

UCSF

UC San Francisco Electronic Theses and Dissertations

Title

REFORM BY NUMBERS: Accountability and the Sociotechnical Transformation of American Medicine

Permalink

<https://escholarship.org/uc/item/5s32p1vh>

Author

Cruz, Taylor Marion

Publication Date

2018

Peer reviewed|Thesis/dissertation

REFORM BY NUMBERS
Accountability and the Sociotechnical Transformation
of American Medicine

by

Taylor Marion Cruz

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Sociology



in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Copyright 2018
by
Taylor Marion Cruz

ACKNOWLEDGEMENTS

As a young undergraduate student, one of my first serious engagements with health care began through an internship at a large public safety-net health system. It was quite the eye-opening experience. While most of the other pre-medical students only really paid attention to the doctors and residents we shadowed, I found myself fascinated with everything else but the clinical care. In navigating hallways and waiting rooms, I marveled at the tight orchestration of time and space through the scheduling of patient visits; in delivering supplies and materials to various departments, I lost track of the expansive list of overlapping specialty classifications; in providing health education lessons, I experienced firsthand the challenges of working across barriers of language, literacy, and culture in pursuit of health for all; in visiting the accounting department, I learned about the complexities of financial payment policy and witnessed the behind-the-scenes labor needed to keep the organization running. Each individual clinical encounter depended upon a tremendous amount of “work” to make it all happen, and yet most of my fellow students treated it all as background noise against the real prize of physician shadowing and clinical care.

Looking back upon the completion of my doctorate, so much “work” has been carried out beyond the writing of the dissertation to make it all come together. I am humbled by the amount of care and attention others have shown in my career and overall success, and it would be impossible to recount all of the time, resources, and guidance others have invested in me before and during my time in graduate school. Nonetheless, my goal here is recognize several key people who helped make it all work for me. It is because of my mentors, colleagues, family and friends that I have been able to arrive at this point at such a young age, and I truly appreciate this moment, like the clinical encounters above, as a thoroughly social accomplishment.

My family has cared for me my entire life, and their care continues to shape the way I move throughout the world. From investing in my early education to having patience with my unconventional career path, they have continued to show me an immense amount of love and support over the years. My mother taught me to have compassion and solidarity with other people: her generous spirit continues to remind me of the value of living life in the service of others. My father encouraged me to pursue my dreams, and gave me the work ethic and support I needed to achieve them. My sisters Juliet and Carissa have always found ways to overcome our differences for the sake of the family. Thank you all for being a part of my life today.

I never would have pursued graduate study in sociology without the guidance of my undergraduate mentor, Mary Ingram-Waters at Arizona State University. From our very first meeting so many years ago, she convinced me that I was “already” a sociologist with important things to say about medicine, science, and society. Thank you for re-inviting me to the project of sociology after I had all but abandoned the social sciences. I am so proud to finally join you now in the ranks as a fellow sociologist.

I would not be the scholar I am today without the careful mentorship of my doctoral advisor, Janet Shim. She single-handedly encouraged me to take up STS, medical sociology, and social theory within my work, and worked tirelessly to help me craft a scholarly persona. Thank you for challenging me to do better, and for respecting my decisions in navigating the academic world. You will always serve as a role model and source of inspiration for me within academia and beyond. Howard Pinderhughes pushed me to balance policy relevance with scholar activism, and showed me the importance of working across social worlds. Thank you for believing in the value of my work. Both of you taught me how to “do” sociology, and I will always carry your training with me as I embark upon a lifetime of research, teaching, and mentorship.

I consider myself fortunate to have encountered so many kind mentors and colleagues beyond my main academic advisors. Catherine Bliss, Adele Clarke, and Carolyn Wiener provided helpful advice in completing this project, introducing me to new bodies of literature, theoretical arguments, and methodological devices. Shari Dworkin never wavered in her support of my development as a scholar, even as my research interests changed course early on in my graduate career. Tooru Nemoto's research team showed me how to weave research, service, and activism together in a community setting, and taught me the importance of doing hard, undervalued work. Carl Stempel provided continuous guidance in my initiation as teacher-scholar, which ultimately proved invaluable when navigating the academic job market. Thank you all for being a part of this process with me. My colleagues at UC San Francisco, especially Sara Rubin, Ariana Thompson-Lastad, Sonia Rab Alam, Natalie Ingraham, Michael Levesque, Florencia Rojo, Meredith Van Natta, Kate Darling, Jarmin Yeh, and Chuck Cloniger, deserve special recognition. Thank you for sharing your insight, experience, and friendship with me during our time together in graduate school.

The financial support I received throughout my undergraduate and graduate studies provided me with unparalleled freedom to discover new fields and explore a range of topics, resulting in tremendous personal and intellectual growth. I must acknowledge the generous support I received through the Forsythe Dissertation Award from UCSF Anthropology, History, and Social Medicine, the Harrington and Newcomer Health Policy Scholar Funds from UCSF Social and Behavioral Sciences, the Graduate Research Fellowship from the National Science Foundation, the Graduate Dean's Health Sciences Fellowship from UCSF Social and Behavioral Sciences, and the National Hispanic Recognition Program from College Board and Arizona State University. I am further grateful to the Andrew W. Mellon Foundation and American Council of

Learned Societies, the Ford Foundation, and the Society for the Study of Social Problems for recognizing and honoring my work.

My life mentor, Breonna McCree, has made me the person that I am today outside of the academic life I lead. In her own way, she taught me more than I could ever learn through research or theory. Thank you for taking me in and showing me how to chart my own path in life. My partner, Edward Ramirez, watched this dissertation move from initial project proposal all the way through writing and defense, always knowing when to offer support and when to step back. Thank you for being a part of this process with me, and for living within, beyond, inside and outside of all of the categories with me. You inspire me every day to remember what “truly matters” in life. You both have done the most important work of all over these past six years: you have made me a better person.

REFORM BY NUMBERS

Accountability and the Sociotechnical Transformation of American Medicine

Taylor M. Cruz

Over the past two decades, numerous political actors have called for greater accountability from health care providers on the “value” of the care they provide. Against the backdrop of variation in health outcomes, rising medical expenditures, and persistent health inequities, these actors have mobilized behind the use of quantified quality measures to hold providers to account. Through new mechanisms of public accountability, quantified knowledge has emerged as a central node in the reconfiguration of medical-social relations in the twenty-first century. I examine this widespread phenomenon of “reform by numbers” through a three-pronged analysis. First, I trace the historical drivers of contemporary delivery system reform efforts, highlighting the role of public critique in making sense of the *origin* of quantification. Second, I analyze the deliberation process of three institutionalized quality measures via a national public-private forum, identifying the values and priorities that inform the *design* of quantification. Third, I examine the local politics within new “data-driven care” strategies, emphasizing the emerging *consequences* of quantification. In our current era of cultural anxiety over the value of prominent social institutions, this turn to quantified knowledge, data, and technology is widely expected to address longstanding social problems. However, I argue that this phenomenon merely displaces old conflicts onto emerging technical domains while creating new sources of contention across the rapidly evolving social landscape. This results in the widespread expansion of the institution itself as well as an intensification of the political nature of its activity, ushering in what I term *the sociotechnical transformation of American medicine*.

TABLE OF CONTENTS

1 INTRODUCTION	
Health Care Reform by Numbers.....	1
2 CRITIQUE, REFORM, ACCOUNTABILITY	
The Rise of Quantification in United States Health Care Delivery Policy, 1965-2015.....	23
3 POLITICS AND THE STATE CALCULUS	
Consensus, Conflict, and Controversy in the New Knowledge Democracy.....	53
4 ACCOUNTING FOR DIFFERENCE	
Population Health and the Emerging Technopolitics of Data-Driven Care.....	101
5 CONCLUSION	
The Sociotechnical Transformation of American Medicine.....	131
6 REFERENCES.....	147
7 APPENDICES	
Appendix A: Data Sources.....	160
Appendix B: Tables and Figures.....	164

LIST OF TABLES

Table 3.1. Three Quality Measures.....69

Table 4.1. Four Forms of Difference in Population Health Management (PHM).....114

LIST OF FIGURES

Figure 3.2. Controlling High Blood Pressure (NQF 0018).....	70
Figure 3.3. Ischemic Vascular Disease (IVD): Use of Aspirin or Another Antithrombotic (NQF 0068).....	79
Figure 3.4. Hospital 30-Day, All-Cause, Risk Standardized Readmissions Rate (RSRR) following Heart Failure (HF) Hospitalization (NQF 0330).....	89
Figure 3.5. Final List of Risk Adjustment Variables for NQF 0330 (NQF CV10).....	92
Figure 5.1. The Complexity of Sociotechnical Transformation.....	134

INTRODUCTION

Health Care Reform by Numbers

Taylor M. Cruz

Department of Social and Behavioral Sciences

University of California, San Francisco

INTRODUCTION

Health Care Reform by Numbers

Every day, more and more reports come out concerning the state of the United States health care system. A news story covers the individual health insurance market, focusing on rising premiums and the consequences for taxpayers. Policy briefs and infographics circulate online, flashing estimates of coverage gains and long-term savings. A professional medical society releases an official statement on new recommended guidelines, and other organizations take note and form their own positions on the topic. Doctors affiliated with elite institutions write op-eds on how to best “fix” health care, urging policymakers to include physicians in proposal design. Patients create public posts detailing their frustration in accessing care, and advocates call for additional change building on existing political momentum. An established company announces a strategic investment in the health care industry, claiming to offer yet another business solution within an already crowded space. Comprehensive analytic reports compare the country’s health care with the care received in other Western democracies, identifying serious deficiencies with the most expensive health system in the world.

It would be difficult to overemphasize the significance of health care reform for understanding medicine in society today. The signing of the Patient Protection and Affordable Care Act by President Obama on March 23, 2010 catapulted long-standing tensions in health care squarely onto the national stage. Much of the Affordable Care Act has attracted significant amounts of attention from policy makers, researchers, and the media in addition to the key stakeholder groups of providers, patients, payers, and purchasers. Political debates continue to surround the insurance provisions of the Act as the new administration takes up the challenge of balancing access, cost, and quality within the United States health care system.

But underneath these public considerations of health insurance, party politics, and coverage mandates, there is another set of reforms that have quietly institutionalized the use of data, technology, and quantified knowledge within health care. Beyond the clinical, administrative, and political realms of health that have long been established and co-constituted through each other, an emerging *technical* domain is reconfiguring well-worn territory to generate a new kind of politics. Consider for example the following three occurrences, each taking place after the Affordable Care Act:

September 2010: California Medical Association v. Blue Shield of California

Just a few months after the signing of President Obama's signature health care law, a lawsuit appears in a California court. The subject of the lawsuit is a new "Blue Ribbon" recognition program created by Blue Shield of CA: under the program, the commercial insurer calculates provider performance based on physician-submitted quality data and publishes Blue Ribbons next to certain providers on its website. The goal of the program is to convey quality of care information to patients broken down by individual physician, resulting in a quality rating and ranking system. The California Medical Association claims that the data and published results do not accurately reflect quality of care and result in economic profiling, directing patients towards particular physicians who charge Blue Shield less for services. The judge dismisses the lawsuit. The Blue Ribbon program draws from aggregate data collected by the California Physician Performance Initiative, managed by the large employer coalition the Pacific Business Group on Health.

November 2014: CMS Office of Enterprise Data and Analytics and Chief Data Officer

Deep into the implementation of the 2010 law, the Centers for Medicare and Medicaid Services (CMS) takes the time to restructure its many offices and divisions. In addition to the creation of the new Center for Medicare and Medicaid Innovation (CMMI), the agency announces the opening of the Office of Enterprise Data and Analytics and appoints its first ever Chief Data Officer to head the division. In an accompanying press release, the agency highlights the growing centrality of data for public decision-making in health care, citing the "new responsibilities, including stewardship of the EHR [Electronic Health Records] Incentive Programs, more expansive quality measurement programs, and the establishment of the Health Insurance Marketplaces, [that] have expanded the scope of data that CMS collects." With the new appointment, CMS "signals to the industry that there is no turning back from the health care data agenda," as the then-Chief Administrator of the agency related in the press release. The Chief Data Officer "will help make sure CMS leads the way."

April 2015: Medicare Access and CHIP Reauthorization Act (PL 114-10)

In the midst of heightened partisan feuds over the future of the Affordable Care Act and an impending 2016 election, a rare bipartisan bill passes both chambers of Congress and is quickly signed into law. The enactment of the Medicare Access and CHIP Reauthorization Act (MACRA) marks a significant overhaul to Medicare physician payment policy. The new law aims to move away from fee-for-service to "reward clinicians for providing value, not volume-based care." Combining several existing quality reporting initiatives and funding the development of new performance measures, MACRA modifies provider reimbursement by tying payment to

performance as assessed by standardized metrics. A joint effort from the heads of both parties in the House (Speaker John Boehner, R-OH, and Minority Leader Nancy Pelosi, D-CA), the passage of MACRA reflects a widespread, unifying turn to quality measurement as an expected solution to address longstanding problems in health care.

With sustained attention on certain visible aspects of the Affordable Care Act, such as the individual mandate and the health insurance exchanges, scarce attention has been given to each of these and related events. And yet, these kinds of occurrences reflect a more profound transformation of medical-social relations beyond the politics of insurance coverage reform. Each represent emerging activity around the use of quantified knowledge, data, and technology within health care, and these developments have fully reconfigured the contours of health politics. As data and technology reverberate across the health care landscape, they create new alliances, objects, responsibilities, and conflicts while destroying or modifying previously-established relations. This reconfiguration of existing arrangements, and the sociotechnical transformation it constitutes, is the subject of this dissertation.

THE SOCIOTECHNICAL TRANSFORMATION OF AMERICAN MEDICINE

This *sociotechnical transformation of American medicine*, as I approach it here, comprises three core defining characteristics as quantified knowledge, data, and technology are incorporated into the social landscape of health care:

- (1) The explosion of expected and actual production of *technical knowledge*, including the creation of new records and accounts in pursuit of multiple co-existing objectives;
- (2) The building and institutionalization of a *technical infrastructure* need to generate, share, and disseminate this knowledge to both internal and external entities;
- (3) The disbanding and realignment of people, objects, goals and desires through *technical activity*, including activity that may not take the content of technical knowledge into account.

Because this transformation has been so expansive in a relatively short period of time, in some ways it is difficult to appreciate the depth and profundity of its effects. In the everyday life of providing care, braving political battles, and crunching numbers, every new change feels like

more of the same. And yet, the main argument of this dissertation is that this series of changes have ushered in a new era of health care. If the age of accountable care is the continuation of longstanding dynamics in health care (Starr 1987; Timmermans and Oh 2010), I take this work as my point of departure to consider the turn to quantified knowledge, data, and technology within medical-social relations, as outlined by the characteristic I present above.

I approach this transformation as not merely “social,” “political,” or “technical,” but as *sociotechnical*. Early scholarship in science and technology studies (STS) designates this “sociotechnical” space as not simply the addition of the “social” and the “technical” but as an altogether new type of phenomenon. Consider, for example, Wiebe Bijker’s (1995: 274) early definition:

The sociotechnical is not to be treated merely as a combination of social and technical factors. It is *sui generis*. Instead of technical artifacts, our unit of analysis is now the “sociotechnical ensemble.” Each time “machine” or “artifact” is written as shorthand for “sociotechnical ensemble,” we should, in principle, be able to sketch the (socially) constructed character of that machine. Each time “social institution” is written as shorthand for “sociotechnical ensemble,” we should be able to spell out the technical relations that go into stabilizing that institution. Society is not determined by technology, nor is technology determined by society. Both emerge as two sides of the sociotechnical coin during the construction processes of artifacts, facts, and relevant social groups.

There is a “seamless web” between scientists and engineers, material objects, goals and objectives, interactive end-users, and key stakeholders. Economics, organizations, devices, politics, advocates, and ambitions are all present and co-constitute one another. “Purely social relations are to be found only in the imagination of sociologists, and purely technical relations are to be found only in the wilder reaches of science fiction. The technical is socially constructed, and the social is technically constructed” (Bijker 1995: 273). This is a wholly new domain that not only *modifies* existing arrangements, but actively *transforms* them.

In this new configuration of medical-social relations, many of the commonly recognized players continue to assert a stake in matters of health and health care. Social science scholarship,

for example, has outlined the specific roles played by providers (predominantly physicians), payers (insurers), patients (“consumers”), purchasers (large employers), and policymakers (the state) (Starr 1982; Starr 2011; Light 1993; Hafferty and Light 1995; McKinlay and Stoeckle 1988; Quadagno 2006; Skocpol 1997; Timmermans and Oh 2010; Mayes and Berenson 2006). Professional medical societies, insurance groups, the state’s public agencies, and Congress have all long been involved in the business of health politics. And with the rise of quantified knowledge, data, and technology in medicine, each of these stakeholders has reorganized its political activity around these new technical developments. This is easily illustrated by considering the three post-ACA occurrences presented above. The Centers for Medicare and Medicaid Services creates data-driven programs in pursuit of making care “better,” continuing its longstanding role in shaping the administrative infrastructure of United States health care. The speedy bipartisan passage of MACRA appears as rare Congressional approval to temper physician power by specifically linking reimbursement to “performance” according to quantified metrics, making the shift from “volume” under fee-for-service to “value” through new payment arrangements. Meanwhile, a physician group sues a commercial insurer over its use of a rating and ranking system; the system itself, in turn, is managed by a group of large employers seeking to guide its employees as patient-consumers towards “high-value” care. Few could deny these shifting dynamics brought on by new technical objects, as providers, payers, patients, purchasers, and policymakers all continue decades-long political conflicts in the evolving health care arena. But relationships *across* stakeholders, objects, and practices have also been reconfigured in unprecedented fashion. Social relations are not only displaced to new and emerging technical domains, *but actively remade and transformed with the rise of new forms of knowledge, infrastructure, and activity.*

NEW NODES OF POWER

The sociotechnical transformation of American medicine does not merely consist of the displacement of previous social relations and political conflicts onto rapidly growing technical domains (although this phenomenon has clearly taken place, as illustrated above), but also generates *new kinds of arrangements* within the already crowded and complex arena of health care. If the sociotechnical is *sui generis* as Bijker (1995) claims, then there is now *a new kind of health politics*.

Consider, for example, several emerging nodes of power I identify where new apparatuses, gatherings, and positions each come to play a central role in the reconfiguration of previous social relations. The key nodes of the National Quality Forum, Electronic Health Records, Chief Medical Informatics Officers, and the data scientists and analysts all facilitate particular kinds of work that refract and diffuse established countervailing powers. First, as quality measures have proliferated across the health care landscape, and more and more accountability programs are announced each year, the need for evaluation and review of the measures themselves becomes ever more important. Congress has mandated that a consensus-based entity, the National Quality Forum (NQF), gather different stakeholders (including providers, large employers, patients, and insurers), experts, and users to conduct this evaluation work. NQF reviews measures, votes on endorsement, and provides key input to federal agencies in the use of certain measures within specific programs to meet objectives. Once measures receive NQF-endorsement, they are rapidly integrated into a wide range of accountability mechanisms including national assessments, public reporting, and pay-for-performance contracts. As a key site of decision-making over the use of measures within health care, NQF serves as a critically powerful node in unfolding medical-social relations.

Second, electronic health records (EHRs) have also emerged as a changing force within health care. The paper-to-digital transition was predominantly catalyzed by the Meaningful Use Program from Centers for Medicare and Medicaid Services, which provided financial incentives for the adoption and use of EHRs within clinical practice. But EHR systems do more than eliminate paper records within patient-provider interactions: one technical consultant I spoke with referred to them as “data buckets,” holding valuable documentation of health care encounters that open up the clinical space to new audiences. A programmer I interviewed characterized the local EHR system as “the backbone” of delivery system reform, facilitating all of the coding, data analytics, and report-building that have been absorbed as a part of regular health system operations. The EHR is also critical for external reporting to the state, as data scientists program the technical specifications of each measure to draw from “the single source of truth” (the EHR) to arrive at final performance numbers. Both the technical records and the vendors themselves play a growing role in health care, as the emerging issues of interoperability and data-sharing illustrate.

Third, a new C-suite level position has quietly appeared within health systems across the country. The Chief Medical Informatics Officer (CMIO), a “new career path for young doctors,” as one CMIO I interviewed put it, is responsible for developing local data and technical strategy for the health system. They merge clinical and technical knowledge, interface with administration and physician leaders, and create data policies concerning issues such as governance, integrity, and role definition. The position is bestowed with a tremendous amount of power and authority, as existing powerful groups within each health system come to greatly depend on the CMIO. Quality measures, EHRs, and data reports attract ire and frustration as medicine is reconfigured to absorb these new objects, and CMIOs play a key role in driving

successful social, technical, and clinical integration amidst growing demands from outside the health system.

And finally, the kind of people working in health care has also expanded. In addition to the typical slew of physicians, nurses, managers, executives, receptionists, janitors, and accountants that have long been found in hospitals and clinics, an emerging class of data scientists, technical consultants, and analysts have joined the ranks of health staff. Housed in various units within health systems, ranging from departments of quality, finance, information technology, and analytics, these employees are critical in navigating the ever-evolving health care landscape. They implement EHRs and program customized builds, gather data and conduct statistical analysis, lead meetings disseminating local “findings,” translate the technical specifications of metrics to various audiences, and prepare reports for external entities. They come from a range of backgrounds, including nursing, business, non-profit management, public health, and data science, and find creative ways to integrate into the larger health systems and clinics. They provide much of the behind-the-scenes labor needed to make the metrics, data reports, and EHRs “work,” connecting local systems to the outside world.

This is not an exhaustive list of the new nodes of power appearing in medicine, but it illustrates the full extent of the transformation taking place within health care today. Furthermore, it demonstrates the new arrangements that are neither solely “social” nor “technical,” but *sociotechnical*. New roles are created, and responsibilities are remade and redefined; sites of powerful decision-making shift to new locations; and newly incorporated objects and artifacts facilitate communication across established and new actors. Data, technology, and quantified knowledge have facilitated health care “reform by numbers,” a sociotechnical phenomenon with far-reaching consequences for understanding medical-social

relations in the twenty-first century. It is these widespread changes that are the subject of this dissertation.

THEORETICAL JUSTIFICATION

This project is grounded in the sociology of quantification and evaluation (Espeland and Sauder 2008; Espeland and Stevens 1998; Lamont 2012; Dahler-Larsen 2012; Desrosières 1998), and as such I approach quantification as a site of sociological inquiry. Despite this being a project about numbers, the reader will not find any quantitative analyses in the contents of this or any future draft of this work. I do not aim to make the case for or against a particular program's "effectiveness" or "cost savings" for the purposes of public policy. Nor do I provide an exhaustive review of the several programs and initiatives that comprise this work (in fact, any attempt to do so would quickly become obsolete, as these initiatives are constantly updated and revised much to the frustration of the many actors involved). Rather, my aim here is to document and understand the social and political dynamics within this unfolding arena in order to explain the growing activity taking place within medicine in society today. I draw from Max Weber's (1968 [1922]) orientation towards *interpretivism*, a long-established scholarly tradition within the social sciences, and share his inquisitive spirit towards the unceasing expansion of rationalization and calculation across more and more domains of modern life. This project also emerges out of my longstanding personal interest in the relationship between legitimate knowledge, power, and social change: in our current era of health care reform and political uncertainty, the pursuit of quantified accountability within medicine has proven itself particularly opportune for studying these dynamics.

As a scholarly project, this dissertation also emerged out of my sustained engagement with two contemporary theoretical programs: Bruno Latour's actor-network theory (Latour 1987;

1988; 2005) and Luc Boltanski's sociology of critique (Boltanski 2011; Boltanski and Chiapello 2006; Boltanski and Thévenot 2006). Both seek to rework classical theoretical arguments, and fully incorporate the significance of the technical within their proposed versions of sociology. Each served as a guiding framework not solely informing field site selection and initial points of data collection, but also provided a backdrop against which to consider my analytic findings and present my final write up. I briefly outline both theoretical programs below, focusing on the elements that are most relevant for the purposes of this project.

Bruno Latour's actor-network theory (ANT) (Latour 1987; 1988; 2005) is an established and contested tradition often closely associated with birth of science and technology studies (STS). Alongside the scholarship of Michel Callon and John Law, and later the empirical philosopher Annemarie Mol (Callon 1986; Mol 1999; Law and Mol 2002), ANT has called for social science and STS scholarship to attend to both human and non-human actors, the role of the material world in shaping activity, and the building of institutionalized networks through heterogeneous assemblages. Latour (2005) contrasts this approach to more traditional forms of sociology that claim dominion over "the social," harkening back to Emile Durkheim's (1982 [1895]) *Rules of Sociological Method*. ANT, instead, calls for the sociologist to "reassemble the social" not as a particular material (with its respective placement alongside the "political," the "economic," the "cultural," or the "legal," for example) but as a series of *associations*; in this light, ANT is presented as a means of renewing the very task of sociology itself. There are strong parallels here between ANT and Bijker's (1995) approach to the "sociotechnical" as *sui generis*; indeed, "the social" is not merely one domain to be considered alongside "the technical," but a new seamless web of associations across people, devices, non-human objects, political goals, and knowledge practices.

