88	
	Excellent	**0.119**	**0.021**		0.022	0.67	
Perceived Healthcare Access Index							
	Perceived Healthcare Access	**0.204**	**$6.0*10^{-10}$**	**0.041**	**0.154**	**$3.8*10^{-06}$**	**0.024**
Privacy Concerns							
	Privacy Concerns	**-0.260**	**$1.9*10^{-14}$**	**0.068**	**-0.264**	**$5.7*10^{-14}$**	**0.070**

b* = standardized beta

2.4.6 Stepwise regression models

Demographic predictors of comfort with sharing health data with third-party commercial companies ***for patient purposes*** show that in the multivariable model, 11% of the variability can be explained by demographic differences, perceived healthcare access, and attitudes towards privacy in the Bonferroni-corrected stepwise regression model. Five variables remained in the final regression model for comfort with sharing health data with third-party commercial companies for patient purposes: sex, age, education, perceived healthcare access, and privacy concerns. Possession of a bachelor's degree was the strongest demographic predictor of increased comfort with sharing health data with third-party commercial companies for patient purposes (b*=0.205, p=0.0009). Examination of demographic predictors of public comfort with sharing health data with third-party commercial companies ***for business purposes*** show that in the multivariable model, 10% of variability can be explained by demographic differences, perceived healthcare access, and attitudes towards privacy. In the Bonferroni-corrected stepwise regression model, five variables remained in the final business purpose model: sex, education, employment, perceived healthcare access, and privacy concerns, which displayed the strongest association with comfort with sharing health data with third-party commercial companies for business purposes. As privacy concerns decreased, comfort with sharing health data with third-

party commercial companies increased for both patient purposes (b*=-0.223, p=$6.9 * 10^{-10}$) and business purposes (b*=-0.246, p=$4.5 * 10^{-12}$).

Table 2-6 Stepwise regression modeling of predictors of comfort with sharing health data with third-party commercial companies for patient and business purposes

Table 2.6 Stepwise regression modeling of predictors of comfort with sharing health data with third-party commercial companies for patient purposes and comfort with sharing health data with third-party commercial companies for business purposes (N=1841)					
	Patient Purposes Multivariable stepwise Bonferroni corrected ($\alpha = 0.002$)			**Business Purposes** Multivariable stepwise Bonferroni corrected ($\alpha = 0.002$)	
	Model R^2	0.1117		Model R^2	0.0978
	b*	p-value		b*	p-value
Sex					
Male	ref			ref	
Female	-0.056	0.062		-0.064	0.037
Age					
18-29	ref				
30-44	-0.104	0.02			
45-59	-0.154	0.0012			
60+	-0.117	0.012			
Education					
Less than High School	ref			ref	
High School	0.089	0.16		-0.040	0.62
Some college	0.133	0.021		-0.069	0.37
BA or above	0.205	$9.0*10^{-4}$		-0.145	0.079
Employment					
Employed				ref	
Not employed				0.071	0.034
Retired				-0.037	0.22
Disabled/Other				-0.060	0.053
Perceived Healthcare Access Index					
Perceived Healthcare Access	0.140	$5.4*10^{-5}$		0.070	0.051
Privacy Concerns Index					
Privacy Concerns	-0.223	$6.9*10^{-10}$		-0.246	$4.5*10^{-12}$

b* = standardized beta

2.5 Discussion

To better understand the public's comfort with sharing health data with third-party commercial companies, this study sought to examine differences in comfort with sharing health data for patient purposes and comfort with sharing health data for business purposes, and identify the demographic variables that contribute to increased or decreased public comfort with sharing health data with third-party commercial companies. We also examined the relationship between privacy attitudes and comfort with sharing health data with commercial companies. Survey results revealed significantly less comfort with sharing health data with commercial companies for business purposes than patient purposes.

Although demographic factors had a modest effect on the public's comfort with sharing health data with third-party commercial companies; sex, age, education, and employment emerged as significant demographic variables when modeled using stepwise regression. Notably, self-reported health status did not persist in the final model presented here, despite the significance of healthcare status in other studies (K. K. Kim et al., 2017; Tikoo, 2014; Weitzman et al., 2010). One of the most significant findings of this study is that comfort with sharing health data with commercial companies for patient purposes ***increased*** with educational attainment, and that comfort with sharing health data with commercial companies for business purposes ***decreased*** with educational attainment. Although previous studies have identified an inverse relationship between willingness to share information and education, this study reveals that communicating the patient-centered motives for sharing health information with third-party commercial companies may reverse that trend. Perceived healthcare access was strongly associated with comfort with sharing health data with third-party commercial companies for patient purposes, likely indicating that ease of access, and the overall sense that one's own needs are being adequately met, increases personal motivation to extend that care to others as well. It is also likely that individuals with high healthcare satisfaction are more empowered consumers of healthcare resources and feel a greater sense of agency and control over their health data and healthcare experience.

Decreased privacy concerns or decreased worry about how healthcare data is used was associated with increased comfort with sharing health data with third-party commercial

companies and remained in the final stepwise regression model. In the privacy concerns index, we asked patients whether it was very or not true that the healthcare system respected [their] privacy. Interestingly, only half of participants indicated that it was either fairly or very true that the healthcare system respected [their] privacy. Patient privacy is of paramount importance in healthcare, with HIPAA representing one of the few examples of comprehensive privacy law worldwide. That only 52.69% participants feel respected in this way indicates the gap between what we have been able to provide patients with regard to their privacy and what they require now amidst this rapidly developing computing and big data environment. One third of participants (35.58%) indicated their concern that private information about their health could be used against them, and 40.96% of participants indicated concern that their health information was available to people who had no business seeing it. Alleviation of these fears of abuse can decrease privacy concerns and increase comfort with sharing health data with third-party commercial companies.

2.5.1 Implications for research

Healthcare research is rapidly becoming dependent on the large data sets provided by electronic personal health information (ePHI). Data partnerships with companies like Google are increasingly being sought in order to expand the data processing and research capabilities of healthcare systems. At a minimum, this research indicates the importance of promoting the patient-centered benefits of these partnerships at not only their announcement, but at their inception. Although healthcare systems anticipate tremendous benefits to their patients in the creation of these partnerships, it should not be assumed that the public automatically perceives these partnerships to be beneficial. While improvements to business efficiency and processes will in turn likely benefit patient care, this research indicates that these connections are too abstract for patients and should be made more explicit. The large difference between comfort with sharing data with commercial companies for business versus patient purposes suggests a need for further interrogation of the different predictors of these variables and deeper examination of the meaning the public has ascribed to commercial companies.

2.5.2 Implications for policy and practice

In this research we found that a little over half of respondents (52.69%) felt that "[their] healthcare system respects [their] privacy", and that the total privacy concerns index accounted for just over 20% of public comfort with sharing health data with third-party commercial companies for both patient and business purposes in the stepwise regression model presented in this analysis. That the public feels so inadequately protected signals an urgent need to reassess the privacy laws and regulations of healthcare, and to take quick steps to differentiate the manner in which healthcare systems use personal health information from the manner in which personal data is used in other industries. One possibility is to consider nationwide adoption of the Texas Medical Privacy Act, which is one of the broadest and most strict medical privacy laws in the United States. Under the Texas Medical Privacy Act, 1) any organization that assembles, collects, stores, or transmits PHI, or 2) comes into possession of PHI, is subject to HIPAA (Solove & Schwartz, 2019). Texas adds the additional prohibition of re-identifying de-identified data under any circumstance (Luna, 2011). As third-party partnerships proliferate, application of HIPAA to only covered entities requires reexamination.

2.5.3 Limitations

As with any survey, this study is merely a snapshot of patient beliefs and preferences, limited due to the nature of survey questions – other aspects of patient and public privacy concerns, perceived healthcare access, and health that may provide a more complete portrait of the public's comfort with sharing health data with third-party commercial companies may not be captured here. Additionally, a stepwise regression model is a conservative model that eliminates factors that might be important to understanding patient and public comfort with sharing health data with third-party commercial companies.

The circumstances of data sharing and the privacy context in which that sharing will occur will continue to evolve as laws, expectations, and experiences of healthcare data sharing change. Longitudinal studies that evaluate changes in comfort with sharing health data with third-party commercial companies would be superior, especially in light of changing media coverage of these partnerships. In subsequent research, we will examine a sampling of media

events and their potential effect on comfort with sharing health data with third-party commercial companies.

2.5.4 Conclusion

This study revealed that educational attainment is associated with increased comfort with sharing health data with third-party commercial companies for patient purposes and decreased comfort with sharing health data with third-party commercial companies for business purposes, and privacy concern is strongly associated with less comfort with sharing health data with third-party commercial companies for both patient and business purposes. This study also revealed differences in comfort with sharing health data with third-party commercial companies explicitly patient centered purposes versus business purposes with no explicitly stated patient benefit. Healthcare systems embarking on new third-party data partnerships to expand their ability to process and analyze health data can benefit from early identification and communication of the patient-centered benefits that will result from their third-party commercial partnerships. Healthcare systems can do more to provide reassurances that healthcare privacy will be protected, for example: communicating data protection efforts to the public at the time a new third-party partnership is announced, proactive acknowledgement of privacy concerns as privacy breaches unfold, and frequent communication of what healthcare systems are doing to mitigate privacy risks. More research is needed on attitudinal dimensions related to privacy (trust, comfort with researchers, quality analysts, and commercial companies, protections and notification of data access) to better understand comfort with sharing health data with third-party commercial companies for patient and business purposes.

2.6 Acknowledgements

Research reported in this manuscript was supported by the National Cancer Institute of the National Institutes of Health under award number 5 R01 CA214829-03.

2.7 References

Anderson, C., & Agarwal, R. (2011). The Digitization of Healthcare: Boundary Risks, Emotion, and Consumer Willingness to Disclose Personal Health Information. *Information Systems Research*, *22*(3), 469–490. https://doi.org/10.1287/isre.1100.0335

Bishop, L. "Sam," Holmes, B., & Kelley, C. (2005). National Consumer Health Privacy Survey 2005. *California Health Care Foundation*. https://www.chcf.org/publication/national-consumer-health-privacy-survey-2005/

Blank, G., Bolsover, G., & Dubois, E. (2014). A New Privacy Paradox: Young People and Privacy on Social Network Sites. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2479938

Boulware, L. E., Cooper, L. A., Ratner, L. E., LaVeist, T. A., & Powe, N. R. (2003). Race and Trust in the Health Care System. *Public Health Reports*, *118*, 8.

Brodkin, J. (2019, November 13). *Google: You can trust us with the medical data you didn't know we already had [Updated]*. Ars Technica. https://arstechnica.com/tech-policy/2019/11/google-you-can-trust-us-with-the-medical-data-you-didnt-know-we-already-had/

Castell, S., & Evans, H. (2016). The One-Way Mirror: Public attitudes to commercial access to health data. *Ipsos MORI Social Research Institute*, 161.

CooperKatz. (2018, February 5). *Paige.AI Created to Transform Cancer Diagnosis and Treatment by Applying Artificial Intelligence to Pathology*. Business Wire. https://www.businesswire.com/news/home/20180205005557/en/Paige.AI-Created-Transform-Cancer-Diagnosis-Treatment-Applying

Copeland, R., & Needleman, S. (2019, November 13). WSJ News Exclusive | Google's 'Project Nightingale' Triggers Federal Inquiry. *Wall Street Journal*.

https://www.wsj.com/articles/behind-googles-project-nightingale-a-health-data-gold-mine-of-50-million-patients-11573571867

Damschroder, L. J., Pritts, J. L., Neblo, M. A., Kalarickal, R. J., Creswell, J. W., & Hayward, R. A. (2007). Patients, privacy and trust: Patients' willingness to allow researchers to access their medical records. *Social Science & Medicine*, *64*(1), 223–235. https://doi.org/10.1016/j.socscimed.2006.08.045

Doukas, D. J., & Hardwig, J. (2014). Patient Informed Choice for Altruism. *Cambridge Quarterly of Healthcare Ethics*, *23*(4), 397–402. https://doi.org/10.1017/S0963180114000073

Griggs, M. B. (2019, November 11). *Google reveals 'Project Nightingale' after being accused of secretly gathering personal health records*. The Verge. https://www.theverge.com/2019/11/11/20959771/google-health-records-project-nightingale-privacy-ascension

Guzman, G. G. (2019). American Community Survey Briefs—Household Income: 2018. *United States Census Bureau*, 13.

Kim, J., Kim, H., Bell, E., Bath, T., Paul, P., Pham, A., Jiang, X., Zheng, K., & Ohno-Machado, L. (2019). Patient Perspectives About Decisions to Share Medical Data and Biospecimens for Research. *JAMA Network Open*, *2*(8), e199550–e199550. https://doi.org/10.1001/jamanetworkopen.2019.9550

Kim, K. K., Sankar, P., Wilson, M. D., & Haynes, S. C. (2017). Factors affecting willingness to share electronic health data among California consumers. *BMC Medical Ethics*, *18*(1), 25. https://doi.org/10.1186/s12910-017-0185-x

Landi, H. (2020, March 4). *Google defends use of patient data on Capitol Hill among scrutiny of Ascension deal*. FierceHealthcare. https://www.fiercehealthcare.com/tech/senators-pressing-ascension-google-data-deal-as-tech-giant-defends-its-use-patient-records

LaVeist, T. A., Isaac, L. A., & Williams, K. P. (2009). Mistrust of Health Care Organizations Is Associated with Underutilization of Health Services. *Health Services Research*, *44*(6), 2093–2105. https://doi.org/10.1111/j.1475-6773.2009.01017.x

Lee, H., Wong, S. F., & Chang, Y. (2016). Confirming the Effect of Demographic Characteristics on Information Privacy Concerns. *PACIS 2016 Proceedings*, 8.

Luna, J. (2011). *Texas Medical Privacy Act, Health Law & Policy Institute*. University of Houston Law Center. https://www.law.uh.edu/healthlaw/perspectives/Privacy/010830Texas.html

O'Neil, D. (2001). Analysis of Internet Users' Level of Online Privacy Concerns. *Social Science Computer Review*, *19*(1), 17–31. https://doi.org/10.1177/089443930101900103

Ornstein, C., & Thomas, K. (2018, September 20). Sloan Kettering's Cozy Deal With Start-Up Ignites a New Uproar. *The New York Times*. https://www.nytimes.com/2018/09/20/health/memorial-sloan-kettering-cancer-paige-ai.html

Platt, J. E., Jacobson, P. D., & Kardia, S. L. R. (2018). Public Trust in Health Information Sharing: A Measure of System Trust. *Health Services Research*, *53*(2), 824–845. https://doi.org/10.1111/1475-6773.12654

Reynolds, W. W., & Nelson, R. M. (2007). Risk perception and decision processes underlying informed consent to research participation. *Social Science & Medicine*, *65*(10), 2105–2115. https://doi.org/10.1016/j.socscimed.2007.06.021

Shavers, V. L., Lynch, C. F., & Burmeister, L. F. (2001). Factors that influence African-Americans' willingness to participate in medical research studies. *Cancer*, *91*(S1), 233–236. https://doi.org/10.1002/1097-0142(20010101)91:1+<233::AID-CNCR10>3.0.CO;2-8

Sheehan, K. B. (1999). An investigation of gender differences in on-line privacy concerns and resultant behaviors. *Journal of Interactive Marketing*, *13*(4), 15.

Singer, N., & Wakabayashi, D. (2019, November 11). Google to Store and Analyze Millions of Health Records. *The New York Times*. https://www.nytimes.com/2019/11/11/business/google-ascension-health-data.html

Solove, D. J., & Schwartz, P. M. (2019). *Privacy Law Fundamentals* (SSRN Scholarly Paper ID 1790262). Social Science Research Network. https://papers.ssrn.com/abstract=1790262

Stockdale, J., Cassell, J., & Ford, E. (2019). "Giving something back": A systematic review and ethical enquiry into public views on the use of patient data for research in the United Kingdom and the Republic of Ireland. *Wellcome Open Research*, *3*, 6. https://doi.org/10.12688/wellcomeopenres.13531.2

Teixeira, P. A., Gordon, P., Camhi, E., & Bakken, S. (2011). HIV patients' willingness to share personal health information electronically. *Patient Education and Counseling*, *84*(2), e9–e12. https://doi.org/10.1016/j.pec.2010.07.013

Tikoo, P. M. (2014). Evaluating Connecticut's Health Information Technology Exchange. *Connecticut Department of Public Health*, 58.

US Census Bureau. (2019). *Health Insurance Coverage in the United States: 2018*. The United States Census Bureau. https://www.census.gov/library/publications/2019/demo/p60-267.html

Vincent, J. (2019, June 27). *Google accused of inappropriate access to medical data in potential class-action lawsuit*. The Verge. https://www.theverge.com/2019/6/27/18760935/google-medical-data-lawsuit-university-of-chicago-2017-inappropriate-access

Wakabayashi, D. (2019, June 26). Google and the University of Chicago Are Sued Over Data Sharing. *The New York Times*. https://www.nytimes.com/2019/06/26/technology/google-university-chicago-data-sharing-lawsuit.html

Weitzman, E. R., Kaci, L., & Mandl, K. D. (2010). Sharing Medical Data for Health Research: The Early Personal Health Record Experience. *Journal of Medical Internet Research*, *12*(2). https://doi.org/10.2196/jmir.1356

Willison, D. J., Steeves, V., Charles, C., Schwartz, L., Ranford, J., Agarwal, G., Cheng, J., & Thabane, L. (2009). Consent for use of personal information for health research: Do people with potentially stigmatizing health conditions and the general public differ in their opinions? *BMC Medical Ethics*, *10*(1), 10. https://doi.org/10.1186/1472-6939-10-10

Chapter 3 Trust in the Health System, Privacy Concerns, and The Public's Comfort with Sharing Health Data with Third-Party Commercial Companies

3.1 Abstract

Background: Healthcare systems are partnering with third-party commercial companies to provide novel data processing and analysis capabilities that are critical to realizing the goals of learning health systems. These partnerships, however, provide personal health information to a for-profit industry that may store and sell data and the public may not be comfortable with such transactions.

Objective: To describe the public's comfort sharing health data with third-party commercial companies for patient purposes versus sharing of health data with third-party commercial companies for business purposes, and how comfort with both purposes is associated with attitudes towards **trust in the health system**, **trust in healthcare providers**, **altruism**, **privacy concerns**, **past experience with a data breach**, and **concern about recent events involving data breaches or third-party data partnerships in healthcare**.

Methods: We analyzed the results of a nationally representative sample of US residents responding to a survey (n = 1841) containing questions about their comfort sharing health data with third-party commercial companies for patient or business purposes. We assessed whether there was a difference between comfort with these two types of health data sharing and then used univariable and stepwise regression modeling to estimate the relationship between comfort with third-party commercial companies for patient and business purposes (outcomes) and trust in the health system, trust in providers, altruism, privacy, personal experience regarding a data breach, and concern about recent news events (independent variables). Also included are demographic factors, perceived healthcare access, and self-reported health status.

Results: Trust in the health system exhibited the strongest association with comfort sharing health data with third-party commercial companies for both patient purposes ($b^* = 0.367, p = 4.6 * 10^{-20}$) and business purposes ($b^* = 0.326, p = 4.7 * 10^{-17}$), followed by trust in providers (patient purposes: $b^* = 0.139, p = 1.6 * 10^{-4}$); business purposes: $b^* = 0.218, p = 6.4 * 10^{-9}$) and privacy concerns (patient purposes: $b^* = -0.110, p = 0.002$); business purposes: $b^* = -0.115, p = 0.001$). Education was strongly associated with comfort sharing health care data with third party commercial companies for patient purposes only (possession of college degree: $b^* = 0.298, p = 2.6 * 10^{-6}$). Among our experience variables and concerns about recent events, past experience with a data breach was not associated with either comfort sharing data with third-party commercial companies for patient or business purposes. Concern about the Memorial Sloan Kettering Cancer Center start up, Paige.AI, however, was negatively associated with comfort sharing health data with third-party commercial companies for business purposes ($b^* = -0.139, p = 5.4 * 10^{-6}$).

Conclusions: For both patient and business purposes, trust in the health system was the strongest predictor of comfort with sharing health data with third-party commercial companies. This finding suggests that regardless of the privacy controls and assurances put in place, if the health system at large is not trusted by the patient, increased comfort with and widespread support of third-party commercial company partnerships will not be forthcoming.

3.2 Introduction

3.2.1 Background

The healthcare transition to electronic health records (EHRs) and health information exchanges (HIE) has opened the possibility of big data insights into health outcomes and rare diseases that were not before conceivable. To accelerate the progress of data-driven discoveries, health systems are increasingly turning to third-party commercial companies to expand on limited in-house data analysis capabilities and take advantage of the expertise and processing power of large, established data companies like Amazon and Google. These partnerships, however, have been met by the general public with reluctance to participate (Huang, 2018) and concerns about health data privacy (Griggs, 2019).

In September 2018, Memorial Sloan Kettering Cancer Center encountered public and media opposition when their health data startup partnership, Paige.AI, was announced to the public (Ornstein & Thomas, 2018). A former Sloan Kettering patient wrote an open letter in response, expressing the "betrayal of trust" and unease that patient tissue was "being commercialized without our consent" (Petrow, 2018). One year later, in November 2019, Google announced "Project Nightingale"—a large scale data partnership with Ascension, one of the largest healthcare systems in the United States (Singer & Wakabayashi, 2019). As of this writing, Project Nightingale has been stymied by bi-partisan senatorial demands for the details of their data sharing partnership and an investigation by the Health and Human Services' (HHS) Office for Civil Rights (OCR) to examine the "mass collection of individuals' medical records with respect to the implications for patient privacy under HIPAA" (Cohen, 2019).

Both of these health data partnerships were done in compliance with existing HIPAA regulations – both healthcare systems submitted their projects to their respective Institutional Review Boards, and both were approved (Ornstein & Thomas, 2018; Shaukat, 2019). At the time the Ascension partnership was announced, Google was already in the midst of a lawsuit brought by a University of Chicago medical student for the use of patient data that was allegedly not de-identified (Moon, 2019) and was grappling with the growing perception that the company cared little about people's privacy (Porter, 2019; Turow, 2017). However, large healthcare data projects such as these are neither unusual nor novel. HIPAA has long permitted healthcare systems and EHR vendors to share and sell de-identified data to third-party commercial companies. Microsoft holds established data partnerships with healthcare systems and insurers including the University of Pittsburgh Medical Center and Premera Blue Cross intended to advance artificial intelligence (AI) healthcare initiatives for cancer care (P. Lee, 2017; Roach, 2019; Vincent, 2016). Since 2012, IBM has had access to 1.5 million Sloan Kettering patient records to explore better cancer treatments (Chen, 2018; Memorial Sloan Kettering Cancer Center, 2014; Vincent, 2016). These partnerships were not subjected to the same scrutinizing media response as was Project Nightingale or Paige.AI—news coverage and presentation of these projects expressed hope for progress and discovery as opposed to concerns about data security and violations of patient data use.

The backlash endured by Google, Sloan Kettering, and Ascension Healthcare follow on the heels of data breaches in other sectors – in the years preceding Google and Ascension's partnership announcement, the Equifax data breach in 2017 compromised personal and financial data of 143 million people (Fruhlinger, 2020); the Marriot data breach compromised the data of 500 million people (Gressin, 2018); and Facebook was in the midst of the Cambridge Analytica scandal in which the data of 87 million people was used for enhanced political profiling without their consent (Chang, 2018). These events may have fundamentally shifted the public's privacy concerns, in turn affecting patient attitudes toward data sharing with third-party commercial companies beyond the purposes of direct clinical care.

Third-party commercial companies are those companies that fall outside of usual "covered entities", covered entities being healthcare systems or providers who transmit any health information, health care plans, and health care clearinghouses (billing services, community health information systems, etc.) (Office for Civil Rights (OCR), 2013). Traditionally, "third-party" referred to administrators or payors of healthcare expenses—intermediaries that network with other providers and systems to fairly price medical billing claims or facilitate in some other capacity tasks critical to the administration of a health system. More recently, however, "third-party" has also come to refer to companies that purchase de-identified health data from providers, payors, and pharmacies in order to gain market strategy insights (Arndt, 2018). The business of third-party medical data trading has exploded, with some companies leading this field making over $2 billion in revenue in one year alone (Tanner, 2016). Once a data transaction is completed between a covered entity and a third-party commercial company, that company is then able to re-sell this de-identified information on the secondary data market. The chosen partners of these healthcare systems may simply the least trusted of all and will continue to endure criticism and greater concern. Media coverage of Amazon, Facebook, and Google and the general public's trust or mistrust of these companies have included variations on "tax-avoiding, soul-sucking machine[s]" (Galloway, 2018), evidencing growing fissures in the public's trust towards these companies and possibly the healthcare systems that partner with them.

3.2.2 Problem Statement

The consequences of violating or being perceived to violate patient privacy can include decreased organizational trust, decreased customer loyalty, and decreased patient candor with their provider (Agaku et al., 2014; Shen et al., 2019). Research in pursuit of privacy approaches that assuage or meet the concerns of patients has not resulted in conclusive recommendations about how to manage data privacy expectations in healthcare. Studies have found that patients and the public are generally of two minds: there is widespread willingness to share health data for the benefit of others, yet concerns persist about the security of health data that is used for research and further, that the data will be used for profit-driven motives instead of patient care (Stockdale et al., 2019). Deeper exploration of the public's comfort with sharing health data with third-party companies and the relationship of this comfort with the public attitudes - privacy, trust, altruism - and the impact on recent events (data breach, concern about recent events) is needed to better understand how healthcare systems can mitigate, manage, or prevent negative reactions towards existing or future data partnerships.

3.2.3 Conceptual Model

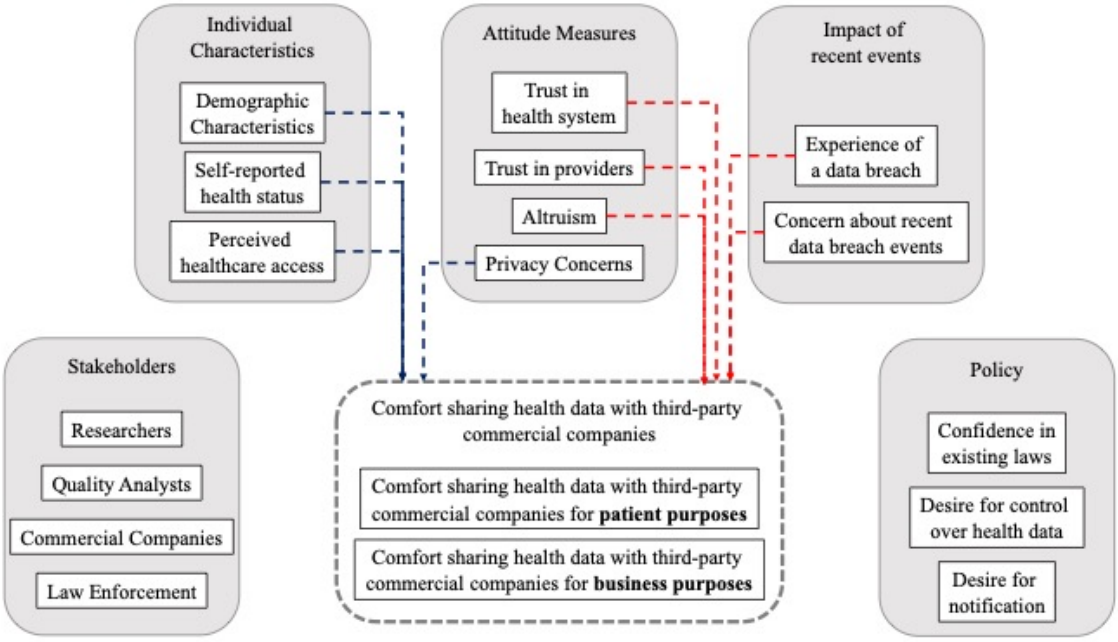

Figure 3-1 Conceptual Model of Dissertation Research – this analysis is focused on attitude measures and impact of recent events

Outcomes of interest: comfort with sharing data with commercial companies for business and patient purposes

In this analysis we examine the public's comfort with sharing health data with third-party commercial companies, investigating comfort when data is shared for patient purposes, i.e., to improve care, diagnosis, or treatment, versus comfort when data is shared for business purposes such as selling de-identified data. This division of patient-focused versus business-focused health data sharing allows for examination of the effect of stated purpose on the public's comfort with sharing health data with third-party commercial companies. Previous research indicates patients are more willing to provide access to their health information if the potential health benefits are clear (Anderson & Agarwal, 2011), and are less willing to provide access if their health data will be used for profit-generating research (Willison et al., 2009). In a study of US veterans, participants expressed to the study team that research studies must have "high value with an 'overall impact on society' and not be 'an academic exercise'" (Damschroder et al., 2007). We thus separate sharing of health data with third-party companies for patient-focused purposes from more general business purposes to examine and then compare participant's comfort sharing health data with third-party commercial companies.

In this paper, we consider the attitudes (trust, altruism, privacy concerns) and experiences (e.g., previous data breach) that may be associated with comfort sharing health data with third-party commercial companies.

Attitudes: Trust in the Health System and Trust in Providers

Trust, or the willingness to accept vulnerability to the actions of others (Richards & Hartzog, 2015), has been shown to increase the likelihood patients will participate in research (Kim et al., 2017) and is a strong predictor of attitudes and behaviors in other areas such as online shopping (Wang & Tseng, 2011), customer loyalty (Kantsperger & Kunz, 2010), and self-disclosure on social media (Fogel & Nehmad, 2009; Taddei & Contena, 2013). While there are a number of approaches to defining and examining trust, four high level categories have been most frequently used to capture the dimensions of trust: benevolence, integrity, competency, and predictability or fidelity (McKnight & Chervany, 2001). Benevolence means "caring and being motivated to act in one's interest rather than acting opportunistically"; integrity means "making

good faith agreements, telling the truth, and fulfilling promises"; competence means "having the ability or power to do for one what needs to be done"; and predictability means "trustee actions that are consistent enough to be forecasted" (McKnight & Chervany, 2001). Existing research on trust, privacy, and control of information have differed in their results on the mediating or moderating relationship between each. Some have found that trust is the commanding variable, reducing privacy concerns and increasing willingness to share personal information overall (Fogel & Nehmad, 2009), while others have found that privacy is actually the commanding variable, increasing trust and leading to greater disclosure of personal information (Krasnova et al., 2010).

Attitudes: Privacy Concerns

Surveys of the public's general attitude toward privacy have consistently found that people are concerned about their personal privacy. Many of these same studies have also found, however, that if the stated benefits of disclosure are attractive enough, individuals are willing to suspend privacy concerns for the benefits promised in exchange for disclosure. In the consumer sector, loyalty points, shopper discount cards, and customer accounts are reliant on this suspension of privacy concerns in exchange for monetary benefits or for convenience. This tradeoff calculus has also been shown in healthcare. While privacy is of paramount concern in healthcare, patients are more willing to share health information for research purposes if the benefits of participation for themselves or for the greater good are clear (Damschroder et al., 2007). This willingness to share health information is, however, still modified by general privacy concerns; even when the patient benefits are clear, individuals with greater privacy concerns express greater reluctance to share data than those with less privacy concerns (Anderson & Agarwal, 2011).

Attitudes: Altruism

Altruism is the prioritization of the needs of others even though no direct benefit may be conferred onto the individual. Individuals who consider themselves altruistic place high value on how their efforts may contribute to the well-being of others. As was stated earlier, studies on patient willingness to participate in healthcare research find that individuals are more willing to participate in research studies if their participation would help a friend or relative or has the

potential to benefit society (Doukas & Hardwig, 2014; Reynolds & Nelson, 2007; Shavers et al., 2001).