Latour's privileging of technical accounts is of particular significance for this project. For Latour, numbers serve as "immutable mobiles" that *stabilize* reality and *transport* it to arenas outside of its making, creating new forms of "centers of calculation" that foster *action at a distance* while *acting in multiple places at once* (Latour 1987; Latour 1988). This approach to technical objects is useful for understanding the construction of social-material relations across different physical sites, actors, objects, and goals and desires, and has been used as a starting place for studying clinical standards and decision-tools, infrastructure, the economic forum, and knowledge practices (Berg 1997; Berg and Timmermans 2003; Bowker and Star 2000; Callon, Méadel, and Rabeharisoa 2002; Mol 2002). However, this theoretical account has rarely considered the use of numbers explicitly for political aims, and there is substantial space for considering stabilization and transportation as it appears through on-the-ground material practices. While this theoretical paradigm is certainly not without its critics, its challenge to the sociological project and its privileging of the sociotechnical provide a fruitful opportunity in conducting scholarly inquiry, with considerable space for critique and engagement within social science and STS scholarship.

The theoretical program of Luc Boltanski is no less revolutionary, as he also seeks to rework classical sociological formulations under the banner of a renewed pragmatic sociology. A former student of Pierre Bourdieu, Boltanski's (2011) sociology of critique starts by reworking the very foundation of critique itself. The tradition of "critical sociology," extending back to Marx's 1998 [1845] classic treatise in *The German Ideology* concerning the false conceptions that people have of themselves and social relations, strives to unmask forces of domination for the purposes of social emancipation. Bourdieu is the most prominent contemporary example of this, as he places primacy on the "misrecognition" that common people have of the arbitrary

basis of social classifications while privileging the incorporation of the primary principle of division (e.g., social class) within social analysis, thus striving for liberation through a critical social science. Critical sociology is also closely aligned with the Frankfurt school (Benjamin 2008 [1936]; Horkheimer and Adorno 2007 [1944]; Marcuse 2008 [1964]; Habermas 1981), which has overall experienced significant decline but continues to serve as a place marker within the history of social thought for bringing issues of power and domination back into sociology. The issue, for Boltanski, is critical sociology's strong separation between the everyday actor and the all-knowing social scientist. While both may live in society and experience the effects of structured social relations, Boltanski charges that critical sociology only recognizes critiques as developed from the social scientist by dismissing those placed forth by everyday actors (e.g., by treating the accounts of everyday actors as reflective of ideology and thus rendering them suspect).

Drawing from a pragmatic political philosophy, Boltanski (2011) instead centers his version of sociology around the actual critiques developed by everyday actors, paying particular attention to their ordinary sense of justice. He demonstrates this approach in several important works. The *New Spirit of Capitalism* (Boltanski and Chiapello 2006), written with Eve Chiapello, considers emergent critiques of capitalism surrounding events of civil unrest in France in May 1968: these critiques ultimately led to subsequent changes to capitalist social structure, effectively disarming critique through absorption. The May 1968 events were characterized by widespread student protests and worker strikes drawing attention to “unacceptable forms of reality” that then created the impetus for successive social change. An *artistic* critique that denounced capitalism for (a) fueling relations of inauthenticity and (b) oppressing the autonomy and creative potential of human beings, and a *social* critique that condemned capitalism for (c)

producing massive socioeconomic inequalities and (d) destroying social solidarity and community bonds between the rich and the poor, developed alongside one another, the former by elites, intellectuals, and artists, the latter by radicals, Marxists, and activists. The task here is to link these critiques, which identifies widespread societal problems, to successive social action and reform:

In this book we shall not directly tackle a question dealt by political science and social history: the conditions on which the degree of effectiveness of critique in determinate historical situations depends. Although we shall not ignore the set of factors that condition the vigour and effectiveness of critique, we have focused predominantly on its specifically ideological dimension – that is to say, on the way in which the formulation of indignation and the condemnation of contraventions of the common good operates... [this] highlights an essential part of the work of critique: the codification of ‘what is not going well’ and the search for the causes of this situations, with the aim of proceeding to solutions (Boltanski and Chiapello 2006: 41).

Boltanski and Chiapello argue that these critiques brought about a profound transformation in “the spirit of capitalism,” as evidenced by new management literature of the 1990s, shifting away from older notions of “hierarchy,” “security,” and “structure” and towards newer ones of “flexibility,” “autonomy,” and “networks.” This in turn resulted in the reorganization and structuring of the overarching social order, reconfiguring the distribution of power and ultimately neutralizing further critique. In *On Justification: Economies of Worth* (Boltanski and Thévenot 2006) with Laurent Thévenot, Boltanski examines practices of dispute resolution to uncover the multiple principles mobilized by actors in justifying their social actions. The authors recognize six logics or “orders of worth,” including civic, market, industrial, domestic, inspiration, and fame, that each may serve as alternative frames of reference. By linking critique to reform, and the mobilization of the use of stabilized forms to drive further social action, Boltanski’s (2011) sociology of critique provides ample space for further theoretical examination as it pertains to quantification and social power.

For the purposes of the project, the relationship between critique and social action is of paramount importance. If critique serves to render reality unacceptable by highlighting “what is not going well” (in the case of the May 1968 events, capitalist relations that fueled inauthenticity and stifled creative potential, while perpetuating massive inequalities and compromising communal relationships), social conflict and disagreement is then resolved by “reality tests” (Boltanski and Thévenot 2006). Through these stabilized tests, social heterogeneity is coordinated together for the purposes of evaluation and further social action. Technical means of assessment, such as quality measures, EHRs, and formal accounts may thus serve as key sites filled with rich empirical data. Furthermore, linking critique to subsequent social action – both in terms of concrete policy action as well as the reorganization of social structure itself – provides a model of understanding the dynamics of social change. This approach has experienced much success particularly in economic and cultural sociology scholarship examining processes of judgment, valuation practices, and constructions of quality (Beckert and Musselin 2013; Fourcade 2011; Karpik 2010; Lamont 2009; Lamont 2012; Vatin 2013). Given the contested nature of the site of inquiry, the increasing centrality of technical accounts within the arena, and the pervasive organized activity in pursuit of “value” and quality, this new approach holds particular promise for understanding the dynamics of critique and reform.

My task in this dissertation is to take up both of these theoretical calls, to consider what each has to offer in furthering our understanding of the ever-changing world, and to develop critiques and renewals of these new directions in sociology. Because both centralize the importance of “the technical” – for Latour, as scientific activity and stabilized numbers, and for Boltanski, as evaluative tests used in disputes – this project is well-suited to examine these new theoretical advances through empirical fieldwork. I also aim to make contributions to science and

technology studies and medical sociology more generally, particularly as these fields interface with issues of quantification, power, and state legitimacy.

METHODS AND DATA SOURCES

Methodologically, I draw heavily from an interactionist-inspired grounded theory and situational analysis, closely following the principles placed forth by Kathy Charmaz and Adele Clarke (Charmaz 2007; Clarke 2005). There are several key principles that guide data collection and analysis within these methodological approaches. Overall, grounded theory is an inductive qualitative research method that is characterized by intentional sampling, an emphasis on exploring heterogeneity, the development of categorical themes through coding and memoing, and the simultaneous pursuit of data collection and analysis. Preliminary data analysis may inform future data collection efforts by guiding the researcher in new directions both conceptually and substantively. This results in theoretical sampling, or the identification and intentional selection of additional data sources to gain further understanding of the phenomenon of interest. These methods descend from Weber's (1968 [1922]) interpretivist approach, as presented in the previous section.

As is typical in qualitative research, I draw from multiple heterogeneous sources of data to fully examine the *range* of viewpoints, arrangements, and activities found within the site of inquiry. This results in a deeper, fuller account of social phenomena that yields considerable insight beyond single data sources and standard analytic approaches. Indeed, I not only share my informants' critical eye towards numbers and unmarked data that are expected to serve as final, authoritative accounts (after all, there is not only a politics *of* numbers [Chapter 2], but a politics *in* numbers [Chapter 3] and a politics *with* numbers [Chapter 4]), but empirically demonstrate through rigorous qualitative research the social and political nature of working with, producing,

and acting on quantified accounts. This project reveals the particular strengths of qualitative research and the methodological principles needed when examining quantified knowledge, data, and technology in society.

I draw from several sources of data, presented here in three phases (see Appendix A for full list of data sources). In the first phase, I conducted an in-depth content analysis of technical documents pertaining to quality measurement, delivery system reform, and value-based care. These documents include final reports, event proceedings, transcripts of public meetings, and technical evaluation reports. In selecting national-level policy reports, I prioritized the work published by the Centers for Medicare and Medicaid Services, the National Quality Forum, and legislation and reports to Congress. During initial data collection and analysis, I reviewed eleven reports from the Department of Health and Human Services and the Centers for Medicare and Medicaid Services, six bills recently passed by Congress and four Congressional Research Service reports, and seven NQF reports to Congress and four NQF final project reports, resulting in thirty-two documents. This analysis led me to consider additional sources, including a broad range of historical sources (presented in Chapter 2) and an additional thirty-six NQF documents ranging from meeting transcripts and memos, technical evaluations, public comments, and letters of support from targeted NQF measure projects. These sources provide key insights as to policy-level understanding of health care, the use of measurement, and expected goals of reform.

In the second phase, I conducted in-depth interviews and conference fieldwork to document expert perspectives on these developments. A total of thirteen interviews were conducted with representatives and actors from key organizations, with potential participants identified from the content analysis from the first phase, online searches, other participant referrals, and observations at health care quality related events (described below). Each interview

was recorded and transcribed. I also conducted observations and informal interviews at a total of fifteen health care quality related events that took place between January 2016 and March 2018. These national-level gatherings took place in Washington DC, Sacramento, and the San Francisco Bay Area, with policymakers, experts, executives, and physician leaders in attendance and as presenters. I took fieldnotes of the public statements made at these events during conference panel sessions, keynote addresses, and question and answer sessions, and noted the presence of certain actors and organizations and the interactions between different parties.

In the third phase, I conducted seven months of ethnographic fieldwork at a local health system to examine the implementation of these policy developments on the ground. I shadowed a quality team responsible for the implementation of a large pay-for-performance program within a local health system. The health system is a large, public organization that serves as the geographic area's "safety-net" hospital. I attended meetings led by the team with outpatient clinics, executives, analytics staff, and medical staff, and became involved in their everyday work of reporting writing, data collection, and meeting facilitation. I also conducted observations at select outpatient clinics to examine the work that takes places at the clinical setting to work with the data. In total, I conducted observations twice a week between September 2017 and March 2018, totaling approximately 400 hours of observations. I also conducted a total of twenty-one interviews with the quality team members, medical directors, managers and executives, and data analytics staff to gain perspective on doing this work within the organization.

All data, including documentary materials, interview transcripts, and field notes, were coded through a grounded, inductive process. Segments of data were assigned "codes" or meaningful labels as guided by the researcher's interpretation (Charmaz 2007). This line-by-line

coding process initially resulted in hundreds of codes, which were then organized and clarified through inductive, free-writing memos. Through the process of memoing, I drew connections between different points of data, examined the relationship between different codes, and made decisions that guided further data analysis. These early codes were then reclassified into larger categories of more meaningful codes, eventually building towards the creation of themes. The simultaneous pursuit of data collection and analysis allowed me to change research directions, seek out new sources of data, and verify initial findings throughout the course of the project. All analysis was supported with qualitative analysis software (Atlas.ti).

I also relied on situational analysis mapping techniques to further explore the social arena of interest (Clarke 2005). Mapping creates visual representations of qualitative data in order to examine the connections between different kinds of social phenomena, including (non)human entities, sociopolitical and symbolic elements, and historical and contemporary discourses. Throughout the project, I used an “ordered” situational map to keep a running list of the multiple components implicated in the arena of interest while creating several “messy” maps during the course of analysis. These maps served as a guide to identify missing perspectives within the data, leading me to pursue additional interviews and conduct fieldwork at new events.

CHAPTER OUTLINE

After careful analysis of these sources, I then assembled the emerging phenomenon that I have thus labeled *the sociotechnical transformation of American medicine*. In this write-up of the dissertation, I present my findings in three data chapters, each examining the *origin*, *design*, and *consequences* of quantification, respectively. Taken together, I examine different facets of this emerging social phenomenon: from a macro perspective, I look to the history of health care delivery system reform; from a meso perspective, I consider national-level decision-making and

the politicization of health care within programs of accountability; and from a micro perspective, I examine the shifting locus of control within local data-driven care practices. I present a chapter outline of these components below.

The second chapter, “Critique, Reform, Accountability: The Rise of Quantification in United States Health Care Delivery Policy, 1965-2015,” outlines the role of public critique in making sense of the *origin* of quantification through an examination of the historical drivers of contemporary delivery system reform. By drawing on key policy documents, legislation, and public reports, and the theoretical framework provided by Boltanski (2011), I identify three emergent critiques of medical-social relations (critiques of effectiveness, efficiency, and equity, respectively) that led to the subsequent demand for quantification. Over the years, multiple social groups have become involved in the business of “critique and reform” within health care, and contemporary efforts of quantification seek to induce accountability to make health care more scientific, more market-like, and more equitable. I show that this work of “reform by numbers” has contributed to the expansion of medicine, as more and more people are drawn into the fold in increasingly partial ways. This chapter highlights the role of public critique in making sense of the historical turn to quantification that emerged at the turn of the century.

The third chapter, “Politics and the State Calculus: Consensus, Conflict, and Controversy in the New Knowledge Democracy,” considers the evaluation and design of three institutionalized quality measures within a public-private national forum. By examining the endorsement work that takes place through the National Quality Forum (one of the new nodes of power identified above), I detail the tension between consensus, conflict, and controversy under what I term “the new knowledge democracy” as national decisions are made about what and how to measure through technical measure specifications. I compare findings from national-level

technical reports, meeting transcripts, and public comments with on-the-ground fieldwork at conferences around the country and at one local health system. I demonstrate that in contrast to Latour's "immutable mobiles" characterized by closure and the stabilization of reality, performance measures *generate* politics across the national landscape, even after rigorous stakeholder review and forum endorsement. This chapter considers an emerging "distrust in numbers" as accountability metrics conceal the decision-making behind their technical specifications, creating new kinds of controversy even after achieving "consensus."

The fourth chapter, "Accounting for Difference: Population Health and the Emerging Technopolitics of Data-Driven Care," studies the politics involved in the implementation of data, technology, and quantified knowledge on the ground, revealing the emerging *consequences* of quantification as the local locus of discretion is actively remade. Using the case study of Population Health Management (PHM), a new model of care delivery characterized by the sorting of patients into "populations," I reveal diffuse organizational decision-making over the collection, use, and dissemination of data that builds off on the infrastructure instituted through federal accountability programs. By identifying four population forms widely discussed across the health care delivery landscape – citizens, categories, classifications, and consumers – I demonstrate that the use of data and technology alone cannot address the needs of providers, patients, insurers, and the state, as key decisions must be made regarding the type and kind of data available. This creates the new dilemma of getting the "right data," a new sociopolitical point of contention emerging within local health systems across the country.

In a way that most would have never expected, quantified knowledge, data, and technology have brought about unprecedented change to American medicine. In the conclusion of this dissertation, I provide an overview of the project's theoretical contributions, policy

implications, future research directions, and some closing thoughts on the evolving nature of medical social power in the twenty-first century.

CRITIQUE, REFORM, ACCOUNTABILITY

The Rise of Quantification in United States Health Care Delivery Policy, 1965-2015

Taylor M. Cruz

Department of Social and Behavioral Sciences

University of California, San Francisco

CRITIQUE, REFORM, ACCOUNTABILITY

The Rise of Quantification in United States Health Care Delivery Policy, 1965-2015

In recent years there have been growing calls for “accountability” regarding the performance of the United States health care delivery system. With tremendous variation in service utilization and the production of optimal health outcomes, continued rising health care costs, and widespread treatment disparities across the lines of social difference, health care providers have increasingly found themselves under the scrutiny of various actors, including commercial insurers, public officials, patients and consumer groups, health equity advocates, and policymakers. The health care delivery system depends upon these external entities for financing, patient volume, and political support, but has historically possessed a tremendous amount of power and autonomy in controlling matters of health care. Calls for accountability seek to challenge these social relations in pursuit of health care “value.”

The primary means through which external actors seek to produce accountability for health care delivery performance is through quantification. By measuring provider performance and its connection to health, health care encounters are transformed into data points, and data are used by external entities for further social action. Several mechanisms of accountability have proliferated over the past decade or so: public reporting (the public dissemination of performance information disaggregated by hospital or provider), pay for performance (tying financial payment to hospital or provider performance; also referred to as value-based payment), and disparity identification and monitoring (stratifying performance scores along a point of difference, usually race and ethnicity). This pursuit of accountability in health care delivery is fundamentally dependent upon measurement, as this external evaluation depends on quantification. The

production of these numbers has become a key node in contested efforts to transform the landscape of health care delivery.

This chapter examines the rise of quantification within United States health care delivery policy since the enactment of Medicare and Medicaid in 1965 to understand the social and historical conditions that gave rise to these new forms of social knowledge and action. In alignment with the theoretical program of Boltanski (2011), which recognizes and extends the critical capacity of everyday actors that seek to “render reality unacceptable,” I identify three separate critiques of health care delivery that developed from a wide-ranging group of actors, including physician leaders, public officials, patient-consumers, economists, and health equity advocates. Through distinct critiques of effectiveness, efficiency, and equity, various actors sought to construct health care delivery as a social problem in need of actionable reform: some drew attention to the underlying basis of medical practice, while others noted unusual economic conditions between acting entities, and others linked population health gaps to health care treatment differentials. Actors now seek accountability for the problem of health care through quantification. The sociotechnical project of quantification emerged at the end of the twentieth century as a means of addressing all three of these competing demands: numbers are expected to make health care more clinically effective, more market-like, and more equitable. By considering the virtues of quantified knowledge – its seeming ability to render things transparent, communicate to a wide range of audiences, facilitate standardized comparison, and claim a cultural and political authority due to its association with objectivity – sociological scrutiny of these knowledge-action relations within delivery system reform furthers understanding of how social problems are recognized, collectively defined, and addressed (Blumer 1971). I refer to this sociotechnical phenomenon of critique, knowledge, and social action as “reform by numbers.”

THE CRISIS OF AMERICAN MEDICINE

Paul Starr's (1982) classic *The Social Transformation of American Medicine* traces the institutionalization of medicine within the United States since the late 1700s, with a focus on the medical profession's development of professional sovereignty. For the purposes of this chapter, Starr's historical work on the post-World War II era is of particular relevance. Following World War II, the medical profession and health care industry grew with renewed investment in scientific research, the expansion of Veterans Administration and community hospitals, and successive changes to academic medicine (Starr 1982:338-359). This period, often referred to as "the Golden Age of Doctoring" (McKinlay and Stoeckle 1988; Hafferty and Light 1995), is widely characterized as a time of high levels of public confidence and trust in medical care; it is often taken as a sign of the profession's power and autonomy (Freidson 1970; McKinlay and Stoeckle 1988). The increased uptake of private insurance as a benefit for middle-class workers also contributed to these developments by changing the way the public related to health care.

The passage of Medicare and Medicaid in 1965, however, ushered in a significant change in the federal government's role in health care. The development of these two programs was controversial at the time among health care providers: Starr (1982:375) characterizes these conditions under the "politics of accommodation" as public agencies sought to establish support from doctors and hospitals in initial program implementation. The government initially made limited provisions on payment policy and oversight of health care delivery, adopting the role of payer for the health care of the elderly (and to a lesser extent, low-income people, and later on people with disabilities). But with a new public entity as payer, the federal government came to develop a significant interest in the operation, administration, and performance of the health care delivery system. While many scholars consider the successive loss of general confidence in the

medical profession as part of an on-going corporatization (Navarro 1986; McKinlay and Stoeckle 1988) or deprofessionalization (Haug 1988) of medicine, a growing public interest in the costs and outcomes of health care delivery came to construct health care delivery as an arena of actionable policy concern. It is these latter concerns that are the focus of the chapter.

This chapter contributes to this scholarship through a closer examination of the social and historical conditions of contemporary delivery system reform efforts, with particular focus on the development of quantified metrics with which to hold providers and hospitals accountable for the “value” of care they provide. Most social historical scholarship on health care reform emphasizes policy attempts at establishing national health insurance or achieving universal coverage (Starr 2011; Quadagno 2005; Skocpol 1997; Hoffman 2012). While this work on *insurance reform* is of critical importance, it has left the domain of *delivery system reform* relatively understudied. And while some scholarship has examined the administration of health care programs (particularly for Medicare; Marmor 2000 and Oberlander 2003), relatively little links the social significance of critique to knowledge and action. By studying the issue of “value” across the knowledge-action nexus – that is, the central concern of “what truly matters” in health care, as well as how we know, work towards, and achieve it – social scientific study of delivery system reform offers the possibility of understanding the future of health care as these developments continue forward. Studying the dynamic between critique and reform as an instance of social change also extends the applicability of this work beyond the field of health care. These developments are of particular social significance given our period of widespread social change.

UNDOING CRITIQUE

Social critique has a long history within sociology, extending back to the early writings of Karl Marx. Marx’s (1998 [1932]) critique of ideology for its failure to account for the social

and material conditions of its very form serves as a negative denunciation: to provide a social critique in this sense means to unveil dominant power relations and expose them for all to see. The all-knowing social scientist is able to see beyond the illusory and deceptive practices of ideology that everyday actors are subjected to – and she alone is the central figure integral in bringing about social change. Successive scholars have since critiqued a wide array of phenomena, ranging from the media industry, the mechanical production of art, and cultural taste, utilizing similarly critical approaches that depend upon distance between everyday actors and the critical social scientist (Benjamin 2008 [1936]; Horkheimer 2002 [1972]; Bourdieu 1984). This tradition strives for an emancipatory sociology that challenges dominant power relations through a particular form of critique carried out by the social scientist.

Luc Boltanski's (2011) sociology of critical capacity questions this privileging of the all-knowing scientist over the accounts of everyday actors. In this alternate take, Boltanski displaces the position of the social theorist to recognize the capacity for critique among ordinary actors, whereby everyday actors themselves put forth their own critiques to "render reality unacceptable." Through their critical capacities, actors render certain phenomena as social problems and worthy of attention and amelioration. The accounts of everyday actors are not seen as reflective of dominant ideology: rather, this work of critique is instead to be taken as part and parcel of social change itself (Boltanski 2011). For widespread, far-reaching social change to occur, arenas must be constructed as problems and evaluated, often through competing criteria and conflicting orders of worth (Boltanski and Thévenot 2006). Existing scholarship on the policy process has noted that not all problems are ultimately addressed through political action and policy (Kingdon 1984; Stone 1989), nor are all forms of "dysfunction" collectively deemed

social problems (Blumer 1971). Critique may indeed be used to construct social problems and drive change: the critique, however, may be developed from actors other than the social scientist.

This work on the social foundation of critique also highlights the problem of knowledge in facilitating further social action. Quantified knowledge offers an attractive means of identifying, documenting, and addressing social problems, and may serve as the central node linking critique to actionable reform (Boltanski 2016; Porter 1995). Desrosières' (1998) groundbreaking work in this area notes two possibilities in relating critique to quantification. On the one hand, social actors become involved in critiquing existing forms of measurement (e.g., the construction of the category of unemployment), as measurement practices draw together wide-ranging heterogeneous phenomena across time, space, and entities. Scholarship that “critiques” new forms of health care evaluation falls closer to this line of thinking, with critical scholars denouncing the ranking and rating systems that quantify phenomena within matters of health and health care (Navarro 2001; Lynch 2015). But on the other hand, quantification may also be understood as being driven *by* wider social critiques for the purposes of enacting social change. If quantified knowledge creates objects of permanence so that they might be acted upon for the purposes of reform, then social measurement cannot be easily dismissed when studying social problems and collective efforts to address them. For Desrosières (1998), quantitative knowledge may be simultaneously treated as the *product* as well as the *subject* of social critique, facilitating both social order *and* social change (see also Boltanski 2016).

Drawing from this work, the sociology of quantification has become a growing area of study (Espeland and Stevens 2008), with many scholars studying the proliferation of publicly available quantified knowledge within the realms of law, education, public administration, and human rights and the accompanying set of politics within these arrangements (Espeland and

Sauder 2016; Espeland and Vannebo 2007; Merry 2016; Radin 2006; Cruz 2017). Some scholars link this social activity to the imperative for a society to evaluate itself (Dahler-Larsen 2012), but there are particular features of quantification that render this work of additional sociological interest. Porter (1995) notes that quantification may lend objectivity to public officials when their social authority is questioned, and this “trust in numbers” cuts across traditional political party and ideological lines. Quantitative ways of knowing are bestowed with a tremendous amount of power in the modern world, particularly because of their seeming ability to promote transparency and communicate to all (Porter 1995). Through public dissemination, this knowledge becomes actionable at many levels of consideration as the social phenomenon in question is created as a knowable object (Latour 1987; Callon, Méadel, and Rabeharisoa 2002). The availability of quantified knowledge reconfigures relations between actors and facilitates new kinds of social action, thereby transforming the social landscape.

This chapter contributes to this scholarship on the relationship between quantification and critique by examining the social and historical conditions that gave rise to performance measurement as a critical component of modern-day delivery system reform. I identify three emergent critiques of American medicine and subsequent policy efforts to reform health care delivery social relations. Through separate critiques of effectiveness, efficiency, and equity, respectively, various groups sought to transform health care delivery into an actionable concern: some actors seek to render health care more rational and scientific, others aim to bring the health care arena closer to the market form, and still others strive to make health care treatment equitable and just for all patients and communities. I show how professionals, experts, public officials, and other actors with institutional legitimacy use critique to outline particular kinds of concerns that shape successive policy action. As quantified knowledge aimed at measuring

health care performance is now expected to be used by external entities in multiple ways to reform health care delivery, I argue that the production of numbers is as much the *product* as well as it may be the *subject* of critique, as quantified knowledge is used to induce accountability from health care providers and transform the health care delivery landscape. This suggests that sociological scholarship cannot afford to dismiss the role of quantified knowledge for the conduct of social life through the extension of traditional forms of critique. Rather, it requires that sociologists attend to the production process of quantified knowledge as certain values become institutionalized within measurement practices themselves (Espeland and Vannebo 2007; Cruz 2017), as these commitments are difficult to uncover once knowledge is institutionalized and used for reform and further social action.

THE CRITIQUE OF EFFECTIVENESS

In this section I consider three main developments that contributed to a sustained critique of the effectiveness of medical practice. Various actors, including physician leaders, public health officials, and critical scholars developed critiques of medicine for its missing “scientific basis,” its fundamental role in matters of patient safety and iatrogenic disease, and the prioritization of disease treatment over prevention (Feinstein 1967; Illich 1976; United States Department of Health, Education, and Welfare 1979). These developments eventually gave way to the creation of clinical standards as part of the evidence-based medicine movement; however, this movement has experienced limited success in reforming medical practice, as individual providers have tended to follow guidelines selectively (Gro1 2001; Timmermans 2005; Burstin et al. 1999). The current development of performance measures and their expected use within public reporting and value-based payment initiatives strive to promote adherence to these guidelines by linking performance to status and reimbursement for service. Many of these

measures are developed from the clinical standards created under evidence-based medicine and aim to connect individual practice to standards of evidence. Quantified knowledge is thus expected to reform social action within care delivery to make health care more “effective.”

One of the early primary arguments against the effectiveness of health care treatment was medical practice’s lack of scientific basis. Physician leaders Alvan Feinstein and Lawrence Weed developed critiques of the “art” of medical practice, arguing that clinical care should follow rational, scientific principles. Medical practice was charged as lacking “the scientific qualities of valid evidence, logical analyses, and demonstrable proofs” (Feinstein 1967; originally quoted in Berg 1997). This was a shift in the question of “scientific medicine,” as Berg (1997) notes: the “Golden Years” were widely regarded as such because of the period’s development of a scientific knowledge base for medicine (e.g., the standardization of records and terminology) to be artfully applied to individual patients, whereas the late 1960s were characterized by attention to medical action itself (e.g., the application of medical knowledge in a scientific, rational manner). Komaroff (1982) contends this critique of clinical practice stemmed from the overall perception that medical decision making was made in an unpredictable and irresponsible manner that could negatively impact quality of care. This lower quality of care could then have undesirable societal consequences, such as excessive patient harm or resource waste. A missing scientific basis was seen as the primary issue at stake: how could logical claims be made that health care delivery produced optimal health without an underlying rational practice? By calling for a science of “effective” medicine (Berg 1997), these physician-advocates invited the possibility of scrutiny and judgment of medical practice through critique.

In a similar but distinct vein, others pushed the issue of effectiveness to encompass medical errors and the relationship of health care to patient well-being. Variation in individual

provider behavior could potentially harm individual patients and serve as a source of injury, illness, or infection. Ivan Illich (1976) advanced this problem of iatrogenic disease – illness resulting from medical treatment – in his classic work *Medical Nemesis: The Expropriation of Health*. Illich's work furthered the critique of effectiveness by introducing the possibility that medicine could work against its own claims of producing patient health:

Futile but otherwise harmless medical care is the least important of the damages a proliferating medical enterprise inflicts upon society. The pain, dysfunction, disability, and anguish resulting from technical medical intervention now rival the morbidity due to traffic and industrial accidents and even war-related activities, and make the impact of medicine one of the most rapidly spreading epidemics of our time (Illich 1976:26).

Medical care places patients at risk of disease and illness; thus, health care can produce sickness as well as health. Illich furthered this critique concerning the connection between health care and health, and invited additional scrutiny to the problem of unexamined medical care.

But perhaps the strongest critique of organized medicine emerged from a renewed movement within public health. Thomas McKeown's *The Modern Rise of Population* (1976a) and *The Role of Medicine: Dream, Mirage, or Nemesis?* (1976b) represent a substantial body of work that sought to understand the decrease in mortality and subsequent increase in life expectancy that occurred in the second half of the nineteenth century. In contrast to the conventional wisdom of the time – that these developments were to be attributed to the availability and receipt of improved medical care – McKeown instead charged that these improvements took place through changing environmental conditions. Higher standards of living including improved hygiene and diet driven by economic growth should instead be given proper credit for these population health improvements. This critique has had a lasting effect that continues to be found to this day, despite later research that questioned the specific details of McKeown's research (Colgrove 2002; Link and Phelan 2002). If medical care is not the only

contributor to the attainment of health and longevity, then perhaps environmental factors and social conditions should receive greater attention from those seeking to promote health.

The McKeown thesis found significant currency among researchers and policymakers in the United States (McKinlay and McKinlay 1977; Fuchs 1974; Colgrove 2002). The argument supported a growing movement within public health seeking to challenge curative medicine's emphasis on treatment over prevention. The first *Surgeon General's Report on Health Promotion and Disease Prevention* (1979) served as a landmark publication from the U.S. Department of Health, Education, and Welfare (the historical counterpart to the Department of Health and Human Services) for its marked effort to revisit national health priorities. This report was partly built around a critique of medical effectiveness in achieving health, and indirectly references a version of McKeown's argument:

The health of the American people can and will be achieved – not alone through increased medical care and greater health expenditures – but through a renewed national commitment to efforts designed to prevent disease and promote health... much of the credit [to recent gains in health status] must go to earlier efforts at prevention, based on new knowledge which we have obtained through research (1-1-4).

Citing concern over resource use for health care delivery, the report outlined other means beyond medical care to promote health, thereby challenging medicine's claim to the health of the nation.

In the following year, the Department of Health and Human Services released a national strategy for health improvement, *Promoting Health/Preventing Disease: Objectives for the Nation* (1980), serving as Healthy People's ten year goals for 1990. While Healthy People goals focus primarily on disease prevention, physicians and other providers are also called upon as a significant component in health promotion (e.g., patient education, screening and testing, individual patient behavior change). Healthy People's ten year goals have since been published each successive decade, acknowledging the continued need of change for health promotion.

These early developments gave rise to the evidence-based medicine movement in the 1990s and early 2000s. The medical profession (and the broader health care industry, including hospitals and private insurers) recognized the value of establishing the “effectiveness” of medical care. Through the development of clinical standards, medicine could be made scientific, addressing concerns of variation in practice and patient safety. These standards could also specify provider activities for the monitoring and prevention of disease. Scholars in medical sociology and science and technology studies have elsewhere written about these developments, noting the general unpopularity of these changes among health care providers (Berg 1997; Timmermans and Berg 2003). While evidence-based initiatives stressed the importance of demonstrating results and justifying medical practice, particularly around the problem of variation in clinical practice, evidence-based medicine was met with limited success as providers tend to follow clinical standard guidelines selectively (Grol 2001; Timmermans 2005; Burstin et al. 1999). But the critique of effectiveness firmly established sustained attention to uneven utility and quality of services and the potential result of compromised health outcomes, and many social actors sought to implement clinical guidelines and evidence-based standards in new ways. Performance measurement through standardized quality measures could track whether individual providers comply with evidence-based standards of care for public reporting or payment policy purposes. By measuring health outcomes, comparative health services researchers could also assess which kinds of treatment result in the optimal state of health. Quantification thus came to be valued by external entities in evaluating and promoting the effectiveness of medicine, carrying out the work of previous unsuccessful attempts at reform.

THE CRITIQUE OF EFFICIENCY

A second critique emerged alongside the critique of effectiveness in the latter half of the twentieth century: this section considers growing attention to concerns of medical efficiency. With the development of new hospitals and sites of care, public investment in scientific research, increasing uptake of private insurance among employers, and renewed expansion of health professions training, the health care industry grew tremendously in the post-WWII era with the new influx of financial resources. Clinical medicine also emphasized the patient-provider relationship, often calling for an ethical imperative to provide as much treatment and intervention as possible to promote the health of the individual patient. Emergent critiques from health economists (a then-new group of professional experts) and public officials noted the power of health care providers in price-setting, and highlighted concerns over demand inducement and limited patient-consumer ability to hold providers accountable due to asymmetries of market information. This led to a series of payment reforms in the 1980s and 1990s that sought to control price increases from hospitals and physicians. These reforms, however, did little to control the *volume* of services provided, and some scholars suggest that attempts at cost control resulted in provider increase in the number of services charged to offset potential income losses (Mayes and Berenson 2006; Timmermans 2005). Recent movements toward paying for “value” aim to address this problem by linking financial payment to performance measures and health outcomes: insurers, public officials, and beneficiaries all seek to “pay for what works” in health care, and strive to shift financial risk to providers in pursuit of doing more (e.g., produce health) with less (e.g., fewer resources). These measures are also expected to provide comparative quality information to patient-consumers, thereby potentially creating a market of health care provision and inducing provider competition on “value.” Quantification is expected to specify

and define “value” to communicate to economic entities external to the medical profession, and thus links social knowledge to action under reform efforts to make health care more “efficient.”

The primary issue of health care efficiency concerns the use of financial resources to pay health care providers. Since 1965, total health expenditures have skyrocketed. While most experts contend that the rise in health expenditures is due to a complex combination of factors (Catlin and Cowan 2015; Feldstein 2007), sharp price increases charged to the Medicare program from physicians and hospitals led to an early critique of these providers for their power in influencing overall health care costs under a fee-for-service system. From the initial enactment of Medicare to the early 1970s, rising health expenditures were attributed to non-price factors as more people were covered, with many accessing continuous health care for the first time (Catlin and Cowan 2015). But starting in the early 1970s, price increases played a significant role in rising expenditures. As the federal government was now responsible for the health care costs of public insurance recipients, these price increases served as a significant policy problem. Public officials recognized the need for provider accountability for the growing social problem of cost.

Medicare’s early payment arrangements received immediate attention from health services researchers and policymakers, and fueled the critique of medicine’s unchecked economic power and behavior. Under Medicare’s initial implementation, hospitals were reimbursed on a costs incurred, fee-for-service basis through Medicare Part A; physicians were reimbursed under similar arrangements through Medicare Part B (Starr 1982; Mayes and Berenson 2006). Policymakers and public officials initially sought to allow hospitals and physicians to charge their “customary, prevailing, and reasonable” fees to the Medicare program to assuage potential opposition from health care providers in the early years of implementation. As Robert Ball, former commission of Social Security from 1962 to 1973, recounts:

By and large, our posture at the beginning was one of paying full costs and not intervening very much in how hospitals, at least the better ones, conduct their business... we believed in paying fully. We opposed shifting costs to other payers, and we avoided discounts beyond what our contractors might have secured for their own insured persons (Ball 1995; originally cited in Rosenthal 2015).

This position was driven by the desire to achieve physician and hospital acceptance of public insurance, a particularly contentious issue at the time (Marmor 2000; Starr 1982). But this policy preserved provider autonomy in price setting, which quickly led to rapid escalation of medical expenditures in the years that followed (Catlin and Cowan 2015; Mayes and Berenson 2006). Providers could easily increase the price of their services, as they would simply be reimbursed for all costs charged. Without external evaluation, hospitals and providers had little incentive to keep their prices “reasonable” as this would conflict with their own economic interests.

These financial arrangements drew additional attention from health economists, a new professional group of experts, who further critiqued medicine as the primary driver of rising health care costs under the unusual economic arrangements between providers, patients, and payers under health insurance. Under traditional market mechanisms of supply and demand, consumers and producers have reciprocal roles that maximize market efficiency through the mechanism of price; however, there are several characteristics of health care arrangements that complicate the applicability of this ideal market model to medical care. Kenneth Arrow’s seminal article “Uncertainty and Welfare Economics of Medical Care” is often heralded as the birth of health economics by pointing out these unique considerations. Arrow (1963) identified at least three relevant points that subsequently led to a collective critique of efficiency:

- Neither providers nor patients have an incentive to limit use of health care services or question the utility of services, as true costs are covered by insurers;
- Physicians control the terms of patient patterns of care-seeking, as patient-consumers are dependent on physicians for both diagnosis and recommended treatment;
- Patient-consumers lack the knowledge to evaluate different providers, rendering market logics poorly equipped to drive out lower quality and/or higher cost providers.

The first point highlights the concern of total use of services and health care costs under fee-for-service, in which both “producers” and “consumers” are shielded from the costs of consumption and production (whereas the earlier critique reviewed above concerned price-setting, rather than volume). The presence of a third party payer modifies the behavior of both patients and providers: providers are incentivized to maximize the number of services provided per patient, as payment from the insurer shields patients from the “true” cost of care per service.

The latter two points eventually came to form a separate critique of physicians for independently controlling the terms, content, and volume of care. Aptly referred as “physician-induced demand” by health economists, this critique highlighted the limited ability of patient-consumers to hold providers accountable through traditional market transactions (Fuchs 1978; Hay and Leahy 1982; Dranove 1988). Because of the unique expertise of the medical provider, physicians could order services that would otherwise be refused by patients had they the same knowledge and expertise (Hay and Leahy 1982). The physician’s power in defining, knowing, and evaluating medicine and health care delivery results in an asymmetry of information that challenges the applicability of traditional supply and demand. Furthermore, providers that provide inferior quality services, or that charge abnormally high rates for the same set of services, could not be managed through changes in consumer decision-making. Missing comparative quality and cost information on hospitals and providers limited the ability of patients to behave as informed and engaged consumers: patients have limited opportunity to switch to a new competitor, and have little means of assessing which competitor might offer higher quality and/or lower cost services. According to economists, these market imperfections gave providers a tremendous amount of power in determining overall health care costs.

These concerns materialized into attempts at provider payment reform. Most significantly, Congress sought to establish firm cost controls through centralized price-setting mechanisms. With the implementation of the Prospective Payment System (PPS) in the early 1980s and the Resource-Based Relative Value Scale (RBRVS) in the early 1990s, both hospitals and physicians were no longer reimbursed on a per-charge basis. The Medicare program, operating as a central vehicle of communication between federal policymakers and health care providers, established yearly fee schedules for standard diagnostic codes (for hospitals) and services (for physicians). However, Mayes and Berenson (2006) suggest that the establishment of national fee schedules and caps on yearly growth did little to address the problem of volume – hospitals and providers continue to have little incentive to control the *quantity* of services provided per patient. There was also concern over the limited ability of patients to balance out the power of the providers through the market. These developments expanded the number of actors with a stake in health care delivery: in addition to the policy concerns of Medicare, economists, private insurers, and patient-consumers also gained a significant interest and role in addressing the issue of inefficiency. These actors came to value knowledge of quality generated through measurement for the purposes of driving market-based social action. If quality measures specify “value” by assessing a proper course of conduct (e.g., compliance with clinical standards) or the delivery of desired outcomes, then payers could now “pay for value” by tying financial reimbursement to quality scores, thereby reorganizing economic relations between providers and payers. Through public reporting, quality information could also be disseminated to patient-consumers so that they might make informed decisions about where to seek out care, transforming the problem of information asymmetry and inducing competition between

providers. Quantified knowledge could be used by external entities to create a market of health care, addressing the efficiency of medicine through the linking of knowledge to social action.

THE CRITIQUE OF EQUITY

This section considers the third critique, focused on equity, which further drove concerns of accountability and the subsequent demand for quantification. Concerns of equity developed in relation to the Civil Rights movement and other social activism from the 1950s to the 1970s. The critique of equity within health care treatment drew attention to persistent differences in health across lines of social difference, the role of clinical autonomy in contributing to these differences, and the significance of standard-setting in justifying national policy goals for various people. Increased access to care following hospital desegregation as a condition of participation in the Medicare program did not necessarily result in equally better health across social groups, and public officials, advocates, and activists came to consider the problem of treatment, noting a relationship between receipt of proper health care and optimal health for equity purposes. This led to the development of new Offices of Minority Health within existing public agencies to encourage research on this phenomenon. But while general knowledge of the existence of health care disparities has grown, these inequities have persisted and continue to require actionable reform. Performance measurement could be used to identify and locate differential treatment within sites of care, as quantification allows for the evaluation of treatment across patients and thus facilitates the standard comparison needed to document and act upon inequity. Quantified knowledge here is expected to link critique to reform through social action to make health care more “equitable.”

In 1985 the Department of Health and Human Services’ Task Force on Black and Minority Health released the Heckler Report, a watershed event for policy concern of minority

health. Even after the desegregation of hospital facilities in the 1960s, racial and ethnic minorities continued to have poorer overall health compared to their white counterparts.¹ The Heckler Report emerged as a particularly strong critique of medicine, as it suggested that the persistence of health disparities serves as “an affront both to our ideals and to the ongoing genius of American medicine” (x). The value of medicine could not be heralded as universally benefiting the public, as this “value” takes place against a backdrop of inequity and accrues unevenly along the lines of social difference. The report noted racial and ethnic disparities in chronic conditions such as heart disease, cancer, and diabetes; minorities also had higher rates of homicide, infant mortality, and low birthweight. The existence of these inequities reflected poorly on medical social relations: who was to be held accountable for these conditions?

Much of the report focused on non-clinical factors as an explanation for the population health differential between whites and racial and ethnic minorities, but the report also addressed the problem of health care services and quality of care. While health care itself is not the only determinant of health (see discussion under critique of effectiveness), difference in service use was highlighted as a concern for patients that have already incurred illness, particularly for those living with chronic conditions. But extending the argument beyond the issue of access, the report notes the problem of quality and provider performance once patients do access care:

The narrowing of the disparity in reported use of health services between minorities and non-minorities is an encouraging trend... [however] the indicators themselves do not reflect delays between the onset of problems and the seeking of medical attention, severity of the problem when care is sought, quality of the care received, and whether appropriate referrals are made to specialists (16).

¹ For an excellent overview of health equity initiatives that took place before 1965, including hospital desegregation under *Simkins v. Moses H. Cone Memorial Hospital* and *Eaton v. Grubbs*, see Gamble and Stone (2006), Thomas, Benjamin, Almario, and Lathan (2006), and Quadagno (2000). Hospital desegregation was particularly significant in addressing some of the problems with access to care for racial and ethnic minorities, but it did not address the problem of differential treatment once linked to sites of care. The Heckler Report is significant in first highlighting quality of care as a concern for health equity.

By drawing attention to uneven quality of care, the Heckler Report directly highlighted the problem of unexamined medical practice. If providers provide differential treatment to different patient groups, medical autonomy could be seen as potentially responsible for persistent health inequities, fueling a critique of equity and call for reform:

Appropriate medical care is a major determinant of morbidity and mortality due to cardiovascular disease. Under optimal medical care conditions, for example, a patient with essential hypertension can achieve blood pressure control and reduce the risk of cardiovascular sequelae. However, with *variations in physician behavior* and patient care-seeking behavior, optimal medical care circumstances are difficult to achieve for most minority populations. Simultaneous attention to all the elements of interaction in the medical care setting, *including both patient and physician behavior*, is necessary (emphasis added; 119).

By referencing physician behavior as a potential problem, variation in health care delivery across different providers and patients received renewed attention. Whereas previous critiques viewed variation as stemming from the non-scientific basis of medical practice, here variation is understood through the lens of injustice. Evaluation came to be valued in the pursuit of equity as well as effectiveness, with different actors seeking accountability for health care “value.”

While methods to capture individual patient-provider interactions to quantify gaps in treatment equity may have been difficult to imagine given available technology at the time, the Heckler Report was groundbreaking in calling for HHS to “develop methods to monitor coronary heart disease events that occur in the community, such as: sudden death, hospital admissions and discharges of patients diagnosed as having heart disease, and emergency room visits for patients with chest pains” (44). By tracking statistics such as these, policymakers hoped to be able to identify where health care remained inequitable and track progress in closing treatment gaps. Health care delivery was no longer to be left under the control of hospitals and providers, but to be monitored, tracked, and researched to address the problem of health inequity.

The Heckler Report eventually gave way to the development of the Office of Minority Health at the Department of Health and Human Services in 1986, and other agencies quickly

followed suit: the Centers for Disease Control and Prevention established the Office of Minority Health in 1988, and the National Institutes of Health established the Office of Research on Minority Health in 1990. The creation of offices in these two sub-agencies speaks to the significance of knowledge production in addressing the problem of disparity: health and health care inequity was to be addressed and defined through data surveillance and the production of health information (Centers for Disease Control and Prevention) and through biomedical and public health research (National Institutes of Health). Both of these sub-agencies routinely conduct or fund research on health and health care, including work on disparities.

Another critical development took place through the Healthy People reports, as the logic of disparity itself evolved through the setting of national goals. This version of disparity is conceived as the quantitative difference across a commonly shared covering value. Healthy People 2000 (released in 1990) continued the historical practice of establishing separate target objectives for majority and minority populations. For example, infant mortality rates were to be reduced to 7 per 1,000 births among whites, but targeted to be reduced to 11 per 1,000 births among blacks. In the following report of Healthy People 2010 (released in 2000), separate goals were eliminated, effectively “closing the gap” between different groups (Welsh 1998; Thomas, Benjamin, Almario, and Lathan 2006). With a single target set for all people, differences in common covering values – such as infant mortality rates – across groups came to characterize the problem of inequity. The health and goals of the majority population came to be taken as a reference point, and disparity became understood as the quantitative difference between different social groups (U.S. Department of Health and Human Services 2005).² Disparity is a technical

² In a recent definition of disparity, the National Center for Health Statistics defines the concept thusly: “The quantity that separates a group from a specified reference point on a particular measure of health that is expressed in terms of a rate, percentage, mean, or some other quantitative measure... This definition provides the basis for the direct measurement of the disparities in indicators of health between

form of inequity that depends upon commensuration, as different social entities are ultimately compared against a standard common metric (Espeland and Stevens 1998).

These developments came to inform a movement for health equity that has focused on the elimination of disparities through data collection, research, and public awareness. Much of this activism has variously involved advocacy organizations, public agencies, and health services research to address the problem of health inequity. Concern over disparities and differential treatment eventually attracted the attention of the American Medical Association: its Council of Ethical and Judicial Affairs acknowledged in 1990 that “despite the progress of the past 25 years, racial prejudice has not been entirely eliminated in this country. The health care system, like all other elements of society, has not full eradicated this prejudice” (Council on Ethical and Judicial Affairs 1990:2346). While the movement has had some success in spreading awareness of health and health care disparities, drawing upon the research evidence confirming their existence, additional work remains on how to best eliminate quality of care disparities. General research did not suffice, as it does not link provider accountability to individual patient-provider treatment encounters. Performance measurement, as a means of holding hospitals and providers accountable for treating different patients differently, has attracted attention and support from this movement. Advocates see promise in numbers that provide the standard comparison needed to identify disparity, monitor progress towards disparity elimination, and locate the source of the problem within community across different providers and patients. Quantified knowledge could be used by external entities to address disparity, pursuing equity through actionable reform.

groups. It also provides the basis for monitoring changes in disparities over time, and for making comparisons of disparities across health-related indicators and across geographic areas or populations” (Department of Health and Human Services 2005).