Impact of Data Breaches and Concern about Recent Events

In a study of healthcare data breaches, healthcare systems lost a reported average of $408 per record due to detection efforts, notification, legal expenditures and fines, and lost business (Ponemon Institute, 2018). This same study found that healthcare organizations worldwide lost customers as a result of data breaches. Studies of data breaches in other sectors have found similar effects. Research on online shopping behavior found that in the wake of a data breach, customers engage in protective behaviors that include avoiding the online store entirely or doubling-down on protective monitoring efforts to mitigate any issues that might arise due to compromised personal information (M. Lee & Lee, 2012). Similar patterns of decreased customer retention were also identified in a 2002 examination of hotel data breaches – a data breach event resulted in changes to the offline behavior of customers who reported being less willing to revisit and recommend that hotel (Belanger et al., 2002). Yet another study on the effect of data breaches found a negative and statistically significant impact on the market value of the breached companies (Acquisto et al., 2006). Research examining the effects of data breach events on comfort with sharing data with third-party commercial companies specifically has not yet been conducted to our knowledge.

3.2.4 Study Objective

The aims of this study are to examine the relationship between the public's comfort with sharing health data with third-party commercial companies and the public's attitudes regarding trust in the health system, trust in providers, privacy concerns, altruism, past experience with a data breach, as well as concern about recent data breaches or third-party data partnerships in healthcare.

3.3 Methods

Respondent comfort with sharing health data with third-party commercial companies was captured using a 20-minute online survey of US adults. In this section I explain how the concepts

described above are operationalized in this survey, followed by an explanation of the statistical methods used to analyze this data.

3.3.1 Participants

Respondents were surveyed using the National Opinion Research Center's (NORC) probability-based, nationally representative sample of US adults, based on 2010 Census Information. NORC's national sample frame employs a two-stage probability sample design to select a representative sample of households in the United States, oversampling African American, Hispanic populations, as well as households 200% below the federal poverty level. Survey recruitment and deployment was done in May 2019. Data collection was completed by June 2019. Eligible participants (at least 21 years old and able to read and write in English) were contacted via email to participate in the online survey, resulting in a total of 2,157 participants (66% response rate). The first component of the survey was a short (90 seconds) animated video describing how health data of a fictional patient is shared through the duration of care—to insurers, billers, and analysts learning from the outcomes of treatment. Definitions of important terms such as "healthcare system", "healthcare providers", "electronic health record", "de-identified health information [or biospecimens]", and "commercial companies" were provided to survey participants wherever those terms appeared. "Commercial companies" was defined for respondents to this survey as "third-party companies that are not part of a hospital. For example, a third-party commercial company may conduct genetic tests and analyze information for a hospital or healthcare provider for a fee when a hospital is not able to conduct the test on their own." "De-identified [health information or biospecimens]" was defined for respondents in the following manner: "de-identified means that "identifying information" about you is removed from your health information. Identifying information includes things like your name, address, date of birth, etc."

NORC calculated post-stratification weights according to US Census demographic benchmarks for age, sex, household income, education, as well as race and ethnicity to reduce sampling bias. For the purposes of this paper, records with missing responses to one or more of the questions used in this analysis were not included, resulting in a final analyzed sample of

1,841 responses. This study protocol was approved by the University of Michigan Health Sciences Institutional Review Board.

3.3.2 Survey Design

Variables used in this study were derived from a 20-minute, 164-item survey created to examine knowledge, attitudes, and beliefs about data sharing. Trust measures were adapted from the work of Mark Hall and colleagues (Hall, Camacho, et al., 2002; Hall et al., 2001; Hall, Zheng, et al., 2002). Altruism measures were adapted from the General Social Survey (Smith et al., 2019), the National Election Survey (Feldman & Steenbergen, 2001), and the General Self-Efficacy Scale (Schwarzer & Jerusalem, 1992). Privacy measures were adapted from Anderson's work on consumer willingness to disclose personal health information and the California Health Foundation's 2005 National Consumer Health Privacy survey (Anderson & Agarwal, 2011; Bishop et al., 2005). Privacy measures also include questions about deception and medical mistrust (Boulware et al., 2003; LaVeist et al., 2009) and have been used in previous studies (Platt et al., 2018).

3.3.3 Measures used in this study

The Public's Comfort with Sharing Health Data with Third-Party Commercial Companies for Patient and Business Purposes

To explore public comfort with sharing health data with third-party commercial companies for patient purposes, respondents answered questions about "how comfortable" they were with three statements regarding data sharing with third-party commercial companies, each along a 4-point Likert scale. Participants were asked "How comfortable are you with a third-party commercial company using your DNA and health information to improve the diagnosis and treatment of cancer in other patients?" and "How comfortable are you with a third-party commercial company developing predictions about how you will respond to a particular cancer treatment?: "not at all comfortable" (1), "somewhat comfortable" (2), "fairly comfortable" (3), and "very comfortable" (4). Participants were also asked "how true" it was that "The organizations that have my health information and share it can use large amounts of data to improve patient care": "not true" (1), "somewhat true" (2), "fairly true" (3), and "very true" (4).

To examine participant comfort with sharing health data with third-party commercial companies for business purposes, participants were asked "How comfortable are you with a third-party commercial company storing your DNA and health information?"; "How comfortable are you with a third-party commercial company sharing predictions about how you will respond to cancer treatment with insurance companies?"; and "How comfortable are you with a third-party commercial company selling de-identified health information to a pharmaceutical company?". "Business purpose" in this research is understood as storage of health data beyond the purposes of clinical care and sharing information with third-party commercial companies to improve their own business processes without explicitly stated direct benefit to patients. Respondents were provided with the options "not at all comfortable" (1), "somewhat comfortable" (2), "fairly comfortable" (3), and "very comfortable" (4). Indices for data use for patient purposes and business purposes were then calculated as the sum of participant responses to the three questions in each index divided by the number of questions.

The Cronbach's alpha for these questions was 0.766 for comfort with sharing health data with third-party commercial companies for patient purposes and 0.786 for comfort with sharing health data for business purposes.

Privacy Concerns

To measure individual privacy concerns, respondent privacy attitudes were evaluated using a 4-item index. These questions assessed respondent's beliefs about the privacy measures used by their healthcare system and whether they have concerns that personal health information about themselves is being misused or could be used in a way that is harmful. The component questions for the privacy index are: "a) My healthcare system respects my privacy; b) I worry that private information about my health could be used against me; c) I worry my health information is available to people who have no business seeing it; d) There are some things I would not tell my healthcare providers because I can't trust them with the information". Each item asked respondents to rate "how true" each was for themselves on a Likert-type scale ranging from 1 (not true) to 4 (very true). The final privacy index score reflects the average of

each participant's response to these four questions. The Cronbach's alpha was 0.771 for the questions used in the privacy concerns index.

Trust in the Healthcare System

Trust in the healthcare system, or system trust, is measured in this study using a modified version of the Platt et al. (2018) System Trust Index. The index was originally conceptualized as a composite index of three dimensions of system trust – competency, fidelity, and integrity (Platt et al., 2018). Component questions for the index asked participants to rate "how true" the following statements were "about the organizations that have your health information and share it": "a) …try hard to be fair in dealing with others; b) …would try to hide a serious mistake they made; c) …would never mislead me about how my health information is used; d) … have specialized capabilities that can promote innovation in health; e) … can be trusted to use my health information responsibly; f) … think about what is best for me; and g) …act in an ethical manner". Respondents rated "how true" each statement was for themselves on a Likert-scale ranging from 1 ("not true") to 4 ("very true"). The final system trust index score reflects the average of the participant's responses to each of these seven questions. The Cronbach's alpha was 0.845 for questions contributing to the system trust index.

Provider Trust

Provider trust is evaluated here using a 4-item index assessing respondent's trust in their healthcare providers. Component questions for this index are: a) "Health care providers care most about making money for themselves"; b) "Health care providers do not care about helping people like me"; c) "I trust health care providers to use my health information responsibly"; d) "All things considered, health care providers in this country can be trusted". Each item asked respondents to rate "how true" each question was for themselves on a Likert scale ranging from 1 ("not true) to 4 ("very true"). The final provider trust score is the total of each response divided by the number of questions for the index (4). The Cronbach's alpha for the questions contributing to the provider trust index is 0.697. Provider trust questions were used in a previous study by Platt et al. (2019).

Altruism

Respondent altruism is measured in this study using a 4-item index that asks "how true" the following questions are for the respondent on a Likert scale ranging from 1 ("not true") to 4 ("very true"): a) "I find ways to help others less fortunate than me"; b) "The dignity and well-being of all should be the most important concern in any society"; c) "One of the problems of today's society is that people are often not kind enough to others"; and, d) "All people who are unable to provide for their own needs should be helped by others". Respondents' altruism index scores are calculated as the average of their responses to these four questions. The Cronbach's alpha is 0.711 for questions contributing to the altruism index.

Experience of a past data breach

To find out whether respondents had their data compromised in the past, and explore the impact this experience might have on the respondent's comfort with sharing health data with third-party commercial companies, respondents were asked "Have you ever experienced problems with stolen or misused information?". Response options to this question were a) "Yes – I am currently experiencing problems"; b) "Yes – but all problems have been resolved"; and c) "No – I have not experienced any problems within the past five years". A second question was used to capture experience of a data breach: "I believe my financial information has been compromised as the result of a data breach or hacking". For this question respondents could answer "yes" or "no".

Concern about recent events and breaches

To evaluate the effect of recent data breaches or misuse of health information on comfort with sharing health data with third-party commercial companies, we selected a sample of recent events varying in industry and scale of impact and asked respondents "how concerned" they were about the following selection of events: a) "Concern about Facebook sharing information with Cambridge Analytica for political purposes"; b) "Concern about data breach of people's social security numbers and driver's license numbers at Equifax"; c) "Concern about Memorial Sloan Kettering hospital executives using hospital data for their own startup company"; d) "Concern about Marriot data breach of passport numbers and credit card numbers." Respondents answered on a Likert-scale from 1 ("not at all concerned") to 4 ("very concerned").

3.3.4 Data Analysis

For this analysis we first generated summary descriptive statistics on respondent characteristics (demographics), privacy concerns, system trust, provider trust, altruism, experience of a data breach and concern about recent events. A paired t-test examining the difference between comfort sharing health data with commercial companies for patient purposes and comfort sharing health data with commercial companies for business purposes was conducted to determine whether the difference between the two means is statistically significant.

Weighted Ordinary Least Squares (OLS) Regression analysis was used to estimate the linear relationship between comfort with third-party commercial companies for patient and business purposes and each demographic and health variable separately. We then estimated a multivariable model with all demographic and health variables and conducted a stepwise regression model to identify a parsimonious set of variables that explained the greatest amount of variability in the two outcomes – comfort with sharing data with commercial companies for business or patient purposes. For the stepwise regression model, we set statistical significance at $\alpha=0.05$ ($p<0.002$) for inclusion and $\alpha=0.01$ for exclusion, applying a Bonferroni correction to minimize Type I error. To enable comparison of effect sizes, regression coefficients were normalized (mean = 0, SD = 1).

3.4 Results

To examine predictors of comfort with third-party commercial companies using health information for patient and business purposes, we first examined the descriptive statistics for each of the independent variables and then conducted univariable and multivariable stepwise regression to identify predictors. Statistical significance was set at the 0.05 level.

3.4.1 Demographic descriptive statistics

The resulting weighted sample of 1,841 participants shows a near even split between male and female participants (49% male). Approximately 12% of participants were under the age of 29, and 31% of participants were over the age of 60. Nearly 60% of participants identified as white non-Hispanic, 15% as black, non-Hispanic, 19% as Hispanic, 3% as Asian, non-Hispanic,

2% of participants identified race and ethnicity as "other", and 3% identified as multiethnic, consistent with 2016 data from the US Census Bureau (12.3% of the US population identifies as black or African-American, non-Hispanic). Nearly half of participants completed some college (46%), and 33% of participants have a bachelor's degree. While the proportion of participants with a bachelor's degree is consistent with national percentages (30%, 2016 census data), the proportion of participants with some college, no degree is much higher in this study than national percentages (21%, 2016 census data). Just over half of participants (59%) made an income less than $60,000, consistent with the median household income for 2018 (Guzman, 2019). Over half of participants (60%) had employment. Of the health questions included in this analysis, 89% of study participants reported having health insurance of some type, which is slightly lower than reported national percentages - 92% of the US population according to the 2018 US Census (US Census Bureau, 2019). The mean self-reported health score of participants was 3.08, suggesting that on average, the respondents were of "good" health.

Table 3-1 Demographic descriptive statistics

Table 3.1 Demographic descriptive statistics (N = 1841)			
		N	Frequency
Sex			
	Male	903	49.05%
	Female	938	50.95%
Age			
	18-29	227	12.33%
	30-44	554	30.09%
	45-59	483	26.24%
	60+	577	31.34%
Race/Ethnicity			
	White	1086	58.99%
	Black, NH	273	14.83%
	Other, NH	30	1.63%
	Hispanic	358	19.45%
	Multiracial, NH	47	2.55%
	Asian, NH	47	2.55%
Education			
	Less than High School	73	3.97%
	High School	317	17.22%
	Some college	841	45.68%
	BA or above	610	33.13%
Income			
	Less than $60,000	1082	58.77%

	$60,000 or greater	759	41.23%
Employment			
	Employed	1112	60.40%
	Not employed	87	4.73%
	Retired	373	20.26%
	Disabled/Other	269	14.61
Insured			
	Is insured	1638	88.97%
	Is not insured	203	11.03%
Self-reported health			
	Range: 1 (Poor) to 5 (Excellent)		Mean: 3.08 (SD=0.92)

2.4.2 Public Comfort with Sharing Healthcare Data with Third-party Commercial Companies

Public comfort with sharing health data with third-party commercial companies was evaluated using two indices: 1) comfort with sharing health data with third-party commercial companies for patient purposes (for themselves and for others), and 2) comfort with sharing health data with third-party commercial companies for business purposes (Table 2.2). The resulting mean of comfort with sharing data with third-party commercial companies for patient purposes was 2.54 (SD = 0.81) or between "somewhat comfortable" and "fairly comfortable". Roughly half of participants indicated that they were either fairly for very comfortable sharing data with third-party commercial companies for patient purposes as expressed by the component questions of the index: (53.39% are comfortable with a third-party commercial company using their DNA and health information to improve the diagnosis and treatment of cancer in other patients , 49.16% are comfortable with third-party commercial companies developing predictions about how they will respond to a particular cancer treatment, and 47.80% believe that the organizations that have their health information and share it can use large amounts of data to improve patient care). Comfort with sharing health data with third-party commercial companies for business purposes had a resulting mean of 1.93 (SD = 0.85) or "somewhat comfortable". One quarter to one third of participants indicated they were either fairly or very comfortable with each of the component questions in comfort sharing health data third-party commercial companies for business purposes (28.90% are comfortable with a third-party commercial company storing their DNA and health information, 31.02% are comfortable with a third-party

company sharing predictions about how they will respond to cancer treatment with insurance companies, and 24.39% are comfortable with a third-party commercial company selling de-identified health information to a pharmaceutical company). Figure 3.2 shows the distributions of the two indices.

A paired t-test was conducted on both comfort indices, the results of which show that there is a statistically significant difference between comfort with sharing health data with third-party commercial companies for patient purposes and comfort with sharing health data with third-party commercial companies for business purposes only, paired $t = 39.84, p < 0.001$.

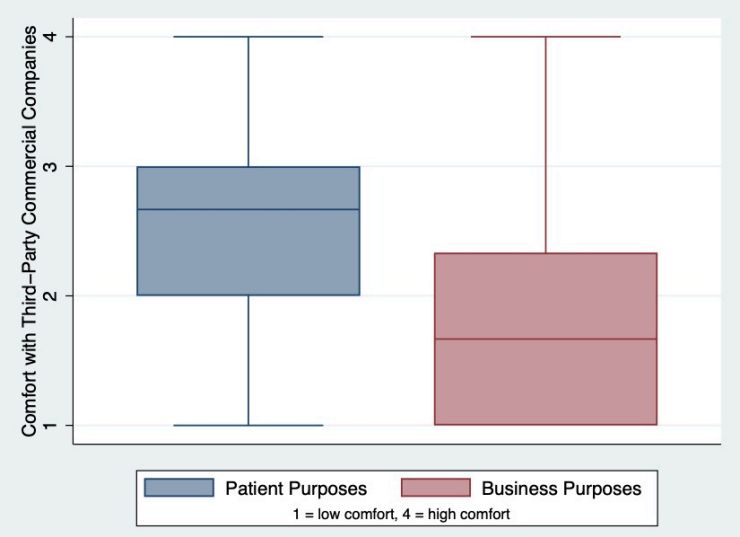

Figure 3-2 Box plot distributions of indices measuring Comfort Sharing Health Data with Third-Party Commercial Companies for Patient Purposes and Business Purposes

Table 3-2 Descriptive statistics for survey questions used in indices measuring comfort sharing health data with third-party commercial companies for patient and business purposes

Table 3.2 Descriptive statistics for survey questions used in indices measuring comfort sharing health data with third-party commercial companies for patient purposes and comfort sharing health data with third-party commercial companies for business purposes. (N = 1841)		
	Frequency (% fairly or very comfortable/ fairly or very true)	Mean (SD)
Comfort Sharing Health Data with Third-Party Commercial Companies for **Patient Purposes**		

How comfortable are you with a third-party commercial company using your DNA and health information to improve the diagnosis and treatment of cancer in other patients? (q27a)	53.39%	2.58 (1.05)
How comfortable are you with a third-party commercial company developing predictions about how you will respond to a particular cancer treatment? (q27b)	49.16%	2.48 (1.02)
The organizations that have my health information and share it can use large amounts of data to improve patient care (q35c)	47.80%	2.56 (0.86)
Comfort sharing health data with third-party commercial companies for patient purposes index (Cronbach's α=0.769)	Median: 2.67	2.54 (0.81)
Comfort Sharing Health Data with Third-Party Commercial Companies for **Business Purposes**		
How comfortable are you with a third-party commercial company storing your DNA and health information? (q27c)	28.90%	1.98 (1.01)
How comfortable are you with a third-party commercial company sharing predictions about how you will respond to cancer treatment with insurance companies? (q27d)	31.02%	2.00 (1.04)
How comfortable are you with a third-party commercial company selling de-identified health information to a pharmaceutical company? (q27e)	24.39%	1.81 (1.01)
Comfort sharing health data with third-party commercial companies for business purposes index (Cronbach's α=0.786)	Median: 1.67	1.93 (0.85)

** Range of indices: 1 = not comfortable sharing health data with third-party commercial companies; 2 = somewhat comfortable sharing health data with third-party commercial companies; 3 = fairly comfortable sharing health data with third-party commercial companies; 4 = very comfortable sharing data with third-party commercial companies*

3.4.3 Privacy Concerns

Participant attitudes toward privacy were assessed using a four-item index (Table 1.3) examining various facets of privacy in healthcare. Just over half of participants (52.69%) responded that it was fairly or very true that their healthcare system respected their privacy, 35.58% responded that it was fairly or very true that they were worried health information could be used against them, 40.96% of participants indicated that it was fairly or very true that they worried their health information is being inappropriately accessed, and 24.12% responded that it was fairly or very true that they would withhold certain types of information from their care providers because of a lack of trust. One item in the index, "my healthcare system respects my privacy" was reversed for inclusion in the index so that higher Privacy Index scores consistently indicated greater privacy concerns. The resulting privacy attitudes index mean was 2.22 (SD=0.78), indicating overall privacy concerns of this study sample are "somewhat true" based on these questions.

Table 3-3 Descriptive statistics for survey questions measuring privacy concerns

Table 3.3 Descriptive statistics for survey questions measuring privacy concerns (N=1841)		
	Frequency (% fairly or very true)	Mean (SD)
Privacy Index*		
My healthcare system respects my privacy (q39d)	52.69%	2.63 (0.91)
I worry that private information about my health could be used against me (q41a)	35.58%	2.22 (1.07)
I worry my health information is available to people who have no business seeing it (q41b)	40.96%	2.38 (1.05)
There are some things I would not tell my healthcare providers because I can't trust them with the information (q41c)	24.12%	1.89 (1.00)
Privacy index (Cronbach's α=0.771)	Median: 2.25	2.22 (0.78)

* Range: 1 = "not true"; 2 = "somewhat true"; 3 = "fairly true"; 4 = "very true"

3.4.4 Trust in The Health System

Participant trust in their health system (system trust) was assessed using a seven-item index examining various aspects of the trustee/trustor relationship. Only 40.14% felt that it was fairly or very true that "the organizations that have my health information and share it try to be fair in dealing with others", 42.26% responded that it was fairly or very true that these organizations "would try to hide a serious mistake", and only 36.06% responded that it was fairly or very true that these organizations "tell me how my health information is used". Just over a quarter of participants (29.11%) felt that it was fairly or very true that the organizations that have their health information and share it "would never mislead me about how my health information is used". Over 40% of respondents (43.24%) thought it was fairly or very true that these same organizations "have specialized capabilities that can promote innovation in health". Only 36.12% responded that it was fairly or very true that these organizations "can be trusted to use my health information responsibly", 31.12% felt these organizations "think about what is best for me", and just over 40% (41.12%) responded that the organizations that have their health information and share it "act in an ethical manner". One item in the index, "The organizations that have your health information and share it … would try to hide a serious mistake they made" was reversed for inclusion in the index so that higher scores would indicate greater system trust. In this study, the average system trust index score was 2.30 (SD=0.64), which corresponds to an evaluation by

our participants that it is "somewhat true" that the organizations that have their health information and share it can be trusted.

Table 3-4 Descriptive statistics for survey questions used in indices measuring trust in the health system (System Trust)

Table 3.4 Descriptive statistics for survey questions used in indices measuring System Trust (N=1841)		
"For you, how true are the following statements about the organizations that have your health information and share it?"	Frequency (% fairly or very true)	Mean (SD)
System Trust Index		
…Try hard to be fair in dealing with others (q34a)	40.14%	2.37 (0.86)
…Would try to hide a serious mistake they made (q34b)	42.26%	2.43 (1.05)
… Tell me how my health information is used (q34c)	36.06%	2.24 (0.98)
… Would never mislead me about how my health information is used (q34d)	29.11%	2.06 (0.92)
… Have specialized capabilities that can promote innovation in health (q35b)	43.24%	2.44 (0.85)
… Can be trusted to use my health information responsibly (q36a)	36.12%	2.25 (0.88)
… Think about what is best for me (q36b)	31.12%	2.11 (0.93)
… Act in an ethical manner (q36c)	41.12%	2.37 (0.87)
System Trust index (Cronbach's α=0.845) *q34b has been reverse coded for this index*	Median: 2.25	2.30 (0.64)

* Range: 1 = "not true"; 2 = "somewhat true"; 3 = "fairly true"; 4 = "very true"

3.4.5 Trust in Healthcare Providers

Participant's trust in their healthcare providers (provider trust) was examined using a four-item index that examined provider intention and care. Only 38.62% of respondents indicated that it was fairly or very true that "health care providers care most about making money for themselves", and 13.63% of respondents believe that it is fairly or very true that "health care providers do not care about helping people like me". These two questions were reverse coded for inclusion in the provider trust index so that higher score indicate greater trust in health care providers. Just over 40% of participants responded both that it was fairly or very true that they "trust health care providers to use my health information responsibly" (44.98%) and that it was fairly or very true that, "all things considered, health care providers in this country can be trusted" (41.34%). The provider trust index score for this population sample was 2.20 (SD=0.46), which corresponds to an evaluation that it is "somewhat true" that providers can be trusted.

Table 3-5 Descriptive statistics for survey questions used in indices measuring trust in healthcare providers (Provider Trust)

Table 3.5 Descriptive statistics for survey questions used in indices measuring trust in healthcare providers (Provider Trust) (N=1841)		
	Frequency (% fairly or very true)	Mean (SD)
Provider Trust Index		
Health care providers care most about making money for themselves (q38a)	38.62%	2.37 (1.00)
Health care providers do not care about helping people like me (q38b)	13.63%	1.62 (0.85)
I trust health care providers to use my health information responsibly (q38c)	44.98%	2.46 (0.92)
All things considered, health care providers in this country can be trusted (q38e)	41.34%	2.37 (0.81)
Provider Trust index (Cronbach's α=0.697)	Median: 2.25	2.20 (0.46)

* Range: 1 = "not true"; 2 = "somewhat true"; 3 = "fairly true"; 4 = "very true"

3.4.6 Altruism

Participant altruism was evaluated using four questions. Over half of participants (59.27%) responded that it is fairly or very true that "I find ways to help others less fortunate than me", and 76.04% responded that it is fairly or very true that "the dignity and well-being of all should be the most important concern in any society". Nearly three-quarters of participants (73.06%) responded that it was fairly or very true that "one of the problems of today's society is that people are often not kind enough to other" and over half (55.19%) responded that it is fairly or very true that "all people who are unable to provide for their own needs should be helped by others". The resulting altruism index score across participants is 2.96 (SD=0.66), which corresponds to an evaluation that it is "fairly true" that this study population is altruistic.

Table 3-6 Descriptive statistics for survey questions used in indices measuring altruism

Table 3.6 Descriptive statistics for survey questions used in indices measuring altruism (N=1841)		
	Frequency (% fairly or very true)	Mean (SD)
Altruism Index		
I find ways to help others less fortunate than me (q16a)	59.27%	2.78 (0.86)
The dignity and well-being of all should be the most important concern in any society (q16b)	76.04%	3.19 (0.90)

One of the problems of today's society is that people are often not kind enough to others (q16c)	73.06%	3.12 (0.91)
All people who are unable to provide for their own needs should be helped by others (q16d)	55.19%	2.75 (0.94)
Altruism index (Cronbach's α=0.711)	Median: 3.0	2.96 (0.66)

* Range: 1 = "not true"; 2 = "somewhat true"; 3 = "fairly true"; 4 = "very true"

3.4.7 Experience of a past data breach

The majority of participants (67.79%) responded "No-I have not experienced any problems [with stolen or misused information] within the past 5 years". Of the respondents that indicated they had experienced problems with stolen or misused information, only 5.11% were currently experiencing problems. The remaining 27.10% of participants that had experienced problems with stolen or misused information considered those problems resolved. A similar majority of participants (65.73%) did not believe their financial information had been compromised due to a data breach or hacking.

Table 3-7 Descriptive statistics for personal experience with a data breach

Table 3.7 Descriptive statistics for personal experience with a data breach (N=1841)			
		n	Frequency (%)
Have you experienced problems with stolen or misused information (q53)			
	Yes – I am currently experiencing problems	94	5.11%
	Yes – but all problems have been resolved	499	27.10%
	No – I have not experienced any problems within the past five years	1248	67.79%
I believe my financial information has been compromised as the result of a data breach or hacking (q54)			
	Yes	631	34.27%
	No	1210	65.73%

3.4.8 Concern about recent events and breaches

For each of the recent data breach or data misuse events selected for this study, at least 60% of participants indicated that they were either fairly or very concerned. The Marriot data breach was the least concerning for respondents, with 62.79% indicating that they were fairly or very concerned, and the Equifax data breach was the most concerning, with 76.48% of respondents indicating that they were either fairly or very concerned. Facebooks involvement with Cambridge Analytica was fairly or very concerning to 69.53% of respondents, and Sloan

Kettering's use of hospital data for a start-up company was fairly or very concerning for 65.07% of participants.

Table 3-8 Descriptive statistics for concern about recent events

Table 3.8 Descriptive statistics for concern about recent events (N=1841)		
	Frequency (% fairly or very concerned)	Mean (SD)
Concern about Facebook sharing information with Cambridge Analytica for political purposes (q55a)	69.53%	3.08 (1.06)
Concern about data breach of people's social security numbers and driver's license numbers at Equifax (q55b)	76.48%	3.27 (0.94)
Concern about Sloan Kettering hospital executives using hospital data for their own startup company (q55c)	65.07%	2.95 (1.10)
Concern about Marriot data breach of passport numbers and credit card numbers (q55d)	62.79%	2.90 (1.13)

* Range: 1 = "not at all concerned"; 2 = "somewhat concerned"; 3 = "fairly concerned"; 4 = "very concerned"

3.4.9 Univariate Linear Regression

Univariate examination of comfort sharing health data with third-party commercial companies for patient and business purposes show strong associations with system trust ($patient\ purposes\ b^* = 0.453, p = 5.5 * 10^{-42}, business\ purposes\ b^* = 0.452, p = 4.0 * 10^{-43}$) and provider trust ($patient\ purposes\ b^* = 0.234, p = 3.7 * 10^{-10}, business\ purposes\ b^* = 0.298, p = 1.7 * 10^{-13}$) as well as privacy concerns ($patient\ purposes\ b^* = -0.260, p = 1.9 * 10^{-14}, business\ purposes\ b^* = -0.264, p = 5.7 * 10^{-14}$) for both patient and business data sharing purposes.

Altruism was strongly associated with comfort with third-party commercial companies sharing data for patient purposes only ($b^* = 0.141, p = 5.5 * 10^{-05}$). Belief that financial information has been compromised as well as concern about recent events showed significant associations with comfort with sharing health data with third-party commercial companies for business purposes only: Facebook ($b^* = -0.118, p = 8.7 * 10^{-04}$), Equifax ($b^* = -0.131, p = 8.0 * 10^{-05}$), Sloan Kettering ($b^* = -0.185, p = 7.4 * 10^{-08}$), and Marriot ($b^* = 0.092, p = 0.0068$). Perceived healthcare access was significantly associated with comfort with sharing

health data wit third-party commercial companies for business purposes only ($b^* = 0.154, p = 3.8 * 10^{-06}$). Amongst the demographic factors examined here, being between the ages of 45 and 59 ($b^* = -0.102, p = 0.032$), possessing a college degree ($b^* = 0.197, p = 0.002$), and a self-reported health of "excellent" ($b^* = 0.119, p = 0.021$), showed strong associations with comfort with sharing health data with third-party commercial companies for patient purposes only .

Table 3-9 Univariate associations for attitudes (provider trust, system trust, privacy concerns, altruism), experience of a data breach, concern about recent events, and demographic data with comfort sharing health data with third-party commercial companies

Table 3.9 Univariate associations for attitudes (provider trust, system trust, privacy concerns, altruism), experience of a data breach, concern about recent events, and demographic data with comfort sharing health data with third-party commercial companies for patient purposes and business purposes (N=1841)

		Patient Purposes (univariate)			**Business Purposes (univariate)**		
		b*	p-value	R^2	b*	p-value	R^2
Attitudes							
Trust							
	System Trust	**0.453**	**$5.5*10^{-42}$**	**0.205**	**0.452**	**$4.0*10^{-43}$**	**0.204**
	Provider Trust	**0.234**	**$3.7*10^{-10}$**	**0.055**	**0.298**	**$1.7*10^{-13}$**	**0.089**
Altruism Index							
	Altruism Index	**0.141**	**$5.5*10^{-05}$**	**0.020**	0.054	0.12	0.003
Privacy Index							
	Privacy Concerns	**-0.260**	**$1.9*10^{-14}$**	**0.068**	**-0.264**	**$5.7*10^{-14}$**	**0.070**
Experience of Data Breach							
Have you experienced problems with stolen or misused information (q53)							
	Yes – I am currently experiencing problems	ref			ref		
	Yes – but all problems have been resolved	**0.131**	**0.026**	0.004	-0.015	0.78	0.006
	No – I have not experienced any problems within the past five years	**0.131**	**0.028**		0.062	0.23	
I believe my financial information has been compromised as the result of a data breach or hacking (q54)							
	Yes	ref			ref		
	No	0.044	0.16	0.002	**0.121**	**7.2e-05**	**0.015**
Concern about recent events							
	Concern about Facebook sharing information with Cambridge Analytica for political purposes (q55a)	-0.012	0.75	0.0001	**-0.118**	**8.7e-04**	**0.014**

	Concern about data breach of people's social security numbers and driver's license numbers at Equifax (q55b)	-0.018	0.60	0.0003	**-0.131**	**8.0e-05**	**0.017**
	Concern about Sloan Kettering hospital executives using hospital data for their own startup company (q55c)	-0.046	0.18	0.002	**-0.185**	**7.4e-08**	**0.034**
	Concern about Marriot data breach of passport numbers and credit card numbers (q55d)	-0.011	0.75	0.0001	**-0.092**	**0.0068**	**0.008**
Demographics							
Sex							
	Male	ref			ref		
	Female	-0.037	0.25	0.001	-0.042	0.20	0.002
Age							
	18-29	ref			ref		
	30-44	-0.078	0.085	0.007	-0.035	0.49	0.005
	45-59	**-0.102**	**0.032**		-0.091	0.087	
	60+	-0.029	0.53		-0.027	0.61	
Race/Ethnicity							
	White	ref			ref		
	Black, NH	-0.028	0.37	0.005	0.034	0.31	0.005
	Other, NH	-0.029	0.38		-0.014	0.64	
	Hispanic	-0.062	0.067		0.021	0.55	
	Multiracial, NH	-0.031	0.29		-0.039	0.15	
	Asian, NH	0.004	0.90		0.039	0.28	
Education							
	Less than High School	ref			ref		
	High School	0.098	0.14	0.014	0.001	0.99	0.011
	Some college	**0.126**	**0.034**		-0.045	0.55	
	BA or above	**0.197**	**0.002**		-0.117	0.14	
Income							
	Less than $60,000	ref			ref		
	$60,000 or greater	0.059	0.069	0.003	-0.028	0.41	0.001
Employment							
	Employed	ref			ref		
	Not employed	0.032	0.32	0.003	0.087	0.13	0.009
	Retired	0.022	0.44		0.014	0.63	
	Disabled/Other	-0.037	0.29		-0.034	0.28	
Insured							
	Has insurance	ref			ref		
	Does not have insurance	-0.060	0.057	0.004	0.021	0.46	0.001
Self-reported health							
	Poor	ref			ref		
	Fair	0.013	0.84	0.012	0.010	0.88	0.003
	Good	0.043	0.57		0.062	0.38	
	Very Good	0.073	0.29		0.010	0.88	
	Excellent	**0.119**	**0.021**		0.022	0.67	

Perceived Healthcare Access Index							
	Perceived Healthcare Access	0.204	6.0e-10	0.041	**0.154**	**3.8e-06**	**0.024**

b* = standardized beta

3.4.10 Stepwise regression modeling

In the Bonferroni-corrected stepwise regression model, 26% of variability of comfort with sharing health data with third party commercial companies for ***patient purposes*** can be explained by trust in the health system (system trust) and trust in healthcare providers (provider trust), privacy concerns, and education ($R^2 = 0.257$). For the ***business purposes*** model, 28% of variability in comfort with third-party commercial companies can be explained by system and provider trust, privacy concerns, and concern about the Sloan Kettering event only ($R^2 = 0.281$).. System trust is the most strongly associated variable with comfort sharing health data with third-party commercial companies for patient purposes ($b^* = 0.367, p = 4.6 * 10^{-20}$) and business purposes ($b^* = 0.326, p = 4.7 * 10^{-17}$). As system trust increased, comfort with sharing health data with third-party commercial companies also increased. Provider trust was also strongly associated with comfort sharing health data with third-party commercial companies for patient purposes ($b^* = 0.139, p = 1.6 * 10^{-4}$) and business purposes ($b^* = 0.218, p = 6.4 * 10^{-9}$). Privacy concerns showed less strong but still significant associations with comfort with third-party commercial companies for patient purposes ($b^* = -0.110, p = 0.002$) and business purposes ($b^* = -0.115, p = 0.001$). Having some college or a college degree was significantly associated with comfort sharing health data with third-party commercial companies for patient purposes only (college educated: $b^* = 0.298, p = 2.6 * 10^{-6}$). Past experience of a data breach did not remain in this final model, nor did concern about Facebook and Cambridge Analytica, Equifax, or Marriot. Concern about Sloan Kettering hospital executives using health data for their own start up, however, remained in the final model for business purposes only ($b^* = -0.139, p = 5.4 * 10^{-6}$). Greater concern about the Sloan Kettering event was associated with less comfort with sharing health data with third-party commercial companies for business purposes.

Table 3-10 Stepwise regression modeling of predictors of comfort sharing health data with third-party commercial companies for patient purposes and comfort sharing health data with third-party commercial companies for business purposes

Table 3.10 Stepwise regression modeling of predictors of comfort sharing health data with third-party commercial companies for patient purposes and comfort sharing health data with third-party commercial companies for business purposes (N=1841)					
	Patient Purposes Multivariable stepwise Bonferroni corrected ($\alpha = 0.002$)			**Business Purposes** Multivariable stepwise Bonferroni corrected ($\alpha = 0.002$)	
	Model R^2	0.257		Model R^2	0.281
	b*	p-value		b*	p-value
Attitudes					
System Trust	**0.367**	**4.6e-20**		**0.326**	**4.7e-17**
Provider Trust	**0.139**	**1.6e-04**		**0.218**	**6.4e-09**
Altruism	0.060	0.092			
Privacy Concerns	**-0.110**	**0.002**		**-0.115**	**0.001**
Recent events concern					
Sloan Kettering				**-0.139**	**5.4e-06**
Demographic Factors					
Education					
Less than High School	ref				
High School	**0.149**	**0.022**			
Some college	**0.201**	**6.7e-04**			
BA or above	**0.298**	**2.6e-06**			
Employment					
Employed				ref	
Not employed				**0.089**	**0.016**
Retired				0.005	0.83
Disabled/Other				-0.052	0.11
Self-Reported Health					
Poor				ref	
Fair				-0.059	0.31
Good				-0.001	0.99
Very Good				-0.048	0.44
Excellent				-0.073	0.11

b* = standardized beta

3.5 Discussion

Previous research on the effect of data breaches on consumers has shown that trust in the organization mitigates the negative effects of data breaches (Chakraborty et al., 2016; McKnight et al., 2002) and is of greater importance than privacy and security features (Belanger et al., 2002). These conclusions from other studies comport with the main finding of this research: trust in the health system has a far greater association with comfort with sharing health data with

third-party commercial companies than does concern about privacy or experience of a past data breach. It is thus possible that regardless of the privacy protections put in place, privacy controls alone will not be enough to assuage concerns about the use of health data if trust is not already secured.

Research on the neuroscience of trust shows that "reputational priors", or prior information about a given individual or company, reduced individual uncertainty about any subsequent decisions. Moreover, in this same study, participants continued to rely on these reputational priors even when the behavior of the individual or company that followed was not consistent with that given reputation (Fouragnan et al., 2013). In the wake of a breach event, trusted organizations with strong "reputational priors" will bounce back more quickly after the breach event, and the breach itself will take a smaller role in the mind of the public or patient. Fouragnan's study and the results of this study agree with Ari Waldman's assertion that "seeing trust as antecedent to privacy judgments is a step in the right direction". Waldman argues that users make privacy-related decisions based on their existing perception of the trustworthiness of the companies or individuals they transact with, and that "seeing trust as a byproduct of a functioning privacy regime misses the fact that sharers tend to expect privacy protection where trust exists already" (Waldman, 2016). The results of this study suggest that without trust in the health system already secured, privacy controls will be insufficient in assuaging concerns. The importance of trust in the decision to reveal or conceal information can explain why there is little clarity about the meaning or importance of privacy to the public, or guidance on what tactics can be employed to meet privacy needs.

This study also found significant associations between concern about recent data breach events and comfort with sharing health data with third-party commercial companies for business purposes in our univariate analysis, but these associations did not remain in the final multivariable model. These results suggest that although privacy violations are occurring in other sectors, their effect on health data privacy concerns is limited. This results also shows that while some may believe that "privacy is dead" or that it is impossible to maintain one's privacy in today's digitized experience, health data, and data use in health systems, remains separate and distinct.

Concern about the Memorial Sloan Kettering Paige.AI startup, however, did remain in the final multivariable model, suggesting that perceived data violation events in healthcare specifically may have deleterious effects for other health systems. It may be considered that individuals in this study who expressed more concerned about the use of health data for a start-up venture may have had pre-existing reservations about the use of health data by commercial companies, or became warier about the use of their personal health information in the wake of the event. More research is needed to examine whether other perceived data misuse events similar to Sloan Kettering's start-up do in fact have a deleterious after-effect on the public's subsequent comfort sharing health data with third-party commercial companies for health systems at large.

3.5.1 Implications for policy and practice

Healthcare research is rapidly becoming dependent on the large data sets provided by electronic personal health information (ePHI) and data partnerships with companies like Google and Amazon are increasingly being sought in order to expand the data processing and research capabilities of healthcare systems. The results of this research underscore the importance of maintaining patient and public trust in the organization and in their providers while conducting research on electronic health information. Trust has long been a prerequisite in healthcare – patients, who are often vulnerable when they seek medical attention and care, must trust in the competence and integrity of their physician (Goold, 2002). In health research, trust can be increased with the use of community-based participatory research (CBPR), which engenders trust by communicating research findings back to study participants and through engaging community members and patient representatives in every step of the research process (Israel et al., 2005; McDavitt, 2016).

In this research we found that trust in the health system is associated with comfort with sharing health data with third-party commercial companies, and that greater trust indicates greater comfort. However, 42.26% of participants believe that it is fairly or very true that these organizations "would try to hide a serious mistake they made" and that only 29.11% of participants in this study indicated that it was either fairly or very true that the organizations that

have their health information and share it "would never mislead me about how my health information is used". Policies can be put in place that create clear expectations for communication of serious mistakes as well as reporting the use of healthcare data. Adverse event reporting is already mandated by many states, Minnesota being the first state in 2003 (Minnesota Department of Health, n.d.). Health data use reporting, however, has neither an existing policy nor a clear forward progression. HIPAA has long managed the use of data, but as deidentified data is increasingly packaged and sold and resold, the guidelines provided by HIPAA are no longer sufficient. Future research is needed to understand how reporting can be done from the perspective of both providers and patients

3.5.2 Implications for research

Trust in the health system displayed stronger associations with comfort with sharing health data with third-party commercial companies than did privacy concerns. As health systems negotiate the possibilities of big data research with the potential risks to patient privacy, provision of health data use controls to patients has been offered as a solution. The results of this study suggest that the effect of privacy controls may be limited if trust in the health system is already low. As discussed in earlier chapters, individuals constantly negotiate and renegotiate their privacy boundaries without their knowing, and as a result find it difficult to tune preferences when presented with the option to control their health data. Research on what privacy control or notification structures are most valued by patients is important to establish next steps to protect patient privacy and engender trust in the health system.

3.5.3 Limitations

As with any survey, this study is merely a snapshot of patient beliefs and preferences, limited due to the nature of survey questions – different aspects of the healthcare experience that may provide a more complete portrait of the public's comfort with sharing health data with third-party commercial companies may not be captured here. Additionally, a stepwise regression model is a conservative model that eliminates factors that might be important to understanding patient and public comfort with sharing health data with third-party commercial companies.

The circumstances of data sharing and the privacy context in which that sharing will occur will continue to evolve as laws, expectations, and experiences of healthcare data sharing change. Longitudinal studies that evaluate changes in comfort with sharing health data with third-party commercial companies would be superior, especially in light of changing media coverage of these partnerships. In subsequent research, we will examine the public's confidence in current laws and policies that protect their data, and explore the public's preferences for control and notification of information sharing.

3.5.4 Conclusion

This study revealed that trust in the health system, more than any other variable examined here, is associated with increased comfort with sharing health data with third-party commercial companies for patient as well as business purposes. This study also opened the possibility that data breach and data misuse events might have lasting consequences on the public's overall comfort with sharing health data with third-party commercial companies. The magnitude of these consequences may also be impacted by trust in the health system. More research is needed on how much trust may impact the deleterious effects of data breaches.

As was stated in previous research (Chapter 2), healthcare systems embarking on new third-party data partnerships to expand their ability to process and analyze health data can benefit from early identification and communication of the patient-centered benefits that will result from their third-party commercial partnerships. Engaging patients at the inception of these partnerships – taking a community based participatory research approach to partnerships with third-party commercial companies – can engender trust in the partnership and strengthen patient trust in the organization overall.

3.6 Acknowledgements

Research reported in this manuscript was supported by the National Cancer Institute of the National Institutes of Health under award number 5 R01 CA214829-03.

3.7 References

Acquisto, A., Friedman, A., & Telang, R. (2006). Is There a Cost to Privacy Breaches? An Event Study. *Twenty-Seventh International Conference on Information Systems, Milwaukee*, 19.

Agaku, I. T., Adisa, A. O., Ayo-Yusuf, O. A., & Connolly, G. N. (2014). Concern about security and privacy, and perceived control over collection and use of health information are related to withholding of health information from healthcare providers. *Journal of the American Medical Informatics Association*, *21*(2), 374–378. https://doi.org/10.1136/amiajnl-2013-002079

Anderson, C., & Agarwal, R. (2011). The Digitization of Healthcare: Boundary Risks, Emotion, and Consumer Willingness to Disclose Personal Health Information. *Information Systems Research*, *22*(3), 469–490. https://doi.org/10.1287/isre.1100.0335

Arndt, R. Z. (2018, April 7). *How third parties harvest health data from providers, payers and pharmacies*. Modern Healthcare. https://www.modernhealthcare.com/article/20180407/NEWS/180409938/how-third-parties-harvest-health-data-from-providers-payers-and-pharmacies

Belanger, F., Hiller, J. S., & Smith, W. J. (2002). Trustworthiness in electronic commerce: The role of privacy, security, and site attributes. *The Journal of Strategic Information Systems*, *11*(3), 245–270. https://doi.org/10.1016/S0963-8687(02)00018-5

Bishop, L. "Sam," Holmes, B., & Kelley, C. (2005). National Consumer Health Privacy Survey 2005. *California Health Care Foundation*. https://www.chcf.org/publication/national-consumer-health-privacy-survey-2005/

Boulware, L. E., Cooper, L. A., Ratner, L. E., LaVeist, T. A., & Powe, N. R. (2003). Race and Trust in the Health Care System. *Public Health Reports*, *118*, 8.

Chakraborty, R., Lee, J., Bagchi-Sen, S., Upadhyaya, S., & Raghav Rao, H. (2016). Online shopping intention in the context of data breach in online retail stores: An examination of older and younger adults. *Decision Support Systems*, *83*, 47–56. https://doi.org/10.1016/j.dss.2015.12.007

Chang, A. (2018, March 23). *The Facebook and Cambridge Analytica scandal, explained with a simple diagram*. Vox. https://www.vox.com/policy-and-politics/2018/3/23/17151916/facebook-cambridge-analytica-trump-diagram

Chen, A. (2018, July 26). *IBM's Watson gave unsafe recommendations for treating cancer*. The Verge. https://www.theverge.com/2018/7/26/17619382/ibms-watson-cancer-ai-healthcare-science

Cohen, J. (2019, November 13). *Google, Ascension data partnership sparks federal probe*. Modern Healthcare. https://www.modernhealthcare.com/information-technology/google-ascension-data-partnership-sparks-federal-probe

Damschroder, L. J., Pritts, J. L., Neblo, M. A., Kalarickal, R. J., Creswell, J. W., & Hayward, R. A. (2007). Patients, privacy and trust: Patients' willingness to allow researchers to access their medical records. *Social Science & Medicine*, *64*(1), 223–235. https://doi.org/10.1016/j.socscimed.2006.08.045

Doukas, D. J., & Hardwig, J. (2014). Patient Informed Choice for Altruism. *Cambridge Quarterly of Healthcare Ethics*, *23*(4), 397–402. https://doi.org/10.1017/S0963180114000073

Feldman, S., & Steenbergen, M. R. (2001). The Humanitarian Foundation of Public Support for Social Welfare. *American Journal of Political Science*, *45*(3), 658–677. JSTOR. https://doi.org/10.2307/2669244

Fogel, J., & Nehmad, E. (2009). Internet social network communities: Risk taking, trust, and privacy concerns. *Computers in Human Behavior*, *25*(1), 153–160. https://doi.org/10.1016/j.chb.2008.08.006

Fouragnan, E., Chierchia, G., Greiner, S., Neveu, R., Avesani, P., & Coricelli, G. (2013). Reputational Priors Magnify Striatal Responses to Violations of Trust. *Journal of Neuroscience*, *33*(8), 3602–3611. https://doi.org/10.1523/JNEUROSCI.3086-12.2013

Fruhlinger, J. (2020, February 12). *Equifax data breach FAQ: What happened, who was affected, what was the impact?* CSO Online. https://www.csoonline.com/article/3444488/equifax-data-breach-faq-what-happened-who-was-affected-what-was-the-impact.html

Galloway, S. (2018, February 8). *The Case for Breaking Up Amazon, Apple, Facebook and Google*. Esquire. https://www.esquire.com/news-politics/a15895746/bust-big-tech-silicon-valley/

Goold, S. D. (2002). Trust, Distrust and Trustworthiness. *Journal of General Internal Medicine*, *17*(1), 79–81. https://doi.org/10.1046/j.1525-1497.2002.11132.x

Gressin, S. (2018, December 4). *The Marriott data breach*. Consumer Information. https://www.consumer.ftc.gov/blog/2018/12/marriott-data-breach

Griggs, M. B. (2019, November 11). *Google reveals 'Project Nightingale' after being accused of secretly gathering personal health records*. The Verge. https://www.theverge.com/2019/11/11/20959771/google-health-records-project-nightingale-privacy-ascension

Guzman, G. G. (2019). American Community Survey Briefs—Household Income: 2018. *United States Census Bureau*, 13.

Hall, M. A., Camacho, F., Dugan, E., & Balkrishnan, R. (2002). Trust in the Medical Profession: Conceptual and Measurement Issues: Trust in the Medical Profession: Conceptual and Measurement Issues. *Health Services Research*, *37*(5), 1419–1439. https://doi.org/10.1111/1475-6773.01070

Hall, M. A., Dugan, E., Zheng, B., & Mishra, A. K. (2001). Trust in Physicians and Medical Institutions: What Is It, Can It Be Measured, and Does It Matter? *The Milbank Quarterly*, *79*(4), 613–639. https://doi.org/10.1111/1468-0009.00223

Hall, M. A., Zheng, B., Dugan, E., Camacho, F., Kidd, K. E., Mishra, A., & Balkrishnan, R. (2002). Measuring Patients' Trust in their Primary Care Providers. *Medical Care Research and Review*, *59*(3), 293–318. https://doi.org/10.1177/1077558702059003004

Huang, B. (2018, September 25). *LVHN patient data not shared with for-profit company in Sloan Kettering trials*. Mcall.Com. https://www.mcall.com/health/mc-nws-lvhn-msk-paigeai-20180924-story.html

Kantsperger, R., & Kunz, W. H. (2010). Consumer trust in service companies: A multiple mediating analysis. *Managing Service Quality: An International Journal*, *20*(1), 4–25. https://doi.org/10.1108/09604521011011603

Kim, K. K., Sankar, P., Wilson, M. D., & Haynes, S. C. (2017). Factors affecting willingness to share electronic health data among California consumers. *BMC Medical Ethics*, *18*(1), 25. https://doi.org/10.1186/s12910-017-0185-x

Krasnova, H., Spiekermann, S., Koroleva, K., & Hildebrand, T. (2010). Online Social Networks: Why We Disclose. *Journal of Information Technology*, *25*(2), 109–125. https://doi.org/10.1057/jit.2010.6

LaVeist, T. A., Isaac, L. A., & Williams, K. P. (2009). Mistrust of Health Care Organizations Is Associated with Underutilization of Health Services. *Health Services Research*, *44*(6), 2093–2105. https://doi.org/10.1111/j.1475-6773.2009.01017.x

Lee, M., & Lee, J. (2012). The impact of information security failure on customer behaviors: A study on a large-scale hacking incident on the internet. *Information Systems Frontiers*, *14*(2), 375–393. https://doi.org/10.1007/s10796-010-9253-1

Lee, P. (2017, February 16). *Microsoft and partners combine the cloud, AI, research and industry expertise to focus on transforming health care*. The Official Microsoft Blog. https://blogs.microsoft.com/blog/2017/02/16/microsoft-partners-combine-cloud-ai-research-industry-expertise-focus-transforming-health-care/

McKnight, & Chervany. (2001). Trust and Distrust Definitions: One Bite at a Time. In R. Falcone, M. Singh, & Y.-H. Tan (Eds.), *Trust in Cyber-societies* (pp. 27–54). Springer. https://doi.org/10.1007/3-540-45547-7_3

McKnight, D. H., Choudhury, V., & Kacmar, C. (2002). Developing and Validating Trust Measures for e-Commerce: An Integrative Typology. *Information Systems Research*, *13*(3), 334–359. https://doi.org/10.1287/isre.13.3.334.81

Memorial Sloan Kettering Cancer Center. (2014, April 11). *Memorial Sloan Kettering Trains IBM Watson to Help Doctors Make Better Cancer Treatment Choices | Memorial Sloan Kettering Cancer Center*. Memorial Sloan Kettering Cancer Center. https://www.mskcc.org/blog/msk-trains-ibm-watson-help-doctors-make-better-treatment-choices

Minnesota Department of Health. (n.d.). *Adverse Events Reporting—Minnesota Dept. Of Health*. MDH Division of Health Policy. Retrieved May 18, 2020, from https://www.health.state.mn.us/facilities/patientsafety/adverseevents/index.html

Moon, M. (2019, June 27). *Google and University of Chicago face lawsuit over shared patient data*. Engadget. https://www.engadget.com/2019-06-27-google-university-of-chicago-lawsuit-patient-data.html

Office for Civil Rights (OCR). (2013). *190-Who must comply with HIPAA privacy standards* [Text]. HHS.Gov. https://www.hhs.gov/hipaa/for-professionals/faq/190/who-must-comply-with-hipaa-privacy-standards/index.html

Ornstein, & Thomas. (2018, September 20). *Sloan Kettering's Cozy Deal With Start-Up Ignites a New Uproar*. ProPublica. https://www.propublica.org/article/sloan-kettering-cozy-deal-with-start-up-paige-ai-ignites-new-uproar

Petrow, S. (2018, October 3). Memorial Sloan Kettering, you've betrayed my trust. *STAT*. https://www.statnews.com/2018/10/03/memorial-sloan-kettering-betrayed-my-trust/

Platt, J. E., Jacobson, P. D., & Kardia, S. L. R. (2018). Public Trust in Health Information Sharing: A Measure of System Trust. *Health Services Research*, *53*(2), 824–845. https://doi.org/10.1111/1475-6773.12654

Ponemon Institute. (2018). *2018 Cost of Data Breach Study: Global Overview*. IBM Security. https://www.ibm.com/downloads/cas/861MNWN2

Porter, J. (2019, January 21). *Google fined €50 million for GDPR violation in France*. The Verge. https://www.theverge.com/2019/1/21/18191591/google-gdpr-fine-50-million-euros-data-consent-cnil

Reynolds, W. W., & Nelson, R. M. (2007). Risk perception and decision processes underlying informed consent to research participation. *Social Science & Medicine*, *65*(10), 2105–2115. https://doi.org/10.1016/j.socscimed.2007.06.021

Richards, N. M., & Hartzog, W. (2015). Taking Trust Seriously in Privacy Law. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2655719

Roach, J. (2019, October 28). *Microsoft + The Jackson Laboratory: Using AI to fight cancer*. The AI Blog. https://blogs.microsoft.com/ai/jackson-lab-project-hanover/

Schwarzer, R., & Jerusalem, M. (1992). General Self-Efficacy- Schwarzer (GSES). *Statistics Solutions*. https://www.statisticssolutions.com/general-self-efficacy-schwarzer-gses/

Shaukat, T. (2019, November 11). *Our partnership with Ascension*. Google Cloud Blog. https://cloud.google.com/blog/topics/inside-google-cloud/our-partnership-with-ascension/

Shavers, V. L., Lynch, C. F., & Burmeister, L. F. (2001). Factors that influence African-Americans' willingness to participate in medical research studies. *Cancer*, *91*(S1), 233–236. https://doi.org/10.1002/1097-0142(20010101)91:1+<233::AID-CNCR10>3.0.CO;2-8

Shen, N., Bernier, T., Sequeira, L., Strauss, J., Silver, M. P., Carter-Langford, A., & Wiljer, D. (2019). Understanding the patient privacy perspective on health information exchange: A systematic review. *International Journal of Medical Informatics*, *125*, 1–12. https://doi.org/10.1016/j.ijmedinf.2019.01.014

Singer, N., & Wakabayashi, D. (2019, November 11). Google to Store and Analyze Millions of Health Records. *The New York Times*. https://www.nytimes.com/2019/11/11/business/google-ascension-health-data.html

Smith, T., Davern, M., Freese, J., & Morgan, S. (2019). General Social Surveys, 1972-2018. *NORC*, 11.

Stockdale, J., Cassell, J., & Ford, E. (2019). "Giving something back": A systematic review and ethical enquiry into public views on the use of patient data for research in the United Kingdom and the Republic of Ireland. *Wellcome Open Research*, *3*, 6. https://doi.org/10.12688/wellcomeopenres.13531.2

Taddei, S., & Contena, B. (2013). Privacy, trust and control: Which relationships with online self-disclosure? *Computers in Human Behavior*, *29*(3), 821–826. https://doi.org/10.1016/j.chb.2012.11.022

Tanner, A. (2016, February 1). *How Data Brokers Make Money Off Your Medical Records*. Scientific American. https://doi.org/10.1038/scientificamerican0216-26

Turow, J. (2017, June 28). *Google Still Doesn't Care About Your Privacy*. Fortune. https://fortune.com/2017/06/28/gmail-google-account-ads-privacy-concerns-home-settings-policy/

US Census Bureau. (2019). *Health Insurance Coverage in the United States: 2018*. The United States Census Bureau. https://www.census.gov/library/publications/2019/demo/p60-267.html

Vincent, J. (2016, September 20). *Microsoft announces new AI-powered health care initiatives targeting cancer*. The Verge. https://www.theverge.com/2016/9/20/12986314/microsoft-ai-healthcare-project-hanover-cancer

Waldman, A. E. (2016). Privacy, sharing, and trust: The Facebook study. *Case Western Reserve Law Review*, *67*(1), 193-. Academic OneFile.

Wang, T.-L., & Tseng, Y. F. (2011). A Study of the Effect on Trust and Attitude with Online Shopping. *International Journal for Digital Society*, *2*(2), 433–440. https://doi.org/10.20533/ijds.2040.2570.2011.0052

Willison, D. J., Steeves, V., Charles, C., Schwartz, L., Ranford, J., Agarwal, G., Cheng, J., & Thabane, L. (2009). Consent for use of personal information for health research: Do people with potentially stigmatizing health conditions and the general public differ in their opinions? *BMC Medical Ethics*, *10*(1), 10. https://doi.org/10.1186/1472-6939-10-10

Chapter 4 Desire for Control Over Data or Data Use Notification and the Public's Comfort with Sharing Health Data with Third-Party Commercial Companies

4.1 Abstract

Background: Healthcare partnerships with third-party commercial companies, while critical to the goal of realizing a learning health system, have been met with reservations from the public about the privacy of health information and concerns about how this data is used. While research points to the need to provide patients greater control over the use of their data, or notification of data use, it is not yet clear how to move forward with this effort while balancing the needs of researchers for quality data sets.

Objective: To examine the relationships between the public's comfort sharing health data with third-party commercial companies for patient and business purposes with the public's comfort with researchers, quality analysts, commercial companies, and law enforcement, the public's confidence in existing laws and policies, and desire for control over their health data and notification of data use.

Methods: We surveyed the US public (n = 1841) to assess comfort with sharing health data with third-party commercial companies for patient or business purposes. Weighted Ordinary Least Squares (OLS) Regression analysis was used to first estimate the linear relationship between comfort with third-party commercial companies for patient and business purposes (dependent variables) and comfort with researchers, quality analysts, commercial companies and law enforcement, confidence in existing laws and policies, and desire for control over data and notification of use (independent variables), followed by stepwise regression modeling to estimate a full model of contributing factors to the public's comfort with sharing health data with third-party commercial companies—demographic variables, trust in the health system and in providers, privacy concerns, altruism, experience of a data breach, and concern about recent events and data breaches are included in this final analysis.

Findings: In this analysis, three variables were strongly associated with respondent's comfort with sharing health data with third-party commercial companies for both patient and business purposes: trust in the health system, confidence in existing laws and policies, and desire for notification. Trust in the health system persisted in this full model of the public's comfort third-party commercial companies for both patient purposes ($b^* = 0.235, p = 2.6 * 10^{-12}$) and business purposes ($b^* = 0.159, p = 3.0 * 10^{-08}$). Confidence in the existing laws and policies governing health data was also strongly associated with comfort with third-party commercial companies for both patient purposes ($b^* = 0.136, p = 0.0003$) and business purposes ($b^* = 0.102, p = 0.00012$). Notably, desire for control over how health information was shared did not persist in the final model—instead, desire for notification displayed a positive association with comfort with sharing health data with third-party commercial companies for patient purposes ($b^* = 0.118, p = 0.00064$), and an inverse association with comfort sharing health data with third-party commercial companies for business purposes ($b^* = -0.096, p = 1.6 * 10^{-05}$).

Implications: The results of this study suggest that increasing trust in the health system may have a greater impact on the public's comfort than efforts to address privacy concerns alone. Desire for notification was also more significant than the desire for control over health data. Patients may be better served by focusing on efforts to build trust in the health system and provide notification of health data use instead of providing granular control over data use.

4.2 Introduction

4.2.1 Background

Google, Facebook, Amazon, Comcast, and Apple all have their eyes on healthcare. Driven by an aging population and a rise in prevalence of chronic conditions, demand for healthcare is growing faster than any industry in the United States and worldwide—by 2022, the global healthcare market is expected to grow to $11.9 trillion (Business Wire, 2019). Given this enduring growth and market size, it is no surprise that companies like Google and Apple have aggressively expanded their businesses into health services. These commercial entrants into

healthcare, however, pose huge challenges for current healthcare data policy. As these companies push their own healthcare products to meet the very real demands of patients—products that offer better telehealth services (Quil Health by Comcast, Google, and Facebook), better heart monitoring devices and wearables (Apple and Google), and applications and data platforms for patients and providers (IBM, Google, Amazon, and Facebook)—they upturn existing models of patient data management and privacy.

The Health Insurance Portability and Accountability Act (HIPAA) provides most of the regulatory restrictions on the use of health data, and third-party companies typically come under the purview of HIPAA if and when they became a business associate (BA) of a healthcare provider or payer. However, by providing healthcare services directly to the patient consumer and bypassing healthcare systems and payors, companies may fall outside of that regulatory system. Additionally, HIPAA provisions apply only to identified health data—deidentified data, or data that has been stripped of identifiers, is not restricted. Deidentified data is aggregated, packaged, and sold by healthcare systems, insurers, and now technology companies, resulting in a worldwide data market estimated at $67 billion (Merken & Elfin, 2018). It is important to understand whether the public feels comfortable with the shifting landscape of health data organizations and the policy that regulates them.