THE MEASUREMENT MANDATE

These three critiques crystallized into several national efforts of performance measurement at the turn of the 21st century. In 1997, President Clinton created the Advisory Commission on Consumer Protection and Quality in the Health Care Industry. The Commission's final report, *Quality First: Better Health Care for All Americans*, called for a public-private forum to develop a comprehensive plan for quality measurement and reporting. The report draws attention to the significance of measurement around each of the critiques reviewed above: through measurement, practice variability could be reduced to have providers comply with standards of evidence (in pursuit of effectiveness), resulting comparative quality information could facilitate improved purchaser and consumer decision-making across providers (in pursuit of efficiency), and measures could drive quality improvements for “vulnerable populations” including racial and ethnic minorities and low-income people (in pursuit of equity). Measurement was thus called upon to continue various programs of reform, as the three critiques each identified the need of knowledge for social action, despite the multiple goals different social actors had to make health care “better.” The report led to the founding of the National Quality Forum, a consensus-based public-private entity that plays a significant role in the use of quality measures for delivery system reform. The Forum evaluates and endorses measures before they are used by external parties, and provides input to the Department of Health and Human Services regarding measure selection and implementation within accountability mechanisms such as public reporting and value-based payment. The Forum draws together different social actors, including health care providers, private and public insurers, employers and purchasers, consumers and patients, public officials and administrators, and health services researchers and statisticians. Across competing critiques and concerns over the “value” of health care delivery,

quantification through performance measurement serves as a central organizing node. It is through quantification that social actors seek knowledge for future social action, thereby bringing about “reform by numbers” (National Quality Forum 2012).

The Commission’s final report also led to the passage of the Healthcare Research and Quality Act of 1999, which reauthorized the Agency for Healthcare Research and Quality (AHRQ) and defined its role in measurement and evaluation. Among other activities, the Act charged AHRQ with primary responsibility for:

- (i) The identification and assessment of methods for the evaluation of the health of [health care recipients];
- (ii) The ongoing development, testing, and dissemination of quality measures, including measures of health and functional outcomes;
- (iii) The compilation and dissemination of health care quality measures developed in the private and public sector;
- (iv) Assistance in the development of improved health care information systems;
- (v) The development of survey tools for the purpose of measuring participant and beneficiary assessments of their health care; and
- (vi) Identifying and disseminating information on mechanisms for the integration of information on quality into purchaser and consumer decision-making processes.

Evaluation, particularly through quantified measures, emerged as a means of knowing and assessing health care delivery and became central to the work of the public agency.

During the same time period, Congress also called for performance measures to assess the problem of disparity. In the Minority Health and Health Disparities Research and Education Act of 2000, AHRQ was charged with responsibility to ensure that measurement work also address disparities in health care quality between different social groups, including between racial and ethnic minority and white patients:

To ensure that health disparity populations, including minority health disparity populations, benefit from the progress made in the ability of individuals to measure the quality of health care delivery, the Director shall support the development of quality of health care measures that assess the experience of such populations with health care systems, such as measures that assess the access of such populations to health care, the cultural competence of the care provided, the quality of the care provided, the outcomes of care, or other aspects of health care practice that the Director determines to be important.

By collecting additional racial and ethnic identification data, treatment and quality differences could be identified by stratifying measure scores across social groups. Available knowledge resulting from measurement could be used to evaluate care and address established problems.

AHRQ was also required to produce national reports of health care quality and disparities, and has since published these two reports annually (National Healthcare Quality Report and National Healthcare Disparities Report; first published in 2003 and since combined into one report in 2014). These developments institutionalized evaluation and measurement work as part of AHRQ's public activity, as the reports serve as a means of ensuring national progress in improving health care value. These critical developments established the basic infrastructure for measurement and subsequent delivery system reform initiatives that foreground the pursuit of accountability under reform by numbers.

THE TURN TO ACCOUNTABILITY

In our current period of health care reform, there has been little social scientific attention to the significance of health care delivery quantification in bringing about system change. But with the proliferation of these quality measures and the work various actors do to create them and put them to use, measurement has come to play a critical role in bringing reform to health care delivery. Our contemporary moment of "accountable care" cannot be understood as distinct from quantification: it is through numbers that providers and hospitals are increasingly held accountable for variation in medical practice, rising health care costs, and persistent inequities. Through mechanisms such as public reporting (the dissemination of provider quality information online), value-based payment (the linking of financial payment to performance metric scores), and disparity identification and monitoring (the tracking of the location of differential treatment

at the local level), evaluation has come to transform American medicine.³ These mechanisms all seek to bring about “accountability” to health care for the problems identified through longstanding historical critiques of medicine. Performance measurement is widely expected to address and improve social relations across different actors.

At the surface level, reform by numbers seems to predominantly affect physicians, hospitals, and other providers. But by serving as institutionalized knowledge expected to represent various stakeholder interests, these quality measures and the information they produce have greatly expanded the number of actors implicated in the problem of accountability. Through public reporting, for example, public agencies are held accountable by consumer watchdog groups to release quality data on competing providers. But patients are also expected to use this information as consumers, and are made accountable for making the landscape of health care delivery closer to the market form. Public and private insurers are held to an account by purchasers (employers or taxpayers) to control costs, and they in turn devise new payment contracts that hold providers accountable for “value.” Healthy equity advocates come to define accountability through quantified measures, transforming the kinds of reform available in the name of justice. Measurement has drawn in new kinds of actors into an increasingly contested social arena, expanding the problem of medical accountability in unprecedented fashion.

³ Recent legislation, including the Patient Protection and Affordable Care Act of 2010 and the Medicare Access and CHIP Reauthorization Act of 2015, pursue value-based payment reform for the Medicare program, and also further public reporting through the development of the Physician Compare website (Medicare’s Hospital Compare website was initially developed in 2005). The Department of Health and Human Services’ (2011) *Action Plan to Reduce Racial and Ethnic Health Disparities* and the Centers for Medicare and Medicaid Services’ (2015) recent *Equity Plan for Improving Quality in Medicare* both call for the use of performance measures to locate disparities within the community and develop targeted interventions to “close the gap.” The perceived value of these measures spans across the policy arena in addressing the social problem of medical care.

The pursuit of accountability through quantification is responsible for the transformation of the health care delivery landscape. Through performance metrics, social responsibilities and expectations are altered, new forms of health care organization emerge, relationships between entities are created and destroyed, conceptions of justice and equity are defined, and health care activity is reconfigured across an increasingly greater number of social groups. Yet traditional sociological critiques of quality measurement and evaluative practices fail to appreciate the extent of this social change and its connection to public critiques of medicine by multiple social actors (Navarro 2001; Lynch 2015). In this case, physician leaders, public health officials, health economists, and equity advocates all sought to construct medicine as a social problem, and valued knowledge of health care delivery for different kinds of social action for reform. While Boltanski (2011) privileges the everyday actor in everyday situations in decentering critique, I instead show how professionals, experts, and entities with access to institutional legitimacy develop critiques to outline problems that then shape successive policy action. Critique is indeed not an activity reserved for the all-knowing social scientist, but a fundamental activity central to social life. I also show how the common desire for social knowledge brought competing critiques together, despite quite divergent expectations for resulting social action based on this knowledge.

While quantification must be appreciated as a means of attempting to accommodate wide-ranging critiques of medicine across several decades of health policy and advocacy, few social actors appreciate the full extent of this resulting delivery system transformation. If metrics identify, measure, and encourage “what works” and “what matters” in health care in the name of “value,” then they conceal social and political conflicts in defining these values (Espeland and Sauder 2016; Radin 2006). The numbers are expected to speak for themselves, but impose their own form of social order on medicine (Porter 1995). Because of the cultural and political

authority afforded to numbers, it is difficult to oppose quantification and accountability: one can only suggest that the evaluation be done in a different way, thereby shifting social and political conflicts under technical practices of measurement. How are decisions made about what to measure and how to measure it when there are multiple perspectives on the precise “problem” in health care to be addressed? What kinds of values and priorities become embedded within available quantified metrics, and how do they shape the pursuit of “effectiveness,” “efficiency,” and “equity” under programs of accountability? When metrics are used to implement standard rules, make decisions for people, and communicate to new audiences of actors, the social phenomena under consideration transform in unexpected ways (Espeland and Sauder 2016).

CONCLUSION

In connecting critique to reform, it is clear that quantification is just as much the *product* of social critique as it might be the *subject* of typical sociological critique. How else could competing groups with different kinds of problems and concerns come to value a particular kind of knowledge in making health care “better”? Quantification in this case emerged across several critiques from social groups with conflicting goals, varying ideological and political standpoints, and differing degrees of power. The unexpected alliance across social actors with different concerns of the problem of “value” reveals the powerful allure of numbers in the modern world. Quantification offers an attractive form of knowledge for the purposes of reform: and yet when used in this capacity, it defines and stabilizes reality in a way that renders it susceptible to further interrogation. Alain Desrosières (1998) writes of this tension between social order and change in connecting statistical knowledge to public debate:

We *must* have things that hold up well, independently of particular interests, in order to be able to act on them. These things are categories of action: poverty, unemployment, inflation, the trade deficit, monetary mass, fertility, causes of death. The language used is pragmatic: means toward an end... [but] we *must* [also] open up the black boxes to show what they conceal... [social forms]

must simultaneously remain undebated so that life may follow its course, and debatable, so that life can change its course (Desrosières 1998:336-337).

If critique strives to render reality unacceptable, and quantification creates objects of permanence that stabilize reality so that it might be acted upon, then social order and conflict co-exist under “reform.” Through quantification, social actors establish order through formal “accounts” that draw society together in particular kinds of ways, a controversial practice that remains open to public debate. The issue for a thorough sociology of quantification is a sustained engagement with the multiple *kinds* of quantification that are invariably possible. Measurement decisions are consequential as they shape which features and aspects are deemed important and which are not, exposing certain issues and ignoring others. Quantified knowledge may bring about “reform by numbers,” but this reform may also forgo that which is not counted. Social action that depends upon this knowledge cannot interrogate the underlying basis of this information, as assumptions are difficult to trace once social and political disagreements are concealed behind the numbers.

As social scientists, we must fully appreciate the politics of quantification. Numbers are not simply tools to challenge traditional power arrangements between providers, hospitals, patient-consumers, insurers, public officials, economists, and health equity advocates. Numbers also come to transform health care itself, focusing on that which can be made quantifiable and rendered as a problem. Moving forward, these sociotechnical connections between quantification, power, and social relations will come to dominate the future of health care delivery as they currently shape the present. Performance measurement will come to define medicine itself, and we may all be held accountable for the social problem of medical care.

POLITICS AND THE STATE CALCULUS

Consensus, Conflict, and Controversy in the New Knowledge Democracy

Taylor M. Cruz

Department of Social and Behavioral Sciences

University of California, San Francisco

POLITICS AND THE STATE CALCULUS

Consensus, Conflict, and Controversy in the New Knowledge Democracy

In this chapter I examine the public evaluation of a set of quality measures to trace the politics of quantification for the purposes of political accountability. I consider the work of the National Quality Forum (NQF), and the organization's evaluation of three health care quality measures, all related to heart disease and health but each representing a distinct measure type. The in-depth case studies I present in this chapter draw from policy legislation and legal orders, documentation of technical evaluations of several NQF quantification projects, and informant interviews with actors involved in measurement endorsement and related policy. Through this public evaluation work, I identify consensus, conflict, and controversy related to (a) the interpretation and use of scientific evidence, (b) the distribution and interconnected-ness of responsibility within health care encounters, and (c) the perceived boundary between health systems and remaining society. These issues reflect much deeper political conflicts than mere issues of measure implementation (e.g., issues of data collection burden, federal and state oversight), as the difference in values that is presumed in models of democracy and political deliberation are ultimately flattened as they confront quantification's requirements for standardization. As measures receive public endorsement and are implemented within programs of accountability, however, they do not simply settle controversies or stand in for previous social arrangements, but transport these issues to new sites, thereby reenacting conflict and consensus in multiple local sites at once. This results in a widespread, unprecedented expansion of the political nature of health care, fueling a new kind of politics under what I term "the new knowledge democracy."

This chapter is organized into four sections. First, I review literature on the relationship between the state and social knowledge, tracing the emergence of a new form of political quantification with the rise of benchmarking; this form creates the conditions for emerging form of knowledge democracy. Second, I consider theoretical arguments from Bruno Latour and Luc Boltanski that consider the relationship between consensus, conflict, and controversy as it pertains to technical accounts (in this empirical examination, quantified measures used within health care accountability programs). Third, I present in-depth case studies drawing on national evaluation work facilitated by the National Quality Forum, the consensus-based entity charged with endorsing quality measures before use in accountability programs. Finally, I consider how this national endorsement work of achieving consensus creates conflict and controversy on the local level as measures are implemented across the national landscape, making a difference in multiple places at once. In contrast to Latour, who approaches numbers as “immutable mobiles” that readily stabilize reality to foster action at a distance, and Boltanski, who approaches technical evaluation as a means of resolving conflict through the use of agreed-upon criteria, I suggest that this endorsement work actually *obscures* the complex administrative work that goes into endorsing the measures. Conflicts and points of contention that materialize during evaluation work do not close off even after committee-approved endorsement, but rather reemerge at local sites as on-the-ground actors make sense of implementing the measures. Because these measures are presented as simply as possible, even with complex technical specifications, they create conditions where those working with the measures on the ground struggle to make sense of decisions made in the making of the measure, which further creates conflict as they are implemented within the health system overall. Ultimately, I argue that this

interplay between simplicity and complexity results in the measures being *political by design*, creating a new form of “politics of numbers” under the emerging knowledge democracy.

THE STATE AND SOCIAL KNOWLEDGE

The relationship between social knowledge and the state has long attracted scholarly interest. One of the primary uses of social knowledge is for the purposes of *state-building and administration*. For example, much social science attention has focused on the Census as an official means of “knowing” the population. Anderson’s (1988) historical work in this area suggests that the establishment of the Census served as an integral solution to the balancing of power between states by providing knowledge of the population for the purposes of political representation and taxation. Benedict Anderson’s (1983) *Imagined Communities* also considers the Census alongside the map and the museum as a tool of establishing national identity. Rueschemeyer and Skocpol (1996) have considered the centrality of official social scientific knowledge, including both economic and sociological information, for the project of nation-building across Western democracies. Scott’s (1998) classic work claims that modern states have failed to achieve the large-scale promise of modernity and progress precisely because the state “sees” society through large scale surveys that are ultimately unable to account for the reality of social heterogeneity. As this form of knowledge pertains to health and health care, national health and social surveys provide an overall picture of the state of the nation’s health, and the state collects operational and administrative information on the hospitals and health systems of the country. Many of these scholars see the collection of social knowledge as central to the founding of the state itself.

A second relationship between social knowledge and the state highlights the use of quantified data for the purposes of *social reform*. Numbers play a critical role in facilitating the

construction of policy problems (Stone 2002; Alonso and Starr 1987). Desrosières (1998) argues that national social surveys may be used to address social ills – for example, the calculation of income thresholds based on aggregate individual incomes to determine a national poverty line. In Durkheimian fashion (1982 [1895]), statistics express a generalized *collective form* that may be used by the state to identify problems across the national landscape. Surveillance statistics track mortality and morbidity, births and deaths, and prevalence and incidence of certain diseases; panel surveys on income, education, and occupation provide data of the nation’s workforce, including unemployment; and national social surveys provide information of the changing demographics of particular regions and across the country as a whole. Within health and health care, social scientific analysis of survey data identifies disparities across different patient populations, and economists calculate pricing and assess market conditions to identify imperfections and propose solutions. In this sense, numbers are used for *social reform* in a general policy manner.⁴

Both of these scholarly areas have contributed much to our understanding of the state and social knowledge, but they are unable to account for what I identify as a third form of state-social knowledge relations. In our current era of accountability, knowledge is no longer used merely for *state-building and administration* or *social reform*, but instead is used for *standard-setting*.

Benchmarking and associated comparative practices are quickly emerging within education, law,

⁴ Some scholars would ask for the recognition of a fourth perspective – that of Foucauldian governmentality, particularly in the spirit of Nikolas Rose (Rose and Miller 1992; Rose 1991). It is true that work on state governmentality – the making of a governable population, the spread of self-calculation among individual people, and so forth – is also concerned with the state and calculation. However, this work is more concerned with issues of human *subjectivity* (as Rose puts it, “governing the soul”) that are not the focus of this project. Rather, this project seeks to understand the link between the state, institutionalized knowledge, and subsequent state-society action. The Foucauldian approach, in contrast, is more concerned with reconstitution of personhood than concrete social and political action stemming from institutionalized knowledge.

public administration, and digital platforms (Espeland and Sauder 2016; Espeland and Vannebo 2007; Merry 2016; Scott and Orlikowski 2012; Bruno and Didier 2013). By collecting information on system “performance,” making this information publicly available, and using this knowledge to drive further social action, the state sets standards through the use of social knowledge around which other actors, political groups, organizations reorganize their relationships and activity. This differs from state administration, as it is an *empirical* project of physically building material things, reconfiguring and reorganizing things. That is to say, the state now aims to use social knowledge to actively remake the world it attempts to govern. In health care, this emerging relationship between the state and social knowledge is evidenced by the tracking of individual and organizational performance on health care quality and costs, which is then used to drive further state and social action.

This emerging relationship between social knowledge and the state creates the conditions for what I term “the new knowledge democracy.” If social knowledge is now being used to set standards, monitor performance, compare alternative entities, and inform successive decision-making and social action, then what form should this centralized knowledge take? The stakes for this issue are clearly quite high. As this social knowledge is generated through mandated reporting and data collection, it shapes the way public agencies “see” the social world, focusing on particular aspects and features over others. As quantified knowledge is publicly reported, traditional conceptions of status and reputation are reconfigured through direct comparison and competition across entities. As this information on “performance” is linked to financial payments and reimbursement, organizations risk losing millions of dollars on the line based on metric scoring. To determine the form of this centralized knowledge that has far-reaching consequences across actors, organizations, and institutions within a given social landscape, the democratic

forum has emerged a key location in which decisions are considered about what to measure and how to measure it. The forum consists of the gathering of different stakeholders to take into account their different perspectives on the measures themselves, and to come to “consensus” on the standard-setting in question. Under the aim of making new domains know-able and action-able, more and more people are actively drawn into debates over what “counts” and what “matters” in remaking the social.

EVALUATION IN THE NEW KNOWLEDGE DEMOCRACY

If social knowledge is being used to transform existing relations, and the state is leading many of these efforts, this brings to light new kinds of questions. Under a knowledge democracy, what should these forms of quantification look like? Evaluation, and its subprocesses of comparison, commensuration, and translation, are quickly emerging as core activities in the reconfiguration of prominent social institutions (Lamont 2012; Dahler-Larsen 2012; Espeland and Stevens 1998; Hank and Severi 2014). Under Bourdieu’s influence, various scholars have studied the evaluative cultures of various institutions (Lamont 2009; Beckert and Musselin 2013; Fourcade 2011; Stark 2009), some with particular focus on the sociotechnical apparatuses that serve as “judgement devices” (Karpik 2010; Callon, Méadel, and Rabearisoa 2002). In this section I review two traditions that make serious theoretical contributions to this work. Latour’s actor-network theory (1987; 1988; 2005) from science and technology studies (STS) provides a starting point in making sense of the sociotechnical devices, activities, and frames of reference that revolve around standard quantified accounts, particularly around the “settling of controversy” within the work of building particular apparatuses (in this case, performance metrics). Boltanski and Thévenot’s (2006) post-Bourdieuian work draws from political philosophy to recognize multiple orders of worth within the practice of negotiating “conflict and

consensus” through evaluative *reality tests* (here, the use of performance metrics to assess the health care activity to be evaluated). Each tradition offers a competing approach to the social scientific study of evaluation, but both privilege establishment of technical means for the purposes of further social action.

Latour’s actor-network theory centers on “feeding off of scientific controversies.” Controversy serves as a starting point in opening up the “black box” of science and technology (Latour 2005). In early versions of his work, he privileges the closing off of these controversies – the settling of controversies – as part of the activity of conducting science. His conceptualizations of the transcending of distance – and in particular, *action at a distance*, and *making a difference in multiple places at once* – are frequently cited as among his major contributions to scholarly thinking (Latour 1987; Latour 1988; Rose 1991; Robson 1992). This links the activity within an initial arena (e.g., the laboratory) to the broader world outside of it (e.g., economy, education, infrastructure) through network building, thereby transcending distance and extending widespread pervasive reach across the social world. Other STS scholarship similarly links the issues of transcending particular arenas to the development of authoritative accounts overall, considering scientific activity as fundamentally dependent upon the cultivation of skills so as to produce an aperspectival way of “seeing” the world (Porter 1995; Daston 1992). That is to say, science and technology are organized around standardization and closed off controversies for the purposes of coordination, whether this is done to foster public action at a distance or in the pursuit of scientific knowledge across laboratories, geographies, and cultures. Numbers serve as “immutable mobiles” that transcend local sites, transporting certain things and concealing the work it takes to develop them through the “settling of controversies”

(Latour 1987). As representations, numbers “stand in for other things,” resulting in *strong explanations* the account for heterogeneity within a standard account (Latour 1988).

While Latour builds the case for knowledge-action relations based on the model of science, Boltanski and Thévenot (2006) instead draw from the model of politics. By looking to political philosophy, Boltanski centralizes heterogeneity and difference in values, ideologies, and interests across different social actors, and the work they do to outline particular problems and act upon them (Boltanski 2011). This scholarship begins by considering the resolution of public disputes, taking conflicts across difference seriously. For Boltanski, conflict is resolved through the use of evaluative tests against which social heterogeneity is rendered intelligible and actionable. Evaluation is the means through which we know the world, and may be used to settle “conflicts.” By subjecting social phenomena to tests, actors must agree on the importance of different criteria, the best means of making assessments, and the merit of the different arguments put forward. In other words, legitimacy is both made and sustained through a process of *evaluation* in which both consensus (establishment of standard evaluation tests) and conflict (social heterogeneity to be evaluated) can be found. Through justification and critique, actors work together to create common forms of generality (i.e., categories); these can then be used to compare between competing alternatives (Boltanski and Thévenot 2006). Following this theoretical line of inquiry, the task for the sociologist is it make sense of how different people, groups, and objects with competing goals, desires, and values come together to form a meta-framework against which heterogeneity is compared and valued. The development of this framework is arguably of central concern, since it is how things are made intelligible, recognizable, and actionable.

Evaluation of quantified measures – testing through the use of stabilized forms – lies comfortably at the center of both of these approaches, drawing upon both models of science and politics, the two poles of standardized simplicity and heterogeneous complexity. On the one hand, there is endless conflict within society and social reality. On the other hand, consensus exists in the construction of common *reality tests* that assess activity. Together, these tests highlight new information, shed light on particular issues, and thus work to bring about change to the social world, but the evaluation is a common starting place where consensus and conflict readily co-exist. I consider these tensions within evaluative work by consider the endorsement work of the National Quality Forum, where measures are evaluated on the national level before being implemented programs of accountability and translated to local sites of care.

DATA AND METHODS

This chapter draws from several sources of data, including: (1) policy legislation specifying the expected work of a measurement consensus-based entity (e.g., National Quality Forum) and associated legal orders, (2) technical evaluation documents produced by the National Quality Forum (NQF) for its endorsement work, and (3) informant interviews with actors involved in measurement endorsement and related policy, as well as interviews and fieldwork at a local health care delivery system. I review these sources of data below.

I first reviewed annual reports to Congress from NQF (2009-2015) to outline the scope of work as presented to the state. These reports referred to key legislation and legal documents, and identified issues and concerns from all measurement activities for the respective year. I then tracked down the referred legislation and associated legal documents for additional review. After considering preliminary insights from these reviews and overall project aims, I narrowed the scope of my analysis to consider measures related to cardiovascular disease. I then downloaded

all technical documentation associated with measure endorsement for projects related to heart disease. Between 2009 and 2015, NQF endorsed heart disease measures in four separate projects, each titled “Cardiovascular Measures 2010, 2013, 2014, 2015,” respectively. I then conducted an in-depth examination of these technical documents, starting with a coding of the final reports for each of the four projects. These final reports provided an overview of the measures endorsed, emergent issues in the endorsement process, and suggestions for future work.

Based on this coding and review, I purposively selected three measures (controlling blood pressure, use of aspirin for Ischemic Vascular Disease (IVD) patients, and hospital readmissions following heart failure; see Figures 3.1-3.4 throughout chapter). These measures each represent a different type of measure (outcome, process, administrative), and each are widely used within various federal and state programs of accountability. I then conducted a targeted analysis of the technical documents, focusing on all documentation corresponding to each measure. These documents consisted of technical evaluation reports, committee meeting transcripts and memos, records of public comments and committee responses, and endorsement voting records. I then backtracked and compared my findings with the NQF annual reports to Congress to verify the adequacy and scope of my analysis.