The privacy risks of deidentified data, however, are significant in our current data driven landscape. In October 2019, Facebook announced a health initiative called Preventative Health, which intended to send Facebook users reminders about flu shots, cancer screenings, and blood pressure checks in order to assist physicians in the management of chronic conditions (Ousfar, 2019). True to Facebook's core business of data collection, however, the initiative hoped to "leverage the cache of data users already give [Facebook] – about their education, relationships, habits, spoken languages, employment status, and more, all of which have an enormous impact on health outcomes—to create a sort of subclinical health-care system [that would warn] providers if, for example, a user recovering from surgery had a small support group" (Fussell, 2020). Both Facebook and Google state in their privacy notices that individual user data is never shared or sold to third-parties—which sidesteps the fact that the real monetary value of the data collected is in the behavioral meta-data that belongs to no individual user but aids in the

generation of probability-based health profiles on all users. Machine learning and artificial intelligence techniques have made it possible to create individual-specific health profiles based on browser histories, social media activity, and social connections. Moreover, data that has been deidentified can be reidentified (Hoffman, 2020) through processes like geocoding (Rushton et al., 2006). Further, "anecdotal evidence suggests [that] algorithms already exist that can re-identify patient information with prescription drug information after third-party data mining companies ostensibly de-identify the information" (Gellman, 2011).

Recent changes to the management of patient data make the problems of data privacy and security all the more salient. In March 2020, HHS finalized "two transformative rules that will give patients unprecedented safe, secure access to their health data" (HHS Press Release, 2020). One of those rules "establishes secure, standards-based application programming interface (API) requirements to support a patient's access and control of their electronic personal health information (ePHI)" so that patients can "securely and easily obtain and use their electronic health information […] using *an app of their choice*" (HHS Press Release, 2020). While these rules go far in putting control of health information in the hands of patients as intended, they are out of sync with existing HIPAA regulations. Per HIPAA as it currently stands, if the patient is the initiator of a relationship with a third-party commercial company, then HIPAA regulations no longer apply, and the patient's data is no longer protected (Davis, 2019; Singer, 2019). Furthermore, the individual patient's right of access to their ePHI will prevent healthcare organizations from refusing to share the patient's health data with the application or company, regardless of organizational concerns that the chosen application may be a possible threat to the patient's data security and privacy.

In this data environment, patients are both winners and losers. The healthcare marketplace is expanding to provide new offerings to patients who have been frustrated by poor healthcare access and availability (Batbaatar et al., 2017), but in so doing, companies outside of the dyadic patient-provider relationship are poised to collect even more health data than they already have. Previous research has indicated patients are largely unaware of how their identified and deidentified data can be and is used by third-party commercial companies (C. Smith, 2011), and case studies provided by Google and Facebook, among others, indicate that third-party data

efforts are unfavorably received by the public. Literature examining patient willingness to share health information for research broadly has found that generally patients are unwilling to share information beyond what is necessary to provide care, do not like to be surprised about how their data are used (Jilka et al., 2015; Kim et al., 2019; Shen et al., 2019), and desire greater control over the use and dissemination of their health data (Damschroder et al., 2007; Jilka et al., 2015). Patient reservations about the use of health data also apply to deidentified health data (Cohen & Mello, 2018; C. Smith, 2011). While an abundance of research has been conducted on the privacy concerns of the public, provision of health data privacy controls to patients are still lacking or inadequate. While the reasons for this include both technical and bureaucratic issues, healthcare researchers have also indicated concerns that wholesale control over health data will compromise data sets and limit research efforts that intend to improve healthcare, especially amongst underrepresented populations. Furthermore, privacy expectations differ greatly from person to person. Any substantial effort made is bound to be adequate for some and inadequate for others. Better understanding of the factors that influence patient comfort with commercial companies is needed to guide future approaches to regulatory management, patient consent, and patient notification.

4.2.2 Problem Statement

Current research emphatically recommends providing privacy controls to patients; however, there has not been much clarity on what controls to provide and what impact they may have on patient satisfaction and resolution of privacy concerns. Examination of public attitudes towards these companies and how these attitudes intersect with trust and privacy concerns is needed to help guide privacy efforts as patient data is increasingly shared with and generated by third-party commercial companies.

4.2.3 Conceptual Model

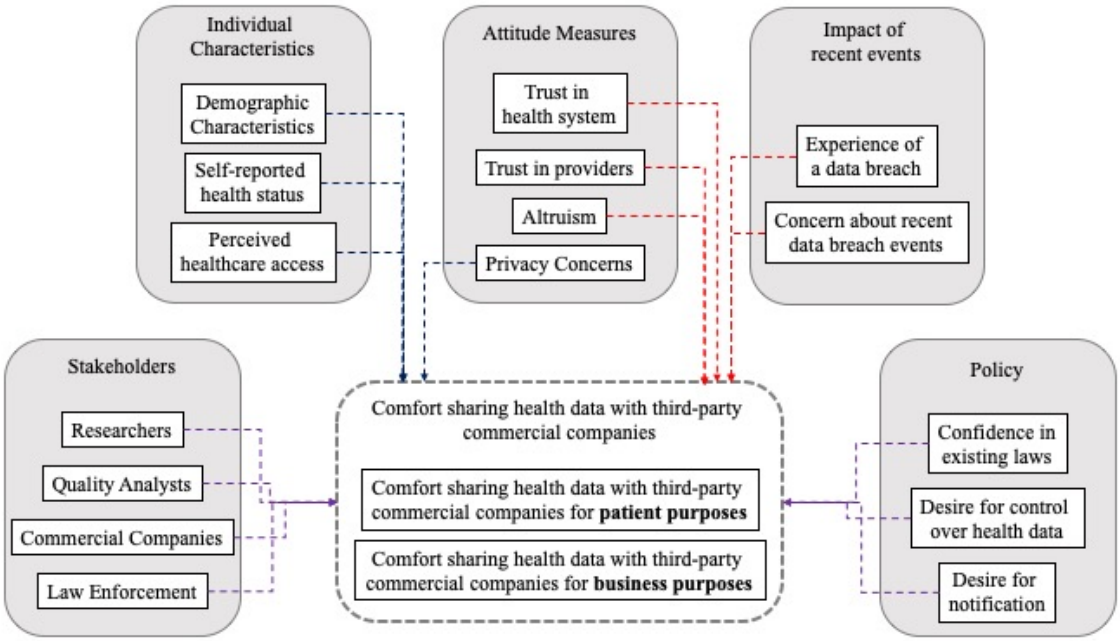

Figure 4-1 Conceptual Model of Dissertation Research – this analysis focuses on stakeholders and policy, all variables are included in the final multivariable regression model

Difference between business and patient purposes

Third party commercial companies are those companies that typically fall outside of usual "covered entities". Covered entities include healthcare systems or providers who transmit any health information, health care plans, and health care clearinghouses (billing services, community health information systems, etc.) (Office for Civil Rights (OCR), 2013). In this analysis we examine the public's comfort sharing health data with third-party commercial companies, investigating comfort when data is shared for *patient purposes*, i.e., to improve care, diagnosis, or treatment, versus comfort when data is shared for *business purposes* such as selling de-identified data. In chapter two we closely examined the division of patient-focused versus business-focused health data sharing and found a statistically significant difference between comfort with sharing data for patient and business purposes, results which will be reviewed and restated here for continuity. The public is more comfortable with sharing health data with third-party commercial companies if the benefits to the patient or to the public are clear, which comports with past research that also concluded patients are more willing to provide access to their health information if the potential health benefits are clear (Anderson & Agarwal, 2011),

and are less willing to provide access if their health data will be used for profit-generating research (Willison et al., 2009). We expand that analysis here.

Comfort with Researchers, Quality Analysts, Commercial Companies

Multiple studies have shown that on the whole, patients are willing to share their health data with researchers (Damschroder et al., 2007; Karampela et al., 2019; Seltzer et al., 2019; Spencer et al., 2016; Teixeira et al., 2011). Despite dissatisfaction about how the results of researcher studies were communicated back to participants, patients were also willing to share data beyond their personal health information for the purposes of research, including tax records and credit card histories (Seltzer et al., 2019). This willingness to share health information is, however, largely confined to research efforts only. As an example of willingness to share health data with third-party commercial companies, a 2019 survey conducted by the Chicago Booth/Kellogg School Financial Trust found that 93% of survey participants were unwilling to share their health data with Facebook (Promarket, 2020). This study limited their question to Facebook only, when the health data industry contains so many other players vying for patient data.

Comfort with Law Enforcement

Third-party access to health data also includes police. A 2019 survey conducted by the Pew Research Center found that 48% of Americans consider it acceptable for DNA testing companies (23andMe, Ancestry.com) to share their customers' data with law enforcement, while one-third (34%) of respondents said sharing with law enforcement was unacceptable, and 18% were unsure whether the practice was acceptable or unacceptable (Perrin, 2020). As third-party companies command ever greater access to health data as well as data that can be considered health data (web histories, fitness applications), law enforcement is also gaining an increasingly robust data set with which to conduct investigations. In 2019, a Florida judge granted a warrant that allowed the police to search the complete genetic database of GEDMatch. The terms of the warrant included all customers of GEDMatch, including those who didn't opt-in to any data sharing agreement. Following this event, GEDMatch updated their terms of agreement to ensure that only those who consent to a search of their genetic information will be access. Notably, as of

November 2019, only 185,000 of the company's total population of 1.3 million users, a mere 7% of the total user population, have provided this consent (Tiller, 2019).

Confidence in existing laws and policies

As healthcare breaches increase year after year, patient confidence in the ability of the health system to protect their health information is decreasing (HIPAA Journal, 2017). In a 2016 survey, 89% of healthcare consumers reported withholding health information during their visit because of privacy and security concerns (Black Book Market Research, 2017). Research shows patients desire clear and consistent consequences for anyone who violates patient privacy, and for researchers to be held accountable for maintaining confidentiality (Damschroder et al., 2007), but changes to regulations governing information sharing have not yet occurred. In 2007, only 25% of participants were aware that researchers could use their medical records without explicit permission from the patient (Damschroder et al., 2007). In 2016, 81% of respondents to a Black Book Market Research survey (Black Book Market Research, 2017) reported concern that their chronic condition data was being shared with retailers, employers, or the government without their knowledge (Gooch, 2017).

Desire for control and notification

Research on willingness to allow for personal health data to be used for research have found that in general, 96% of patients were willing to provide their data for research, yet 78% also indicated their desire for more control over how their information was used (Damschroder et al., 2007). Patients have consistently reported wanting to know how their health data may have contributed to helping others, and who was using their medical records for what purpose (Damschroder et al., 2007; Kim et al., 2019; Weitzman et al., 2010). Desire for notification persists even when there is high institutional trust, as seen by Damschroder et al in their research on veterans (Damschroder et al., 2007).

4.2.4 Study Objective

The aims of this study are to a) examine the relationship between the public's comfort with sharing health data with third-party commercial companies and their comfort with researchers, quality analysts, commercial companies, and law enforcement; b) examine the

public's comfort with sharing health data with third-party commercial companies for patient and business purposes and their confidence in existing health data privacy laws and policies; and c) examine the public's comfort with sharing health data with third-party commercial companies and their desire for control over their health information or desire for notification of health data use.

4.3 Methods

Respondent comfort with sharing health data with third-party commercial companies was captured using a 20-minute online survey of US adults. In this section I explain how the concepts described above are operationalized in this survey, followed by an explanation of the statistical methods used to analyze this data.

4.3.1 Participants

Respondents were surveyed using the National Opinion Research Center's (NORC) probability-based, nationally representative sample of US adults, based on 2010 Census Information. NORC's national sample frame employs a two-stage probability sample design to select a representative sample of households in the United States, oversampling African American, Hispanic populations, as well as households 200% below the federal poverty level. Survey recruitment and deployment was done in May 2019. Data collection was completed by June 2019. Eligible participants (at least 21 years old and able to read and write in English) were contacted via email to participate in the online survey, resulting in a total of 2,157 participants (66% response rate). The first component of the survey was a short (90 seconds) animated video describing how health data of a fictional patient is shared through the duration of care—to insurers, billers, and analysts learning from the outcomes of treatment. Definitions of important terms such as "healthcare system", "healthcare providers", "electronic health record", "de-identified health information [or biospecimens]", and "commercial companies" were provided to survey participants wherever those terms appeared. "Commercial companies" was defined for respondents to this survey as "third-party companies that are not part of a hospital. For example, a third-party commercial company may conduct genetic tests and analyze information for a hospital or healthcare provider for a fee when a hospital is not able to conduct the test on their

own." "De-identified [health information or biospecimens]" was defined for respondents in the following manner: "de-identified means that "identifying information" about you is removed from your health information. Identifying information includes things like your name, address, date of birth, etc."

NORC calculated post-stratification weights according to US Census demographic benchmarks for age, sex, household income, education, as well as race and ethnicity to reduce sampling bias. For the purposes of this paper, records with missing responses to one or more of the questions used in this analysis were not included, resulting in a final analyzed sample of 1,841 responses. This study protocol was approved by the University of Michigan Health Sciences Institutional Review Board.

4.3.2 Survey Design

Variables used in this study were derived from a 20-minute, 164-item survey created to examine knowledge, attitudes, and beliefs about data sharing. Trust measures were adapted from the work of Mark Hall and colleagues (Hall, Camacho, et al., 2002; Hall et al., 2001; Hall, Zheng, et al., 2002). Altruism measures were adapted from the General Social Survey (T. Smith et al., 2019), the National Election Survey (Feldman & Steenbergen, 2001), and the General Self-Efficacy Scale (Schwarzer & Jerusalem, 1992). Privacy measures were adapted from Anderson's research on consumer willingness to disclose personal health information and the California Health Foundation's 2005 National Consumer Health Privacy survey (Anderson & Agarwal, 2011; Bishop et al., 2005). Privacy measures also include questions about deception and medical mistrust (Boulware et al., 2003; LaVeist et al., 2009) and have been used in previous studies (Platt et al., 2018).

4.3.3 Measurements used in this study

The Public's Comfort with Sharing Health Data with Third-Party Commercial Companies for Patient and Business Purposes

To explore public comfort with sharing health data with third-party commercial companies for patient purposes, respondents answered questions about "how comfortable" they

were with three statements regarding data sharing with third-party commercial companies, each along a 4-point Likert scale. Participants were asked "How comfortable are you with a third-party commercial company using your DNA and health information to improve the diagnosis and treatment of cancer in other patients?" and "How comfortable are you with a third-party commercial company developing predictions about how you will respond to a particular cancer treatment?: "not at all comfortable" (1), "somewhat comfortable" (2), "fairly comfortable" (3), and "very comfortable" (4). Participants were also asked "how true" it was that "The organizations that have my health information and share it can use large amounts of data to improve patient care": "not true" (1), "somewhat true" (2), "fairly true" (3), and "very true" (4).

To examine participant comfort with sharing health data with third-party commercial companies for business purposes, participants were asked "How comfortable are you with a third-party commercial company storing your DNA and health information?"; "How comfortable are you with a third-party commercial company sharing predictions about how you will respond to cancer treatment with insurance companies?"; and "How comfortable are you with a third-party commercial company selling de-identified health information to a pharmaceutical company?". "Business purpose" in this research is understood as storage of health data beyond the purposes of clinical care and sharing information with third-party commercial companies to improve their own business processes without explicitly stated direct benefit to patients. Respondents were provided with the options "not at all comfortable" (1), "somewhat comfortable" (2), "fairly comfortable" (3), and "very comfortable" (4). Indices for data use for patient purposes and business purposes were then calculated as the sum of participant responses to the three questions in each index divided by the number of questions.

The Cronbach's alpha for these questions was 0.766 for comfort with sharing health data with third-party commercial companies for patient purposes and 0.786 for comfort with sharing health data for business purposes.

Comfort with Researchers, Quality Analysts, Commercial Companies

Comfort with researchers, quality analysts, and commercial companies was measured using a 4-item index for each group. Respondents were asked "how true" the following

statements were for themselves: a) "I am comfortable with [Researchers/Quality Analysts/Commercial Companies] using my de-identified health information", b) "I am comfortable with [Researchers/Quality Analysts/Commercial Companies] using my identified health information", c) "I am comfortable with [Researchers/Quality Analysts/Commercial Companies] using my de-identified biospecimens", d) "I am comfortable with [Researchers/Quality Analysts/Commercial Companies] using my identified biospecimens. Respondents were provided with options along a Likert scale ranging from 1 ("not true") to 4 ("very true"). The final comfort scores for each group is the total of participant responses for that group divided by the number of questions for that group index. The Cronbach's alpha for questions comprising the comfort with researchers index is 0.791, the comfort with quality analysts index is 0.809, and the comfort with commercial companies index is 0.813.

Comfort with Law Enforcement

In this study we focus on law enforcement access to health information and genetic data specifically, consistent with recent events whereby law enforcement collected genetic data from direct-to-consumer companies such as 23andMe and Ancestry.com. Respondents were asked "how true" they found the following statement: "It is okay for law enforcement to access health information". Choices for this question ranged from 1 ("not true") to 4 ("very true"). Respondents were asked a second question regarding law enforcement access to genetic data: "Are you comfortable with law enforcement using genetic and ancestry data from companies like 23andMe and AncestryDNA?". Options for this latter question were "yes" (1) and "no" (2).

Confidence in Existing Laws and Policies

Two questions were used to assess respondent confidence in existing laws and policies. Respondents were asked to evaluate "For you, how true are the following statements": a) "Existing laws provide a reasonable level of protection for the privacy of patients" and b) "I am confident that electronic health information is sufficiently protected by current law and regulation". Available responses were provided along a Likert scale ranging from 1 ("not true") to 4 ("very true").

Desire for Control and Notification

Desire for greater control over the use of health data and notification of the use of health data was examined using three questions. Each of the three questions asked respondents to evaluate "For you, how true are the following statements": a) "I should have more control over how my health information is used", b) "It is important I know who has health information about me", and c) "I should be able to find out how my health information is shared". All three questions were accompanied by response options along a Likert scale ranging from 1 ("not true") to 4 ("very true").

4.3.4 Data Analysis

For this analysis we first generated summary descriptive statistics on each question included in this analysis – comfort with researchers, quality analysts, commercial companies, law enforcement, confidence in existing laws and policies, and desire for control over the use of health data and desire for notification of the use of health data.

Weighted Ordinary Least Squares (OLS) Regression analysis was used to estimate the linear relationship between comfort with third-party commercial companies for patient and business purposes and each demographic and health variable separately. We then estimated a multivariable model with all demographic and health variables and conducted a stepwise regression model to identify a parsimonious set of variables that explained the greatest amount of variability in the two outcomes – comfort with sharing data with commercial companies for business or patient purposes. For the stepwise regression model, we set statistical significance at α=0.05 ($p<0.002$) for inclusion and α=0.01 for exclusion, applying a Bonferroni correction to minimize Type I error. To enable comparison of effect sizes, regression coefficients were normalized (mean = 0, SD = 1).

4.4 Results

To examine predictors of comfort with third-party commercial companies using health information for patient and business purposes, we first examined the descriptive statistics for each of the independent variable and then conducted univariable and multivariable stepwise regression to identify predictors. Statistical significance was set at the 0.05 level.

4.4.1 Sample Demographics

The resulting weighted sample of 1,841 participants shows a near even split between male and female participants (49% male). Approximately 12% of participants were under the age of 29, and 31% of participants were over the age of 60. Nearly 60% of participants identified as white non-Hispanic, 15% as black, non-Hispanic, 19% as Hispanic, 3% as Asian, non-Hispanic, 2% of participants identified race and ethnicity as "other", and 3% identified as multiethnic, consistent with 2016 data from the US Census Bureau (12.3% of the US population identifies as black or African-American, non-Hispanic). Nearly half of participants completed some college (46%), and 33% of participants have a bachelor's degree. While the proportion of participants with a bachelor's degree is consistent with national percentages (30%, 2016 census data), the proportion of participants with some college, no degree is much higher in this study than national percentages (21%, 2016 census data). Just over half of participants (59%) made an income less than $60,000, consistent with the median household income for 2018 (Guzman, 2019). Over half of participants (60%) had employment. Of the health questions included in this analysis, 89% of study participants reported having health insurance of some type, which is slightly lower than reported national percentages - 92% of the US population according to the 2018 US Census (US Census Bureau, 2019). The mean self-reported health score of participants was 3.08, suggesting that on average, the respondents were of "good" health.

Table 4-1 Demographic descriptive statistics

Table 4.1 Demographic descriptive statistics (N = 1841)			
		N	Frequency
Sex			
	Male	903	49.05%
	Female	938	50.95%
Age			
	18-29	227	12.33%
	30-44	554	30.09%
	45-59	483	26.24%
	60+	577	31.34%
Race/Ethnicity			
	White	1086	58.99%
	Black, NH	273	14.83%
	Other, NH	30	1.63%
	Hispanic	358	19.45%
	Multiracial, NH	47	2.55%

	Asian, NH	47	2.55%
Education			
	Less than High School	73	3.97%
	High School	317	17.22%
	Some college	841	45.68%
	BA or above	610	33.13%
Income			
	Less than $60,000	1082	58.77%
	$60,000 or greater	759	41.23%
Employment			
	Employed	1112	60.40%
	Not employed	87	4.73%
	Retired	373	20.26%
	Disabled/Other	269	14.61
Insured			
	Is insured	1638	88.97%
	Is not insured	203	11.03%
Self-reported health			
	Range: 1 (Poor) to 5 (Excellent)		Mean: 3.08 (SD=0.92)

4.4.2 Public Comfort with Sharing Healthcare Data with Third-Party Commercial Companies for Patient and Business Purposes

Public comfort with sharing health data with third-party commercial companies was evaluated using two indices: 1) comfort with sharing health data with third-party commercial companies for patient purposes, and 2) comfort with sharing health data with third-party commercial companies for business purposes (Table 2.2). The resulting mean of comfort with sharing data with third-party commercial companies for patient purposes was 2.54 (SD = 0.81) or between “somewhat comfortable” and “fairly comfortable”. Roughly half of participants indicated that they were either fairly for very comfortable sharing data with third-party commercial companies for patient purposes as expressed by the component questions of the index: (53.39% are comfortable with a third-party commercial company using their DNA and health information to improve the diagnosis and treatment of cancer in other patients , 49.16% are comfortable with third-party commercial companies developing predictions about how they will respond to a particular cancer treatment, and 47.80% believe that the organizations that have their health information and share it can use large amounts of data to improve patient care).

Comfort with sharing health data with third-party commercial companies for business purposes had a resulting mean of 1.93 (SD = 0.85) or "somewhat comfortable". One quarter to one third of participants indicated they were either fairly or very comfortable with each of the component questions in comfort sharing health data third-party commercial companies for business purposes (28.90% are comfortable with a third-party commercial company storing their DNA and health information, 31.02% are comfortable with a third-party company sharing predictions about how they will respond to cancer treatment with insurance companies, and 24.39% are comfortable with a third-party commercial company selling de-identified health information to a pharmaceutical company). Figure 2.1 shows the distributions of the two indices.

A paired t-test was conducted on both comfort indices, the results of which show that there is a statistically significant difference between comfort with sharing health data with third-party commercial companies for patient purposes and comfort sharing health data with third-party commercial companies for business purposes only, paired $t = 39.83, p < 0.001$.

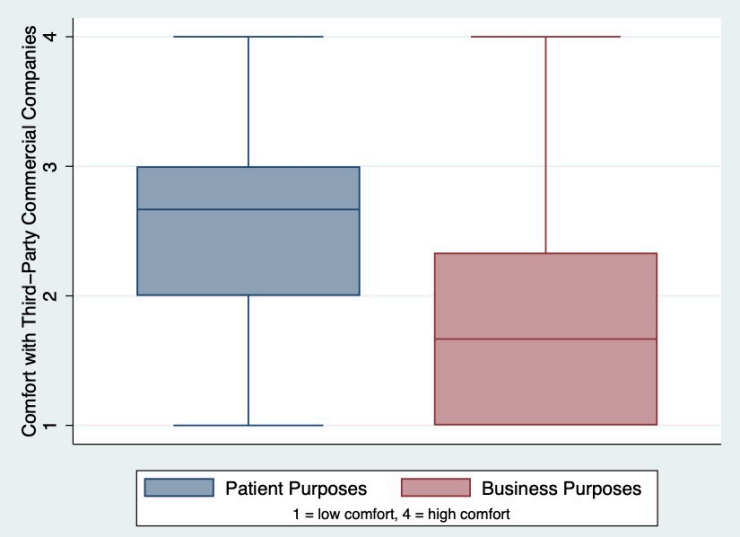

Figure 4-2 Box plot distributions of indices measuring Comfort Sharing Health Data with Third-Party Commercial Companies for Patient Purposes and Business Purposes

Table 4-2 Descriptive statistics for survey questions used in indices measuring comfort sharing health data with third-party commercial companies for patient and business purposes

Table 4.2 Descriptive statistics for survey questions used in indices measuring comfort sharing health data with third-party commercial companies for patient purposes and comfort sharing health data with third-party commercial companies for business purposes. (N = 1841)		
	Frequency (% fairly or very comfortable/ fairly or very true)	Mean (SD)
Comfort Sharing Health Data with Third-Party Commercial Companies for **Patient Purposes**		
How comfortable are you with a third-party commercial company using your DNA and health information to improve the diagnosis and treatment of cancer in other patients?	53.39%	2.58 (1.05)
How comfortable are you with a third-party commercial company developing predictions about how you will respond to a particular cancer treatment?	49.16%	2.48 (1.02)
The organizations that have my health information and share it can use large amounts of data to improve patient care	47.80%	2.56 (0.86)
Comfort sharing health data with third-party commercial companies for patient purposes index (Cronbach's α=0.769)	Median: 2.67	2.54 (0.81)
Comfort Sharing Health Data with Third-Party Commercial Companies for **Business Purposes**		
How comfortable are you with a third-party commercial company storing your DNA and health information?	28.90%	1.98 (1.01)
How comfortable are you with a third-party commercial company sharing predictions about how you will respond to cancer treatment with insurance companies?	31.02%	2.00 (1.04)
How comfortable are you with a third-party commercial company selling de-identified health information to a pharmaceutical company?	24.39%	1.81 (1.01)
Comfort sharing health data with third-party commercial companies for business purposes index (Cronbach's α=0.786)	Median: 1.67	1.93 (0.85)

** Range of indices: 1 = not comfortable sharing health data with third-party commercial companies; 2 = somewhat comfortable sharing health data with third-party commercial companies; 3 = fairly comfortable sharing health data with third-party commercial companies; 4 = very comfortable sharing data with third-party commercial companies*

4.4.3 Comfort with Researchers, Quality Analysts, Commercial Companies

Comfort with researchers, quality analysts, and commercial companies was assessed using a four-item index for each group. Over sixty percent of respondents indicated that it was either fairly or very true that they were comfortable with university researchers using their *de-identified health information* (65.24%) and their *de-identified biosp*ecimens (65.73%). Comparatively, however, just over a quarter of respondents indicated that it was either fairly or

very true that they were comfortable with university researchers using their *identified health information* (28.36%) and their *identified biospecimens* (29.93%). The total comfort with researchers index score was 2.44 (SD=0.85), corresponding with an assessment between "somewhat" and "fairly true" that respondents were comfortable with researchers.

Table 4-3 Descriptive statistics for survey questions used in indices measuring Comfort with Researchers using patient health information and biospecimens

Table 4.3 Descriptive statistics for survey questions used in indices measuring Comfort with Researchers using patient health information and biospecimens (N=1841)		
	Frequency (% fairly or very true)	Mean (SD)
Comfort with Researchers Index		
I am comfortable with university researchers using my de-identified health information	65.24%	2.93 (1.06)
I am comfortable with university researchers using my identified health information	28.36%	1.93 (1.06)
I am comfortable with university researchers using my de-identified biospecimens	65.73%	2.91 (1.09)
I am comfortable with university researchers using my identified biospecimens	29.93%	1.99 (1.10)
Comfort with researchers index (Cronbach's α=0.791)	Median: 2.5	2.44 (0.85)

* Range: 1 = "not true"; 2 = "somewhat true"; 3 = "fairly true"; 4 = "very true"

Similar to comfort with researchers, over sixty percent of respondents indicated that it was either fairly or very true that they were comfortable with quality analysts using their *de-identified health information* (64.10%) and their *de-identified biosp*ecimens (60.89%). Just over a quarter of respondents indicated that it was either fairly or very true that they were comfortable with quality analysts using their *identified health information* (30.53%) and their *identified biospecimens* (28.68%). The total comfort with quality analysts index score was 2.42 (SD=0.86), also corresponding with an assessment between "somewhat" and "fairly true" that respondents were comfortable with quality analysts, again, similar to university researchers.

Table 4-4 Descriptive statistics for survey questions used in indices measuring Comfort with Quality Analysts using patient health information and biospecimens

Table 4.4 Descriptive statistics for survey questions used in indices measuring Comfort with Quality Analysts using patient health information and biospecimens (N=1841)		
	Frequency	

	(% fairly or very true)	Mean (SD)
Comfort with Quality Analysts Index		
I am comfortable with quality analysts using my de-identified health information	64.10%	2.89 (1.08)
I am comfortable with quality analysts using my identified health information	30.53%	2.00 (1.07)
I am comfortable with quality analysts using my de-identified biospecimens	60.89%	2.82 (1.10)
I am comfortable with quality analysts using my identified biospecimens	28.68%	1.96 (1.06)
Comfort with Quality Analysts index (Cronbach's α=0.809)	Median: 2.5	2.42 (0.86)

* Range: 1 = "not true"; 2 = "somewhat true"; 3 = "fairly true"; 4 = "very true"

Only 40% of respondents indicated that it was either fairly or very true that they were comfortable with commercial companies using their *de-identified health information* (40.47%) and their *de-identified biosp*ecimens (41.01%). Only 16.03% of respondents indicated that it was either fairly or very true that they were comfortable with commercial companies using their *identified health information* and only 17.11% of respondents indicated that it was either fairly or very true that they were comfortable with commercial companies using their *identified biospecimens* (28.68%). The total comfort with quality analysts index score was 1.95 (SD=0.84), corresponding with an assessment of "somewhat true" that respondents were comfortable with commercial companies.

Table 4-5 Descriptive statistics for survey questions used in indices measuring Comfort with Commercial Companies using patient health information and biospecimens

Table 4.5 Descriptive statistics for survey questions used in indices measuring Comfort with Commercial Companies using patient health information and biospecimens (N=1841)		
	Frequency (% fairly or very true)	Mean (SD)
Comfort with Commercial Companies Index		
I am comfortable with commercial companies using my de-identified health information	40.47%	2.30 (1.16)
I am comfortable with commercial companies using my identified health information	16.03%	1.59 (0.94)
I am comfortable with commercial companies using my de-identified biospecimens	41.01%	2.31 (1.15)
I am comfortable with commercial companies using my identified biospecimens	17.11%	1.62 (0.95)
Comfort with Commercial Companies index (Cronbach's α=0.813)	Median: 2.00	1.95 (0.84)

* Range: 1 = “not true”; 2 = “somewhat true”; 3 = “fairly true”; 4 = “very true”

4.4.4 Comfort with Law Enforcement

Comfort with law enforcement was assessed using two questions. Only 17.22% of respondents indicated that it was either fairly or very true that “it is okay for law enforcement to access health data” (mean 1.73, SD=0.88) while 37.26% of respondents responded “yes”, they were comfortable with law enforcement using genetic and ancestry data from sites like 23andMe and AncestryDNA. This latter question has been reverse-coded in the univariate and multivariable models that follow so that higher scores consistently indicate greater comfort.

Table 4-6 Descriptive statistics for comfort with law enforcement access to health data

Table 4.6 Descriptive statistics for comfort with law enforcement access to health data (N=1841)		
	Frequency (% fairly or very)	Mean (SD)
Comfort with law enforcement access to health data		
It is okay for law enforcement to access health information*	17.22%	1.73 (0.88)
Comfortable with law enforcement using genetic and ancestry data	37.26% (yes)	1.63 (0.48) Yes = 1 No = 2

* Range: 1 = “not true”; 2 = “somewhat true”; 3 = “fairly true”; 4 = “very true”

4.4.5 Confidence in existing laws and policies

Less than half of participants indicated they had confidence in existing health privacy laws and policies for both questions. Only 42.02% of respondents responded that it was “fairly or very true” that existing laws provided a reasonable level of protection for the privacy of patient information (mean 2.43, SD=0.90), and only 33.78% (mean 2.19, SD=0.92) responded that it was “fairly or very true” that they were confident that electronic health information was sufficiently protected by current laws and regulations.

Table 4-7 Descriptive statistics for confidence in current privacy law

Table 4.7 Descriptive statistics for confidence in current privacy law (N=1841)

	Frequency (% fairly or very)	Mean (SD)
Confidence in Current Privacy Law		
Existing laws provide a reasonable level of protection for the privacy of patient information	42.02%	2.43 (0.90)
I am confident that electronic health information is sufficiently protected by current law and regulation	33.78%	2.19 (0.92)

* Range: 1 = "not true"; 2 = "somewhat true"; 3 = "fairly true"; 4 = "very true"

4.4.6 Desire for control and notification

The majority of respondents indicated their desire for both control over and notification of the use of their health data. Over 60% of respondents (63.17%) stated that it was either fairly or very true that they should have more control over how their health information is used (mean 2.93, SD=0.98). Desire for notification was even greater with 80.12% of respondents indicating that it was either fairly or very true that it was important they know who has health information about themselves (mean 3.38, SD=0.87) and 84.31% indicating that it was fairly or very true that they should be able to find out how their health information was shared (mean 3.51, SD=0.82).

Table 4-8 Descriptive Statistics for desire for control and notification of health information sharing

Table 4.8 Descriptive Statistics for desire for control and notification of health information sharing (N=1841)		
	Frequency (% fairly or very)	Mean (SD)
Desire for control and notification of health information sharing		
I should have more control over how my health information is used	63.17%	2.93 (0.98)
It is important I know who has health information about me	80.12%	3.38 (0.87)
I should be able to find out how my health information is shared	84.31%	3.51 (0.82)

* Range: 1 = "not true"; 2 = "somewhat true"; 3 = "fairly true"; 4 = "very true"

4.4.7 Univariate linear regression

Univariate examination of comfort sharing health data with third-party commercial companies for patient and business purposes show statistically significant relationships with comfort with researchers ($patient\ purposes\ b^* = 0.517, p = 7.9 * 10^{-67}$, $business\ purposes\ b^* = 0.520, p = 2.0 * 10^{-67}$), quality analysts ($patient\ purposes\ b^* = 0.505, p = 3.9 * 10^{-56}$, $business\ purposes\ b^* = 0.533, p = 4.2 * 10^{-59}$), and commercial companies ($patient\ purposes\ b^* = 0.479, p = 1.6 * 10^{-52}$, $business\ purposes\ b^* = 0.589, p = 8.9 * 10^{-81}$). Comfort with law enforcement access to health data also showed significant associations with comfort sharing health data with third-party commercial companies for both access to and use of genetic data provided by 23andMe and AncestryDNA ($patient\ purposes\ b^* = 0.209, p = 4.7 * 10^{-11}$, $business\ purposes\ b^* = 0.249, p = 6.4 * 10^{-14}$) and access to health information more broadly ($patient\ purposes\ b^* = 0.118, p = 0.001$, $business\ purposes\ b^* = 0.253, p = 1.7 * 10^{-11}$).

Confidence in existing laws and policies was also significantly associated with comfort with commercial companies. Belief that existing laws provide a reasonable level of protection associated with greater comfort sharing health data with third-party commercial companies ($patient\ purposes\ b^* = 0.318, p = 4.4 * 10^{-23}$, $business\ purposes\ b^* = 0.338, p = 1.7 * 10^{-25}$), as did confidence that electronic health information was sufficiently protected ($patient\ purposes\ b^* = 0.273, p = 6.3 * 10^{-16}$, $business\ purposes\ b^* = 0.372, p = 9.1 * 10^{-28}$).

Desire for control over and notification of the use of health data displayed strong associations with respondent comfort with sharing health data with third-party commercial companies for both patient and business purposes, but for business purposes especially. Desire for more control over health information generally ("I should have more control over how my health information is used") was significantly associated with comfort with sharing with sharing health data with commercial companies for both patient purposes and business purposes ($patient\ purposes\ b^* = -0.126, p = 0.00014$, $business\ purposes\ b^* = -0.162, p = 2.5 * 10^{-06}$). Desire to know who has their health information ($business\ purposes\ b^* = -0.211, p = 2.9 * 10^{-10}$) and how that information is shared ($business\ purposes\ b^* =$

$-0.173, p = 4.2 * 10^{-07}$) was strongly associated with comfort with sharing health data with third-party commercial companies for business purposes only.

Table 4-9 Univariate associations for comfort (with researchers, quality analysts, commercial companies, and law enforcement), confidence in existing privacy laws and protections, and desire for control and notification of health data sharing

Table 4.9 Univariate associations for comfort (with researchers, quality analysts, commercial companies, and law enforcement), confidence in existing privacy laws and protections, and desire for control and notification of health data sharing with comfort sharing health data with third-party commercial companies for patient purposes and business purposes (N=1841)							
	Patient Purposes (univariate)				**Business Purposes (univariate)**		
	b*	p-value	R^2		b*	p-value	R^2
Comfort							
Comfort with Researchers	0.517	$7.9*10^{-67}$	0.268		0.520	$2.0*10^{-67}$	0.271
Comfort with Quality Analysts	0.505	$3.9e*10^{-56}$	0.255		0.533	$4.2*10^{-59}$	0.284
Comfort with Commercial Companies	0.479	$1.6*10^{-52}$	0.229		0.589	$8.9*10^{-81}$	0.347
Comfort with Law Enforcement Access							
It is okay for law enforcement to access health information	0.118	0.001	0.014		0.253	$1.7*10^{-11}$	0.064
Comfortable with law enforcement using genetic and ancestry data	**0.209**	**$4.7*10^{-11}$**	**0.044**		**0.249**	**$6.4*10^{-14}$**	**0.062**
Confidence in Current Privacy Law							
Existing laws provide a reasonable level of protection	**0.318**	**$4.4*10^{-23}$**	**0.101**		**0.338**	**$1.7*10^{-25}$**	**0.115**
I am confident that electronic health information is sufficiently protected	**0.273**	**$6.3*10^{-16}$**	**0.075**		**0.372**	**$9.1*10^{-28}$**	**0.138**
Desire for control and notification of health information sharing							
I should have more control over how my health information is used	**-0.126**	**0.00014**	**0.016**		**-0.162**	**$2.5*10^{-06}$**	**0.026**
It is important I know who has health information about me	**-0.073**	**0.029**	**0.005**		**-0.211**	**$2.9*10^{-10}$**	**0.045**
I should be able to find out how my health information is shared	0.028	0.40	0.001		**-0.173**	**$4.2*10^{-07}$**	**0.030**

b* = standardized beta

4.4.8 Stepwise regression model

In the Bonferroni-corrected stepwise regression model, both comfort with researchers and quality analysts remained in the final model, accounting for 25% and 13% of variability in comfort with sharing health data with third-party commercial companies for patient and business purposes, respectively. Comfort with law enforcement accessing genetic data also remained in this model for both patient purposes ($b^* = 0.102, p = 5.8 * 10^{-05}$) and business purposes ($b^* = 0.095, p = 7.6 * 10^{-05}$). Desire to find out how health information is shared ("I should be able to find out how my health information is shared") accounted for 12% of the variability in comfort with sharing health data with third-party commercial companies for patient purposes ($b^* = 0.118, p = 0.00064$) and 10% of the variability in comfort with sharing health data with third-party commercial companies for business purposes ($b^* = -0.096, p = 1.6 * 10^{-05}$).

As was mentioned in the methods section of this chapter, demographic variables and attitude measures from previous chapters were included in this final, complete model. Amongst the demographic variables explored in Chapter 2, education remained a significant variable in this final model of comfort with sharing health data with third-party commercial companies for patient purposes (high school: $b^* = 0.148, p = 0.007$; some college: $b^* = 0.177, p = 0.00041$; BA or above: $b^* = 0.252, p = 3.1 * 10^{06}$) and unemployment remained a significant variable for comfort with sharing health data with third-party commercial companies for business purposes (not employed: $b^* = 0.108, p = 0.0002$). Trust in the health system (system trust), which was discussed in Chapter 3, accounted for 24% of variability in the final model of comfort with sharing health data with third-party commercial companies for patient purposes ($b^* = 0.235, p = 2.6 * 10^{-12}$) and 16% of the variability in comfort sharing health data with third-party commercial companies for business purposes ($b^* = 0.159, p = 3.0 * 10^{-08}$).

Concern about the Sloan Kettering Paige.AI event, originally discussed in Chapter 3, remained in this final model as well, accounting for 8% of variability in comfort sharing health data with third-party commercial companies for business purposes only ($b^* = -0.079, p = 0.0018$).

Table 4-10 Stepwise regression modeling of predictors of comfort sharing health data with third-party commercial companies for patient and business purposes (full model)

Table 4.10 Stepwise regression modeling of predictors of comfort sharing health data with third-party commercial companies for patient purposes and comfort sharing health data with third-party commercial companies for business purposes (N=1841)					
	Patient Purposes Multivariable stepwise Bonferroni corrected ($\alpha = 0.002$)			**Business Purposes** Multivariable stepwise Bonferroni corrected ($\alpha = 0.002$)	
	Model R^2	0.417		Model R^2	0.473
	b*	p-value		b*	p-value
Comfort					
Comfort with Researchers	**0.250**	**$6.1*10^{-12}$**		**0.134**	**0.0023**
Comfort with Quality Analysts				**0.101**	**0.023**
Comfort with Commercial Companies	**0.207**	**$4.0*10^{-08}$**		**0.301**	**$5.1*10^{-15}$**
Comfort with Law Enforcement					
It is okay for law enforcement to access health information	**-0.055**	**0.037**			
Comfortable with law enforcement using genetic and ancestry data	**0.102**	**$5.8*10^{-05}$**		**0.095**	**$7.6*10^{-05}$**
Confidence in Current Privacy Law					
Existing laws provide a reasonable level of protection	**0.136**	**0.0003**		**0.102**	**0.00012**
I am confident that electronic health information is sufficiently protected	-0.067	0.082			
Desire for control and notification of information sharing					
It is important I know who has health information about me	**-0.094**	**0.012**			
I should be able to find out how my health information is shared	**0.118**	**0.00064**		**-0.096**	**$1.6*10^{-05}$**
Concern about recent events					
Concern about Sloan Kettering startup and conflict of interest				**-0.079**	**0.0018**
Attitudes					
Trust					
System Trust	**0.235**	**$2.6*10^{-12}$**		**0.159**	**$3.0*10^{-08}$**
Altruism					
Altruism	**0.062**	**0.036**			
Demographics					
Age					
18-29	ref				
30-44	-0.046	0.26			
45-59	**-0.098**	**0.023**			
60+	-0.080	0.058			

Education						
	Less than High School	ref				
	High School	**0.148**	**0.007**			
	Some college	**0.177**	**0.00041**			
	BA or above	**0.252**	**$3.1*10^{-06}$**			
Income						
	Less than $60,000	ref				
	$60,000 or greater	0.045	0.088			
Employment						
	Employed				ref	
	Not employed				**0.108**	**0.0002**
	Retired				-0.013	0.56
	Disabled/Other				-0.034	0.19

b* = standardized beta

4.5 Discussion

To better understand the public's comfort with sharing health data with third-party commercial companies, this study sought to examine differences in comfort sharing health data for patient purposes and comfort sharing health data for business purposes. Our analysis concludes with three main findings. First, trust in the health system had a greater impact on comfort with sharing health data with third-party commercial companies for patient or business purposes did than any other variable—including privacy concerns. Second, respondent's confidence in existing laws and policies was significantly associated with comfort with sharing health data. Third, desire for notification of health data sharing events was strongly associated with comfort with sharing health data with third-party commercial companies for both patient and business purposes and persisted in the multivariable model. Desire for control over how health data is shared, although significant in the univariate analysis, did not persist in the final multivariable model.

That trust in the health system showed a stronger association with comfort with sharing health data with third-party commercial companies, comports with research that positions trust as an antecedent to privacy concerns (Bijlsma & Koopman, 2003; Waldman, 2016). In a study of trust in the workplace, the "higher the level of trust, the lower the costs of monitoring and other control mechanisms" (Bijlsma & Koopman, 2003). In the context of health data sharing, patients with more trust in the health system will be less concerned about the privacy and security of their

health data. That trust in the health system has so great of an effect on comfort with third-party commercial companies reinforces the numerous downstream effects of trust on the healthcare experience – patients with a high degree of trust in the health system are more likely to adhere to their treatment regimen (Sweeney, 2018), perceive their care as high quality (Cunningham, 2009), and, based on the results of this study, confer the benefits of that trust onto third-party partners outside of the health system.

Respondent comfort with sharing health information with third-party commercial companies was strongly associated not only with respondent comfort with commercial companies having access to both identified and de-identified data, but was also strongly associated with respondent's comfort with both researchers and quality analysts having access to identified and de-identified data. Only 40% of participants reported being fairly or very comfortable sharing de-identified health information with commercial companies, while 60-65% of participants were fairly or very comfortable sharing de-identified health information with researchers and quality analysts. As the nature of healthcare research shifts to a greater number of partnerships with commercial companies, patients' willingness to share health information with researchers may come to resemble their reluctance to share their health information with third-party commercial companies, leading to even greater withholding and compromised data sets. Although we cannot conclude from this study whether the Sloan Kettering event had an impact on subsequent comfort with sharing health information with third-party commercial companies, or if those who were already least comfortable with third-party commercial companies indicated that lack of comfort in their concern about the Sloan Kettering/Paige.AI event. That this event persisted in the final model leads one to speculate that events such as these may have a lasting effect on patients' willingness to share their health information with commercial companies, and, by extension, health system researchers and quality analysts (table 2.1.5: "65.07% of respondents were fairly or very concerned about Sloan Kettering hospital executives using hospital data for their own startup company").

Confidence in existing laws that protect health information is important to assuaging patient concerns about the use of personal health information. As advances in medical technologies outpace the ability of regulators to respond, it is increasingly difficult for patients to

be confident that their health data is protected, as well as increasingly difficult for health systems to balance the demand for innovation with a changing legal landscape out of sync with patient expectations. That only 42.02% of respondents thought that it was fairly or very true that "existing laws provide a reasonable level of protection for the privacy of patient information" underscores the inadequacy of current regulation and/ or that patients are unaware, and if they find out may be unpleasantly surprised.

One of the most significant findings in this study was that *desire for notification* remained strongly associated with comfort with sharing health information with third-party commercial companies in the final multivariable model, while *desire for control over health data* did not. Privacy management is both context dependent and individual specific (Nissenbaum, 2009) and efforts to provide privacy controls in healthcare are likely to be met with criticism by the more privacy-minded public, and apathy by others who may have no patience for managing and tuning their individual privacy preferences (Guo & Chen, 2012). Notification, however, while providing patients with information about where their data is being shared, can also provide the transparency necessary to enable greater trust in health data partnerships and confidence that the health system is acting with the patient's own best interests in mind.

4.5.1 Implications for research

Healthcare research is rapidly becoming dependent on the large data sets provided by electronic personal health information (ePHI), and increasingly partnering with third-party commercial companies to leverage their data analysis capabilities and computing power. As it currently stands, patients are by and large willing to share their personal health information for research purposes. As research efforts become increasingly involved with commercial companies, it is possible that that the high degree of comfort with researchers reported in this study (65.24% of respondents were fairly or very comfortable with university researchers using their de-identified health information) may gradually shift towards the lower degree of comfort respondents reported having with commercial companies (41.01% of respondents were fairly or very comfortable with commercial companies using their de-identified health information). In the face of diminishing trust in the health system, involvement of patient stakeholders at the

inception of these partnerships is critical in order to provide transparency about the nature of data sharing and the potential benefits to patient care, and in so doing, mitigate fears of data abuse.

4.5.2 Implications for policy and practice

Federal adoption of a comprehensive law governing medical privacy is likely to go much further than the efforts of any one health system to improve trust. As previously stated, HIPAA is unable to sufficiently protect patient privacy in the era of big data. Adding to the difficulty inherent in regulatory management of both de-identified and identified health data, each state has their own laws in addition to HIPAA—but most do not have their own comprehensive laws governing medical privacy. Two states have added wide-reaching obligations to protect patient confidentiality: California and Texas. The California Medical Information Act (CMIA) explicitly covers mobile health apps, software, and hardware designed to maintain medical information. Failure to obtain authorization from the patient or violation of patient privacy can lead to penalties including civil lawsuits, administrative fines, or civil penalties (Solove & Schwartz, 2019). The Texas Medical Privacy Act is, as of this writing, the broadest and strictest medical privacy laws in the United States, in short, expanding application of HIPAA to not only covered entities, but any organization that collects and stores protected health information (Solove & Schwartz, 2019), and prohibiting the re-identification of deidentified data under any circumstance (Luna, 2011). Adoption of either of these two standards would raise the floor of medical privacy protections and give both providers and patients a clear standard to which they must adhere.

4.5.3 Implications for design

Health systems navigating consumer-facing data for the first time may glean insight on how to handle designing for trust and privacy from the work done by researchers of electronic commerce. In their research on trustworthiness of electronic commerce, Belanger et al. (2002) concluded that winning public trust was the primary hurdle to e-commerce growth, and, based on their research and the research of others, offer six web design features were key to increasing consumer's ratings of trustworthiness: 1) safeguard assurances, 2) company reputation, 3) ease of navigation, 4) robust order fulfillment, 5) professionalism of the website, and 6) use of state-of-

the-art web design (Belanger et al., 2002). While these features were designed for online commerce sites, very similar criteria can be applied to health systems. Application to health systems can mean 1) providing "plain language" explanation of technical and regulatory privacy protections to patients, as well as the consequences of violation; 2) maintenance of reputation through transparency – companies have been shown to recover trust more quickly in the wake of a data breaches if customers are notified of the breach in a timely manner (Bansal & Zahedi, 2015); 3) designing electronic health information portals and notification dashboards to be easy to navigate for patients; 4) responding quickly to patient inquiries about their data; 5) maintaining professional websites; 6) using state-of-the art web design. In future research we will conduct more in-depth research on these design features as they apply to healthcare.

4.5.4 Limitations

As with any survey, this study is merely a snapshot of patient beliefs and preferences, limited due to the nature of survey questions – other demographic variables, attitude measures, and questions exploring comfort with data sharing generally that may provide a more complete portrait of the public's comfort with sharing health data with third-party commercial companies may not be captured here. Additionally, a stepwise regression model is a conservative model that eliminates factors that might be important to understanding patient and public comfort with sharing health data with third-party commercial companies.

4.5.5 Conclusion

The results of this study, as well as the results of the chapter that precedes it, underscore the importance of trust in the health system as healthcare systems navigating increasing numbers of partnerships with third-party commercial companies and patients' comfort with sharing health data with those third-party commercial companies. This study also provided more insight into the relationship between comfort with third-party commercial companies and notification of data use versus control over data use. Healthcare systems can begin patient privacy efforts by focusing not on data control, which is fraught with a myriad number of technological issues, but on data use notification. That confidence in existing laws remained significantly associated with comfort with sharing health data with third-party commercial companies adds to the growing body of

research and commentary that states current laws and regulations are inadequate for our current data landscape and demand revision.

4.6 Acknowledgements

Research reported in this manuscript was supported by the National Cancer Institute of the National Institutes of Health under award number 5 R01 CA214829-03.

4.7 References

Anderson, C., & Agarwal, R. (2011). The Digitization of Healthcare: Boundary Risks, Emotion, and Consumer Willingness to Disclose Personal Health Information. *Information Systems Research*, *22*(3), 469–490. https://doi.org/10.1287/isre.1100.0335

Bansal, G., & Zahedi, F. M. (2015). Trust violation and repair: The information privacy perspective. *Decision Support Systems*, *71*, 62–77. https://doi.org/10.1016/j.dss.2015.01.009

Batbaatar, E., Dorjdagva, J., Luvsannyam, A., Savino, M. M., & Amenta, P. (2017). Determinants of patient satisfaction: A systematic review. *Perspectives in Public Health*, *137*(2), 89–101. https://doi.org/10.1177/1757913916634136

Belanger, F., Hiller, J. S., & Smith, W. J. (2002). Trustworthiness in electronic commerce: The role of privacy, security, and site attributes. *The Journal of Strategic Information Systems*, *11*(3), 245–270. https://doi.org/10.1016/S0963-8687(02)00018-5

Bijlsma, K., & Koopman, P. (2003). Introduction: Trust within organisations. *Personnel Review*, *32*(5), 543–555. https://doi.org/10.1108/00483480310488324

Bishop, L. "Sam," Holmes, B., & Kelley, C. (2005). National Consumer Health Privacy Survey 2005. *California Health Care Foundation*. https://www.chcf.org/publication/national-consumer-health-privacy-survey-2005/

Black Book Market Research. (2017, January 3). *Healthcare's Digital Divide Widens, Black Book Consumer Survey*. Black Book Market Research. https://blackbookmarketresearch.newswire.com/news/healthcares-digital-divide-widens-black-book-consumer-survey-18432252

Boulware, L. E., Cooper, L. A., Ratner, L. E., LaVeist, T. A., & Powe, N. R. (2003). Race and Trust in the Health Care System. *Public Health Reports*, *118*, 8.

Business Wire. (2019, June 25). *The $11.9 Trillion Global Healthcare Market: Key Opportunities & Strategies (2014-2022) - ResearchAndMarkets.com*. Businesswire. https://www.businesswire.com/news/home/20190625005862/en/11.9-Trillion-Global-Healthcare-Market-Key-Opportunities

Cohen, I. G., & Mello, M. M. (2018). HIPAA and Protecting Health Information in the 21st Century. *JAMA*, *320*(3), 231–232. https://doi.org/10.1001/jama.2018.5630

Cunningham, P. J. (2009). High Medical Cost Burdens, Patient Trust, and Perceived Quality of Care. *Journal of General Internal Medicine*, *24*(3), 415–420. https://doi.org/10.1007/s11606-008-0879-3

Damschroder, L. J., Pritts, J. L., Neblo, M. A., Kalarickal, R. J., Creswell, J. W., & Hayward, R. A. (2007). Patients, privacy and trust: Patients' willingness to allow researchers to access their medical records. *Social Science & Medicine*, *64*(1), 223–235. https://doi.org/10.1016/j.socscimed.2006.08.045

Davis, J. (2019, April 15). *Third-Party Vendors Behind 20% of Healthcare Data Breaches in 2018*. HealthITSecurity. https://healthitsecurity.com/news/third-party-vendors-behind-20-of-healthcare-data-breaches-in-2018

Feldman, S., & Steenbergen, M. R. (2001). The Humanitarian Foundation of Public Support for Social Welfare. *American Journal of Political Science*, *45*(3), 658–677. JSTOR. https://doi.org/10.2307/2669244

Fussell, S. (2020, January 8). *The Sneaky Genius of Facebook's New Preventive Health Tool*. The Atlantic. https://www.theatlantic.com/technology/archive/2020/01/facebook-launches-new-preventative-health-tool/604567/

Gellman, R. (2011). The Deidentification Dilemma: A Legislative and Contractual Proposal. *Fordham Intellectual Property, Media and Entertainment Law Journal*, *21*(1), 31.

Gooch, K. (2017, January 3). *Privacy issues drive health IT consumer skepticism: 10 Black Book survey findings*. Becker's Health IT. https://www.beckershospitalreview.com/healthcare-information-technology/privacy-issues-drive-health-it-consumer-skepticism-10-black-book-survey-findings.html

Guo, S., & Chen, K. (2012). Mining Privacy Settings to Find Optimal Privacy-Utility Tradeoffs for Social Network Services. *2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Confernece on Social Computing*, 656–665. https://doi.org/10.1109/SocialCom-PASSAT.2012.22

Guzman, G. G. (2019). American Community Survey Briefs—Household Income: 2018. *United States Census Bureau*, 13.

Hall, M. A., Camacho, F., Dugan, E., & Balkrishnan, R. (2002). Trust in the Medical Profession: Conceptual and Measurement Issues: Trust in the Medical Profession: Conceptual and Measurement Issues. *Health Services Research*, *37*(5), 1419–1439. https://doi.org/10.1111/1475-6773.01070

Hall, M. A., Dugan, E., Zheng, B., & Mishra, A. K. (2001). Trust in Physicians and Medical Institutions: What Is It, Can It Be Measured, and Does It Matter? *The Milbank Quarterly*, *79*(4), 613–639. https://doi.org/10.1111/1468-0009.00223

Hall, M. A., Zheng, B., Dugan, E., Camacho, F., Kidd, K. E., Mishra, A., & Balkrishnan, R. (2002). Measuring Patients' Trust in their Primary Care Providers. *Medical Care Research and Review*, *59*(3), 293–318. https://doi.org/10.1177/1077558702059003004

HHS Press Release. (2020, March 6). *HHS Finalizes Historic Rules to Provide Patients More Control of Their Health Data* [Text]. HHS.Gov. https://www.hhs.gov/about/news/2020/03/09/hhs-finalizes-historic-rules-to-provide-patients-more-control-of-their-health-data.html

HIPAA Journal. (2017, January 5). Patients Holding Back Health Information Over Data Privacy Fears. *HIPAA Journal*. https://www.hipaajournal.com/patients-holding-back-health-information-over-fears-of-data-privacy-8634/

Hoffman, S. (2020). Citizen Science: The Law and Ethics of Public Access to Medical Big Data. *Berkeley Tech. LJ*, *30*(3), 66.

Jilka, S. R., Callahan, R., Sevdalis, N., Mayer, E. K., & Darzi, A. (2015). "Nothing About Me Without Me": An Interpretative Review of Patient Accessible Electronic Health Records. *Journal of Medical Internet Research*, *17*(6), e161. https://doi.org/10.2196/jmir.4446

Karampela, M., Ouhbi, S., & Isomursu, M. (2019). Connected Health User Willingness to Share Personal Health Data: Questionnaire Study. *Journal of Medical Internet Research*, *21*(11). https://doi.org/10.2196/14537

Kim, J., Kim, H., Bell, E., Bath, T., Paul, P., Pham, A., Jiang, X., Zheng, K., & Ohno-Machado, L. (2019). Patient Perspectives About Decisions to Share Medical Data and

Biospecimens for Research. *JAMA Network Open*, *2*(8), e199550–e199550. https://doi.org/10.1001/jamanetworkopen.2019.9550

LaVeist, T. A., Isaac, L. A., & Williams, K. P. (2009). Mistrust of Health Care Organizations Is Associated with Underutilization of Health Services. *Health Services Research*, *44*(6), 2093–2105. https://doi.org/10.1111/j.1475-6773.2009.01017.x

Luna, J. (2011). *Texas Medical Privacy Act, Health Law & Policy Institute*. University of Houston Law Center. https://www.law.uh.edu/healthlaw/perspectives/Privacy/010830Texas.html

Merken, S., & Elfin, D. (2018, October 31). *What's Your Health Data Worth? Startups Want to Help You Sell It*. Bloomberg Law. https://news.bloomberglaw.com/tech-and-telecom-law/whats-your-health-data-worth-startups-want-to-help-you-sell-it

Nguyen, G. C., LaVeist, T. A., Harris, M. L., Datta, L. W., Bayless, T. M., & Brant, S. R. (2009). Patient Trust-in-Physician and Race Are Predictors of Adherence to Medical Management in Inflammatory Bowel Disease. *Inflammatory Bowel Diseases*, *15*(8), 1233–1239. https://doi.org/10.1002/ibd.20883

Nissenbaum, H. (2009). *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford University Press.

Office for Civil Rights (OCR). (2013). *190-Who must comply with HIPAA privacy standards* [Text]. HHS.Gov. https://www.hhs.gov/hipaa/for-professionals/faq/190/who-must-comply-with-hipaa-privacy-standards/index.html

Ousfar, E. (2019, October 29). *Facebook announces new health tool that urges users to get preventive care*. News Center Maine. https://www.newscentermaine.com/article/news/health/facebook-announces-new-health-

tool-that-urges-its-users-to-get-preventive-care/97-93a45c7f-03e6-47d2-b608-60732a2f64d8

Perrin, A. (2020, February 4). About half of Americans are OK with DNA testing companies sharing user data with law enforcement. *Pew Research Center*. https://www.pewresearch.org/fact-tank/2020/02/04/about-half-of-americans-are-ok-with-dna-testing-companies-sharing-user-data-with-law-enforcement/

Platt, J. E., Jacobson, P. D., & Kardia, S. L. R. (2018). Public Trust in Health Information Sharing: A Measure of System Trust. *Health Services Research*, *53*(2), 824–845. https://doi.org/10.1111/1475-6773.12654

Promarket. (2020, February 7). The Real Price of Health Data: Americans Don't Want to Share Their Records for Free. *Pro Market*. https://promarket.org/2020/02/07/the-real-price-of-health-data-americans-dont-want-to-share-their-records-for-free/

Rushton, G., Armstrong, M. P., Gittler, J., Greene, B. R., Pavlik, C. E., West, M. M., & Zimmerman, D. L. (2006). Geocoding in cancer research: A review. *American Journal of Preventive Medicine*, *30*(2 Suppl), S16-24. https://doi.org/10.1016/j.amepre.2005.09.011

Schwarzer, R., & Jerusalem, M. (1992). General Self-Efficacy- Schwarzer (GSES). *Statistics Solutions*. https://www.statisticssolutions.com/general-self-efficacy-schwarzer-gses/

Seltzer, E., Goldshear, J., Guntuku, S. C., Grande, D., Asch, D. A., Klinger, E. V., & Merchant, R. M. (2019). Patients' willingness to share digital health and non-health data for research: A cross-sectional study. *BMC Medical Informatics and Decision Making*, *19*(1), 157. https://doi.org/10.1186/s12911-019-0886-9

Shen, N., Bernier, T., Sequeira, L., Strauss, J., Silver, M. P., Carter-Langford, A., & Wiljer, D. (2019). Understanding the patient privacy perspective on health information exchange: A

systematic review. *International Journal of Medical Informatics*, *125*, 1–12. https://doi.org/10.1016/j.ijmedinf.2019.01.014

Singer, N. (2019, September 3). When Apps Get Your Medical Data, Your Privacy May Go With It. *The New York Times*. https://www.nytimes.com/2019/09/03/technology/smartphone-medical-records.html

Smith, C. (2011). Somebody's Watching Me: Protecting Patient Privacy in Prescription Health Information Constitutional Constraints on State Health Care & Privacy Regulation after Sorrell v. IMS Health. *Vermont Law Review*, *36*(4), 931–994.

Smith, T., Davern, M., Freese, J., & Morgan, S. (2019). General Social Surveys, 1972-2018. *NORC*, 11.

Solove, D. J., & Schwartz, P. M. (2019). *Privacy Law Fundamentals* (SSRN Scholarly Paper ID 1790262). Social Science Research Network. https://papers.ssrn.com/abstract=1790262

Spencer, K., Sanders, C., Whitley, E. A., Lund, D., Kaye, J., & Dixon, W. G. (2016). Patient Perspectives on Sharing Anonymized Personal Health Data Using a Digital System for Dynamic Consent and Research Feedback: A Qualitative Study. *Journal of Medical Internet Research*, *18*(4), e66. https://doi.org/10.2196/jmir.5011

Teixeira, P. A., Gordon, P., Camhi, E., & Bakken, S. (2011). HIV patients' willingness to share personal health information electronically. *Patient Education and Counseling*, *84*(2), e9–e12. https://doi.org/10.1016/j.pec.2010.07.013

Tiller, J. (2019, November 12). *If you've given your DNA to a DNA database, US police may now have access to it*. The Conversation. http://theconversation.com/if-youve-given-your-dna-to-a-dna-database-us-police-may-now-have-access-to-it-126680

US Census Bureau. (2019). *Health Insurance Coverage in the United States: 2018*. The United States Census Bureau. https://www.census.gov/library/publications/2019/demo/p60-267.html

Waldman, A. E. (2016). Privacy, sharing, and trust: The Facebook study. *Case Western Reserve Law Review*, *67*(1), 193-. Academic OneFile.