I also conducted several informant interviews with various actors involved with the endorsement process, including NQF staff, committee advisors, and voting members, which further confirmed the insights from the documentary analysis. Additional fieldwork and interviews with actors at a large health system that continues to carry out implementation work for the selected measures further identified how the issues considered in the national evaluation were then transported to the local level. I also draw on fieldwork at national health policy and quality measurement conferences where representatives discussed issues of implementing

measures within their respective health systems. All of the qualitative data presented above were analyzed using the methodological principles of grounded theory and situational analysis (Charmaz 2007; Clarke 2005). For this phase of the project, theoretical sampling emerged as a strong guiding force, as initial findings and leads informed the selection of successive documents and sources of data. I also purposively selected the three measures with the intention of conveying a wide range of heterogeneity within the presentation here.

NQF AND ENDORSEMENT

To carry out the activities required for these forms of evaluation under the new knowledge democracy, Congress has specified the use of a “consensus-based” entity. This entity, which in this case is the National Quality Forum, must create specific conditions to do this measurement-based work while serving under the guiding rubric of “consensus.” NQF as “consensus-based entity” refers to voluntary consensus standards setting organizations as specified in the National Technology Transfer and Advancement Act of 1995 and the Office of Management and Budget’s Revised Circular A-119. The Circular defines consensus thusly:

Consensus, which is defined as general agreement, but not necessarily unanimity, and includes a process for attempting to resolve objections by interested parties, as long as all comments have been fairly considered, each objector is advised of the disposition of his or her objection(s) and the reasons why, and the consensus body members are given an opportunity to change their votes after reviewing the comments. (Office of Management and Budget, Revised Circular A-119)

One National Quality Forum staff member offered her own working definition of “consensus” to me during an interview:

We are a consensus-based entity, which means we have to follow the Circular from the Office of Management and Budget. Now, the Circular is pretty broad in language, but it means all decisions must be reached through consensus. Here at the National Quality Forum, consensus, say for example, when an endorsement committee is reviewing and voting on measure and the evaluation criteria, consensus does not mean unanimity, but it does mean that every voice needs to be heard. When measures are viewed in a multi stakeholder environment, as with every other decision in health care that’s done in a multi stakeholder environment, there are differing perspectives. We see these differences playing out frequently, as there are different perspectives about when a measure is good enough, where someone is to be held accountable. (Interview, NQF Staff Member)

Consensus, both in its technical definition and in the working definition provided above, specifies a few conditions that are of particular relevance for the purposes of the chapter. First, the issue of consensus refers to the existing of multiple parties, stakeholders, and actors, each possessing a particular voice on the issue. Second, consensus does not refer to widespread unanimity, but to the establishment of general agreement among all parties involved. Finally, consensus is established through a *process* that aims to express the multiple voices, providing the opportunity for voting members to reconsider their own positions. This concerns the conditions of the new knowledge democracy and the centrality of the forum as decisions are made for the purposes of standards setting.

These conditions – the gathering of stakeholders, the deliberation process of hearing multiple perspectives, and arrival at a final decision that reflects agreement but not unanimity – drive the National Quality Forum’s primary work of measure endorsement.⁵ Endorsement work is the evaluative process through which individual measures are reviewed in a multi-stakeholder environment to establish consensus. Similar to the legal binding for “consensus,” Congress also specifies the nature of this endorsement work:

ENDORSEMENT OF MEASURES – The entity shall provide for the endorsement of standardized health care performance measures. The endorsement process under the preceding sentence shall consider whether a measure –

- (A) Is evidence-based, reliable, valid, verifiable, relevant to enhanced health outcomes, actionable at the caregiver level, feasible to collect and report, and responsive to variations in patient characteristics, such as health status, language capabilities, race or ethnicity, and income level; and

⁵ The National Quality Forum also plays a key role in strategy development and measure use recommendation for federal programs. By assembling the National Priorities Partnership (NPP), NQF facilitated input to the Department of Health and Human Services (HHS) in developing a National Quality Strategy (also known as the Triple Aim of the Affordable Care Act). By assembling the Measure Application Partnership (MAP), NQF continues to facilitate input to HHS in the use of specific measures within federal accountability programs. While these forms of work go beyond the scope of the paper, they demonstrate the centrality of the NQF in facilitating national decisions on quality measurement, public accountability programs, and health care improvement.

(B) Is consistent across types of health care providers, including hospitals and physicians.
(Medicare Improvements for Patients and Providers Act of 2008)

After certain organizations, also known as measure stewards,⁶ submit a developed measure for endorsement, the NQF convenes committees that then carry out this review work as spelled out above. In the process, measure stewards explain the design of the measure, the committee examines the submitted testing information to meet the specific criteria outlined above, members ask questions to the steward and each other and express their perspectives, and standing members vote to recommend for endorsement. This is an involved process that results in a tremendous amount of documentation, including meeting transcripts, technical evaluations, voting records, public comments, and final reports. This documentation is then made publicly available through the National Quality Forum's website; this documentation provided the majority of the data for my analysis in this chapter.

NQF endorsement is widely considered as the gold standard across the health care industry, and accordingly, it is heavily funded to support its work. One measurement advocate noted the degree of influence that NQF endorsement has for the Centers for Medicare and Medicaid Services (CMS), as these endorsed measures are implemented within public accountability programs:

I think [NQF] has a lot of influence, and it's deliberate. CMS actually gives them a lot of money to do this, to advise, to bring together stakeholders and advise them. I think that when CMS is looking at measures to include in these programs they do listen closely to what comes out of the NQF process... Sometimes they'll do what they want but they do rely on it heavily, so it has a lot of influence. (Interview, Measurement and Patient Advocate)

⁶ Measure stewards are responsible for submitting measures to NQF for review and continuing the measure maintenance process, including the submission of additional measure information as necessary. Most stewards develop the measures themselves, but some entities (notably CMS) contract with other organizations for measure development and then serve as stewards for the NQF endorsement process. An in-depth examination of measure development goes beyond the scope of this chapter.

Congress awarded nearly \$150 million to the National Quality Forum through 2015 to develop a measurement strategy, evaluate and endorse quality measures, and recommend measures for use in federal programs.⁷ These conditions – the gathering of different stakeholders, the establishment of “consensus,” and the significance of decision-making and endorsement for future implementation – reflect similar conditions as identified by the theoretical scholars above and as I outline under “the new knowledge democracy.” As a consensus-based standard-setting organization, NQF provides the forum through which actors are able to express their perspectives, take each other into account, and vote for the recommendation of measure endorsement. All of this work takes place prior to the selection of measures within accountability programs, and it is done with the goal of achieving consensus across stakeholders, actors, organizations, and entities regarding what to measure and how to measure. This forum is key site of decision-making over what “counts” and what “matters” in health care, especially since NQF-endorsed measures are widely used throughout the entire health care landscape in the spirit of “accountability.”

THREE MEASURES

In this section I present three measures that are used within programs of quantified accountability. In discussing the turn to accountability in general terms, one advocate noted the general alignment and agreement across stakeholder groups regarding these new public policy programs:

⁷ NQF funding for strategy, endorsement, recommendation work provided under Medicare Improvements for Patients and Providers Act of 2008 (PL 110-275), Patient Protection and Affordable Care Act of 2010 (PL 111-148), American Taxpayer Relief Act of 2012 (PL 112-240), Bipartisan Budget Act of 2013 (PL 113-67), Protecting Access to Medicare Act of 2014 (PL 113-93), and Medicare Access and CHIP Reauthorization Act of 2015 (PL 114-10). See also recent Congressional Research Service Reports for additional information (CRS 2009, 2010, 2013).

If you saw the broad consensus that was in the meeting that was on the Hill on Wednesday, almost all the stakeholders, all the provider groups that supported the MACRA⁸ legislation, all the consumer groups that are supporting this, the employers very strongly support it, we want accountability because accountability is a tool to drive improvement. If you're holding people accountable, that gives them a powerful incentive to make the improvements they need to make... there's very little debate anymore that this is something we want to do. (Interview, Measurement Advocate)

But this widespread agreement hardly settles the conflict: instead, it merely displaces it to new domains, as I will show in this section. By analyzing technical documents produced through the NQF endorsement process, I uncovered just how political and contested measurement decisions are. Another prominent advocate with years of involvement with NQF endorsement related this to me during an interview:

There's really bipartisan support for payment reform,⁹ for new ways of thinking about how you pay people. We think it'll go forward. But value-based purchasing isn't going to work if it's just we all think it's a good idea, let's just kumbaya, do it. It's not going to work...

The undercurrent and the opposition come when the details come out. If you say, "Well, here's how we're going to measure your performance. Here are some specifics. Here is how we're going to risk adjust." For whatever the details are of how we're going to do that, the opposition can be fierce. (Interview, Measurement and Patient Advocate)

As this quote so succinctly conveys, there is disagreement underneath the layer of "broad consensus" the first advocate suggested in the previous interview. Conflict and consensus are both present in the technical details of evaluative tests. This creates new politics that are not found surrounding the measures, but are found inside their very design.

⁸ MACRA, or the Medicare Access and CHIP Reauthorization Act of 2015, reforms Medicare Part B payment policy. Under two different tracks (Merit-based Incentive Payment System, MIPS, or Alternative Payment Models, AMPs), physicians will now be required to submit quality data and receive payment under straight pay-for-performance (MIPS) or newly designed models of care that link payment to performance (APMs). In both instances, accountability is pursued through quality measurement.

⁹ MACRA (previous footnote) is frequently cited as one of the most visible components of "value-based payment" under payment reform. For the purposes of this chapter, the payment reform examined here is principally concerned with measurement-based initiatives. The quote presented directly links the technical specifications of measurement to political goals of payment reform.

Based on my analysis of the technical documents in the data and methods section described above, I purposively selected three measures to highlight these new kinds of politics brought on by quantification. Table 3.1 presents an overview of the three selected measures, each of which is widely used within high-profile federal and state accountability programs, thus increasing their significance compared to newer or exploratory measures. In this section of the chapter, I provide details on each measure, and highlight how consensus, controversy, and conflict are to be found in the “endorsement” of these measures for use in programs of political accountability¹⁰. I argue that these case studies highlight three different kinds of politics – a politics of evidence, record, and boundaries – each of which confronts the standardization (e.g., “consensus,” whether as unanimity or general agreement) required by quantification. This standardization requires a “flattening” of accounts through the implementation of technical rules, rendering the conditions for political conflict against the social heterogeneity of the world. These politics, while in reality are closely connected, are here presented as Weberian ideal types.

Table 3.1. Three Quality Measures

Name	Controlling High Blood Pressure	Ischemic Vascular Disease (IVD): Use of Aspirin or Another Antithrombotic	Hospital 30-Day, All-Cause, Risk Standardized Readmissions Rate (RSRR) following Heart Failure (HF) Hospitalization
NQF No.	NQF 0018	NQF 0068	NQF 0330
Steward	National Committee for Quality Assurance (NCQA)	National Committee for Quality Assurance (NCQA)	Centers for Medicare and Medicaid Services (CMS)
Type	Outcome	Process	Administrative
Politics	<i>Evidence</i>	<i>Record</i>	<i>Boundaries</i>

¹⁰ Each measure is also presented in further technical detail in Appendix B of the dissertation.

The Politics of Evidence

One of the most common measures used within health care for accountability purposes is NQF 0018: Controlling High Blood Pressure (see Figure 3.2). This measure is used within many programs ranging from public reporting, value-based payment, and disparity monitoring. But beneath the use of the measure for accountability purposes is what I call an underlying *politics of evidence*. Quality measures are created and evaluated according to scientific evidence, usually as specified within clinical guidelines. Recent

Figure 3.2. Controlling High Blood Pressure (NQF 0018)

<p>Measure Steward: National Committee for Quality Assurance</p> <p>Description: The percentage of patients 18 to 85 years of age who had a diagnosis of hypertension (HTN) and whose blood pressure (BP) was adequately controlled (<140/90) during the measurement year.</p> <p>Numerator: The number of patients in the denominator whose most recent BP is adequately controlled during the measurement year. For a patient's BP to be controlled, both the systolic and diastolic BP must be <140/90 (adequate control). To determine if a patient's BP is adequately controlled, the representative BP must be identified.</p> <p>Denominator: Patients 18 to 85 years of age by the end of the measurement year who had at least one outpatient encounter with a diagnosis of hypertension (HTN) during the first six months of the measurement year.</p> <p>Exclusions: Exclude all patients with evidence of end-stage renal disease (ESRD) on or prior to the end of the measurement year. Documentation in the medical record must include a related note indicating evidence of ESRD. Documentation of dialysis or renal transplant also meets the criteria for evidence of ESRD. Exclude all patients with a diagnosis of pregnancy during the measurement year. Exclude all patients who had an admission to a nonacute inpatient setting during the measurement year.</p> <p>Risk Adjustment: No</p>

controversy regarding changing blood pressure standards reflect deeper tensions regarding the nature and source of expertise and “evidence.” As the measure represents the translation of “best possible evidence,” the development and endorsement of a particular metric over another reveals the decision-making over what constitutes the best evidence, which then generates conflicts on the ground as local actors struggle with implementing measures against the backdrop of multiple clinical standards.

NQF 0018 is one of the oldest measures, and originally drew from then-current clinical standards from the Joint National Committee (JNC). The Joint National Committee on Detection,

Evaluation, and Treatment of High Blood Pressure (JNC) was formed in 1977 to create national standards for the domain of heart disease. These clinical standards are widely expected to translate new medical knowledge into clinical practice. Under the direction of the National Institutes of Health, the JNC released seven reports between 1977 and 2003. Each report examines the then-current state of the evidence and makes recommendations for treating hypertension (see Kotchen 2014 for overview). While there have been modifications to the classification of disease severity (e.g., ‘normal,’ ‘mild,’ ‘moderate,’ ‘severe’ hypertension) over the seven guidelines, the blood pressure thresholds themselves have not experienced much change. For example, JNC 1 set a diastolic blood pressure goal of <90 mm Hg; JNC 4 also added a systolic threshold, creating an overall standard of <140/90 mm Hg. This threshold had been kept at this level for the three successive reports.¹¹ NQF 0018 was designed and endorsed against the then-current standards – JNC 7 – to specify blood pressure control at the <140/90 mm Hg threshold.

After several years of work on new updated guidelines, JNC 8, the National Heart, Lung, and Blood Institute (NHBLI) of the NIH stepped down from the committee and turned responsibility of standard setting over to select professional medical societies, the American Heart Association and American College of Cardiology. The in-progress work was then published in *JAMA* under the title of “2014 Evidence-Based Guidelines for the Management of High Blood Pressure in Adults: Report from the Panel Members Appointed to the Eighth Joint National Committee (JNC 8)” (James et al. 2014). The most significant of the proposed changes consisted of the relaxing of blood pressure goals for non-diabetic patients sixty years or older to <150/90 mm Hg, while keeping the established level of <140/90 for the rest of the adult

¹¹ One minor exception to the maintenance of the <140/90 mm Hg goal was the creation of a <130/80 threshold for patients with diabetes or renal disease, as added in 2003 under the JNC 7.

population. This set of recommendations drew from the evidence of several randomized control trials.¹²

However, the standards were not sanctioned by the NHLBI, and were met with significant controversy. There were conflicting interpretations of the evidence to support changing thresholds for the first time in several decades, with different panel members interpreting the scientific evidence differently. The written text accompanying the specifications of the clinical standards is suggestive here, indicating disagreement among the panel as to the recommendation to change the systolic blood pressure (SBP) threshold for older adults:

While all panel members agreed that the evidence supporting [the recommendation] is very strong, the panel was unable to reach unanimity on the recommendation of a goal SBP of lower than 150 mm Hg. Some members recommended continuing the JNC 7 SBP goal of lower than 140 mm Hg for individuals older than 60 years old based on expert opinion. These members concluded that the evidence was insufficient to raise the SBP target from lower than 140 to lower than 150 mm Hg in high-risk groups (JNC 8: 511).

Some of those members of the panel who disagreed with the JNC 8 proposal to increase the SBP goal for older adults from <140 to <150 published a dissent in the *Annals of Internal Medicine* (Wright et al. 2014). Thus, even when there is general agreement as to what constitutes the best source of evidence (e.g., findings from specific randomized controlled clinical trials), there is no guarantee that experts will interpret the evidence in the same way. “The best possible evidence” does not stand alone, as heterogeneous interpretation suggests alternative thresholds and courses of action. If two options are equally valid, then which should serve as the basis for implementing standard rules within the technical specifications of the quality measure?

This JNC 8 standard, however, was released precisely with the goal of providing scientific evidence (according to the committee) to then be used to redesign or modify measures

¹² These include HYVET, Syst-Eur, SHEP, JATOS, VALISH, and CARDIO-SIS. Results from four of these six trials were published after the release of JNC 7 in 1997.

used within programs of accountability. Indeed, the concluding sentence of the report directly addresses the use of evidence to create quality measures: “The strong evidence of this report should inform quality measures for the treatment of patients with hypertension” (519). It is not surprising then that the measure steward of the widely used blood pressure measure, the National Committee for Quality Assurance (NCQA), created a modified measure that took into account this different threshold for older patients. Measures stewards themselves are generally not involved in the business of evaluating the scientific evidence itself, but instead translate clinical standards into metric specifications. In this sense, NCQA does not evaluate the merits of the threshold increase, but takes evidence and expertise at face value.

Once a modified measure was created, it could be used in political arguments over the technical details of accountability as an alternative to NQF 0018’s original specification. For example, during the National Quality Forum’s review of the new measure, the committee decided not to endorse the modification, highlight inconclusive evidence and expert disagreement. And despite NQF’s refusal to endorse the measure, the modified measure resurfaced during an attempt to arrive at public-private measurement alignment. During the Core Quality Measures Collaborative – an effort to align measures across public and private entities¹³ – both the original and modified measure appeared as “options” under the primary care measure set, resulting in additional statements from cardiologists and professional experts. Including both measures, with different thresholds drawing upon different interpretations of the evidence, was done to achieve “consensus” surrounding the measure set. Without the inclusion of both

¹³ Centers for Medicare and Medicaid Services (CMS) recently convened the Core Quality Measures Collaborative with the goal of aligning the use of measures with the same technical specifications across public and private entities. Multiple measures may be created within a given domain, and while they may be assessing similar items, their means of measurement and assessment may be slightly different, resulting in challenges when comparing across entities or when implementing measures for a particular domain (e.g., blood pressure) on the ground.

measures of controlling high blood pressure – one at <140/90 for the whole population, another with a raised threshold for the elderly – public and private entities could not agree on the terms through which accountability is to be pursued.

These issues do not just appear at particular fora or national efforts to discuss the technical details of measurement. As measures are implemented on the ground, national concerns and debates over “the best possible evidence” are no longer contained to private meetings between specialty societies or public fora dedicated to the topic. Providers, quality staff, and administrators struggle to make sense of the multiple measures used across various accountability programs. At a large conference on the use of ratings within health care, the president of NCQA discussed the significance of measurement alongside the challenges incurred with multiple measures:

“People used to say that quality is not a business issue – well, congratulations to Medicare for making it a business issue.” The president of the National Committee for Quality Assurance (NCQA) goes on to detail some of the issues with the current state of affairs in health care. “Fee-for-service has given us a lot of what we are unhappy with: overutilization, the use of high technology when low-scale interventions work just as fine. We need to do better.”

Now that Medicare is involved in the use of quality measurement to drive change, she relates to the use of quality measures to evaluate the performance of different payment arrangements. “I have this point of view with Medicare, where we want to compare across traditional [fee-for-service], Medicare Advantage [managed care], and Accountable Care Organizations, but with different measures, it makes it difficult to compare across them.” She continues, “I just heard a presenter from John Hopkins, ‘I have three different ways that I am measured in terms of how I’m doing on blood pressure!’” Multiple measures produce lots of challenges, not just for evaluators, but for providers as well. (Field note, National Health Care Ratings Summit)

Each measure is supposed to translate the best evidence and outline a particular goal or outcome. But the existence of more than one measure, each claiming to be doing “the best,” brings out new challenges and questions to the project of quantified accountability. If measures are based in the best possible scientific evidence and there are conflicting accounts as to the evidence itself, the measures proliferate as different groups align themselves with one measure over another. In practice however, this results in conflicting measures used within the same health system,

providing challenges not just for direct frontline providers and quality staff but for those who work with the data at the aggregate level.

At the national level, the controversial release of the JNC 8 resulted in additional review and evaluation activity. In late 2017, the American College of Cardiology (ACC) and the American Heart Association (AHA) released a separate set of new standards, in this case *lowering* the blood pressure threshold even from the historical standard of <140/90 mm Hg to <130/80 mm Hg for *all* patients, including those over 60 years of age. As professional societies representing cardiology, the ACC and AHA lay claim to being the truly qualified experts with regard to the evaluation of scientific evidence for heart disease. In response to this new clinical standard, NCQA proposed reverting the measure back to the <140/80 mm Hg threshold for older adults.

But cardiologists, of course, are not the only providers concerned with controlling high blood pressure. In actual clinical practice, blood pressure control commonly falls under the domain of the primary care, and it is primary care providers who are typically evaluated by measures within these accountability programs. As a result, the American Academy of Family Physicians (AAFP) – a professional medical society representing family medicine – responded to the new ACC/AHA guidelines (lowering the threshold to <130/80), releasing its own set of guidelines for treatment of hypertension among adults over 60 years of age. The AAFP guidelines drew heavily on the JNC 8’s recommendations, by including a threshold increase to <150/90 mm Hg for those older than 60. Their move furthered the divide over what constitutes “best evidence.” Some specialty cardiologists criticized the AAFP for making this move, arguing the notion of expertise itself was being undermined: while the members of the AAFP panel absolutely had *clinical* expertise with regard to primary care, they lacked *scientific* expertise as

determined by publications on hypertension and involvement with clinical trials (Meserli et al. 2017; O'Brien 2017). Thus this national disagreement revealed a deeper question over who has authority to evaluate and determine the best evidence. For the measure steward seeking to implement “best practices” through the technical design of the quality measure, these issues of the source of authority and expertise produce considerable challenges, as heterogeneous interpretation conflicts with the standardization required by quantification.

Taken altogether, since the release of JNC 7 in 2003, standards for blood pressure control underwent three successive guideline releases: First, the JNC 8, published in *JAMA*, raised the threshold for the elderly and suggested that the evidence be used to modify NQF 0018 (then at <140/90, to <150/90 for the elderly subpopulation). The ACC and AHA, claiming specialty expertise in turn, evaluated the evidence and suggested that the threshold be lowered for all people to <130/80. The AAFP, representing a particular subset of primary care providers, responded by releasing its own set of recommended guidelines that reflected the controversial JNC 8 proposal. These national issues represent disagreements over *what* constitutes the best evidence, as well as *who* is qualified to assess it. The measure stewards who create and submit measures for NQF endorsement are unable to address these issues and capture these dynamics within standard metrics.

On the ground, the use of blood pressure measures within programs of accountability reanimates these national disagreements at health systems across the country. Quantified measures seek to implement stable and standard rules through technical specifications, and when they do so, they introduce inflexibilities that are poorly equipped to handle heterogeneity in expertise and evidence. I was in the middle of conducting fieldwork at a local hospital system when the new 2017 blood pressure guidelines were released by the ACC and AHA (lowering the

threshold from <140/90 to <130/80 mm Hg). This change in the standard blood pressure thresholds resulted a flurry of social activity at the local site:

A project manager responsible for a large state pay-for-performance program's primary care delivery system reform initiatives receives a flurry of email messages. Medical directors, health center managers, and frontline providers all want to know: how will the new standards affect us? The health system's many outpatient clinics are currently evaluated by NQF 0018 under an accountability program that links payment to "performance." Will we still be evaluated according to the <140/90 standard?

The project manager, a non-clinical employee, responds to each email inquiry, assuring staff that the team is working on making sense of the change. She in turn writes to an administrator for the state's program, asking for future guidance and whether the program will continue to use the existing measure based on previous clinical standards. The administrator, based out of a separate organization that works closely with state agencies, responds that she has already received several emails from other health systems. Each health system participating in the program has its own quality team facing similar issues as frontline providers and clinic managers ask for clarification on the technical details of the evaluation. She in turn assures the project manager that organization is working on making sense of the change as well.