Weitzman, E. R., Kaci, L., & Mandl, K. D. (2010). Sharing Medical Data for Health Research: The Early Personal Health Record Experience. *Journal of Medical Internet Research*, *12*(2). https://doi.org/10.2196/jmir.1356

Willison, D. J., Steeves, V., Charles, C., Schwartz, L., Ranford, J., Agarwal, G., Cheng, J., & Thabane, L. (2009). Consent for use of personal information for health research: Do people with potentially stigmatizing health conditions and the general public differ in their opinions? *BMC Medical Ethics*, *10*(1), 10. https://doi.org/10.1186/1472-6939-10-10

Chapter 5 Conclusion

5.1 Summary of findings

Health systems and patients are navigating new challenges to privacy and trust as the number of third-party commercial companies operating in the healthcare space continues to increase. My dissertation research sought to understand the public's comfort with sharing health data with third party commercial companies and explore the issues and concerns healthcare systems must anticipate. Anticipation of these concerns, proactive engagement with patients, and protection of patient privacy in the absence of regulatory guidance can mitigate the backlash suffered by other health systems at the announcement of their third-party commercial partnerships.

In the analysis comparing the public's comfort with sharing health data with third-party commercial companies for *patient purposes* with the public's comfort with sharing health data with third-party commercial companies for *business purposes*, a statistically significant difference emerged between the two purposes. Respondents reported greater comfort with sharing health data with third-party commercial companies for patient purposes.

Factors **positively** associated with the public's comfort with sharing health data with third party commercial companies for patient purposes were: having some college education or a college degree, system trust, provider trust, comfort with researchers and commercial companies, comfort with law enforcement accessing genetic information, confidence in existing laws, desire to know *how* health information is shared, and altruism. Factors positively associated with the public's comfort with sharing health data with third-party commercial companies for business purposes were system trust, provider trust, employment status, comfort with researchers, quality analysts, and commercial companies, comfort with law enforcement accessing genetic information, and confidence in existing laws. Factors **negatively** associated with the public's comfort with sharing health data with third-party commercial companies for patient purposes

were respondent age between 45-59, privacy concerns, comfort with law enforcement accessing health information, desire to know *with whom* health information is shared. Factors negatively associated with the public's comfort with sharing health data with third-party commercial companies for business purposes were privacy concerns, concern about Memorial Sloan Kettering's start-up company, Paige.AI, and desire to know *how* health information is shared.

Amongst these significant variables, three emerged as the greatest predictors of comfort with sharing health data with third-party commercial companies for both patient and business purposes: trust in the health system, confidence in existing laws, and desire for notification. Trust in the health system displayed the strongest association with comfort with sharing health data with third-party commercial companies. Privacy concern, although strongly associated with comfort with third-party commercial companies in Chapters 2 and 3 of this dissertation, did not persist in the final model presented in Chapter 4. As was discussed in Chapter 4, legal scholarship in privacy and research in consumer behavior have positioned trust as an antecedent to privacy concern – when trust between parties is high, the need for mechanisms of control (laws, regulation, monitoring) decreases. This finding also indicates that regardless of the privacy controls provided to consumers or to patients, comfort with the third-party commercial partners of health systems will remain low if trust is not secured. The public's low confidence in current health data privacy laws and regulations is likely also a product of diminished trust in the health system. Only 42.02% of respondents felt that "existing laws provided a reasonable level of protection for the privacy of patient information" and only 33.78% were "confident that electronic health information is sufficiently protected by current law and regulation". Desire for notification remained significant in the multivariable model while desire for control over how health information is used did not. Research on usage rates of available privacy controls in social media might provide insight into this.

Education persisted in the final model but displayed a relationship with comfort with sharing health data with third-party commercial companies for patient purposes that is contrary to existing research (Blank et al., 2014; Kim et al., 2019; Sheehan, 1999). These studies have found that greater educational attainment is associated with decreased comfort with sharing health data with third-party commercial companies. In this analysis, we found the opposite was

true when the purpose and intent of the health data sharing partnership was differentiated. Educational attainment, which includes having some college experience, was negatively associated with greater comfort with sharing health data with third-party commercial companies for business purposes, but was positively associated with greater comfort with sharing health data with third-party commercial companies for patient purposes.

Respondent age between 45-59 also remained in the final model of the public's comfort with sharing health data with third-party commercial companies for patient purposes. The negative association seen here potentially reflects a product of what has been referred to as the "grey" digital divide – individuals who came of age in the time before the explosion of the internet and digital technologies use the internet and digital technologies at lower rates than digital natives (Zhou & Salvendy, 2017) and are thus less comfortable with online transactions and are also the age most concerned about online fraud (Milewski, 2016).

Perceived healthcare access displayed strong association with the public's comfort with sharing health data for patient purposes only, but this association did not persist in the multivariable regression conducted in chapter 3 (table 3.10). The reasons for this are likely the introduction of measures evaluating trust in the health system (system trust) and measures evaluating altruism. That perceived healthcare access did not persist in the multivariable model at the introduction of the latter two is consistent with research on trust in healthcare systems. Trust in the health system is high when communication between patient and provider is positive and patients feel able to access sufficient health services reliably (Thiede, 2005).

Notably, while past experience of a data breach and concern about recent data breach events were significantly associated with comfort with sharing health data with third-party commercial companies for business purposes in the univariate analysis, these variables did not remain in the final multivariable model (chapter 3). As discussed in chapter 3, these results suggest that although the number of data breach events in other sectors that magnify the salience of privacy issues is increasing, their effect on the public's concerns about health data is limited. That concern about the Memorial Sloan Kettering Paige.AI startup did persist in the final model underscores this distinction. Health data violations, real or perceived, may in turn affect other

health systems, diminishing trust and increasing privacy concerns for all patients in all health systems. As the number of healthcare partnerships with commercial companies and the number of commercial entrants into the healthcare market continue to increase, these results suggest that failure to address privacy and health data protection in the ways most meaningful to patients will only accelerate the deterioration of trust in the health system.

Comfort with researchers, quality analysts, and commercial companies was positively associated with comfort with sharing health data with third-party commercial companies. This intuitive finding confirms consistency with the measures created to evaluate the public's comfort with sharing health data with third-party commercial companies for patient and business purposes. Examination of the descriptive statistics of comfort with commercial companies having access to both identified and de-identified data shows low rate of comfort when compared to researchers and quality analysts. In chapter 4 I argue that the increasing number of commercial partnerships might create drag on the relatively high rate of comfort the public current has with both researchers and quality analysts.

Examination of the measures used to evaluate the public's comfort with law enforcement displays an interesting contrast: only 17.22% of respondents felt it was fairly or very true that it was "okay for law enforcement to access health information" but 37.26% of these same respondents were fairly or very comfortable with law enforcement using genetic and ancestry data. Further research is needed to illuminate this difference, but it may be a function of ubiquitous advertising and normalization of at-home genetic tests by 23andMe and Ancestry.com. By the start of 2019, more than 26 million consumers had added genetic information to a commercial ancestry database (Regalado, 2019).

From these results, health care systems might do the following to increase the public's comfort with third-party commercial companies and minimize the possibility of public backlash and media outcry:

- Emphasize the intended patient care and improvements that are expected to result from a partnership with a third-party commercial company (practice).

- Improve trust in the health system (practice).
- Advocate for changes to health data privacy laws and rules, modeled after Texas and California health data privacy legislation (policy).
- Focus health privacy efforts on notification rather than on the provision of privacy controls (research and design).

5.2 Implications for practice

From the perspective of the healthcare system, any improvements to the business of healthcare in turn improve patient care and outcomes. For the public, however, this may not be well understood. Chapter 2 of this dissertation describes the statistically significant difference between comfort with sharing health data for patient purposes versus sharing health data for business purposes (data sent to insurers, data that is stored), and the surprising positive association between educational attainment and comfort with sharing health data for patient purposes. These results underscore the importance of emphasizing to the public how patient care, treatment, and delivery might be improved through a partnership with a third-party commercial company, and to proactively communicate both the terms of the partnership and desired insights that might result from the venture.

Trust in the health system emerged in this study as the strongest predictor of comfort with sharing health data with third-party commercial companies. Trust recommendations abound as health and legal researchers realize the extent to which trust in health systems predicts care outcomes, regimen adherence, and patient satisfaction, as discussed in chapters 3 and 4 of this dissertation. In the interest of proactive communication of the terms and desired outcomes of third-party partnerships, health systems may benefit tremendously from the inclusion of patient advocates at the inception of these partnerships.

5.3 Implications for policy

Confidence in existing laws regulating health data sharing was positively associated with comfort with sharing health data with third-party commercial companies, but a minority of participants in this study felt that existing laws provided them “reasonable” or “sufficient”

protection. As detailed in chapter 2, California and Texas are the only states with comprehensive patient privacy laws that go above and beyond the protections currently offered by HIPAA. These policy weaknesses will undermine efforts to improve trust in health systems and de-incentivize investment in the data infrastructure necessary to provide privacy controls or data use notification to patients.

5.4 Implications for Design

That notification persisted in the final model of the public's comfort with sharing health data with third-party commercial companies and desire for control over health data may indicate that patients first need to know how and where their health data is being used, and from that informed starting point, may make more stringent or relaxed privacy decisions based on this information. As stated in chapter 4 of this dissertation, health systems are having to quickly navigate the changing business of healthcare whereby patients expect the same convenience afforded them by other industries. Proactive communication of third-party partnerships can occur via press releases and media reports or can occur through notification dashboards in the patient's health profile. This option is an attractive one for health systems to consider for two reasons: 1) notification can, in theory, be automated; 2) the myriad number of ways patient data is shared is so wide that it would be an inconceivable task for any one health system or department to undertake manually.

Health systems navigating consumer-facing data for the first time may glean insight on how to handle designing for trust and privacy from the work done by researchers of electronic commerce. I restate here the work of Belanger (2002), who concluded that winning public trust was the primary hurdle to e-commerce growth, and, based on their research and the research of others, offer six web design features were key to increasing consumer's ratings of trustworthiness: 1) safeguard assurances, 2) company reputation, 3) ease of navigation, 4) robust order fulfillment, 5) professionalism of the website, and 6) use of state-of-the-art web design (Belanger et al., 2002). While these features were designed for online commerce sites, very similar criteria can be applied to health systems. Application to health systems can mean 1) providing "plain language" explanation of technical and regulatory privacy protections to patients, as well as the consequences of violation; 2) maintenance of reputation through

transparency – companies have been shown to recover trust more quickly in the wake of a data breaches if customers are notified of the breach in a timely manner (Bansal & Zahedi, 2015); 3) designing electronic health information portals and notification dashboards to be easy to navigate for patients; 4) responding quickly to patient inquiries about their data; 5) maintaining professional websites; 6) using state-of-the art web design. In the computing environment, interface design, the quality of a website, grammatical errors, poor navigation structure, and broken links can also influence perceptions of trustworthiness. Based on the results of this study, healthcare systems may consider notification dashboards in lieu of granular control over health data sharing.

5.5 Implications for research

In future studies I will further explore notification design and the accompanying technical issues, visual and design options, and communication techniques on which notification depends. The research team of Professor Jodyn Platt has conducted two focus group deliberation sessions with Michigan residents to explore patient's understanding of how their health data is used inside and outside of the healthcare system. In just the two sessions that have been conducted, we have gained insight into how patients would like to be notified, how they would like to retrieve that information, and what information is most valuable to them. In my subsequent research I will expand on the findings presented here, designing and testing a potential data use notification interface for patients.

5.6 Conclusion

As technology companies increasingly enter the healthcare marketplace, they are doing so in a regulatory environment that is out of sync with our current digital reality –data that is collected, sold, and bartered not only increases the likelihood of re-identification, but also increases the likelihood of a data breach. Patients are aware of these risks, as evidenced by the low comfort with sharing health data with third-party commercial companies and by the increased prevalence of reported information withholding by patients. Healthcare systems can no longer rely on de-identification alone to protect patient privacy. Machine learning techniques have provided university researchers as well as companies with the ability to combine data sets

and glean insights into individual behavior. Healthcare systems can build trust by being transparent at the inception of a third-party commercial partnerships and prevent the public surprise and backlash that has characterized the announcements of recent partnerships with third-party commercial companies. Comprehensive patient privacy protection regulation and efforts to provide data use more transparent to patients is necessary for patients, providers, as well as technology companies, to advance innovation in healthcare and protect patient privacy.

5.7 References

Bansal, G., & Zahedi, F. M. (2015). Trust violation and repair: The information privacy perspective. *Decision Support Systems*, *71*, 62–77. https://doi.org/10.1016/j.dss.2015.01.009

Belanger, F., Hiller, J. S., & Smith, W. J. (2002). Trustworthiness in electronic commerce: The role of privacy, security, and site attributes. *The Journal of Strategic Information Systems*, *11*(3), 245–270. https://doi.org/10.1016/S0963-8687(02)00018-5

Blank, G., Bolsover, G., & Dubois, E. (2014). A New Privacy Paradox: Young People and Privacy on Social Network Sites. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2479938

Kim, J., Kim, H., Bell, E., Bath, T., Paul, P., Pham, A., Jiang, X., Zheng, K., & Ohno-Machado, L. (2019). Patient Perspectives About Decisions to Share Medical Data and Biospecimens for Research. *JAMA Network Open*, *2*(8), e199550–e199550. https://doi.org/10.1001/jamanetworkopen.2019.9550

Milewski, D. (2016, September 13). *One in Three Americans Hacked in the Past Year*. HSB. https://www.munichre.com/hsb/en/press-and-publications/press-releases/2016/2016-09-13-one-in-three-americans-hacked-in-the-past-year.html

Regalado, A. (2019, February 11). *More than 26 million people have taken an at-home ancestry test*. MIT Technology Review. https://www.technologyreview.com/2019/02/11/103446/more-than-26-million-people-have-taken-an-at-home-ancestry-test/

Sheehan, K. B. (1999). An investigation of gender differences in on-line privacy concerns and resultant behaviors. *Journal of Interactive Marketing*, *13*(4), 15.

Thiede, M. (2005). Information and access to health care: Is there a role for trust? *Social Science & Medicine (1982)*, *61*(7), 1452–1462. https://doi.org/10.1016/j.socscimed.2004.11.076

Zhou, J., & Salvendy, G. (2017). *Human Aspects of IT for the Aged Population. Applications, Services and Contexts: Third International Conference, ITAP 2017, Held as Part of HCI International 2017, Vancouver, BC, Canada, July 9-14, 2017, Proceedings, Part II*. Springer.

Appendix

NORC AmeriSpeak Survey

Client	University of Michigan
Project Name	Longitudinal Survey of Epidemiology
Project Number	8478
Survey length (median)	20-minute survey
Population	21+ gen pop
	African American/Hispanic and <=200% FPL oversamples
Pretest	N=100
Main	N=2000
MODE	Web only
Language	English
Incentive	5,000 points
Survey description	Healthcare issues
Eligibility Rate	100%

Standard demographic preloads:

Var Name	**Var Type**	**Var length**	**Variable Label**
S_AGE	Numeric	5	Age
S_GENDER	String	8	Gender
S_RACETH	Numeric	8	Race/ethnicity
S_EDUC	Numeric	6	Education
S_MARITAL	Numeric	9	Marital Status
S_EMPLOY	Numeric	8	Current employment status
S_INCOME	Numeric	8	Household income
S_STATE	String	7	State
S_METRO	Numeric	7	Metropolitan area flag
S_INTERNET	Numeric	10	Household internet access
S_HOUSING	Numeric	9	Home ownership
S_HOME_TYPE	Numeric	11	Building type of panelist's residence
S_PHONESERVC	Numeric	11	Telephone service for the household
S_HHSIZE	Numeric	8	Household size (including children)
S_HH01	Numeric	6	Number of HH members age 0-1
S_HH25	Numeric	6	Number of HH members age 2-5
S_HH612	Numeric	7	Number of HH members age 6-12
S_HH1317	Numeric	8	Number of HH members age 13-17

S_HH18OV	Numeric	8	Number of HH members age 18+
S_file_date	Date	11	
S_GENFRACE	Numeric	8	GenF custom race

These populated as a pre-load when the panelists get sampled into the survey

Standard sample preloads

Variable Name	**Variable Type**	**Variable Label**
Username	Numeric	Analogous to Member_PIN
P_Batch	Numeric	Batch Number (if only one assignment, then everyone will be 1)
Dialmode	Numeric	CATI Dialmode (predictive, preview, etc)
P_LCS	Numeric	Life cycle stage, 0=released but not touched
Y_FCELLP	String	
Surveylength	Numeric	Estimated length of survey
SurveyId	Numeric	Survey ID# in A4S
Incentwcomma	String	Study specific
P_Hold01	Numeric	Prevents dialing cases without phone numbers

Custom survey-specific preloads

Variable Name	**Variable Type**	**Variable Label**
P_PARTYID7	Numeric	1 "Strong Democrat"

		2 "Moderate Democrat" 3 "Lean Democrat" 4 "Don't Lean/Independent/None" 5 "Lean Republican" 6 "Moderate Republican" 7 "Strong Republican" ***only preload responses IF NOT MISSING PARTYID7**
FPL200	Numeric	1 “200FPL” 0 “Not 200FPL”

This survey will use the following RND_xx variables:

Note, these are randomized in the script (NOT preloads)

RND_xx	**Associated survey Qs**
RND_00	
RND_01	
RND_02	
RND_03	
RND_04	
RND_05	
RND_06	

Please include the following options for all questions in CATI:

77 DON'T KNOW

99 REFUSED

Please code refusals in CAWI:

98 IMPLICIT REFUSAL, WEB SKIP

Do not code 77 Don't Know/99 Refused options in CAWI unless written in item response options

---

Text shown in green includes researcher notes and should not be included in the programming.

[START OF SURVEY]

CREATE DATA-ONLY VARIABLE: QUAL

1=Qualified Complete

2=Not Qualified

3=In progress

AT START OF SURVEY COMPUTE QUAL=3 "IN PROGRESS"

CREATE MODE_START

1=CATI

2=CAWI

HOVER TEXT PROGRAMMING: BELOW IS THE HOVER TEXT THAT SHOULD BE DISPLAYED WHERE INDICATED THROUGHOUT THE SURVEY

HT_1: My healthcare system

"Your healthcare system" refers to the healthcare professionals and institutions that you personally interact with when getting health care.

HT_2: The healthcare system
"The healthcare system" refers generally to the healthcare system in this country.

HT_3: Healthcare providers
Health care providers include people such as doctors and nurses who provide medical treatment.

HT_4: Electronic health record
A digital version of your paper chart or medical record. An electronic health record contains your medical and treatment history including diagnoses, medications, treatment plans, immunization dates, allergies, radiology images, and laboratory and test results.

HT_5: Health information
Health information includes information about you and your medical treatment history including diagnoses, medications, treatment plans, immunization dates, allergies, radiology images, and laboratory and test results.

HT_7: De-identified [health information or biospecimens]
De-identified means that "identifying information" about *you* is *removed* from your health information. Identifying information includes things like your name, address, date of birth, etc.

HT_8: University researcher

A university researcher is a person who works for colleges or universities. University researchers might use health information to understand how people use the healthcare system, how healthcare providers treat patients, and a wide variety of other health related topics. University researchers might also use biospecimens for research on how illness works, which treatments are most effective, or how genetics affect illness. University researchers may or may not be connected to your hospital in some way.

HT_9: Biospecimens

Biospecimens include blood from a blood test, or tissue or tumor samples from a biopsy. Your biospecimens contain your DNA.

HT_10: Identified [health information or biospecimens]

Identifying information includes things like your name, address, date of birth, etc. Identified biospecimens are biospecimens that include identifying information about you.

HT_11: Quality analysts

Quality analysts are people who work for hospitals or clinics. They use patient health information at their hospital or clinic to check on, and improve, how their organization is working. They often study the cost of healthcare, their organization's efficiency on things like waiting room times, and the health of patients at their hospital or clinic.

HT_12: Commercial companies

Commercial companies are third-party companies that are not part of a hospital. For example, a third-party commercial company may conduct genetic tests and analyze information for a hospital or healthcare provider for a fee when a hospital is not be able to conduct the test on their own.

(Project name) Draft

Date: (Quex start date)

[DISPLAY – WINTRO_1]

Thank you for agreeing to participate in our new AmeriSpeak survey! To thank you for sharing your opinions, we will give you a reward of [INCENTWCOMMA] AmeriPoints after

completing the survey. As always, your answers are confidential.

Please use the "Continue" and "Previous" buttons to navigate between the questions within the questionnaire. Do not use your browser buttons.

[DISPLAY_1]

This is a survey about your healthcare experience and your opinions about how [HT_5] health information is used and shared.

[SPACE]

The survey includes a short video, which <u>needs to be viewed with sound</u>. If you are not able to have your sound on at this time, feel free to take this survey later when you can. Alternatively, if you will not be able to view the video with sound, we have a transcript of the video that can be read instead.

[DISPLAY_2]

We all use [HT_2] the healthcare system or know people who do. This system includes the healthcare provider, like a doctor or nurse, who you visit when you're sick or for routine visits. It also includes hospitals and people who work on quality improvement, and administrators who make decisions about how clinics and hospitals are run.

[SPACE]

When we ask about "*[HT_2] the healthcare system*" we are asking generally about [HT_2] the healthcare system in this country. When we ask about "*your* healthcare system" we are asking about the healthcare professionals and institutions that you personally interact with.

[SPACE]

We are interested in your thoughts and ideas. There are no "right" or "wrong" answers.

[GRID; SP]

Q1.

Please state if the following are true or false:

GRID ITEMS [RANDOMIZE]:

A. My [HT_3] healthcare provider uses an [HT_4] electronic health record
B. I have used a patient portal to access my [HT_5] health information online
C. I am worried about being able to pay medical bills
D. I am confident my health insurance covers my medical needs

RESPONSE OPTIONS:

1. True
2. False
77. Not sure

PM: PLEASE MAKE SURE THE DATE TIME RULE ALWAYS FOLLOWS FIRST

QUESTION

INSERT ITEM TIMESTAMPS: TIME_FIRST, DATE_FIRST

THE REST OF THE CLIENT INSTRUMENT GOES HERE.

[SHOW IF Q1_2=1]

[SP]

Q1A1

I have more than one patient portal.

RESPONSE OPTIONS:

1. Yes
2. No

[RANDOMIZE RESPONSE OPTIONS]

[MP]

Q2.

In the past 12 months, have you...

[SPACE]

<unbold>*Please select all that apply.*<unbold>

RESPONSE OPTIONS:

1. …seen a [HT_3] healthcare provider?
2. …been seen in the emergency room?
3. …spent one or more nights in the hospital?
4. …been screened for cancer? (mammogram, Pap test, colonoscopy, lung cancer)
5. …been treated for cancer?

[SPACE]

6. None of the above [SP] [ANCHOR]

[RANDOMIZE RESPONSE OPTIONS]

[MP]

Q3.

In the past 12 months, have <u>*any of your loved ones...*</u>

RESPONSE OPTIONS:

1. …been seen in the emergency room?
2. …spent one or more nights in the hospital?
3. …been treated for cancer?

[SPACE]

4. None of the above [SP] [ANCHOR]

[SHOW Q4 AND Q5 ON THE SAME PAGE]

[SP]
Q4.
Have you ever been told by a [HT_3] healthcare provider that you have cancer?

RESPONSE OPTIONS:

1. Yes
2. No

[SP]
Q5.
Do you have a family history of cancer?

RESPONSE OPTIONS:

1. Yes
2. No

[SP]
Q6.
Would you say that in general your health is…

RESPONSE OPTIONS:

1. Poor
2. Fair
3. Good
4. Very Good
5. Excellent

[RANDOMIZE RESPONSE OPTIONS]

[SP]

Q8.

What kind of [HT_3] healthcare provider do you typically go if you are sick or need advice about your health??

RESPONSE OPTIONS:

1. A [HT_3] healthcare provider's office
2. A hospital clinic
3. Other clinic or health center
4. Urgent care
5. Emergency room
6. Other [TEXTBOX] [ANCHOR]
7. I don't have a regular healthcare provider

[SP]

Q9.

Approximately when was the last time you saw a [HT_3] healthcare provider?

RESPONSE OPTIONS:

1. Within the past year
2. Within the past 2 years

3. Within the past 5 years
4. More than 5 years ago
5. Have never seen a healthcare provider

[SP]

Q10.

Are you now covered by any form of health insurance or health plan (this includes Medicare, Medicaid, private health insurance and insurance plans available through healthcare.gov)?

RESPONSE OPTIONS:

1. Yes
2. No

[GRID; SP]

Q11.

For you, how true are the following statements?

GRID ITEMS [RANDOMIZE]:

A. [HT_2] The healthcare system in this country is easy to use
B. I can get the healthcare I need when I need it
C. I get all the information I need about my health from my [HT_3] healthcare provider
D. I could access my [HT_4] electronic health record if I wanted to
E. In general, I am satisfied with the treatment I receive from my [HT_3] healthcare providers

RESPONSE OPTIONS:

1. Not true
2. Somewhat true
3. Fairly true
4. Very true

[GRID; SP]

Q12.

For you, how true are the following statements?

GRID ITEMS [RANDOMIZE]:

A. Most healthcare systems in this country are too big to care about individual patients
B. [HT_1] My healthcare system is too big to care about me
C. Healthcare systems in this country work to prevent harm to their patients
D. I feel respected when I seek health care

RESPONSE OPTIONS:

1. Not true
2. Somewhat true
3. Fairly true
4. Very true

[GRID; SP]

Q13.

For you, how true are the following statements?

GRID ITEMS [RANDOMIZE]:

A. [HT_1] My healthcare system treats me fairly
B. [HT_1] My healthcare system treats me with kindness

RESPONSE OPTIONS:

1. Not true
2. Somewhat true
3. Fairly true
4. Very true

[SP]

Q14.

Have <u>you<u> ever experienced discrimination, or been hassled or made to feel inferior while getting medical care?

RESPONSE OPTIONS:

1. Yes
2. No

[SHOW IF Q14=1]

Q14AA.

[SP]

How often has this happened?

RESPONSE OPTIONS:

1. Once
2. 2 or 3 times
3. 4 or more times

[SHOW IF Q14=1]

Q14A.

[SP]

What do you think was the <u>main<u> reason for this experience?

RESPONSE OPTIONS:

1. Ancestry or national origin
2. Gender
3. Race
4. Age

5. Religion
6. Height
7. Weight
8. Shade of skin color
9. Sexual orientation
10. Education or income level
11. Physical disability
12. Speaking English as a second language
13. Other/ please specify [TEXTBOX]

Q15.

How often do you feel that racial/ethnic groups who are not white, such as African Americans and Latinos, are discriminated against in [HT_2] the healthcare system?

RESPONSE OPTIONS:

1. Never
2. Rarely
3. Sometimes
4. Often

[GRID; SP]

Q16.

For you, how true are the following statements?

GRID ITEMS [RANDOMIZE]:

A. I find ways to help others less fortunate than me
B. The dignity and well-being of all should be the most important concern in any society
C. One of the problems of today's society is that people are often not kind enough to others
D. All people who are unable to provide for their own needs should be helped by others

RESPONSE OPTIONS:

1. Not true
2. Somewhat true
3. Fairly true
4. Very true

[SP]

VIDEO_INT.

In the next section of the survey you will see a two-minute video that explains how [HT_5] health information is used and shared in [HT_2] the healthcare system. There will be questions that follow the video.

[SPACE]

<u>You will need to have your sound on while viewing this video. <u>

[SPACE]

Are you able to turn your sound on now?

RESPONSE OPTIONS:

1. Yes
2. No

[SHOW IF VIDEO_INT=1]

VIDEO1.

<u>Instructions for watching your video</u>:

- The continue button will appear once the video has ended.
- Do not fast forward through the video.
- Do not skip past the video before viewing it once.
- You may re-watch the video multiple times.
- The video is best viewed horizontally if watched on a mobile phone.
- Click on the image to start watching the video. Make sure you hear it.

[PLAY VIDEO HERE: Link to Video]

[DELAY PRESENTATION OF CONTINUE BUTTON FOR 120 SECONDS]

SHOW IF VIDEO_INT=1

VIDEO.

Were you able to <u>see<u> and <u>hear<u> the video?

RESPONSE OPTIONS:

1. Yes
2. No

SHOW IF VIDEO_INT=2,98 OR VIDEO=2,98

[DISPLAY]

TRANSCRIPT.

Today's technologies, from genome sequencing to [HT_4] electronic health records, are turning [HT_5] health information into a valuable resource for answering health questions and improving care.

[SPACE]

Imagine you have a friend, Florence, who was recently diagnosed with breast cancer. The handling of her [HT_5] health information looks very different today than it did 10 years ago. This will impact the care that she and other patients will receive.

[SPACE]

Florence's doctors will collect information about her health history, health behaviors, family history and maybe even her neighborhood and job. Doctors might also collect her genetic information using samples of her tumor as well as her normal blood cells to more precisely tailor or personalize her treatment. Blood left over from those tests might be set aside to be used for research. Using a team approach, dozens of people involved in her treatment may look at her chart to help support her care…

[SPACE]

[HT_5] Health information from patients like Florence also typically travels out to many other users in her health system who may not be involved directly in her care – insurers, billers, and analysts who could learn from the outcomes of her treatment.

[SPACE]

Precision health companies might collect and store archives of health data or use it to develop new drugs or digital tools for improving diagnosis and treatment. There are laws designed to protect Florence's privacy, but some of her information can still be shared after personal "identifiers" like her name and address are removed.

[SPACE]

With all of these users and uses of patients' data, Florence, and patients like her, are able to contribute to the improvement of their own medical care and the care of other patients like them. But these changes in the ways health care is conducted also come with some new questions for Florence and for all of us – questions about trust, privacy, duty, and the tradeoffs that come with sharing data.

[SPACE]

This survey asks you to reflect on what you think about the use and sharing of all of this health data. Thank you for your time!

[GRID; SP]

Q17.

Based on what you saw in the video or read in the transcript, are the following statements true or false?

GRID OPTIONS [RANDOMIZE]:

A. [HT_5] Health information can only be used for treating patients. (F)
B. Blood left over from tests used for treating or diagnosing a disease might be set aside to be used for research. (T)
C. Florence has cancer. (T)

RESPONSE OPTIONS:

1. True
2. False

[DISPLAY]

Now that you've heard Florence's story and seen how health information can be shared, we'd like to know what you think about the use and sharing of health information.

[SPACE]

The next questions are about the use of your health information for research.

[SPACE]

Your <u>health information</u> is information about you and your medical treatment history including diagnoses, medications, treatment plans, immunization dates, allergies, radiology images, and laboratory and test results.

[SPACE]

Your health information can be "*de-identified*." This means that "identifying information" about *you* <u>is</u> *removed* from your health information. Identifying information includes things like your name, address, date of birth, etc. De-identified information can then be given to researchers to study things like healthcare costs, quality, and diseases.

[SPACE]

<u>University researchers</u> are people who work for colleges or universities. These researchers might use health information to understand how people use [HT_2] the healthcare system, how [HT_3] healthcare providers treat patients, and a wide variety of other health related topics. University researchers might also use [HT_9] biospecimens for research on how illness works, which treatments are most effective, or how genetics affect illness. University researchers may or may not be connected to your hospital in some way.

[GRID; SP]

Q18.

For you, how true are the following statements?