But in the meantime, the health systems operations continue, and the topic comes up at regular monthly meetings at most of the outpatient clinics, workgroups comprising of clinicians and managers, and presentations to executives concerned with the organization's financial performance. "We've contacted [overseeing organization], thank you for bringing it to our attention." Due in part to the organization's size, the same topic comes up over and over again, renewing some skepticism towards the measure specifically and the program more generally. (Field note, County health system)

The measure itself specifies a particular goal: "controlling high blood pressure" according to NQF 0018 means keeping blood pressure beneath a threshold of <140/90. But with the ACC/AHA's release of a new guidelines, effectively *lowering* the threshold even further, staff and providers ask how they will be evaluated. But beyond this, some skepticism surfaces regarding the value of the measure and the program more generally overall. The quality team had been working for months to meet the current metric, encouraging staff and frontline providers to lower blood pressure beneath the <140/90 threshold because "it's the right thing to do" and because it is the standard used across the country. But with changing evidence and questions of expertise, there is a question of whether stabilized measures really do the work of specifying "value" through the inducement of accountability.

In an interview I conducted with a quality improvement coordinator, I asked for her thoughts on the new guidelines:

I read an article about that, and what's significant about the change is it seems they did not consult the professional organizations that really have a large stake on this metric. It is really nice to consult all the key stakeholders when you do this change so to minimize or eliminate the pushback. With this new guideline I think there's a lot of pushback because one major association was not consulted and it seems that they consider themselves as the guru when it comes to this benchmarking, and so now there's pushback on this new guideline. (Interview, QI Coordinator)

Here the informant expresses not only a significant amount of understanding into the politics of evidence and expertise at the national level, with different professional societies disagreeing in establishing “consensus” for final guidelines, but also how these conflicts play out *on the local level*. Frontline providers and clinic leadership that she works with in meeting the measures provide “pushback” over the measure, even if the measure has received NQF endorsement. Moreover she relays providers' concern regarding their ability to meet a modified measure, when there are already challenges in meeting the 140/90 standard:

[The new standard] is definitely lower than 139 over 89, and we are already struggling to meet that. The standards are asking us to lower it, but it would be really, really... LOTS of work for our providers and our health care workers to have hypertensive people to reduce their blood pressure even more. We are already struggling with the present threshold. That's why we have a lot of pushback, because we're already struggling with the present 139 over 89 blood pressure, and now you're telling us to lower it even more [laughs]! How are we going to do that? (Interview, QI Coordinator)

This case of the blood pressure control measure reveals an underlying *politics of evidence*, as different professional societies, medical groups, and governance bodies relate differently to “evidence-based medicine.” It demonstrates how the heart of scientific evidence and expertise depends upon the fundamental value of “consensus,” and the challenges that ensue when there is difference in expert opinion. As performance metrics, however, quality measures implement standardized technical specifications. But these numbers also are unable to address deeper conflicts about the role and nature of expertise in health care. While the measures can of course be modified and adapted when the scientific evidence changes, they do not resolve

conflicts surrounding the *interpretation* of that evidence, professional society disputes, and the role of government and authoritative bodies in determining what is “best.” This results in multiple measures that all claim to be measuring the same thing, but with different technical specifications. While their difference may appear to be slight, the implications for local practices of health systems and the stakes for those systems, can be significant. In this case, the challenges of arriving at consensus regarding the scientific evidence reveal deep divisions within the practice of medicine that a single measure is unable to fully accommodate.

The Politics of Record

Another set of politics, what I refer to as *the politics of record*, refers to changing roles and responsibilities within health care provision. This set of politics highlights the increasing

Figure 3.3. Ischemic Vascular Disease (IVD): Use of Aspirin or Another Antithrombotic (NQF 0068)

<p>Measure Steward: National Committee for Quality Assurance</p> <p>Description: The percentage of patients 18 years of age and older who were discharged from an inpatient setting with an acute myocardial infarction (AMI), coronary artery bypass graft (CABG) or percutaneous coronary intervention (PCI) during the 12 months prior to the measurement year, or who had a diagnosis of ischemic vascular disease (IVD) during the measurement year and the year prior to the measurement year and who had documentation of routine use of aspirin or another antiplatelet during the measurement year.</p> <p>Numerator: Patients who had documentation of routine use of aspirin or another antiplatelet during the measurement year.</p> <p>Denominator: Patients 18 years or older by the end of the measurement year discharged from an inpatient setting with an AMI, CABG, or PCI during the 12 months prior to the measurement year or who had a diagnosis of IVD during both the measurement year and the year prior to the measurement year.</p> <p>Exclusions: Patients who had documentation of use of anticoagulant medications during the measurement year. Exclude patients using hospice services any time during the measurement period.</p> <p>Risk Adjustment: No</p>

centrality of the *record* for *distributing responsibility*, as documentation relates to the interconnection between social and technical systems. A common phrase heard everywhere throughout the health care system and at the national policy level is the following: “If it’s not documented, it didn’t happen.” To make sense of this set of politics, I refer to the technical evaluation work of measurement of ischemic vascular disease (IVD) and patient

use of aspirin. Following acute myocardial infarction (heart attack), coronary artery bypass graft, and percutaneous coronary intervention, aspirin is used to prevent clots, thus lowering the risk of strokes following any of these serious heart conditions and procedures. While this seems straightforward enough, there are several distinctions here that the NQF committee identified as being consequential for quality measurement: (a) patient *use* of medication versus provider *prescription*, (b) provider documentation of patient use of aspirin as an over-the-counter drug versus blood thinner medication, (c) provider documentation of patient exclusions, such as internal bleeding, through the use of codes for claims and billing purposes. In each instance, documentation emerges as a central organizing node. I review each point in turn below.

The politics of record draws attention to changing roles and responsibilities, in this case, the responsibility of providers for patient use of medication. The *consumption* of medication is what is recommended by clinical standards and professional experts, but there is a question over the line between provider and patient responsibility. Outlining activities within their control, some providers emphasize *prescription* over *patient use* of aspirin. Prescribing medication is undoubtedly within a provider's control, but "adherence" requires shared responsibility for the course of social action. For accountability programs that monitor provider "performance" through measurement, this technical definition reveals the importance of defining the target goal and its relation to the roles and responsibilities of all entities involved. Documentation emerges as a central activity in ensuring use of aspirin "actually" happened.

The politics of record calls attention to another component here: the increasing centrality of the *technical account* that is used to hold providers accountable, including the work needed to document the clinical encounter. This technical account primarily consists of documentation within the Electronic Health Record (EHR), and it serves as a common record that coordinates

internal (e.g., providers, quality team members) and external entities (e.g., insurers, the state). Thus the record here makes documentation—and the precise thing being documented—of utmost importance. But if the data from the EHR are made available to others for public reporting purposes, then activity that is not documented or that is documented in a way that is not readily communicable to others looks like “low quality” care. National stakeholders thus consider the *work of documentation* – the labor involved in producing the technical account that is needed for social action for purposes of accountability – to be reasonable, feasible, and doable for frontline providers and health care staff. Documentation, including both the labor to generate the record and the merits of the record itself, thus becomes a new site of politics within programs of accountability.

Similar to the coordination cooperation between provider and patient, the technical account also “coordinates” but through an inter-connected system itself, namely the interfacing of the EHR and pharmaceutical prescription. For example, national stakeholders consider how in practice, most providers recommend that patients consumer over-the-counter aspirin. Because this is an over-the-counter medication, this may or may not make it into the documented record, as providers may not claim responsibility over OTC medication or acknowledge the shared communication between patient and provider. But the challenge here is how to make this information in a documented fashion that is accessible to others when medical practice confronts non-prescription medication. Prescribing aspirin alternatives (e.g., blood thinner medications) are easier to document and capture, as many EHR systems interface with pharmacy information. Over-the-counter information, on the other hand, is much more difficult to interface with the technical account.

These issues emerged during local fieldwork at a large health system, whereby the quality team sought to implement NQF 0068 under its main accountability program. Staff related to me that in the earlier stages of the accountability program, there was an issue with documenting the use of aspirin itself within the EHR. This represents what was already considered at the national stage when reaching “consensus,” but that the health system needed to learn how to figure out on its own. I asked one quality improvement coordinator if there had been any challenges working with the IVD measure under consideration:

Coordinator: Well, since aspirin is an over-the-counter medication, doctors just tell patients. “I recommend you use it.” So it’s a matter of a doctor adding it to their regimen. But if there’s another thrombolytic agent that’s prescription medication, other than aspirin, it’s easier to accomplish because the doctor feels it’s their responsibility. But if it’s just aspirin, they’re not inclined to add it into their regimen. Especially if they’re already taking it, “Why should I have to add it in their regimen?” So it’s a matter of telling them, by doing that, [the organization] will receive financial incentives [under pay-for-performance program].

TMC: So it sounds like it’s an issue of documentation.

Coordinator: Yeah, exactly, it’s an issue of documenting it in the EHR. (Interview, QI Coordinator)

Here the quality improvement coordinator relates the issue of documentation as it pertains to provider responsibility, but also specifies that it can be a challenge getting the documentation into the record. There is a deeply intertwined set of politics here of a sense of *social responsibility* alongside the *technical account* of the EHR. Prescription medication is both more easily documented through the interfacing of other technical systems and comes with a provider sense of responsibility. Over-the-counter medication, on the other hand, required its own separate build within the record, and yet providers on the frontlines of care may not see the need to document it accordingly.

At the national level, where national stakeholders convene to endorse the measure, this documentation appears “reasonable” and feasible. But the health system learns on its own what needs to be done to meet the record. One program manager related to me some of the early issues

with getting use of OTC aspirin documented properly within the EHR so that the health system would not receive financial penalties under the accountability program. Even after providers take the time to include it in the record, it must be done in a specific format so that others have access it later on:

[Use of OTC aspirin] was in the EHR, but it was in the EHR in the non-report-extractable, friendly realm. If you think of how physicians and clinicians document, there is free texting and you have check boxes. Then your pharmacy piece with your medications, that's a reportable field of what you put in it. But if you're not doing a prescription of aspirin to go to a pharmacy to collect, and you're verbally informing your patient, "You need to take aspirin every day of your life," that would be a free text area. We had to create a box for OTC, over the counter, aspirin. So that way it changed the practice that they were documenting it closer to being a prescribed medication. (Interview, Project Manager)

Because OTC medication does not readily interface with the technical account in question, the EHR, the health system's programmers had to create a separate build to "capture" the data properly. This is extra labor that the health system had to assess and complete on its own, with little to no guidance on how to implement the measure in practice, even after having gone through rigorous endorsement work. The health system here is not remarkable in this regard, as NQF 0068 is felt across the country providers and administrators grapple with questions over roles and responsibilities through the technical account.

At the national level, the centrality of documentation for the purposes of establishing a record with which to hold providers to account also surfaced with concern over patient exclusions. The National Committee of Quality Assurance, the measure developer and steward, recognized that there be certain exclusions, such as for internal bleeding, that would preclude the use of aspirin among IVD patients. However, NCQA notes the difficult of documenting these exclusions, and thus suggest that a "perfect score" for providers is not the goal of the measure:

Rates of physician added exceptions were quite low, inconsistent in rate, and many had to come from extensive manual chart review event from an EMR... Codes (like CPT-II codes) that might be used to indicate exceptions are not widely used, and at present time cannot be easily audited for accuracy. In addition the measure allows for physician discretion in prescribing alternative oral anti-platelet therapies when aspirin is contraindicated.

Rather than attempt to capture all of these exclusions in the technical account, NCQA proposed that acceptable performance be less than 100% *in order to reduce the work of documentation*.

At the local level, however, this lack of technical exclusions is interpreted differently. The lack of exclusions is not seen as part of an effort of reducing the work of documentation, but is instead seen as a “poorly designed” metric. This feature of the lack of exclusions caused a tremendous amount of headache on the ground: as the quality team staff reported performance numbers back to front line providers, some of the physicians are surprised that their performance scores are as low as they are captured in the report. This caused one physician to do her own deep dive into the EHR system and backtrack the metric, and she noted that most of her patients that are not on aspirin are contra-indicated or have a version of IVD that is not treated through aspirin (e.g., varicose veins). The quality team staff turned to their on-site technical consultant, who in turn consulted with outside entities to understand the NQF 0068 metric. What is lost in the chaos is that the lack of exclusions was put in place with the goal of *reducing* work.

At another health system, the quality team searched for ways to capture all of the exclusions within the record. Even though the metric specifies *no* exclusions need be documented, the health system sought to create its own way of ensuring that external entities working with the record later on could “see” the contraindications or justifications for not being on aspirin. I spoke with one Chief Medical Informatics Officer (CMIO) from this health system about the work of implementing measures on the ground without guidance on how to do so:

CMIO: Yeah, a great example of that is that a lot of the specifications ask us to use these billing codes, which are really zero-dollar charges.¹⁴ They’re things like the patient has a medical

¹⁴ Zero-dollar charges refer to documented information within a billing claim that does not determine the final amount of the claim. These “charges” are not reimbursable but provide information about services

contraindication to being on aspirin for Ischemic Vascular Disease (IVD), maybe they had a severe GI bleed in the last year, and so we don't want to be using aspirin. I'm not aware of anyone having clear workflows, other than manual staff reviewing charts, and adding those codes somehow, I'm not aware of having a good electronic process to ensure that the code ultimately gets on the claim, which gets over to the payer for HEDIS¹⁵ calculations, and ultimately could be usable for our performance.

So it's a great example of how we are going to have to figure out how to get the actual code on the claim. We think we've found a creative way to do that, but this is sort of a core issue. I think those who are writing the specifications, and working with the metric just assume that we have an obvious way to get those billing codes from the point of care to the claim.

Docs know, for the most part, they know CPT codes, but there has never been a systematic – in the two systems I've worked in – we've never had a systematic approach to teaching them how use any of these non-billable administrative codes. And there would not be a lot of enthusiasm for doctors to learn this whole new coding set. You're seeing a patient, you know they shouldn't be on aspirin, you're addressing their clinical needs, and then to jump over and realize, "Oh, I have to put in this administrative code so that everyone else knows there's a contraindication for this patient being on aspirin. It's just not intuitive in the physician workflow at all, and definitely an added cognitive step that's not natural in clinical thinking.

TMC: It sounds like a lot of behind the scenes technical work to make it all happen.

CMIO: It is. It's actually a critical component of quality metric reporting that I think is very much under recognized. I have no idea where these administrative codes originated, I'd never even heard of them before [accountability program]. But clearly, there's an expectation that they're broadly in use, and clearly they're not. So there's a disconnect between those tracking this quality data purely from the claim submission side, and what's really happening on the front lines of medical care.

TMC: Who do you think has those expectations that the administrative codes are available and widely used?

CMIO: Well, this is an NCQA metric. These national metric stewards who developed these specifications, clearly seem to expect that we're all using those codes. I have no idea if they're aware that they're not broadly in use, or how difficult it is to implement them, but it's clearly one of the things that's causing a disconnect between quality data and actual care on the front line. (Interview, Chief Medical Informatics Officer)

rendered based on patient condition. By adding this to the claim, the health system is able to communicate with insurers for billing and non-billing purposes.

¹⁵ Health Effectiveness Data and Information Set (HEDIS) is a set of quality measures administered by the National Committee for Quality Assurance used to rate health plans on "performance." HEDIS contains many measures that are individually used within federal accountability programs and in other measurement-based initiatives. Both NQF 0018: Controlling High Blood Pressure and NQF 0068: Ischemic Vascular Disease (IVD): Use of Aspirin or Another Antithrombotic are included in the measure set. In this interview, the informant referred to NQF 0068's use for HEDIS calculations as well as one of multiple other measures under a separate accountability pay-for-performance program interchangeably.

In this case, the CMIO expresses the politics of record that she considers at her local health system. If providers are being held responsible for the use of aspirin at the local site through NQF 0068, there is the recognition that they should not be penalized for having a patient with contraindications. She also recognizes that there is some issue between the national steward, NCQA, and local implementation in terms of documentation.

However, what is lost here is the national decision that exclusions for things like internal bleeding are *not expected to be counted*, and instead that exclusions can be accounted for by the expectation that performance will be less than 100%. But insofar as these are programs of accountability, few providers want to be seen as having “poor performance.” Documentation added to the *technical account* is necessary lest some be seen as providing substandard quality of care: without the “zero dollar charges” noting medical contraindications for the prescribed use of aspirin, those working with quality data outside of the health system will see poor performance based on the CMIO’s understanding of the problem. There becomes a local politics of record as health systems reconsider the relationship between documentation, technical accounts, and changing responsibility, even after a measure undergoes rigorous evaluation and endorsement, even after decisions are made about exclusions and the work of documentation.

Thus the *politics of record* recognizes changing notions of roles and responsibility among human actors and how they interface with technical accounts. In endorsing the measure, national stakeholders consider the roles of providers and patients as well as the EHR as it interfaces with clinical practice and outside information systems (e.g., non-internal pharmacies). But they also consider what is “reasonable” in terms of not only the *amount* of documentation, but also the *kind* of work it takes to integrate information into the record. By endorsing that providers must document use of aspirin, but may not document exclusions, the committee does the work of

deciding what kind of technical work should take place. These kinds of decisions are difficult to trace once this measure is then implemented on the ground across the country, as the cases of individually EHR program builds to capture OTC aspirin and zero-dollar charges for patient exclusions demonstrate. As the measures are implemented across the health care landscape, more and more people across more and more sites of care come to grapple with decisions that were made at the point of evaluation through the National Quality Forum, and learn how to create workarounds without learning the justification for the technical decisions.

The Politics of Boundaries

A third set of politics, what I term *the politics of boundaries*, brings in to question the very jurisdiction of medicine beyond the walls of the clinic. I examine this through the newly developed readmissions measures, widely recognized through the recent Hospital Readmissions Reduction Program. Standardized measures of hospital readmissions provide information on hospital and patient visit activity, which is then used by external entities to “act” on health care in new kinds of ways. But the topic of readmissions brings new issues to light over boundaries between the clinic and the rest of society, including how the patient interfaces with other health systems and organizations as well as social and economic heterogeneity across different patient populations. In facilitating standardized comparison across sites, the question of social heterogeneity emerges as it relates to the problem of accountability.

Hospitals first began voluntarily reporting readmission rates to the Centers for Medicare and Medicaid Services (CMS) in 2009, which CMS in turn began reporting on its public website Hospital Compare. CMS had previously reported 30-day mortality rates for heart attack and heart failure starting in 2007 (with pneumonia added in 2008), and the expansion of readmissions data on these three conditions marked a distinct shift towards considering quality of care. Data on

readmissions were collected, analyzed, and presented for the conditions of heart attack, heart failure, and pneumonia. The readmissions measures were each endorsed by the National Quality Forum. In 2010, the passage of the Patient Protection and Affordable Care Act created the Hospital Readmissions Reduction Program (HRRP),¹⁶ which not only mandated reporting of readmissions data, but also linked performance against national benchmarks to financial penalties and awards, with penalties starting in 2013. The HRRP is now one of the most prominent federal accountability programs. All hospitals are compared against a standard benchmark (the average of all readmissions), with those below the average penalized and those above receiving payment bonuses.¹⁷

The readmission measures were designed by Yale School of Medicine Center for Outcomes Research and Evaluation (CORE) on behalf of the Centers for Medicare and Medicaid Services. The readmission measures are relatively similar across the conditions of heart failure, heart attack, and pneumonia (Figure 3.4 depicts the measure for heart failure). In each case, a readmission is calculated based on administrative claims for any unplanned returns to a hospital following hospitalization within the past thirty days. As an indicator of “quality,” readmissions represent an administrative calculation that strives to “count” inadequate care: theoretically, some patients may return to the hospital within 30 days if they do not receive ideal care the first time around (e.g., patient does not receive or understand discharge instructions), whereas a

¹⁶ The HRRP is mandated to use measures that are endorsed by a consensus-based entity (as in most cases, the National Quality Forum). NQF also provides recommendation for the use of specific measures within federal accountability programs through the Measures Application Partnership, as noted in Footnote 2 of this chapter.

¹⁷ The HRRP is mandated to operate under budget-neutral policy, which specifies that the program cannot draw upon additional resources beyond the amount that would be used if the program were *not* in place. In practice, the benchmarking utilized in the program not only compares hospitals against each other, but also directly redistributes financial resources through these comparative practices. Financial rewards given to “top performers” do not come from a source of new funds but are directly pulled from the financial penalties enacted upon “low performers,” thus fueling intense competition across organizations.

Figure 3.4. Hospital 30-Day, All-Cause, Risk Standardized Readmissions Rate (RSRR) following Heart Failure (HF) Hospitalization (NQF 0330)

Measure Steward: Centers for Medicare and Medicaid Services

Description: This measure estimates a hospital-level, 30-day RSRR for patients discharged from the hospital with a principal diagnosis of HF. Readmission is defined as unplanned readmission for any cause within 30 days of the discharge date for the index admission. A specified set of planned readmissions do not count as readmissions. The Centers for Medicare & Medicaid Services (CMS) annually reports the measure for patients who are 65 years and older and are Medicare fee-for-service (FFS) beneficiaries hospitalized in non-federal hospitals or Veterans Health Administration (VA) hospitals.

Numerator: The outcome for this measure is 30-day readmission. We define readmission as an inpatient admission for any cause, with the exception of certain planned readmissions, within 30 days from the date of discharge from the index HF admission. If a patient has more than one unplanned admissions (for any reason) within 30 days after discharge from the index admission, only one is counted as a readmission. The measure looks for a dichotomous yes or no outcome of whether each admitted patient has an unplanned readmission within 30 days. However, if the first readmission after discharge is considered planned, any subsequent unplanned readmission is not counted as an outcome for that index admission, because the unplanned readmission could be related to care provided during the intervening planned readmission rather than during the index admission.

Denominator: This claims-based measure can be used in either of two patient cohorts: (1) patients aged 65 years or older or (2) patients aged 18 years or older. The cohort includes admissions for patients discharged from the hospital with a principal discharge diagnosis of HF and with a complete claims history for the 12 months prior to admission. The measure is currently publicly reported by CMS for those patients 65 years and older who are Medicare FFS beneficiaries admitted to non-federal hospitals. This measure can also be used for an all-payer population aged 18 years and older. We have explicitly tested the measure in both patients aged 18+ years and those aged 65+ years.

Exclusions: The HF readmission measure excludes index hospitalizations that meet any of the following exclusion criteria:

1. Admissions without at least 30 days of post-discharge enrollment in Medicare FFS;
2. Discharges against medical advice;
3. Admissions within 30 days of discharge from a prior HF index admission; and
4. With a procedure code for LVAD implantation or heart transplantation either during the index admission or in the 12 months prior to the index admission.

Risk Adjustment: Yes

readmission could be avoided if the quality of the initial hospitalization is of sufficient value.

Planned readmissions, such as follow-up procedures scheduled after the initial hospitalization, are excluded from the calculation of readmissions.

Readmissions measures touch everywhere and everything being evaluated, and hospital administrators and providers feel their significance acutely. At the same time, they push back against surface assumptions that readmission self-evidently indicates lack of quality. At a

national conference of health care ratings, for example, a panel member related the issue of boundaries within the hospital and other key actors when it comes to readmissions:

“I don’t think anyone would say that it’s not important to reduce readmissions,” one quality administrator representing a nationally renowned health system and plan remarks, “and we do have CMS measures for readmissions for 30 days out. But ultimately the ones held accountable are the hospital, and the hospital loses control once that patient leaves.” She further relates, “Sure, we have a role before discharge – have we done everything to assess that they are healthy enough to leave? Have we provided all of the information?”

“But what happens when the hospital does follow the protocol? Or when the patient does come back on their own? Or if the home health agency does not do what they were supposed to do? We need to have a measure where all people are involved.” (Field note, National Health Care Ratings Summit)

This administration highlights how the hospital is ultimately held responsible, but that the outcome of readmissions is actually influenced by other activity, including patient action as well as other health agencies that may take care of the patient. Rather than locate accountability solely within the hospital organization, she identifies how there are limits to what the health system can and cannot do.

At the local health system where I conducted fieldwork, I spoke with the head Director of Quality on the topic of quality measures and accountability programs. While discussing the role of measurement more generally, she brought up readmissions in particular as a means of questioning particular measures in these kinds of programs:

For years now, we’ve been arguing about readmissions and why readmissions is, I don’t want to say a bogus measure, but it doesn’t tell us what we need to know. We need to know is this person coming back because of something that happened related to first admission? Not just did they fall and break an ankle and have to have surgery after an MI? How do we have data that’s more helpful to us and not as general? (Interview, Director of Quality)

In this quote the informant questions whether readmission measures actually provide relevant information for the health system, where this information specifies if the readmissions is related to the initial hospitalization. Breaking an ankle after a heart attack, after all, is clearly unrelated, and yet the health system is still penalized for the readmission. The decision that readmissions is a *relative* measure used to compare health systems against each other, again with the

understanding that no health system will have zero readmissions in an evaluation period, does not show up here. Rather, the health system is predominantly concerned with what it can do and how it looks “on the record.”