GRID ITEMS [RANDOMIZE]:

A. I am comfortable with <u>[HT_8] university researchers</u> using my [HT_7] *de-identified* [HT_5] health information
B. I would like to be notified if <u>[HT_8] university researchers</u> will use my [HT_7] *de-identified* [HT_5] health information

RESPONSE OPTIONS:

1. Not true
2. Somewhat true
3. Fairly true
4. Very true

[GRID; SP]
Q19.

Your [HT_5] health information can be "<u>identified</u>." This means that [HT_10]"identifying information" about you <u>is linked to</u> your health information. Identifying information includes things like your name, address, date of birth, etc. Identified information can then be given to researchers to study things like healthcare costs, quality, and diseases.

[SPACE]

For you, how true are the following statements?

GRID ITEMS [RANDOMIZE]:

A. I am comfortable with [HT_8] <u>university researchers</u> using my *identified* health information.
B. I would like to be notified if [HT_8] <u>university researchers</u> will use my *identified* health information.

RESPONSE OPTIONS:

1. Not true

2. Somewhat true
3. Fairly true
4. Very true

[GRID; SP]
Q20.
Your <u>biospecimens</u> may be collected during the course of your treatment.

Biospecimens include blood from a blood test, or tissue or tumor samples from a biopsy. Your biospecimens contain your DNA. Sometimes when there are biospecimens left over from your healthcare (such as blood or urine left over from a diagnostic test) that otherwise would be thrown away, those leftover biospecimens might be used for research.

[SPACE]

Biospecimens can be "<u>[HT_7]de-identified</u>." This means that [HT_10] "identifying information" about you <u>is not linked to your</u> biospecimens. Identifying information includes things like your name, address, date of birth, etc. De-identified biospecimens can be given to researchers to study things like healthcare costs, healthcare quality, and diseases.

[SPACE]

For you, how true are the following statements?

GRID ITEMS [RANDOMIZE]:

A. I am comfortable with [HT_8] <u>university researchers<u> using my *de-identified* biospecimens.
B. I would like to be notified if [HT_8] <u>university researchers</u> use my *de-identified* biospecimens.

RESPONSE OPTIONS:

1. Not true
2. Somewhat true

3. Fairly true
4. Very true

[GRID; SP]
Q21.
[HT_9] Biospecimens can be "<u>identified</u>." This means that [HT_10] "identifying information" about you <u>is linked to</u> your biospecimens. Identifying information includes things like your name, address, date of birth, etc. Identified [HT_9] biospecimens can be given to researchers to study things like health care costs, quality, and diseases.

[SPACE]

For you, how true are the following statements?

GRID ITEMS [RANDOMIZE]:

A. I am comfortable with [HT_8] <u>university researchers<u> using my *identified* biospecimens.
B. I would like to be notified if [HT_8] <u>university researchers<u> use my *identified* biospecimens.

RESPONSE OPTIONS:

1. Not true
2. Somewhat true
3. Fairly true
4. Very true

[GRID; SP]
Q22.
The next questions are about the use of your [HT_5] health information by <u>quality analysts</u>**.**
[SPACE]

<u>Quality analysts<u> are people who work for hospitals or clinics. They use patient health information at their hospital or clinic to check on, and improve, how their organization is working. They often study the cost of

healthcare, their organization's efficiency on things like waiting room times, and the health of patients at their hospital or clinic.

[SPACE]

For you, how true are the following statements about <u>health information<u>?

GRID ITEMS [RANDOMIZE]:

A. I am comfortable with quality analysts using my [HT_7] *de-identified* <u>health information</u>.
B. I would like to be notified if quality analysts use my [HT_7] *de-identified* <u>health information</u>.
C. I am comfortable with quality analysts using my [HT_10] *identified* <u>health information</u>.
D. I would like to be notified if quality analysts use my [HT_10] *identified* <u>health information</u>.

RESPONSE OPTIONS:

1. Not true
2. Somewhat true
3. Fairly true
4. Very true

[GRID; SP]

Q23.

For you, how true are the following statements about [HT_9] <u>biospecimens</u>?

GRID ITEMS [RANDOMIZE]:

A. I am comfortable with [HT_11] quality analysts using my [HT_7] *de-identified* <u>biospecimens</u>.
B. I would like to be notified if [HT_11] quality analysts use my [HT_7] *de-identified* <u>biospecimens</u>.
C. I am comfortable with [HT_11] quality analysts using my [HT_10] *identified* <u>biospecimens</u>.
D. I would like to be notified if [HT_11] quality analysts use my [HT_10] *identified* <u>biospecimens</u>.

RESPONSE OPTIONS:

1. Not true
2. Somewhat true
3. Fairly true
4. Very true

[GRID; SP;4,4]

Q24.

The next questions are about the use of your [HT_5] health information by <u>commercial companies</u>.

[SPACE]

<u>Commercial companies</u> are third-party companies that are not part of a hospital. For example, a third-party commercial company may conduct genetic tests and analyze information for a hospital or [HT_3] healthcare provider for a fee when a hospital is not be able to conduct the test on their own. Commercial companies may keep the information for their own use.

[SPACE]

For you, how true are the following statements?

GRID ITEMS [RANDOMIZE]:

A. I am comfortable with commercial companies using my [HT_7] *de-identified* <u>health information</u>.
B. I would like to be notified if commercial companies use my [HT_7] *de-identified* <u>health information</u>.
C. I am comfortable with commercial companies using my [HT_10] *identified* <u>health information</u>.
D. I would like to be notified if commercial companies use my [HT_10] *identified* <u>health information</u>.
E. I am comfortable with commercial companies using my [HT_7] *de-identified* [HT_9] <u>biospecimens</u>.
F. I would like to be notified if commercial companies use my [HT_7] *de-identified* [HT_9] <u>biospecimens</u>.
G. I am comfortable with commercial companies using my [HT_10] *identified* [HT_9] <u>biospecimens</u>.
H. I would like to be notified if commercial companies use my [HT_10] *identified* [HT_9] <u>biospecimens</u>.

RESPONSE OPTIONS:

1. Not true
2. Somewhat true
3. Fairly true
4. Very true

[SP]

Q25.

How confident are you that [HT_7] de-identifying [HT_5] health information protects your privacy?

RESPONSE OPTIONS:

1. Not at all confident
2. Somewhat confident
3. Fairly confident
4. Very confident

[GRID; SP]

Q26.

How often would you like to be notified about the use of each of the following in research studies that start at your hospital and are shared with a [HT_12] commercial company?

GRID ITEMS [RANDOMIZE]:

A. Your genetic information or DNA
B. Your [HT_9] biospecimens
C. Your [HT_10] identified [HT_5] health information
D. Your [HT_7] de-identified [HT_5] health information

RESPONSE OPTIONS:

1. Never
2. Just once
3. Once every five years
4. Once a year
5. Every time I visit my [HT_3] healthcare provider
6. Every time the information is used

[GRID; SP]

Q27.

Suppose you are a cancer patient at a leading cancer center and your [HT_3] healthcare provider wants to use your DNA to see if you might be a good candidate for a particular cancer treatment. Your cancer center shares DNA and [HT_5] health information with third-party [HT_12] commercial companies when it is unable to perform the analysis themselves.

[SPACE]

How comfortable are you with a third-party commercial company…

GRID ITEMS [RANDOMIZE]:

A. …using your DNA and health information to improve the diagnosis and treatment of cancer in <u>*other*</u> patients
B. …developing predictions about how you will respond to a particular cancer treatment
C. …storing your DNA and health information
D. …sharing predictions about how you will respond to cancer treatment with insurance companies
E. …selling [HT_7] de-identified health information to pharmaceutical companies

RESPONSE OPTIONS:

1. Not comfortable
2. Somewhat comfortable
3. Fairly comfortable
4. Very comfortable

[GRID; SP]

Q28.

In the future, [HT_3] healthcare providers may be able to treat some diseases by making changes to patients' DNA, which is also called gene-editing. To make this possible, researchers need to use [HT_9] biospecimens (e.g. blood or tissue) donated from research participants.

[SPACE]

How comfortable are you with sharing <u>your own</u> [HT_7] *de-identified* biospecimen for research on gene-editing in the following ways?

GRID ITEMS [RANDOMIZE]:

A. To develop gene-editing methods that treat disease or disability.
B. To develop gene-editing methods that enhance a person physically (make them stronger or faster) or mentally (increase intelligence).

RESPONSE OPTIONS:

1. Not comfortable
2. Somewhat comfortable
3. Fairly comfortable
4. Very comfortable

[GRID; SP]

Q29.

In some cases, researchers would like to use [HT_10] *identified* [HT_9] biospecimens—samples linked to information that would identify you as the donor.

[SPACE]

How comfortable are you with sharing <u>your own</u> *identified* biospecimens for research on gene-editing in the following ways?

GRID ITEMS [RANDOMIZE]:

A. To develop gene-editing methods that treat disease or disability.
B. To develop gene-editing methods that enhance a person physically (make them stronger or faster) or mentally (increase intelligence).

RESPONSE OPTIONS:

1. Not comfortable
2. Somewhat comfortable
3. Fairly comfortable
4. Very comfortable

[SP]

Q30.

In the U.S., the federal government is a major funder of medical research using tax revenue.
[SPACE]
How comfortable are you with your tax dollars being used to support research on gene-editing?

RESPONSE OPTIONS:

1. Not comfortable
2. Somewhat comfortable
3. Fairly comfortable
4. Very comfortable

[GRID; SP]

Q31.

How optimistic are you that gene-editing will have a <u>positive</u> impact:

GRID ITEMS [RANDOMIZE]:

A. On you
B. On your family
C. On society

RESPONSE OPTIONS:

1. Not at all optimistic
2. Somewhat optimistic
3. Quite optimistic
4. Very optimistic

[GRID; SP]

Q32.

How fearful are you that gene-editing will have a <u>negative</u> impact:

GRID ITEMS [RANDOMIZE]:

A. On you
B. On your family
C. On society

RESPONSE OPTIONS:

1. Not at all fearful
2. Somewhat fearful
3. Quite fearful
4. Very fearful

[GRID; SP]

Q33.

The next questions ask for your opinions about whether people should share their [HT_5] health information in general.

[SPACE]

For you, how true are the following statements?

GRID ITEMS [RANDOMIZE]:

A. People have an ethical obligation to allow their health information to be used for <u>healthcare quality analysis</u>
B. People have an ethical obligation to allow their health information to be used for <u>research</u>

RESPONSE OPTIONS:

1. Not true
2. Somewhat true
3. Fairly true
4. Very true

[GRID; SP]

Q34.

For you, how true are the following statements about the organizations that have your [HT_5] health information and share it? Organizations include groups such as [HT_3] healthcare providers' offices, hospitals, insurance companies, and [HT_8] university researchers. (If you are unsure, please make your best guess.)

[SPACE]

The organizations that have my health information and share it...

GRID ITEMS [RANDOMIZE]:

A. Try hard to be fair in dealing with others
B. Would try to hide a serious mistake they made

C. Tell me how my health information is used
D. Would never mislead me about how my health information is used
E. Have a particular interest in collecting my [HT_9] biospecimens compared to other people's

RESPONSE OPTIONS:

1. Not true
2. Somewhat true
3. Fairly true
4. Very true

[GRID; SP]

Q35.

The organizations that have my [HT_5] health information and share it...

GRID ITEMS [RANDOMIZE]:

A. Are not good at their jobs
B. Have specialized capabilities that can promote innovation in health
C. Can use large amounts of data to improve patient care

RESPONSE OPTIONS:

1. Not true
2. Somewhat true
3. Fairly true
4. Very true

[GRID; SP]

Q36.

The organizations that have my [HT_5] health information and share it...

GRID ITEMS [RANDOMIZE]:

C. Can be trusted to use my health information responsibly
D. Think about what is best for me
E. Act in an ethical manner

RESPONSE OPTIONS:

1. Not true
2. Somewhat true
3. Fairly true
4. Very true

[GRID; SP]

Q37.

The organizations that have my [HT_5] health information and share it...

GRID ITEM [RANDOMIZE]:

A. Deliberately withhold important information from me about my medical care
B. Disclose their financial conflicts of interest
C. Experiment on patients without telling them
D. Treat everyone the same, regardless of their race or ethnicity
E. Treat everyone the same, regardless of their income

RESPONSE OPTIONS:

1. Not true
2. Somewhat true
3. Fairly true
4. Very true

[GRID; SP]

Q38.

For you, how true are the following statements about [HT_3] <u>*health care providers*</u>?

GRID ITEMS [RANDOMIZE]:

A. Health care providers care most about making money for themselves
B. Health care providers do not care about helping people like me
C. I trust health care providers to use my [HT_5] health information responsibly
D. Health care providers disclose their conflicts of interest
E. All things considered, health care providers in this country can be trusted

RESPONSE OPTIONS:

1. Not true
2. Somewhat true
3. Fairly true
4. Very true

[GRID; SP]

Q39.

For you, how true are the following statements?

GRID ITEMS [RANDOMIZE]:

A. The privacy policies of [HT_1] my healthcare system are clear to me
B. I am satisfied with the level of access I have to information in my [HT_4] electronic health record
C. I am confident in my ability to manage how my [HT_5] health information is used
D. [HT_1] My healthcare system respects my privacy

RESPONSE OPTIONS:

1. Not true
2. Somewhat true
3. Fairly true
4. Very true

[GRID; SP]

Q40.

For you, how true are the following statements?

GRID ITEMS [RANDOMIZE]:

A. I should have more control over how my [HT_5] health information is used
B. [HT_1] My healthcare system is transparent about how my [HT_5] health information is used
C. [HT_1] My healthcare system would notify me if the security of my [HT_5] health information had been compromised

RESPONSE OPTIONS:

1. Not true
2. Somewhat true
3. Fairly true
4. Very true

[GRID; SP]

Q41.

For you, how true are the following statements?

GRID ITEMS [RANDOMIZE]:

B. I worry that private information about my health could be used against me
C. I worry my [HT_5] health information is available to people who have no business seeing it
D. There are some things I would not tell my [HT_3] healthcare providers because I can't trust them with the information
E. [HT_1] My healthcare system discloses its conflicts of interest

RESPONSE OPTIONS:

1. Not true
2. Somewhat true
3. Fairly true
4. Very true

[SP]

Q42.

Have you ever kept information from your [HT_3] healthcare provider because you were concerned about privacy or security?

RESPONSE OPTIONS:

1. Yes
2. No

[GRID;3,3; SP]

Q43.

What do you think of the following statements: true or false?

GRID ITEMS [RANDOMIZE]:

A. My [HT_3] healthcare provider is the only person who decides how information in my [HT_4] electronic health record is used
B. Health insurance companies are prohibited from using my [HT_5] health information to deny me coverage
C. I own my [HT_5] health information
D. My permission is required for research using my [HT_7] de-identified [HT_5] health information
E. My genetic code is specific to me
F. Health organizations are required to report large-scale data breaches

RESPONSE OPTIONS:

1. True
2. False
77. I don't know

[GRID; SP]

Q44.

For you, how true are the following statements?

GRID ITEMS [RANDOMIZE]:

A. [HT_1] My healthcare system will use my [HT_5] health information how they see fit, regardless of my preferences
B. Healthcare systems in this country respect patients' opinions about how their [HT_5] health information should be used

RESPONSE OPTIONS:

1. Not true
2. Somewhat true
3. Fairly true
4. Very true

[GRID; SP]

Q45.

The next questions ask about your perceptions of the uses of [HT_5] health information and the policies that are in place to protect it.

[SPACE]

For you, how true are the following statements?

GRID ITEMS [RANDOMIZE]:

A. I should have a right to participate in medical research
B. Information about my health should be kept indefinitely in my [HT_4] electronic health record
C. I should be able to delete my [HT_4] electronic health record
D. I should be able to access all of my [HT_3] healthcare providers through a single patient portal

RESPONSE OPTIONS:

1. Not true
2. Somewhat true
3. Fairly true
4. Very true

[GRID; SP]

Q46.

For you, how true are the following statements?

GRID ITEMS [RANDOMIZE]:

A. It is important that I know who has [HT_5] health information about me
B. I should be able to find out how my [HT_5] health information has been shared

RESPONSE OPTIONS:

1. Not true
2. Somewhat true
3. Fairly true
4. Very true

[GRID; SP]

Q47.

For you, how true are the following statements?

GRID ITEMS [RANDOMIZE]:

A. It is okay for law enforcement to access [HT_5] health information
C. Existing laws provide a reasonable level of protection for the privacy of patient information
D. I am confident that electronic [HT_5] health information is sufficiently protected by current law and regulation

RESPONSE OPTIONS:

1. Not true
2. Somewhat true
3. Fairly true
4. Very true

[GRID;3,3; SP]

Q48.

Suppose [HT_1] your healthcare system charges a fee to third-party [HT_12] commercial companies that want to keep [HT_9] biospecimens left over after testing. This makes money for the hospital.

[SPACE]

For you, how true are the following statements?

[SPACE]

If my healthcare system makes money from my biospecimens…

GRID ITEMS [RANDOMIZE]:

A. …they should use the money to provide health care for other people who can't afford it.
B. …they should use the money to provide health care for other people who have the same health problems as me.
C. …they should use the money to improve quality of care in my healthcare system.
D. …they should use the money to support future research on my health problems.
E. …they should use the money to support future research on any kind of health problem.
F. …they should use it however they want.

RESPONSE OPTIONS:

1. Not true
2. Somewhat true
3. Fairly true
4. Very true

[SP]
Q49.
Some companies like 23andMe and AncestryDNA test people's genetic material from saliva to trace their ancestry.

[SPACE]

Given what you know about companies like 23andMe and AncestryDNA, do you generally have a favorable or unfavorable opinion of these companies?

RESPONSE OPTIONS:

1. Very favorable
2. Somewhat favorable
3. Somewhat unfavorable

4. Very unfavorable

[SP]
Q50.
Are you comfortable with <u>drug companies</u> purchasing genetic and ancestry data from companies like 23andMe or AncestryDNA?

RESPONSE OPTIONS:

1. Yes
2. No

[SP]
Q51.
Are you comfortable with <u>law enforcement</u> using genetic and ancestry data from companies like 23andMe or AncestryDNA?

RESPONSE OPTIONS:

1. Yes
2. No

[SP]
Q53.
Have you experienced problems with stolen or misused personal information (e.g., social security number, credit or debit cards, and bank accounts) within the last five years? For example, have you spent time clearing up credit accounts or your credit report because someone stole your personal information?

RESPONSE OPTIONS:

1. Yes, I am currently experiencing problems
2. Yes, but all problems have been resolved
3. No, I have not experienced any problems within the past five years

---

[SP]

Q54.

I believe my financial information has been compromised as the result of a data breach or hacking.

RESPONSE OPTIONS:

1. Yes
2. No

[GRID; SP]

Q55.

How concerned are you about the following recent events:

GRID ITEMS [RANDOMIZE]:

A. Facebook sharing information with Cambridge Analytica for political purposes
B. Data breach of people's social security numbers and driver's license numbers at Equifax
C. Sloan Kettering hospital executives using hospital data for their own startup company
D. Marriott data breach of passport numbers and credit card numbers

RESPONSE OPTIONS:

1. Not at all concerned
2. Somewhat concerned
3. Fairly concerned
4. Very concerned

[SHOW IF MISSING P_PARTYID7]

[SP]

PID1.

Do you consider yourself a Democrat, a Republican, an independent or none of these?

RESPONSE OPTIONS:

1. Democrat
2. Republican
3. Independent
4. None of these

Previously programmed as PIDA from TESS 007

[SHOW IF PID1=1]

[SP]

PIDA.

Do you consider yourself a strong or moderate Democrat?

RESPONSE OPTIONS:

1. Strong Democrat
2. Moderate Democrat

Previously programmed as PIDB from TESS 007

[SHOW IF PID1=2]

[SP]

PIDB.

Do you consider yourself a strong or moderate Republican?

RESPONSE OPTIONS:

1. Strong Republican
2. Moderate Republican

Previously programmed as PIDi from TESS 007

[SHOW IF PID1=3, 4, 77, 98, 99]

[SP]

PIDi.

Do you lean more toward the Democrats or the Republicans?

RESPONSE OPTIONS:

1. Lean Democrat
2. Lean Republican
3. Don't lean

[DOUBLE PROMPT IF REFUSED]
[NUMBERBOX; RANGE 1-20]
HHSIZE2.
<u>Including yourself</u>, how many people live in your household?

[NUMBERBOX; RANGE 0-999,000]

[PROMPT TWICE IF REFUSED WITH CUSTOM MESSAGE "We know that questions about income are sensitive, and understand if you would not like to answer this question. However,

some of our key research questions require us to have this information for respondents of this survey."

INCOME2.

What was your total <u>household,/u. income in 2018?

COMPUTE DOV_FPL BASED ON THE CHART BELOW. IF INCOME2 IS AT OR BELOW A GIVEN VALUE FOR A SPECFIC HHSIZE2 DOV_FPL=1. NOTICE THERE ARE DIFFERENT INCOME2 CRITERIA DEPENDING ON S_STATE VALUES AS INDICATED BY THE 3 INCOME2 COLUMNS

HHSIZE2	INCOME2 FOR S_STATE NE AK (Alaska) or HI (Hawaii)	INCOME2 FOR S_STATE= AK (Alaska)	INCOME2 FOR S_STATE= HI (Hawaii)
1	$24,980	$31,200	$28,760
2	33,820	42,260	38,920
3	42,660	53,320	51,080
4	51,500	64,380	59,240
5	60,340	75,440	69,400
6	69,180	86,500	79,560
7	78,020	97,560	89,720
8	86,860	108,620	99,880
9	95,700	119,680	110,040
10	104,540	130,740	120,200
11	113,380	141,800	130,360
12	122,220	152,860	140,520
13	131,060	163,920	150,680
14	139,900	174,980	160,840
15	148,740	186,040	171,000
16	157,580	197,100	181,160
17	166,420	208,160	191,320
18	175,260	219,220	201,480
19	184,100	230,280	211,640

20	192,940	241,340	221,800

[TEXTBOX]

CLOSE1.

In thinking about [HT_5] health information sharing, do you have any comments you would like to share?

[MEDIUM TEXTBOX]

PM PLEASE ALWAYS HAVE THIS AND THE FOLLOWING LOGIC FOLLOW THE

FINAL SUBSTANTIVE QUESTION OF THE SURVEY, AHEAD OF QFINAL

INSERT ITEM TIMESTAMPS: TIME_END, DATE_END

COMPUTE **TEST_TIME**

TEST_TIME = TIME_END – TIME_START

COMPUTE **TEST_DATE** = DATE_END

DISPLAY TESTING-ONLY SCREEN WITH VALUE FOR **TEST_TIME** & **TEST_DATE**

RE-COMPUTE QUAL=1 “COMPLETE”

SET CO_DATE, CO_TIME, CO_TIMER VALUES HERE

CREATE MODE_END

1=CATI

2=CAWI

SCRIPTING NOTES: PUT QFINAL1, QFINAL2, QFINAL3 in the same screen.

[SINGLE CHOICE]

QFINAL1.

Thank you for your time today. To help us improve the experience of AmeriSpeak members like yourself, please give us feedback on this survey.

[RED TEXT – CAWI ONLY] If you do not have any feedback for us today, please click "Continue" through to the end of the survey so we can make sure your opinions are counted and for you to receive your AmeriPoints reward.

Please rate this survey overall from 1 to 7 where 1 is Poor and 7 is Excellent.

Poor						Excellent
1	2	3	4	5	6	7

[SINGLE CHOICE – CAWI ONLY]

QFINAL2.

Did you experience any technical issues in completing this survey?

1. Yes – please tell us more in the next question
2. No

[TEXT BOX] [CATI version needs "no" option]

QFINAL3.

Do you have any general comments or feedback on this survey you would like to share? If you would like a response from us, please email support@AmeriSpeak.org or call (888) 326-9424.

[DISPLAY]

END.

[CATI version]

Those are all the questions we have. You have earned a reward of [INCENTWCOMMA] AmeriPoints for completing the survey. If you have any questions at all for us, you can email us at support@AmeriSpeak.org or call us toll-free at **888-326-942**4. Let me repeat that again: email us at support@AmeriSpeak.org or call us at **888-326-9424**. Thank you for participating in our new AmeriSpeak survey!

[CAWI version]

Those are all the questions we have. You have earned a reward of [INCENTWCOMMA] AmeriPoints for completing the survey. If you have any questions at all for us, you can email us

at support@AmeriSpeak.org or call us toll-free at **888-326-9424**. Thank you for participating in our new AmeriSpeak survey!

You can close your browser window now if you wish or click Continue below to be redirected to the AmeriSpeak member website.

Bibliography

Acquisti, A. (2004). Privacy and Security of Personal Information. In L. J. Camp & S. Lewis (Eds.), *Economics of Information Security* (Vol. 12, pp. 179–186). Kluwer Academic Publishers. https://doi.org/10.1007/1-4020-8090-5_14

Acquisto, A., Friedman, A., & Telang, R. (2006). Is There a Cost to Privacy Breaches? An Event Study. *Twenty-Seventh International Conference on Information Systems, Milwaukee*, 19.

Agaku, I. T., Adisa, A. O., Ayo-Yusuf, O. A., & Connolly, G. N. (2014). Concern about security and privacy, and perceived control over collection and use of health information are related to withholding of health information from healthcare providers. *Journal of the American Medical Informatics Association*, *21*(2), 374–378. https://doi.org/10.1136/amiajnl-2013-002079

Anderson, C., & Agarwal, R. (2011). The Digitization of Healthcare: Boundary Risks, Emotion, and Consumer Willingness to Disclose Personal Health Information. *Information Systems Research*, *22*(3), 469–490. https://doi.org/10.1287/isre.1100.0335

Anonymous. (2018, November 24). *JOANY - Downtown—Los Angeles, CA*. Yelp. https://www.yelp.com/biz/joany-los-angeles?hrid=N2YrwuHx30f9FczxeKbqng&utm_campaign=www_review_share_popup&utm_medium=copy_link&utm_source=(direct)

Arndt, R. Z. (2018, April 7). *How third parties harvest health data from providers, payers and pharmacies*. Modern Healthcare.

https://www.modernhealthcare.com/article/20180407/NEWS/180409938/how-third-parties-harvest-health-data-from-providers-payers-and-pharmacies

Bandara, R., Fernando, M., & Akter, S. (2017). The Privacy Paradox in the Data-Driven Marketplace: The Role of Knowledge Deficiency and Psychological Distance. *Procedia Computer Science*, *121*, 562–567. https://doi.org/10.1016/j.procs.2017.11.074

Bansal, G., & Zahedi, F. M. (2015). Trust violation and repair: The information privacy perspective. *Decision Support Systems*, *71*, 62–77. https://doi.org/10.1016/j.dss.2015.01.009

Batbaatar, E., Dorjdagva, J., Luvsannyam, A., Savino, M. M., & Amenta, P. (2017). Determinants of patient satisfaction: A systematic review. *Perspectives in Public Health*, *137*(2), 89–101. https://doi.org/10.1177/1757913916634136

Beaver, L. (2018, July 24). *The top 5 startups disrupting healthcare within AI, digital therapeutics, health insurance, and genomics*. Business Insider. https://www.businessinsider.com/7-24-2018-digital-health-startups-to-watch-2018-7

Beckers Hospital Review. (2016, May 18). *98% of digital health startups fail—Here's why*. Becker's Health IT. https://www.beckershospitalreview.com/healthcare-information-technology/98-of-digital-health-startups-fail-here-s-why.html

Belanger, F., Hiller, J. S., & Smith, W. J. (2002). Trustworthiness in electronic commerce: The role of privacy, security, and site attributes. *The Journal of Strategic Information Systems*, *11*(3), 245–270. https://doi.org/10.1016/S0963-8687(02)00018-5

Betts, D., & Korenda, L. (2018, September 25). *Patient Engagement findings—2018 Health Care Consumer Survey | Deloitte Insights*. Deloitte Insights.

https://www2.deloitte.com/us/en/insights/industry/health-care/patient-engagement-health-care-consumer-survey.html

Bijlsma, K., & Koopman, P. (2003). Introduction: Trust within organisations. *Personnel Review*, *32*(5), 543–555. https://doi.org/10.1108/00483480310488324

Bishop, L. "Sam," Holmes, B., & Kelley, C. (2005). National Consumer Health Privacy Survey 2005. *California Health Care Foundation*. https://www.chcf.org/publication/national-consumer-health-privacy-survey-2005/

Black Book Market Research. (2017, January 3). *Healthcare's Digital Divide Widens, Black Book Consumer Survey*. Black Book Market Research. https://blackbookmarketresearch.newswire.com/news/healthcares-digital-divide-widens-black-book-consumer-survey-18432252

Blank, G., Bolsover, G., & Dubois, E. (2014). A New Privacy Paradox: Young People and Privacy on Social Network Sites. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2479938

Boulware, L. E., Cooper, L. A., Ratner, L. E., LaVeist, T. A., & Powe, N. R. (2003). Race and Trust in the Health Care System. *Public Health Reports*, *118*, 8.

Brodkin, J. (2019, November 13). *Google: You can trust us with the medical data you didn't know we already had [Updated]*. Ars Technica. https://arstechnica.com/tech-policy/2019/11/google-you-can-trust-us-with-the-medical-data-you-didnt-know-we-already-had/

Burde, H. (2011). The HITECH Act: An Overview. *AMA Journal of Ethics*, *13*(3), 172–175. https://doi.org/10.1001/virtualmentor.2011.13.3.hlaw1-1103.

Business Wire. (2019, June 25). *The $11.9 Trillion Global Healthcare Market: Key Opportunities & Strategies (2014-2022) - ResearchAndMarkets.com*. Businesswire. https://www.businesswire.com/news/home/20190625005862/en/11.9-Trillion-Global-Healthcare-Market-Key-Opportunities

Caine, K., & Hanania, R. (2013). Patients want granular privacy control over health information in electronic medical records. *Journal of the American Medical Informatics Association*, *20*(1), 7–15. https://doi.org/10.1136/amiajnl-2012-001023

Castell, S., & Evans, H. (2016). The One-Way Mirror: Public attitudes to commercial access to health data. *Ipsos MORI Social Research Institute*, 161.

Chakraborty, R., Lee, J., Bagchi-Sen, S., Upadhyaya, S., & Raghav Rao, H. (2016). Online shopping intention in the context of data breach in online retail stores: An examination of older and younger adults. *Decision Support Systems*, *83*, 47–56. https://doi.org/10.1016/j.dss.2015.12.007

Chang, A. (2018, March 23). *The Facebook and Cambridge Analytica scandal, explained with a simple diagram*. Vox. https://www.vox.com/policy-and-politics/2018/3/23/17151916/facebook-cambridge-analytica-trump-diagram

Chen, A. (2018, July 26). *IBM's Watson gave unsafe recommendations for treating cancer*. The Verge. https://www.theverge.com/2018/7/26/17619382/ibms-watson-cancer-ai-healthcare-science

CMS. (2019, October 29). *Hospital Value-Based Purchasing*. CMS.Gov. https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/HVBP/Hospital-Value-Based-Purchasing.html

CMS Press Release. (2018, March 6). *Trump Administration Announces MyHealthEData Initiative to Put Patients at the Center of the US Healthcare System | CMS*. CMS.Gov. https://www.cms.gov/newsroom/press-releases/trump-administration-announces-myhealthedata-initiative-put-patients-center-us-healthcare-system

Cohen, I. G., & Mello, M. M. (2018). HIPAA and Protecting Health Information in the 21st Century. *JAMA*, *320*(3), 231–232. https://doi.org/10.1001/jama.2018.5630

Cohen, J. (2019, November 13). *Google, Ascension data partnership sparks federal probe*. Modern Healthcare. https://www.modernhealthcare.com/information-technology/google-ascension-data-partnership-sparks-federal-probe

CooperKatz. (2018, February 5). *Paige.AI Created to Transform Cancer Diagnosis and Treatment by Applying Artificial Intelligence to Pathology*. Business Wire. https://www.businesswire.com/news/home/20180205005557/en/Paige.AI-Created-Transform-Cancer-Diagnosis-Treatment-Applying

Copeland, R., & Needleman, S. (2019, November 13). WSJ News Exclusive | Google's 'Project Nightingale' Triggers Federal Inquiry. *Wall Street Journal*. https://www.wsj.com/articles/behind-googles-project-nightingale-a-health-data-gold-mine-of-50-million-patients-11573571867

Council, F. A. (2016, September 19). *12 Ways New Companies Can Build Brand Trust*. Forbes. https://www.forbes.com/sites/forbesagencycouncil/2016/09/19/12-ways-new-companies-can-build-brand-trust/

Cunningham, P. J. (2009). High Medical Cost Burdens, Patient Trust, and Perceived Quality of Care. *Journal of General Internal Medicine*, *24*(3), 415–420. https://doi.org/10.1007/s11606-008-0879-3

Curran, T. (2019, October 29). *Data exchange at UPMC* [Personal communication].