In some deference to these sorts of questions, readmissions measures are “risk adjusted” to take into account clinically relevant patient-level and hospital-level factors. Patient-level factors such as age, sex, and selected clinical covariates are included in the statistical model. These clinical covariates are drawn from Medicare claims data from twelve months prior including the initial hospitalization such as illness severity, age, and sex. Consider, for example, the list of risk-adjustment variables presented in the technical documentation for NQF evaluation (Figure 3.5). By “taking into account” these patient-level differences, the technical strategy of risk adjustment seeks to “even the playing field” so that hospitals are not unfairly penalized for serving an older or sicker population. One could easily imagine how greater illness severity or age might increase the likelihood of patient readmission.

Even after taking into account these important technical considerations, the Hospital Readmissions Reduction Program generated significant controversy, with renewed attention to how readmissions were calculated. The primary controversy concerns the ability of the measures to account for factors that are within the hospital’s control. This emerged as a particularly difficult issue for safety-net hospitals, who argued that there were factors outside of their control that shape their ability to reduce readmissions rates, but that cannot be adjusted for according to the sanctioned list of risk adjustment variables. It is not difficult to see this when considering housing, for example: a safety-net setting that serves a larger share of a homeless population may just as well experience “readmissions” because their patients’ unstable housing conditions as well as “low quality.” A hospital that predominantly serves a relatively affluent, middle-class

Figure 3.5. Final List of Risk Adjustment Variables for NQF 0330 (NQF CV10)

<p>Demographic:</p> <ul style="list-style-type: none"> • Age-65 (years above 65, continuous) for 65 and over cohorts; or Age (years, continuous) for 18 and over cohorts • Male <p>Cardiovascular:</p> <ul style="list-style-type: none"> • History of CABG • Cardio-respiratory failure or shock • Congestive heart failure • Acute coronary syndrome • Coronary atherosclerosis or angina • Valvular or rheumatic heart disease • Specified arrhythmias • Other or unspecified heart disease • Vascular or circulatory disease 	<p>Comorbidity:</p> <ul style="list-style-type: none"> • Metastatic cancer or acute leukemia • Metastatic cancer or acute leukemia • Cancer • Diabetes or DM complications • Protein-calorie malnutrition • Disorders of fluid, electrolyte, acid-base • Liver or biliary disease • Peptic ulcer, hemorrhage, other specified gastrointestinal disorders • Other gastrointestinal disorders • Severe hematological disorders • Iron deficiency or other anemias and blood disease • Dementia or other specified brain disorders • Drug/alcohol abuse/dependence/psychosis • Major psychiatric disorders • Depression • Other psychiatric disorders • Hemiplegia, paraplegia, paralysis, functional disability • Stroke • Chronic obstructive pulmonary disease • Fibrosis of lung or other chronic lung disorders • Asthma • Pneumonia • End stage renal disease or dialysis • Renal failure • Nephritis • Other urinary tract disorders • Decubitus ulcer or chronic skin ulcer
---	--

population will not have to consider issues of homelessness among their patients. The Director of Quality at a local safety-net hospital I spoke with emphasized this point:

The problem with readmissions, with this population, with the underserved, with so many homeless people, their readmission rate is going to be higher because if you're going back out the encampment... I've seen pictures of these leg ulcers that you're living in unsanitary conditions, you're not well nourished, you're not getting dressing changes, and you're going to come back one way or another. Then with all the other complications, opioid addictions and everything else, they're going to bounce back. (Interview, Director of Quality)

Acknowledgement of such socioeconomic heterogeneity among patients led to a call to renew the calculation of readmissions through risk adjustment of "social factors" that might affect performance. The American Hospital Association emphasized this in a recent newsletter:

While hospitals are working to reduce readmissions caused by clinical care practices, there are many other factors beyond hospitals' control – including sociodemographic factors, such as poverty and lack of access to supportive services in the community that aid post-hospitalization recovery – that increase the risk of readmission. Public policy efforts intended to reduce hospital readmissions should target the reduction of only avoidable readmissions. In measuring hospital performance, policies must account for many factors beyond hospitals' control in order to facilitate accurate comparisons of performance (American Hospital Association 2015: 1, TrendWatch Newsletter).

Research shows that economically disadvantaged patients often have limited access to services and resources—such as public transportation to get to follow up appointments, grocery stores to support any special dietary needs, and social support—that can help support their recovery post-hospitalization. Therefore, such patients have a higher likelihood of being readmitted. Excluding important sociodemographic factors, such as income, education, occupation and primary language, creates an inherent disadvantage for hospitals treating patient populations at higher risk for readmission.

But it is unclear (a) which factors should be taken into account, (b) how those factors should be measured and collected, (c) and what the resulting data should be used for. The National Quality Forum convened a panel on this topic, attracting more attention than any other measurement project to date. The Institute of Medicine also released a series of five reports “Taking Social Factors into Account for Medicare Payment.” But these factors often expand beyond the issue of socioeconomic status – the report, for example, includes things beyond housing status, transportation issues, to consider race and ethnicity, sexual orientation and gender identity, and disability.

After all, if these data are collected and used to identify patients with social needs – that advocates and policymakers have already readily acknowledged *affect health outcomes*, justifying the need for social risk adjustment in the first place – then our current era of “accountability” could mean meeting those social needs to produce better health outcomes. This returns us back to the original question: as we grapple with which factors are within and outside

of hospitals' control, we also come to consider whether hospitals and health systems should address issues beyond the walls of the clinic. This is a new kind of politics – a politics of boundaries – as health systems are forced to consider the line between clinic and society under programs of quantified accountability. In the exchange below, the Director outlines how health care is only one part of the patient's entire well-being, eventually leading to the recognition of basic social factors beyond what is typically construed as "health-related":

Director: We know how to address this population. We don't get all bent out of shape because we've got homeless people who have psychiatric issues. We just have to figure out how are we going to better manage them. And that's not just us. It's going to take the whole community.

TMC: What do you mean by that, taking the whole community?

Director: I think that because patients touch so many different agencies, we can't just say it's going to all be about the health care they receive here. We can take care of that leg ulcer while they're here, but if we've got other organizations, if they're touching public health, if they're touching a private clinic, one of the community-based organizations, they're getting their care all over...

There's a real opportunity there. We have a lot of community agencies that can really help. Many years ago, I was on the board of a literacy organization. We can change people's lives if in addition to taking care of their physical woes, we can teach them to read. They can get jobs. If we can teach a person how to read, how to use a computer, these days you can't apply for a job without a computer. How do we make sure that their life is better? This community is particularly touch because of the housing situation. There are so many social issues that we need to pay attention to. (Interview, Director of Quality)

Similarly, the Chief Medical Informatics Officer I spoke with at another safety-net hospital called for the greater recognition of the broader environment in which health systems and populations are embedded:

Health systems are going to have to see themselves as part of the larger community, part of the larger political environment. Some of this gets down to how education is funded, it's so unfair that we have unequal opportunity, that a senior graduating from high school in this district has a very different set of opportunities than one graduating from this district, and ultimately those impact people's health outcomes because their education opportunity impacts their employment opportunity, impacts their ultimate income opportunity, impacts their health outcomes... We really need to be interacting at the macro level of policy change, community health, but I don't think we're there yet. (Interview, Chief Medical Informatics Officer)

In both instances, local staff in powerful positions within safety-net institutions consider how the health system connects to other organizations, and similarly, how patient health is connected to a broader array of factors. Literacy, education, employment, and politics all shape the conditions in which patients and health systems interact. This results in the local consideration of that which was traditionally seen as beyond the bounds of the clinic itself.

Thus the issue of boundaries expands consideration of “health-related” factors within broader society, as each health system considers what is and is not within their control, how their work interfaces with other actor’s activity, and social differences among patients may influence likelihood of readmissions. Health systems find themselves considering how the organization itself fits into a broader network of organizations serving the population. These issues reveal an emerging *politics of boundaries* as health systems grapple with issues that cross traditional boundaries between clinic and society. At most of the fieldwork events I attended, there was widespread discussion on how – for better or worse – the health care system is our means of dealing with “social problems.” At the same time, to expect safety-net institutions, that by definition serve a higher proportion of patients with “socioeconomic challenges” to do the “same” with more affluent hospitals – which really means doing *more* given the populations they each serve with less – brings up the issue of what is within the health system’s control, and what is not. Readmissions measures, and their use within accountability programs, bring up a new kind of politics – a politics of boundaries – as the lines between the health system and the rest of society is redrawn through the technical specifications of measurement.

POLITICIZING THE SOCIAL

If the work conducted by the National Quality Forum, as reviewed above, is conducted to achieve “consensus” across stakeholder groups, then one might expect the NQF to “settle

controversies” for the purposes of further social action. This is often presumed to be a relatively non-controversial process, as numbers represent agreement on what it is that is to be measured and how it is to be done. And indeed, social theorists such as Latour (1987) and Boltanski and Thévenot (2006) see numbers as a means of stabilizing socially heterogeneous reality to drive new kinds of knowledge-action.

For Latour (1987), numbers “settle controversies” that then serve as “immutable mobiles” that transcend local environments, creating powerful new “centers of calculation.” By stabilizing reality in a particular way (e.g., agreed upon technical means of assessment), numbers play a critical role in connecting sites to new sites of activity. In this case, accountability programs fulfill this “action at a distance,” with new sites such as the National Quality Forum critical for understanding the process of rendering reality stable. By examining NQF endorsement work, I have shown that the process of establishing standardized means of assessment is filled with heterogeneity and disagreement; however, “consensus-based” endorsement ultimately results in standard technical specifications that then interface with local sites across the nation, thereby building networks outside of the initial point of making the measure itself.

For Boltanski (2011; Boltanski and Thévenot 2006), quantified ways of knowing may be leveraged within political disagreements to address conflicts across actors, groups, and agencies. Given the heterogeneous nature of social reality, reality tests are used to settle disputes. Similar to Latour, Boltanski also finds “consensus” within these tests, as different actors work towards an agreed-upon overarching framework which with to evaluate everyday occurrences. But his sociology of critique recognizes that consensus and conflict co-exist – insofar as there is disagreement in the social world, standardized accounts are needed to drive further social action. For the case at hand, nation-wide political concerns over health care “quality” and

“performance” reflect deeper struggles over controlling matters of care: to that end, stabilized accounts (e.g., quality metrics) are used to drive further social action for the purposes of reform (e.g., accountability programs).

In this case, the National Quality Forum’s rigorous endorsement work is ostensibly carried out to achieve this “consensus,” not only to resolve political conflict over health care, but also communicate to actors across the national landscape and ultimately foster action at a distance through programs of accountability. But in practice, the measures used in the programs provoke national backlash across the entire health care system, with new complaints and disagreements with each annual change to the measures used in these programs. In this instance, the numbers do not settle controversy, but instead *transport them to new locales*, thus bringing political conflicts out in the open in multiple places at once. This is responsible for the widespread expansion of the political nature of health care, as providers, administrators, quality staff, and insurers try to make sense of the measurement decisions that took place during measure development and endorsement. If for Latour numbers solidify the social, here I suggest that they selectively *obscure* the social work it took to create them. In doing so, they create the conditions whereby actors everywhere look to uncover these decisions, increasingly frustrated as they find the decision-making within the technical specifications as largely outside of their control. What is transported, then, is not a stable object under a ready package of “immutable mobiles,” but fixed technical specifications *along with* the uncertainty found within the measure design (the “agreed upon” but unsettled components of the “controversy”). Even after gathering national stakeholders through the forum, and after carefully planned rigorous endorsement, *the numbers remain political* through this particular arrangement of decision-making and implementation.

In this sense, these measures are political *by design* in that the “consensus” achieved does not permit those measures to travel *without* controversy and politics as they are implemented within programs of accountability. By serving as a simple standard metric, local providers and administrators struggle with making sense of the criteria by which their organizational and individual performance will be evaluated. The numbers, then, create the conditions for new kinds of politics. This set of politics extends beyond concerns of federal oversight or surveillance of activity, as *decision-making* behind struggles over evidence, questions over record, and challenges of boundaries is *displaced from local sites to new positions*, such as national fora (e.g., the National Quality Forum) and local implementation teams (e.g., analysts responsible for working with the numbers). This then is the new kind of “politics of numbers” in the emerging knowledge democracy.

CONCLUSION

The tension between the complexity and simplicity of numbers and measures runs throughout all of the work reviewed in this chapter. Latour’s model of science, on the one hand, outlines the building of entire networks following standard agreement of what is to be done and collected. Boltanski’s centralization of political philosophy, on the other hand, recognizes the heterogeneity of values, interests, disagreements that fuel conflict and the necessity of tests to settle disputes. These two models of social activity reflect competing alternative frames of reference, as evidenced in the chapter: the measures in question are simultaneously simple enough to be implemented across the entire national landscape of health care delivery and yet complex enough to generate technical conflicts beneath layers of statistical modelling, new clinical standards, and hours of labor through documentation and programming.

One prominent policy expert with decades of experience in health care payment reform noted this tension as a critique of quantified accountability overall. After all, he told me, no one disagrees with accountability. No one opposes controlling high blood pressure, use of aspirin for IVD patients, or reducing unnecessary readmissions. And yet unlike the advocate quoted earlier on in the chapter who presented this “consensus” across stakeholders in supporting this political development, he understands this agreement as driven by *simplicity* through the political process and the enactment of these accountability programs:

We were at a meeting in which the topic was delivery system reform, and we were talking about the pros and cons of different payment models, how you base payment, the use of performance measurement, and [a former top health aide’s] comment was – because we kept having all of these controversies – “Congress doesn’t have a clue about what you folks are talking about. They know about health insurance, they know big politics, they know this is arcane stuff that is in your hands to deal with because the politicians really don’t know and don’t want to know any of this detail.” To some extent, I think that is right, and yet now it is the law. It used to be they didn’t pass laws about how to measure an individual doctor’s performance. Now they have done that. (Interview, Policy Expert and Physician)

Political concepts of accountability, quantification, and “value,” attract interest across traditional party lines, stakeholder groups, and ideological viewpoints. This further explains why these initiatives continued forward. And yet, nailing down the technical details of making “accountability” happen is driven by *complexity*. Difference in questions of scientific evidence, issues of documentation and responsibility, and social heterogeneity and the clinic-society boundary demonstrate this oppositional tension. Even when there is relative unity in on more simple questions of the evidence to be considered, or the issues that need to be documented or adjusted for, there can exist strong tension and disagreement in the technical details. In such situations, these social and political conflicts – rather than made “transparent” in an open forum – are actually *obscured* beneath layers upon layers of complex administrative work. This is a deep irony to the turn to quantified accountability: while on the one hand it has drawn in different stakeholders into the ever-expanding domain of medicine, creating democratic

“consensus-based” fora where all entities have a right to be heard in the design of accountability, on the other it has brought them in through more and more partial ways, with compromised engagement due to the ever-increasing complexity of administrative work. The greater the complexity and technical detail, the less likely there can be full representation and “consensual” and “collaborative” dialogue.

This has made health care seem increasingly political, as numbers and administrative power and oversight are simultaneously “everywhere and nowhere,” and it is extremely challenging to communicate this complex work to others outside the arena of decision-making. It is precisely *because* of their supposed “simplicity,” ability to transcend distance, and “communicate” particular messages that they create conflict and controversy, even after rigorous “consensus work.” And it is because of this the power and “trust in numbers” (Porter 1995) is called into question, again and again, renewing conflict, controversy, and consensus in the new knowledge democracy.

ACCOUNTING FOR DIFFERENCE

Population Health and the Emerging Technopolitics of Data-Driven Care

Taylor M. Cruz

Department of Social and Behavioral Sciences

University of California, San Francisco

ACCOUNTING FOR DIFFERENCE

Population Health and the Emerging Technopolitics of Data-Driven Care

This paper analyzes the emerging set of politics within new technical tools that accompany the new Population Health Management (PHM) model and associated practices of “data-driven care.” PHM is a health care delivery model that strives to restructure care around patient *populations* through the use of data analytics as a part of the economic shift from fee-for-service to value-based care. Drawing from over eighteen months of policy event fieldwork in Washington DC and the San Francisco Bay Area (n=15 events), in-depth interviews with key experts on delivery system reform (n=13 interviews), and seven months of local fieldwork at one health system site implementing PHM principles (n=400 hours of observation), I identify an emerging *technopolitics of data-driven care* that reconfigures local health care politics. I outline four ideal types of population *forms* found within this model (citizens, categories, classifications, consumers) to highlight the multiple and alternative frames of reference that bring *difference to medicine*. As forms are mobilized in particular ways for the purposes of such social action as clinical care delivery, administrative and cost containment strategies, and targeted patient intervention, I argue that the new evaluative activities that are increasingly becoming a routine part of “data-driven” care are suffused with simultaneously clinical, technical, and political decision-making. Decisions regarding which kind of data are collected, how it is to be collected, and whether or not it is used in successive decision-making and account-building, as well as the work it takes to carry out this activity, become *diffuse* across local actors thus creating new levels of complexity and multiplicity.

Using the case of “population forms” within the Population Health Management model of care, I thus demonstrate the emerging *consequences* of quantification, as specific forms inform

“data-driven care” to “see” and act upon patients in new ways, redistribute organizational and provider resources, and drive administrative and executive decision-making. While the flood of data and technology accompanying data-driven care certainly reconfigures the social and political landscape of health care delivery, it has also generated new challenges of getting data “right.” As health politics increasingly takes place through technical arenas of measures, I suggest that social scientists must have a seat at the table in this emergent local decision-making.

HEALTH CARE REFORM, PRE- AND POST- AFFORDABLE CARE ACT

The passage of the Patient Protection and Affordable Care Act of 2010 has attracted a tremendous amount of attention from journalists, the general public, and academics alike. The bill is widely recognized as one of the most significant reforms of the health care landscape in the United States, driving a series of insurance reforms through several provisions including the individual and large employer mandates, optional state Medicaid expansion, mandated minimum coverage through Essential Health Benefits, and protections for people living with “pre-existing conditions.” Many of these insurance-based reforms continue to dominate the federal health policy discussion, as evidenced by current Congressional activity seeking to stabilize the individual health insurance marketplaces and the recent repeal of the individual mandate as a part of comprehensive tax reform.

While there has been much public attention to the insurance provisions of the Affordable Care Act, the bill’s symbolic significance extends beyond incremental reform towards increasing access to care and insurance coverage. The law itself represents a societal attempt to make health and health care “better” for the broader population and the country overall: in this light, insurance initiatives are part of a much broader, comprehensive package of reforms. Consider the National Quality Strategy (also referred to as the Triple Aim), a dominant organizing principle

guiding contemporary health reform: the Strategy outlines the simultaneous pursuit of “better care, smarter spending, healthier people” nationwide. This reflects state concern with the overall cost and quality of *health care* (rather than mere *health insurance*), especially for the public programs of Medicare and Medicaid.

These nationwide concerns have serious implications for health care delivery. Acting on behalf of the public, the state can no longer tolerate rising costs and unmonitored quality, and has called for new forms of “accountable care.” Carolyn Clancy, former director of Agency for Healthcare Research and Quality (AHRQ), centers this issue of taking societal concerns into account while working towards quality in an address the Senate Subcommittee on Health Care:

The U.S. leads the world in biomedical advances and innovation. However, we do far less well in getting the right care to the right patient at the right time consistently. Moreover, the U.S. spends far more than any other nation, yet numerous studies have found that there is no relationship between high spending and care quality... Today, we have a window of opportunity made possible by all of the attention that is being paid to changing the health care system. We need to be more engaged and aggressive and completely committed to transforming the health care system, because what we are doing clearly is still not good enough (Clancy 2009).

These national concerns are also being acutely felt at the local level of health care delivery, as providers are expected to reconfigure established forms of practice. Consider these remarks from a physician leader at a policy conference on value-based purchasing and delivery system reform:

Now when I’m treating a patient, I not only have to think about their welfare, but also society’s welfare. Historically, I was always taught as a young doctor that we emphasize the patient’s welfare, and let somebody else deal with societal cost. That is no longer the case today. (Interview, Physician leader)

When the state acts on behalf of the public and deems rising costs and unmonitored outcomes or quality unacceptable, these initiatives then demand societal accountability from the historically powerful and autonomous institution of medicine.

The primary means through which health care providers are being held accountable is through the use of data, technology, and quantified knowledge. This is demonstrated, for example, by the ACA’s new Hospital Value-Based Purchasing Program, the Hospital-Acquired Infections Program, the Hospital Readmissions Reduction Program, and efforts to expand public reporting of performance information. These initiatives themselves, however, have a much longer history, and represent deeper struggles over controlling matters of care. Various legislation since the Medicare Modernization Act of 2003 has required standardized data collection on health care performance and “quality,” with hospitals, providers, and other health care facilities reporting quantified information to the Centers for the Medicare and Medicaid Services under pay-for-reporting and pay-for-performance programs.¹⁸ In this manner, the state has taken on an increasingly active role in knowing and acting upon health care through this data collection. Through these programs, the state compares health care providers, redistributes financial rewards and penalties, and drives quality improvement in pursuit of making health care “better.” This has ushered in a new kind of care strategy – that of “data-driven” care.

The goal of this paper is to make sense of the *local* politics that accompany this transition to data-driven care, increasingly fueled by national political tensions as outlined above, through an examination of the implementation of Population Health Management models of care. To bound the scope of the paper, I further the focus on one particular techno-political issue – that of defining a “population” – to show how new data-driven care strategies simultaneously constitute social and political decision-making through “technical practices” of data collection, analysis, and successive social action, thereby offering new opportunities and challenges in making medicine and society “better.” However, on-the-ground actors that are involved in doing this

¹⁸ For a review of the legislation behind these executive programs, see the Congressional Research Service’s series of reports on health care quality (CRS 2009; CRS 2010; CRS 2013; CRS 2017).

work often implement these strategies within specific, generalized *forms* in mind that have particular consequences. Furthermore, this implementation is *diffuse* across multiple actors that each may or may not use resulting data and technology in the manner expected by those charged with building and facilitating the technical infrastructure. Thus, while data-driven care and new care models may be driven by national policy imperatives, I show how this work has generated a different kind of *local* politics as decisions are made about what to measure, how to measure it, and how the data are to be used to drive successive social action. This has created a new problem – how to get the “right” data – under the pretense of delivering the “right care to the right patient at the right time.”

This project draws from three main sources of qualitative data. The first consists of fieldwork conducted at policy conferences, health care industry events, and public gatherings on delivery system reform over the span of twenty four months (n=15 events; January 2016 to December 2017). The majority of these events featured policymakers and public officials, physician leaders from professional medical societies, C-suite executives from nationally recognized health systems and insurance plans, and elite health services researchers involved in national health policy; these groups participated in the events both through public presentation and general attendance. The second source of data includes targeted in-depth, semi-structured interviews with key informants related to health care delivery transformation, quality measurement, and the use of technical tools within reform efforts (n=13 interviews; November 2016 to December 2017). Informants consisted of policymakers, physician leaders, policy experts and consultants, and prominent advocates. These informants were purposively sampled based on participation in the events detailed above, general prominence and recognition in the field (e.g., recommendations from informal discussions at events), and published statements

related to these topics. The third source of data comprises six months of local ethnographic fieldwork at a large public safety-net health system recognized for its data-driven care efforts (n=400 hours of participant observation; June 2017 to December 2017). By following a team of staff working on a state-based delivery system reform program, I attended meetings at the system's hospital and multiple clinics with chief medical directors and local administrators, information technology specialists, data analysts, programmers, and direct practitioners (including primary care providers, nurses, and other support staff). I also conducted additional interviews (n=21) and informal conversations with the many actors involved with data collection and analysis in health care, including those involved with recent implementation of Population Health Management (PHM) within health care delivery across the system. All qualitative data were collected and analyzed following the principles of grounded theory and situational analysis (Charmaz 2007; Clarke 2005).

DATA ANALYTICS AND POPULATION HEALTH

As reform strives to control costs and improve quality, the stakes for these new technical transformations to medicine have been quite high, and have spread to places not previously foreseen as implicated in these issues. With the introduction of Electronic Health Records (EHRs), technology vendors have developed new data analytic tools to support or complement the expansive technical infrastructure. The connection between EHRs and data emerged during an interview I conducted with a Chief Medical Informatics Officer (CMIO). Charged with the task of EHR implementation, she spent much of her time working to integrate the infrastructure within the large health system's many clinics and departments. Even with the significant amount of work that she did with getting the technology off the ground, she barely anticipated the amount data that these digitized records would provide:

When we first went live with our EHR in 2013, I wasn't thinking about data at all, I literally had a one hour meeting about data before we went live. It is just shocking to me to even think back to that day. It was with an entry level analyst who showed me a list of out-of-the-box EHR reports, and said, "Do you want all of these?" and I looked at them and I said, "Yeah, I guess they look kind of interesting." He goes "Okay," and turns them all on...