Damschroder, L. J., Pritts, J. L., Neblo, M. A., Kalarickal, R. J., Creswell, J. W., & Hayward, R. A. (2007). Patients, privacy and trust: Patients' willingness to allow researchers to access their medical records. *Social Science & Medicine*, *64*(1), 223–235. https://doi.org/10.1016/j.socscimed.2006.08.045

Davis, J. (2019, April 15). *Third-Party Vendors Behind 20% of Healthcare Data Breaches in 2018*. HealthITSecurity. https://healthitsecurity.com/news/third-party-vendors-behind-20-of-healthcare-data-breaches-in-2018

Dhopeshwarkar, R. V., Kern, L. M., O'Donnell, H. C., Edwards, A. M., & Kaushal, R. (2012). Health Care Consumers' Preferences Around Health Information Exchange. *The Annals of Family Medicine*, *10*(5), 428–434. https://doi.org/10.1370/afm.1396

Donilon, T. E. (2016). Report on securing and growing the digital economy. *Commission on Enhancing National Cybersecurity*.

Doukas, D. J., & Hardwig, J. (2014). Patient Informed Choice for Altruism. *Cambridge Quarterly of Healthcare Ethics*, *23*(4), 397–402. https://doi.org/10.1017/S0963180114000073

Durrah, H. (2019). My Child Is Sick; Don't Call Her A 'Consumer.' *Health Affairs*, *38*(3), 502–505. https://doi.org/10.1377/hlthaff.2018.05012

Elliott, M. N., Beckett, M. K., Lehrman, W. G., Cleary, P., Cohea, C. W., Giordano, L. A., Goldstein, E. H., & Damberg, C. L. (2016). Understanding The Role Played By Medicare's Patient Experience Points System In Hospital Reimbursement. *Health Affairs*, *35*(9), 1673–1680. https://doi.org/10.1377/hlthaff.2015.0691

Farr, C. (2014, June 2). *Apple unwraps "Healthkit" alongside Mac, iPhone features | Reuters*. Reuters. https://www.reuters.com/article/us-apple-developers/apple-unwraps-healthkit-alongside-mac-iphone-features-idUSKBN0ED1V820140602

Feldman, S., & Steenbergen, M. R. (2001). The Humanitarian Foundation of Public Support for Social Welfare. *American Journal of Political Science*, *45*(3), 658–677. JSTOR. https://doi.org/10.2307/2669244

Fogel, J., & Nehmad, E. (2009). Internet social network communities: Risk taking, trust, and privacy concerns. *Computers in Human Behavior*, *25*(1), 153–160. https://doi.org/10.1016/j.chb.2008.08.006

Fouragnan, E., Chierchia, G., Greiner, S., Neveu, R., Avesani, P., & Coricelli, G. (2013). Reputational Priors Magnify Striatal Responses to Violations of Trust. *Journal of Neuroscience*, *33*(8), 3602–3611. https://doi.org/10.1523/JNEUROSCI.3086-12.2013

Francis, L. P., & Francis, J. G. (2017). *Privacy: What Everyone Needs to Know*. Oxford University Press.

Fruhlinger, J. (2020, February 12). *Equifax data breach FAQ: What happened, who was affected, what was the impact?* CSO Online. https://www.csoonline.com/article/3444488/equifax-data-breach-faq-what-happened-who-was-affected-what-was-the-impact.html

Fussell, S. (2020, January 8). *The Sneaky Genius of Facebook's New Preventive Health Tool*. The Atlantic. https://www.theatlantic.com/technology/archive/2020/01/facebook-launches-new-preventative-health-tool/604567/

Galloway, S. (2018, February 8). *The Case for Breaking Up Amazon, Apple, Facebook and Google*. Esquire. https://www.esquire.com/news-politics/a15895746/bust-big-tech-silicon-valley/

Gellman, R. (2011). The Deidentification Dilemma: A Legislative and Contractual Proposal. *Fordham Intellectual Property, Media and Entertainment Law Journal*, *21*(1), 31.

Gensler, A. (2015, July 28). *Trust is the most powerful currency in business*. Fortune. https://fortune.com/2015/07/28/trust-business-leadership/

Gooch, K. (2017, January 3). *Privacy issues drive health IT consumer skepticism: 10 Black Book survey findings*. Becker's Health IT. https://www.beckershospitalreview.com/healthcare-information-technology/privacy-issues-drive-health-it-consumer-skepticism-10-black-book-survey-findings.html

Goold, S. D. (2002). Trust, Distrust and Trustworthiness. *Journal of General Internal Medicine*, *17*(1), 79–81. https://doi.org/10.1046/j.1525-1497.2002.11132.x

Gressin, S. (2018, December 4). *The Marriott data breach*. Consumer Information. https://www.consumer.ftc.gov/blog/2018/12/marriott-data-breach

Griggs, M. B. (2019, November 11). *Google reveals 'Project Nightingale' after being accused of secretly gathering personal health records*. The Verge. https://www.theverge.com/2019/11/11/20959771/google-health-records-project-nightingale-privacy-ascension

Grossklags, J., & Acquisti, A. (2007). What Can Behavioral Economics Teach Us about Privacy? In S. D. C. di Vimercati, S. Gritzalis, C. Lambrinoudakis, & A. Acquisti (Eds.), *Digital Privacy* (pp. 363–377). Auerbach Publications. https://doi.org/10.1201/9781420052183.ch18

Guo, S., & Chen, K. (2012). Mining Privacy Settings to Find Optimal Privacy-Utility Tradeoffs for Social Network Services. *2012 International Conference on Privacy, Security, Risk*

and Trust and 2012 International Confernece on Social Computing, 656–665. https://doi.org/10.1109/SocialCom-PASSAT.2012.22

Gusmano, M. K., Maschke, K. J., & Solomon, M. Z. (2019). Patient-Centered Care, Yes; Patients As Consumers, No. *Health Affairs*, *38*(3), 368–373. https://doi.org/10.1377/hlthaff.2018.05019

Guzman, G. G. (2019). American Community Survey Briefs—Household Income: 2018. *United States Census Bureau*, 13.

Hall, M. A., Camacho, F., Dugan, E., & Balkrishnan, R. (2002). Trust in the Medical Profession: Conceptual and Measurement Issues: Trust in the Medical Profession: Conceptual and Measurement Issues. *Health Services Research*, *37*(5), 1419–1439. https://doi.org/10.1111/1475-6773.01070

Hall, M. A., Dugan, E., Zheng, B., & Mishra, A. K. (2001). Trust in Physicians and Medical Institutions: What Is It, Can It Be Measured, and Does It Matter? *The Milbank Quarterly*, *79*(4), 613–639. https://doi.org/10.1111/1468-0009.00223

Hall, M. A., Zheng, B., Dugan, E., Camacho, F., Kidd, K. E., Mishra, A., & Balkrishnan, R. (2002). Measuring Patients' Trust in their Primary Care Providers. *Medical Care Research and Review*, *59*(3), 293–318. https://doi.org/10.1177/1077558702059003004

HHS Press Release. (2020, March 6). *HHS Finalizes Historic Rules to Provide Patients More Control of Their Health Data* [Text]. HHS.Gov. https://www.hhs.gov/about/news/2020/03/09/hhs-finalizes-historic-rules-to-provide-patients-more-control-of-their-health-data.html

Hill, K. (2012, February 6). How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did. *Forbes*. https://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/#2e7213656668

HIPAA Journal. (2017, January 5). Patients Holding Back Health Information Over Data Privacy Fears. *HIPAA Journal*. https://www.hipaajournal.com/patients-holding-back-health-information-over-fears-of-data-privacy-8634/

HIPAA Journal. (2019, January 28). Analysis of 2018 Healthcare Data Breaches. *HIPAA Journal*. https://www.hipaajournal.com/analysis-of-healthcare-data-breaches/

HITECH Act, Pub. L. No. 111–5, 123 Stat. 115 (2009). https://www.hhs.gov/sites/default/files/ocr/privacy/hipaa/understanding/coveredentities/hitechact.pdf

Hoffman, S. (2020). Citizen Science: The Law and Ethics of Public Access to Medical Big Data. *Berkeley Tech. LJ*, *30*(3), 66.

Huang, B. (2018, September 25). *LVHN patient data not shared with for-profit company in Sloan Kettering trials*. Mcall.Com. https://www.mcall.com/health/mc-nws-lvhn-msk-paigeai-20180924-story.html

Jilka, S. R., Callahan, R., Sevdalis, N., Mayer, E. K., & Darzi, A. (2015). "Nothing About Me Without Me": An Interpretative Review of Patient Accessible Electronic Health Records. *Journal of Medical Internet Research*, *17*(6), e161. https://doi.org/10.2196/jmir.4446

Kantsperger, R., & Kunz, W. H. (2010). Consumer trust in service companies: A multiple mediating analysis. *Managing Service Quality: An International Journal*, *20*(1), 4–25. https://doi.org/10.1108/09604521011011603

Karampela, M., Ouhbi, S., & Isomursu, M. (2019). Connected Health User Willingness to Share Personal Health Data: Questionnaire Study. *Journal of Medical Internet Research*, *21*(11). https://doi.org/10.2196/14537

Kim, J., Kim, H., Bell, E., Bath, T., Paul, P., Pham, A., Jiang, X., Zheng, K., & Ohno-Machado, L. (2019). Patient Perspectives About Decisions to Share Medical Data and Biospecimens for Research. *JAMA Network Open*, *2*(8), e199550–e199550. https://doi.org/10.1001/jamanetworkopen.2019.9550

Kim, K. K., Sankar, P., Wilson, M. D., & Haynes, S. C. (2017). Factors affecting willingness to share electronic health data among California consumers. *BMC Medical Ethics*, *18*(1), 25. https://doi.org/10.1186/s12910-017-0185-x

Kohler, C. (2017, October 11). *This Health Insurance Survey Will Pay You $25 for 10 Minutes of Your Time* [Text]. The Penny Hoarder. http://www.thepennyhoarder.com/make-money/joany-paid-health-insurance-survey/

Krasnova, H., Spiekermann, S., Koroleva, K., & Hildebrand, T. (2010). Online Social Networks: Why We Disclose. *Journal of Information Technology*, *25*(2), 109–125. https://doi.org/10.1057/jit.2010.6

Kuchinke, W., Ohmann, C., Verheij, R. A., Veen, E.-B. van, & Delaney, B. C. (2016). Development Towards a Learning Health System—Experiences with the Privacy Protection Model of the TRANSFoRm Project. In *Data Protection on the Move* (pp. 101–134). Springer, Dordrecht. https://doi.org/10.1007/978-94-017-7376-8_5

Landi, H. (2020, March 4). *Google defends use of patient data on Capitol Hill among scrutiny of Ascension deal*. FierceHealthcare. https://www.fiercehealthcare.com/tech/senators-pressing-ascension-google-data-deal-as-tech-giant-defends-its-use-patient-records

LaVeist, T. A., Isaac, L. A., & Williams, K. P. (2009). Mistrust of Health Care Organizations Is Associated with Underutilization of Health Services. *Health Services Research*, *44*(6), 2093–2105. https://doi.org/10.1111/j.1475-6773.2009.01017.x

Lee, H., Wong, S. F., & Chang, Y. (2016). Confirming the Effect of Demographic Characteristics on Information Privacy Concerns. *PACIS 2016 Proceedings*, 8.

Lee, M., & Lee, J. (2012). The impact of information security failure on customer behaviors: A study on a large-scale hacking incident on the internet. *Information Systems Frontiers*, *14*(2), 375–393. https://doi.org/10.1007/s10796-010-9253-1

Lee, P. (2017, February 16). *Microsoft and partners combine the cloud, AI, research and industry expertise to focus on transforming health care*. The Official Microsoft Blog. https://blogs.microsoft.com/blog/2017/02/16/microsoft-partners-combine-cloud-ai-research-industry-expertise-focus-transforming-health-care/

Lohse, G. L., Bellman, S., & Johnson, E. J. (2000). Consumer buying behavior on the Internet: Findings from panel data. *Journal of Interactive Marketing*, *14*(1), 15.

Luna, J. (2011). *Texas Medical Privacy Act, Health Law & Policy Institute*. University of Houston Law Center. https://www.law.uh.edu/healthlaw/perspectives/Privacy/010830Texas.html

Marr, B. (2018, May 21). *How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read*. Forbes. https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/

Matouschek, N. (n.d.). *Trust Economics: An Economist's Perspective*. Retrieved November 1, 2019, from https://www.kellogg.northwestern.edu/trust-project/videos/matouschek-ep-1.aspx

McCreary, L. (2008, October 1). *What Was Privacy?* Harvard Business Review. https://hbr.org/2008/10/what-was-privacy

McKnight, & Chervany. (2001). Trust and Distrust Definitions: One Bite at a Time. In R. Falcone, M. Singh, & Y.-H. Tan (Eds.), *Trust in Cyber-societies* (pp. 27–54). Springer. https://doi.org/10.1007/3-540-45547-7_3

McKnight, D. H., Choudhury, V., & Kacmar, C. (2002a). Developing and Validating Trust Measures for e-Commerce: An Integrative Typology. *Information Systems Research*, *13*(3), 334–359. https://doi.org/10.1287/isre.13.3.334.81

McKnight, D. H., Choudhury, V., & Kacmar, C. (2002b). The impact of initial consumer trust on intentions to transact with a web site: A trust building model. *The Journal of Strategic Information Systems*, *11*(3–4), 297–323. https://doi.org/10.1016/S0963-8687(02)00020-3

Medicare, C. for, Baltimore, M. S. 7500 S. B., & Usa, M. (2013, June 3). *Privacy Act of 1974*. CMS.Gov. https://www.cms.gov/Research-Statistics-Data-and-Systems/Computer-Data-and-Systems/Privacy/PrivacyActof1974.html

Memorial Sloan Kettering. (2018, September 23). *Memorial Sloan Kettering and Paige.AI*. Memorial Sloan Kettering Cancer Center. https://www.mskcc.org/press-releases/msk-and-paige-ai

Memorial Sloan Kettering Cancer Center. (2014, April 11). *Memorial Sloan Kettering Trains IBM Watson to Help Doctors Make Better Cancer Treatment Choices | Memorial Sloan Kettering Cancer Center*. Memorial Sloan Kettering Cancer Center.

https://www.mskcc.org/blog/msk-trains-ibm-watson-help-doctors-make-better-treatment-choices

Merken, S., & Elfin, D. (2018, October 31). *What's Your Health Data Worth? Startups Want to Help You Sell It*. Bloomberg Law. https://news.bloomberglaw.com/tech-and-telecom-law/whats-your-health-data-worth-startups-want-to-help-you-sell-it

Milewski, D. (2016, September 13). *One in Three Americans Hacked in the Past Year*. HSB. https://www.munichre.com/hsb/en/press-and-publications/press-releases/2016/2016-09-13-one-in-three-americans-hacked-in-the-past-year.html

Minnesota Department of Health. (n.d.). *Adverse Events Reporting—Minnesota Dept. Of Health*. MDH Division of Health Policy. Retrieved May 18, 2020, from https://www.health.state.mn.us/facilities/patientsafety/adverseevents/index.html

Moon, M. (2019, June 27). *Google and University of Chicago face lawsuit over shared patient data*. Engadget. https://www.engadget.com/2019-06-27-google-university-of-chicago-lawsuit-patient-data.html

Nguyen, G. C., LaVeist, T. A., Harris, M. L., Datta, L. W., Bayless, T. M., & Brant, S. R. (2009). Patient Trust-in-Physician and Race Are Predictors of Adherence to Medical Management in Inflammatory Bowel Disease. *Inflammatory Bowel Diseases*, *15*(8), 1233–1239. https://doi.org/10.1002/ibd.20883

Nissenbaum, H. (2009). *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford University Press.

O'Brien, E. C., Rodriguez, A. M., Kum, H.-C., Schanberg, L. E., Fitz-Randolph, M., O'Brien, S. M., & Setoguchi, S. (2019). Patient perspectives on the linkage of health data for

research: Insights from an online patient community questionnaire. *International Journal of Medical Informatics*, *127*, 9–17. https://doi.org/10.1016/j.ijmedinf.2019.04.003

Office for Civil Rights (OCR). (2013). *190-Who must comply with HIPAA privacy standards* [Text]. HHS.Gov. https://www.hhs.gov/hipaa/for-professionals/faq/190/who-must-comply-with-hipaa-privacy-standards/index.html

Ohm, P. (2010). Broken Promises of Privacy: Responding to the surprising failure of anonymization. *UCLA LAW REVIEW*, *57*, 77.

ONC. (2015). Guide to Privacy and Security of Electronic Health Information. *Office of the National Coordinator for Health Information Technology*, 62.

O'Neil, D. (2001). Analysis of Internet Users' Level of Online Privacy Concerns. *Social Science Computer Review*, *19*(1), 17–31. https://doi.org/10.1177/089443930101900103

Ornstein, C., & Thomas, K. (2018a, September 8). *Top Cancer Researcher José Baselga Fails to Disclose Corporate Financial Ties in Major Research Journals*. ProPublica. https://www.propublica.org/article/doctor-jose-baselga-cancer-researcher-corporate-financial-ties

Ornstein, C., & Thomas, K. (2018b, September 20). Sloan Kettering's Cozy Deal With Start-Up Ignites a New Uproar. *The New York Times*. https://www.nytimes.com/2018/09/20/health/memorial-sloan-kettering-cancer-paige-ai.html

Ornstein, & Thomas. (2018c, September 20). *Sloan Kettering's Cozy Deal With Start-Up Ignites a New Uproar*. ProPublica. https://www.propublica.org/article/sloan-kettering-cozy-deal-with-start-up-paige-ai-ignites-new-uproar

Ousfar, E. (2019, October 29). *Facebook announces new health tool that urges users to get preventive care*. News Center Maine. https://www.newscentermaine.com/article/news/health/facebook-announces-new-health-tool-that-urges-its-users-to-get-preventive-care/97-93a45c7f-03e6-47d2-b608-60732a2f64d8

Park, A. (2019, July 2). *12 AI initiatives launched by hospitals, health systems in 2019*. Becker's Hospital Review. https://www.beckershospitalreview.com/artificial-intelligence/12-ai-initiatives-launched-by-hospitals-health-systems-in-2019.html

Perrin, A. (2020, February 4). About half of Americans are OK with DNA testing companies sharing user data with law enforcement. *Pew Research Center*. https://www.pewresearch.org/fact-tank/2020/02/04/about-half-of-americans-are-ok-with-dna-testing-companies-sharing-user-data-with-law-enforcement/

Petronio, S. (2013). Brief Status Report on Communication Privacy Management Theory. *Journal of Family Communication*, *13*(1), 6–14. https://doi.org/10.1080/15267431.2013.743426

Petrow, S. (2018, October 3). Memorial Sloan Kettering, you've betrayed my trust. *STAT*. https://www.statnews.com/2018/10/03/memorial-sloan-kettering-betrayed-my-trust/

Platt, J. E., Jacobson, P. D., & Kardia, S. L. R. (2018). Public Trust in Health Information Sharing: A Measure of System Trust. *Health Services Research*, *53*(2), 824–845. https://doi.org/10.1111/1475-6773.12654

Ponemon Institute. (2018). *2018 Cost of Data Breach Study: Global Overview*. IBM Security. https://www.ibm.com/downloads/cas/861MNWN2

Porter, J. (2019, January 21). *Google fined €50 million for GDPR violation in France*. The Verge. https://www.theverge.com/2019/1/21/18191591/google-gdpr-fine-50-million-euros-data-consent-cnil

Promarket. (2020, February 7). The Real Price of Health Data: Americans Don't Want to Share Their Records for Free. *Pro Market*. https://promarket.org/2020/02/07/the-real-price-of-health-data-americans-dont-want-to-share-their-records-for-free/

Ramsey, L. (2019, April 4). *Amazon's Alexa can now schedule doctor's appointments and give you updates on your prescription drug shipments*. Business Insider. https://www.businessinsider.com/amazons-alexa-adds-healthcare-skills-2019-4

Regalado, A. (2019, February 11). *More than 26 million people have taken an at-home ancestry test*. MIT Technology Review. https://www.technologyreview.com/2019/02/11/103446/more-than-26-million-people-have-taken-an-at-home-ancestry-test/

Reynolds, W. W., & Nelson, R. M. (2007). Risk perception and decision processes underlying informed consent to research participation. *Social Science & Medicine*, *65*(10), 2105–2115. https://doi.org/10.1016/j.socscimed.2007.06.021

Richards, N. M., & Hartzog, W. (2015). Taking Trust Seriously in Privacy Law. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2655719

Roach, J. (2019, October 28). *Microsoft + The Jackson Laboratory: Using AI to fight cancer*. The AI Blog. https://blogs.microsoft.com/ai/jackson-lab-project-hanover/

Ross, C., & Swetlitz, I. (2018, July 25). IBM's Watson recommended "unsafe and incorrect" cancer treatments. *STAT*. https://www.statnews.com/2018/07/25/ibm-watson-recommended-unsafe-incorrect-treatments/

Rothstein, M. A. (2011, February 17). *Debate Over Patient Privacy Controls in Electronic Health Records*. The Hastings Center. https://www.thehastingscenter.org/debate-over-patient-privacy-controls-in-electronic-health-records/

Rushton, G., Armstrong, M. P., Gittler, J., Greene, B. R., Pavlik, C. E., West, M. M., & Zimmerman, D. L. (2006). Geocoding in cancer research: A review. *American Journal of Preventive Medicine*, *30*(2 Suppl), S16-24. https://doi.org/10.1016/j.amepre.2005.09.011

Sankar, P., Mora, S., Merz, J. F., & Jones, N. L. (2003). Patient Perspectives of Medical Confidentiality. *Journal of General Internal Medicine*, *18*(8), 659–669. https://doi.org/10.1046/j.1525-1497.2003.20823.x

Schwarzer, R., & Jerusalem, M. (1992). General Self-Efficacy- Schwarzer (GSES). *Statistics Solutions*. https://www.statisticssolutions.com/general-self-efficacy-schwarzer-gses/

Scott, C. F., Bay-Cheng, L. Y., Prince, M. A., Nochajski, T. H., & Collins, R. L. (2017). Time spent online: Latent profile analyses of emerging adults' social media use. *Computers in Human Behavior*, *75*, 311–319. https://doi.org/10.1016/j.chb.2017.05.026

Seltzer, E., Goldshear, J., Guntuku, S. C., Grande, D., Asch, D. A., Klinger, E. V., & Merchant, R. M. (2019). Patients' willingness to share digital health and non-health data for research: A cross-sectional study. *BMC Medical Informatics and Decision Making*, *19*(1), 157. https://doi.org/10.1186/s12911-019-0886-9

Shaukat, T. (2019, November 11). *Our partnership with Ascension*. Google Cloud Blog. https://cloud.google.com/blog/topics/inside-google-cloud/our-partnership-with-ascension/

Shavers, V. L., Lynch, C. F., & Burmeister, L. F. (2001). Factors that influence African-Americans' willingness to participate in medical research studies. *Cancer*, *91*(S1), 233–

236. https://doi.org/10.1002/1097-0142(20010101)91:1+<233::AID-CNCR10>3.0.CO;2-8

Sheehan, K. B. (1999). An investigation of gender differences in on-line privacy concerns and resultant behaviors. *Journal of Interactive Marketing*, *13*(4), 15.

Shen, N., Bernier, T., Sequeira, L., Strauss, J., Silver, M. P., Carter-Langford, A., & Wiljer, D. (2019). Understanding the patient privacy perspective on health information exchange: A systematic review. *International Journal of Medical Informatics*, *125*, 1–12. https://doi.org/10.1016/j.ijmedinf.2019.01.014

Sinek, S. (2009). *Start With Why*. Penguin Group. https://simonsinek.com/product/start-with-why/

Singer, N. (2019, September 3). When Apps Get Your Medical Data, Your Privacy May Go With It. *The New York Times*. https://www.nytimes.com/2019/09/03/technology/smartphone-medical-records.html

Singer, N., & Wakabayashi, D. (2019, November 11). Google to Store and Analyze Millions of Health Records. *The New York Times*. https://www.nytimes.com/2019/11/11/business/google-ascension-health-data.html

Smith, C. (2011). Somebody's Watching Me: Protecting Patient Privacy in Prescription Health Information Constitutional Constraints on State Health Care & Privacy Regulation after Sorrell v. IMS Health. *Vermont Law Review*, *36*(4), 931–994.

Smith, T., Davern, M., Freese, J., & Morgan, S. (2019). General Social Surveys, 1972-2018. *NORC*, 11.

Solove, D. J., & Schwartz, P. M. (2019). *Privacy Law Fundamentals* (SSRN Scholarly Paper ID 1790262). Social Science Research Network. https://papers.ssrn.com/abstract=1790262

Spencer, K., Sanders, C., Whitley, E. A., Lund, D., Kaye, J., & Dixon, W. G. (2016). Patient Perspectives on Sharing Anonymized Personal Health Data Using a Digital System for Dynamic Consent and Research Feedback: A Qualitative Study. *Journal of Medical Internet Research*, *18*(4), e66. https://doi.org/10.2196/jmir.5011

Stanton, J. M., & Stam, K. R. (2002). Information Technology, Privacy, and Power within Organizations: A view from Boundary Theory and Social Exchange perspectives. *Surveillance & Society*, *1*(2), 152–190. https://doi.org/10.24908/ss.v1i2.3351

Stockdale, J., Cassell, J., & Ford, E. (2019). "Giving something back": A systematic review and ethical enquiry into public views on the use of patient data for research in the United Kingdom and the Republic of Ireland. *Wellcome Open Research*, *3*, 6. https://doi.org/10.12688/wellcomeopenres.13531.2

Sullivan, T. (2017, September 20). *Cedars-Sinai kicks off new health tech accelerator class*. Healthcare IT News. https://www.healthcareitnews.com/news/cedars-sinai-kicks-new-health-tech-accelerator-class

Sztompka, P. (1999). *Trust: A Sociological Theory*. Cambridge University Press.

Taddei, S., & Contena, B. (2013). Privacy, trust and control: Which relationships with online self-disclosure? *Computers in Human Behavior*, *29*(3), 821–826. https://doi.org/10.1016/j.chb.2012.11.022

Tanner, A. (2016, February 1). *How Data Brokers Make Money Off Your Medical Records*. Scientific American. https://doi.org/10.1038/scientificamerican0216-26

Teixeira, P. A., Gordon, P., Camhi, E., & Bakken, S. (2011). HIV patients' willingness to share personal health information electronically. *Patient Education and Counseling*, *84*(2), e9–e12. https://doi.org/10.1016/j.pec.2010.07.013

The Privacy Advisor. (2012, November 5). *Electronic Health Records vs. Patient Privacy: Who Will Win?* ID Experts. https://www.idexpertscorp.com/knowledge-center/single/electronic-health-records-vs.-patient-privacy-who-will-win

Thiede, M. (2005). Information and access to health care: Is there a role for trust? *Social Science & Medicine (1982)*, *61*(7), 1452–1462. https://doi.org/10.1016/j.socscimed.2004.11.076

Tifferet, S. (2019). Gender differences in privacy tendencies on social network sites: A meta-analysis. *Computers in Human Behavior*, *93*, 1–12. https://doi.org/10.1016/j.chb.2018.11.046

Tikoo, P. M. (2014). Evaluating Connecticut's Health Information Technology Exchange. *Connecticut Department of Public Health*, 58.

Tiller, J. (2019, November 12). *If you've given your DNA to a DNA database, US police may now have access to it*. The Conversation. http://theconversation.com/if-youve-given-your-dna-to-a-dna-database-us-police-may-now-have-access-to-it-126680

Tsai, J. Y., Egelman, S., Cranor, L., & Acquisti, A. (2011). The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study. *Information Systems Research*, *22*(2), 254–268. https://doi.org/10.1287/isre.1090.0260

Turow, J. (2017, June 28). *Google Still Doesn't Care About Your Privacy*. Fortune. https://fortune.com/2017/06/28/gmail-google-account-ads-privacy-concerns-home-settings-policy/

Uberoi, N., Finegold, K., & Gee, E. (2016). *Health Insurance Coverage and The Affordable Care Act, 2010-2016* [Issue Brief]. Department of Health and Human Services. https://aspe.hhs.gov/system/files/pdf/187551/ACA2010-2016.pdf

Upton, F. (2015, July 13). *H.R.6 - 114th Congress (2015-2016): 21st Century Cures Act* [Webpage]. https://www.congress.gov/bill/114th-congress/house-bill/6

US Census Bureau. (2019). *Health Insurance Coverage in the United States: 2018*. The United States Census Bureau. https://www.census.gov/library/publications/2019/demo/p60-267.html

Vincent, J. (2016, September 20). *Microsoft announces new AI-powered health care initiatives targeting cancer*. The Verge. https://www.theverge.com/2016/9/20/12986314/microsoft-ai-healthcare-project-hanover-cancer

Vincent, J. (2019, June 27). *Google accused of inappropriate access to medical data in potential class-action lawsuit*. The Verge. https://www.theverge.com/2019/6/27/18760935/google-medical-data-lawsuit-university-of-chicago-2017-inappropriate-access

Wakabayashi, D. (2019, June 26). Google and the University of Chicago Are Sued Over Data Sharing. *The New York Times*. https://www.nytimes.com/2019/06/26/technology/google-university-chicago-data-sharing-lawsuit.html

Waldman, A. E. (2016). Privacy, sharing, and trust: The Facebook study. *Case Western Reserve Law Review*, *67*(1), 193-. Academic OneFile.

Wang, T.-L., & Tseng, Y. F. (2011). A Study of the Effect on Trust and Attitude with Online Shopping. *International Journal for Digital Society*, *2*(2), 433–440. https://doi.org/10.20533/ijds.2040.2570.2011.0052

Warren, & Brandeis. (1890). The Right to Privacy. *Harv. L. Rev.*, *IV*(5). http://groups.csail.mit.edu/mac/classes/6.805/articles/privacy/Privacy_brand_warr2.html

Weitzman, E. R., Kaci, L., & Mandl, K. D. (2010). Sharing Medical Data for Health Research: The Early Personal Health Record Experience. *Journal of Medical Internet Research*, *12*(2). https://doi.org/10.2196/jmir.1356

Willison, D. J., Steeves, V., Charles, C., Schwartz, L., Ranford, J., Agarwal, G., Cheng, J., & Thabane, L. (2009). Consent for use of personal information for health research: Do people with potentially stigmatizing health conditions and the general public differ in their opinions? *BMC Medical Ethics*, *10*(1), 10. https://doi.org/10.1186/1472-6939-10-10

Zhou, J., & Salvendy, G. (2017). *Human Aspects of IT for the Aged Population. Applications, Services and Contexts: Third International Conference, ITAP 2017, Held as Part of HCI International 2017, Vancouver, BC, Canada, July 9-14, 2017, Proceedings, Part II*. Springer.