So the difference between that, and where we are now, when I think about [other local county health systems] are getting reading to go live with an EHR system, they're clearly going to be talking about data from the minute that they start their implementation process. (Interview, Chief Medical Informatics Officer)

The record no longer merely serves as a chart that documents patient care, but is viewed as containing valuable insights that can be used in the practice and administration of future care. Serving as "data buckets," in the words of one technical consultant I spoke with, the EHR digitizes paper records, providing new sources of data that then reconfigure action and decision-making. This emerging form of "data-driven care" aims to use data to improve health outcomes, optimize clinical workflow, and create operational efficiencies, thus transforming oversight and monitoring within matters of care. One data analyst I followed during local fieldwork put it to me thusly, when I asked what it meant by "data-driven care":

What we mean is we're making decisions based on what our data is telling us. We're trying to do everything evidence-based. What we try to do is when we make decisions that are going to be broad-based decisions like opening a new service or changing a service, we really try to make sure that we're looking at metrics that are well-defined and specific to that particular problem. And then that are quantifiable enough where it's not anecdotal, not just a flash in the pan, if you will.

That's basically what we mean by being data-driven, that we see in our everyday operations to make better decisions for your health care outcomes... All those data collection pieces upfront help us make those decisions going down the road based on cultural criteria, based on social criteria, economic criteria. All of those, they all kind of serve their little piece of the purpose. (Interview, Data Analyst)

This turn to data analytics also involves the building of a local infrastructure within hospitals and health systems. For example, one Director of Quality noted how demands to comply with national reporting results in the reconfiguration of local infrastructure that then provides a key starting point for localized use of data:

Director: There's [federally mandated] data that we have to collect, period, the end. We collect it, we put it out there. Then, we also really have to be about improvement. For example, we

have just been asked by the operating room to look at turnaround times. We will be looking at having our team here go in and measure turnaround times, subprocesses, which parts of the process are taking the most time, and doing the time studies and look at roles and responsibilities and use that data to make recommendations for improvement...

I think you have to have both [federal reporting and local improvement]. Sometimes when you start doing improvement work, you're looking at what is all the data that I have related to this? Because sometimes you find out people have been measuring something. It didn't get shared, or it wasn't done with the correct amount of rigor, and so if you see a disparity between that and some other data set, that can actually be helpful.

TMC: It sounds like these federal and state-based initiatives are building an entire quality infrastructure by mandating data reporting.

Director: Oh, absolutely. They definitely have built it... now we have to have a team of people who are basically just our data team. We actually even cut up the suite that way, so we try to put all the data analysts on one aisle, QI coordinators on the other aisle, because we're trying to make sure that we're fluid enough that they're working together, but the roles are very different. (Interview, Director of Quality)

By considering nationally reported data and making use of existing infrastructure (e.g., EHRs, personnel), data-driven care is a variant of national political developments. However, it also goes beyond federal concerns to inform new kinds of internal quality improvement and local decision-making.

While collecting data is a time-intensive, expensive, laborious endeavor, even when facilitated with the rise of new technical tools that accompany Electronic Health Record (EHR) systems, it is hugely consequential for local practices of delivering healthcare. One data analyst, working for a large county health system, described it as follows:

We collect a lot of data, and I'm the person that translates the data for the team, for the most part... so we have specific metrics that are part of the program. It's my responsibility to track the metrics, and the data coming through, validate the reports, and then push the information out to the organization, and come up with quality improvement activities to improve the metrics overall, based on operations. So for a lot of the data part of it, the data is what really drives the change that makes the case. So that's where my role comes in too, as I help make the case, if you will...

Based on their operational workflow, they're [frontline providers] putting all the data into the Electronic Health Record, and it's getting timestamped. We'll basically go in on the back end, look at the metrics, and then basically collect it, analyze it, come up with a result, and then display it and say, "Here's what we found. These are the better practices." And then, that's usually what will drive the change, because, you know, you can't really argue with the research when it's right there in your face. That's really the gist of the data part. (Interview, Data Analyst)

This work entails both collecting data for reporting purposes to the state and federal government as well as the use of data for process and quality improvement. As a result of increased mandated reporting and the availability of new technical infrastructure, such as Electronic Health Records (EHRs) and visualization tools, health systems are increasingly expanding data collection efforts to also collect additional data to address local needs.

In doing so, “data-driven care” creates new loci of decision-making over matters of care. It is these social dynamics that are the focus of this paper. In some ways, it is surprising that these momentous changes fail to attract much public or scholarly attention, especially when compared to the public politics of the insurance provisions of the Affordable Care Act, as these changes have widespread consequences for understanding health politics. At the same time, this lack of interest in “technical matters” was also found among many of the actors involved in the conduct of this work.

At a recent state-sponsored summit organized for executives, providers, and staff of the state's safety net health systems, there was a full session dedicated to the theme of data in health care. “Being data-driven,” a prominent consultant opened with, “means using data to inform all decisions, including decisions within health care.” He notes that despite the growing significance of data within health administration and clinical practice, technical issues of data governance are rarely given serious consideration at the organizational level. With increased mandated state reporting, more and more quality measures endorsed and implemented each year, new internal quality improvement programs, and technical tools that accompany the growing uptake of Electronic Health Record (EHR) systems, questions of data management and literacy will be key for understanding how to make health care “better.” The audience members – representing each safety-net system in the state, sitting at their own respective tables – seem largely uninterested in these “technical” issues, despite the fact that the majority work with data every day carrying out “data-driven care” initiatives within their respective organizations. (Fieldnote, Data Summit)

This technical turn towards data-driven care, while relatively new, also extends from a broader trend towards biomedicalization. As medical sociologists and science and technology studies scholars have documented, medicine has become increasingly technical in the latter half of the twentieth century. For example, scholars have considered the incorporation of scientific evidence within health care through evidence-based medicine (Timmermans and Berg 2003;

Weisz et al. 2007; Burstin et al. 1999), the development of protocols and decision-making tools (Berg 1997; Waring et al. 2015), and the standardized classification of medical knowledge and administrative codes for billing purposes (Mayes and Berenson 2007; Wiener 2000; Bowker and Star 1999). Drawing in part on these developments, Clarke and co-authors (2003) describe this era of biomedicalization, or “the increasingly complex, multisited, multidirectional processes of medicalization, both extended and reconstituted through the new social forms of highly technoscientific biomedicine.” As medicine is remade in technoscientific ways, it undergoes a transformation that drives reform from the *inside-out*.

However, there has been relatively little scholarship that links the politics of health reform to biomedicalization, which is surprising considering that one of dominant ways that medicine is being made technical is through the introduction of EHRs, mandated quality data reporting to the state, and the use of data and technology to driven organizational decisions on care practices. This infrastructure links local health care delivery to the state, with local actors doing a tremendous amount of work to collect and extract data for reporting purposes. But it is also used to facilitate new ways of reforming care through local quality improvement, as the above quote with the Director of Quality and data analyst convey. This is further evident in the paradigm of Population Health Management, with links both the societal goals of controlled costs and better care with these new technical tools and data analytic capabilities.

Taken together, national policy goals of improving care and reducing costs as well as the technological advances that have increasingly become a part of routine care crystallize to reconfigure local health care practices through the use of new kinds of data to “know” patients. Population Health Management (PHM) is a growing model of care that is seen as the product of changing federal aims within the national health care landscape, with a shift from fee-for-service

to a value-based population health model, and an accompanying shift from *individual* patients to new population *groups*:

Population health is a priority because of the financial and outcomes pressure inherent in reform. Not only do providers need to concern themselves with *patients* who seek care, they also now must engage with *whole populations* in order to meet expectations. A population-driven, patient-centered model of care can meet the needs of all consumers regardless of where those consumers are on the continuum of health. (Population Health Alliance 2012: 6; emphasis added)

In contrast to the historical structure of medicine around fee-for-service and unmonitored professional autonomy – a configuration of social relations that has resulted in high costs and potentially compromised quality, thus creating the policy imperative described above – PHM transforms care by recognizing difference among patients and demonstrating the value of care. Rather than have each patient understood under the same frame, population health strives to moves towards transformation around the concept of population. This then reorganizes patients along new lines to develop different kinds of interventions, redistribute resources, and manage groups to optimize health production across the entire population:

The value of having a broad range of organizational and tailored population health interventions is the ability to provide the best (or most appropriate) intervention from the right source and delivered in the right way for each patient, depending on where they are on the health continuum, as well as to enable a measurable change in health status (or outcome). (Population Health Alliance 2012: 22)

This shift then brings about a new set of concerns – a fundamental concern with *difference in care*. This turn towards difference depends upon data collection to “know” patients, create different population groups, and ultimately drive new care practices. This emerging kind of “data-driven care” represents a larger set of new local health politics, as clinical and technical decisions are suffused with social and political values, norms, and frames of reference.

Population health must be understood as a means of engaging with patient heterogeneity through the creation of *patient populations* with the assistance of new data analytic tools. This then allows clinicians, administrators, and insurers to engage with the issue of difference. With

the available data, differences are created to do different things to different patients. Consider the following remarks made by an executive representing a prominent insurance company and an elite consultant at a conference on delivery system reform and quality measurement:

Previously we would treat all patients in the same way [under fee-for-service], *now we are thinking about how to treat people in different ways*. You might see all thirty-five patients for fifteen minutes each – that’s fee-for-service. However, if you think about risk, budget, and other constraints, and *if you have the right data*, then you can now consider bigger things – who do you *need* to see? How long do you need to see them? You no longer have just one appointment structured in the same way to address each patient for each problem. (Fieldnote, Chief Medical Officer representing nationally renowned health plan, emphasis added)

There is *so much* heterogeneity among patients. Let’s embrace that heterogeneity and create similar patient populations – *that’s* what patient-centered care is all about! *Treat patients as the people they really are*. (Fieldnote, Prominent consultant and policy expert, emphasis added)

Through the concept of population, PHM fosters the recognition of difference within matters of care. Competing conceptions of population, however, reflect the political nature of the use of data and technology to transform care, as decisions are made over the use of certain forms over others. In the rest of the chapter, I present four population forms found within the field of health care delivery reform and their selective mobilization on-the-ground within one large public safety-net health system to demonstrate the emerging technopolitics of data-driven care.

FOUR FORMS OF POPULATION

In this section I outline four distinct population forms that appear within the realm of population health. In policy conferences, health care industry events, and public gatherings, interviews, and health systems where I conducted fieldwork, actors made various references to the notion of a population as a means of sorting through and making sense of patient heterogeneity. I present four idealized forms – citizens, categories, classifications, and consumers – that serve as generalized frameworks with which to organize difference. The main focus here is the organizing principle, or the very structure of difference itself (see Table 4.1), rather than the specific content or topical area that comprise the background of difference. In the

following section, I demonstrate how these competing population forms both reflect and give shape to the *technopolitics* of data-driven care; that is, rather than data and technology alone settling the question of the “right” population to attend to in efforts to maximize quality of care – and by extension, what constitutes the “right” data to collect – data and technology became the grounds upon which such questions are now fully contested and explored, creating a new kind of politics co-constituted with the technical realm.

Table 4.1. Four Forms of Difference in Population Health Management (PHM)

	Structure of Difference	Paradigmatic Framework	Difference in Care	Visual Representation
Population as Citizens	Similarity; one single homogenous form	Standards (clinical guidelines, common workflows)	Remove difference; treat all patients similarly	
Population as Categories	Difference across, similarity within; multiple homogenous forms	Disparities, social determinants (different group outcomes)	Standardize difference; treat patients categorically	
Population as Classification	Similarity across, difference within; sorting on a single scale	Risk stratification (likelihood of becoming sick or costly)	Hierarchalize difference; treat patients based on ranking (e.g., groupings)	
Population as Consumers	Difference; Loose clusters of individual heterogeneity	Preferences (patient satisfaction)	Individualize difference; treat patients based on consumer requests	

Population as Citizens

The first form – *population as citizens* – recognizes the centrality of a common peoplehood. This form is characterized by the recognition of one central form of homogeneity for all patients, much as the concept of citizenship denotes equal standing before the law (Porter 1995). This type of form generally takes its cue from the rise of clinical standards under the evidence-based medicine movement, with a goal of disciplining heterogeneity in clinical practice (Burstin et al. 1999; Timmermans and Berg 2003). For example, concerns of patient safety, vaccinations, and preventative screenings at an early age should be equally available to each person in the population, regardless of insurance payer, demographic characteristics, and individual preferences. This recognition of treating all patients “the same” provides a basic starting place in making sense of difference within a more general population. This form of population as citizens can be found throughout the domain of population health as inefficiencies within the delivery process are removed and practices become standardized.

During an expert panel for a conference on value-based care, an obstetrician opens up his presentation with figures of the number of births delivered across the country every year. Almost half of these births are paid for by Medicaid, the nation’s public insurance program for low-income people. He advocates for the development of a common standard pathway that all patients experience when receiving care. Recognizing that maternal care is often plagued with “unnecessary services” provided by “unnecessarily expensive” providers, all patients (including those with high-paying commercial insurance) are provided the same care and given the same consideration within the organization’s care practices. No provider is granted the flexibility to stray from these standard pathways: by adhering to pre-determined rules, costs are contained through ensuring that all providers comply with these formal mechanisms of determining the provision of care. (Field note, practicing clinician and hospital administrator)

The form described here is one that see patients as the same, worthy of the same kind of treatment. This consists of standardizing workflows, and ignoring particular kinds of heterogeneity (e.g., income and social status, insurance provider and reimbursement rates). The commonality among patients is seen as greater than the differences among them, thereby facilitating one overall population.

Population as Categories

The second form – *population as categories* – recognizes heterogeneity insofar as it aims to create multiple forms of contrasting homogeneity. This structure of difference is one of complementing similarities, such as that there is difference *across* groups but similarity *within* them (see Figure 4.1, “Visual Representation”). This form draws from parallel developments in health and health care disparities and the social determinants of health. Both of these paradigms centralize the importance of basic categorical differences – with regard to income and financial security, racial power and status, and general well-being and dis/advantage – that organize social heterogeneity along strongly structured lines. The form of difference is characterized by *relationality* in a way that the citizen form elides, and structures this relationality into oppositional terms: some group experiences a particular kind of concerns, and a competing group experiences a different set of concerns.

During the same expert panel with the obstetrician advocating standardized workflows across patients, an oncologist situates the patient population she cares for at her specialty center. Located in rural New Mexico, she stresses the poverty of the geographic area and the difficulty most of her patients have in coming in for regular treatment. With co-pays and transportation costs being prohibitive, especially when money could otherwise go towards food and shelter, she chides advocates of value-based care that argue for the need of patients have shared responsibility for their health care costs. What could be a better case of “skin in the game” when the issue at hand is already about the patient’s own body, well-being, and life itself? She relates to the audience the centrality of economic vulnerability in treating patients: “Once you know how to take care of poor people, the rich ones are easy.” Centering and addressing basic social needs is fundamental to improving the quality of care. (Field note, practicing clinician and board member of elite medical society)

This form does not discipline heterogeneity from medicine, but organizes it along categorical lines and employs it as a resource in understanding, knowing, and acting upon patients.

Centering heterogeneity in this fashion standardizes differences among patients in order to create multiple populations that can be treated differently. In this case, the Medicare population will be seen as having a different set of concerns from the commercial population, and low-income and

middle-class patients will be recognized as distinct *kinds* of patients. Patients who primarily come in for acute treatment are contrasted from those living with chronic conditions, with new administrative and medical recognition of social heterogeneity through the category form. While there may be disagreements as to the *content* of difference (e.g., the dimensions along which to sort patients), the *structure* of difference itself (e.g., multiple homogenous groups) is maintained within the generalized category form.

Population as Classifications

A third form – *population as classifications* – further recognizes the emerging importance of heterogeneity within medicine, but moves beyond the organized form of multiple homogeneities. This form instead recognizes differences *within* the commonality of all patients, and sorts all differences to a ranked order along a single continuum, which may then be classified into *relative* groupings (e.g., “high,” “medium,” “low”). This new organizational form, what some have referred to as *ordinalization* (Fourcade 2016), has emerged within social realms beyond health care in the form of credit ratings, online evaluation schemes, and public ranking practices (Fourcade and Healy 2013; Scott and Orlikowski 2012; Espeland and Sauder 2016). This form relies upon commensuration as different entities are compared and rendered intelligible against a common metric (Espeland and Stevens 1998), transforming difference from a matter of *kind* (as in the category form) to a matter of *degree*. This form of difference is also characterized by relationality, but structures this relative difference in a different way, thereby representing a distinct population form. At a conference on delivery system reform, this form was pervasive within consideration of “differences” among patients:

At the same conference (but not during the expert panel of specialists), both consultants and administrators call for the recognition of heterogeneity among patients via hierarchalized difference. “It’s all about stratifying patients and finding out which ones cost more, which ones bring in the revenue. Does it make sense to focus on an entire total population, or for some providers to focus on a subsegment?” Others speak to the paramount significance of coding for

disease severity, recognizing that there are varying levels of progression even within the “same” population afflicted with a common condition. Another consultant puts forth the Patient Activation Measure as a means of assessing meaningful differences among patients, suggesting that some patients do well at managing their health while others do not. He presents three levels of patient activation: Level 1 “takes no responsibility” and “blames” providers for health status, Level 2 knows a little and is somewhat engaged in care, and Level 3 can engage with evidence-based guidelines and takes full responsibility for their health. In each case, activity is expected to be organized along these lines, treating patients themselves according to difference by degree. (Field note, Consultants and health care administrators)

This form organizes difference within a hierarchy, in which certain patients are deemed to have higher associated costs, risk profiles, and disease severity than others. But this heterogeneity is not merely mapped onto categorical heterogeneities, as there is gradation within each dimension. Just as in the category form, there may be disagreements as to the content (e.g., which dimensions are to be sorted), but for our purposes, here the structure of differences through this population form is of primary relevance. Difference is organized along a scale, with populations sorted under the banners of “low,” “medium,” and “high,” and resources and attention distributed thusly. This form creates hierarchies of difference, thus relating patients to each other in new kinds of ways.

Population as Consumers

A fourth form – *population as consumers* – takes heterogeneity one step further to individualize difference. Drawing heavily from neoclassical economic theory, the patient is seen as a consumer, or the agent that purchases health care goods and services. However, the population form of consumer should not be interpreted as solely referring to the purchasing patterns of the individual patient, but rather denotes a particular orientation to treatment. For example, the turn to patient experience recognizes that patients have distinct “preferences”: some may prefer after-hours clinic times, non-English languages spoken by providers and front desk staff, or care within specific geographic locations. While each individual will have a unique configuration of “preferences” that seemingly denote an impossible heterogeneity that resists

organization by form, in practice individuals may be loosely grouped into clusters of patients.

Unlike the category and classification forms, these clusters need not be situated relationally, but may each stand on their own, often overlapping along certain dimensions.

The final presenter on the expert panel is an orthopedic surgeon, who related the importance of performing well on outcome measures (in his case, hip and knee replacements). Because of his success in providing these services with relatively high optimal health outcomes with few complications and a relatively low final price tag, he has attracted the attention of third party payers all over the country. Insurers would rather send patients to him for this one-time procedure than use contracted specialists within their own networks, as their own providers cannot compete with his performance with regard to both outcome and cost. When the panel moderator asks him to elaborate on how specialists can improve health care value, he speaks to the importance of considering individual patient request and motivation. Cutting back on “unnecessary” hip and knee replacements for certain patients (e.g., near end of life in which these high cost procedures have little economic payoff in terms of quality and duration) will not bode well with patients who take up matters of their own health as something through which they create meaning and purpose in their old age. Specialists must take into account these individual patient preferences when discussing “value” in providing care, he argues, and recognize that reducing or changing services may comprise quality from the perspective of the patient. (Field note, Practicing clinician and board member of elite medical society)

Differences among patients are taken into consideration in a way that recognizes individual concerns and desires, which are then organized around populations with similar kinds of preferences.

THE EMERGING TECHNOLITICS OF DATA-DRIVEN CARE

In outlining these four population forms, I provide the groundwork for understanding a new kind of local health politics – what I refer to as the *technopolitics of data-driven care*. As health care becomes increasingly data-driven with the uptake of Electronic Health Records (EHRs) and new data analytic tools, the successive hiring of coders and technical analysts within hospitals and health systems, and mandated state and federal quality reporting to insurers and public agencies under health reform implementation, medicine is undergoing a widespread transformation through the uptake of data, technology, and quantified knowledge. And while much of this is driven at the federal level in the spirit of the Triple Aim – “better care, healthier people, smarter people” – a new kind of local politics accompanies this uptake of data and

technology as decisions are made about what kinds of information are collected, how data are to be generated, and how knowledge is integrated to transform health care – a technopolitics of data-driven care.

The emergence of multiple population forms in practice, alongside one another, suggests a continued multi-sited and complex technoscientific transformation of medicine under biomedicalization (Clarke et al. 2003). But when individually mobilized, these various forms are implicated in fundamental social processes of sociological concern: they restructure the professional gaze, shaping how those in power “see” those without; redistribute resources, time, and attention along new lines, targeting certain people over others; and justify particular kinds of activity in maintaining local and organizational accountability. Crucially, biomedicalized reform in this case draws upon heterogeneity and difference *in different ways* – thereby creating a new kind of localized health politics. These decisions over the kinds of differences that matter, how they will be considered, and whether they will be acted upon now take place through the use of data and technology in care.

To demonstrate this, I draw from local health system fieldwork including interview data I conducted with actors that work with health care data implementing PHM principles. A data analyst with extensive experience working with the data as a part of PHM, including the creation of new local databases containing specific populations, emphasized the importance of definition to this work:

When you look at the population, you have to separate it out. I hate that word segregation, but really it's segregating specific populations, but in the intent to provide better care going forward. Population Health Management is looking at specific disease population criteria and the people in there in those populations and finding commonalities...

It allows us to not only pull patients, but pull resources too. We create standards and protocols and have resources available. Population health management allows us to strategize and prepare, and almost sort of prevent ongoing continuance of this particular prevalence of whatever it is. That's kind of what I think of when I think of population health management, it's a strategy to not only manage but cease the continuance of the population growing.

For instance, when it comes to race and ethnicity, different rates of heart complications run more prevalent in certain races. Having the ability to research that and understand that can change the way that we deliver health care for a specific population of diabetics, of hypertensives. Reviewing the data and seeing the outcomes from specific interventions for specific populations allows us to fine tune our care so that we can provide it not only more efficiently, but effectively for more people. (Interview, Data Analyst)

In this general overview of PHM, the analyst recognizes the significance of *difference* among patients. This use of difference transforms patients into populations for the purposes of improving care, redirecting resources, and overall making health care “better.” In each instance, the recognition of difference – and thus the formation of populations – depends on data availability, thus demonstrating the connection between PHM and “data-driven” care. This creates a new dilemma as local actors struggle to assess what constitutes the “right” data, taking advantage of mandated data collection from the state and CMS while also exploring other means of creating populations.

This was most readily apparent with new data collection practices on demographic information. The recent push to collect identifying information, namely race and ethnicity, but also sexual orientation and gender identity, within population-based data systems under recent state initiatives, highlighted the centrality of the *category* form of population. The collection of this information was commonly used to justify the splitting of patients into populations overall, although I will show that this is not the only population form mobilized within the health system. In an interview I conducted with a data analyst, he commonly referred to sociodemographic characteristics as the main means of sorting populations:

TMC: So it sounds like Population Health Management depends on organizing patients into different populations. Can you give me any examples of how populations might be organized?

Data Analyst: You can do it by like, as we were just saying, race, sexual identity, gender. Different buckets of groups of people. That’s just the way to look at it is whatever the health focus is, if it’s diabetes, blood pressure, depression screening, you place people into that particular population and you just try to manage that piece of it. (Interview, Data Analyst)

Under new efforts to collect patient data on race, ethnicity, and language (REAL) and sexuality orientation and gender identity (SOGI), race, gender, and sexuality serve as bases from which to form populations. This form of difference is categorical, both in definition and in practice. This data collection did not take place within the health system until recently, and now that data are collected they may serve as a basis of sorting through patients for the purposes of particular kinds of action. Health care now may consider issues of difference and heterogeneity through these technical means, and based on the data and the categories specified within collection practices.

But the categorical form is not necessarily the only population form. One of the clearest examples of the *classification* population form, where patients are ranked by some quantitatively measured dimension, is the use of a complexity score to reduce readmissions and further costly care. In this instance, PHM is used to create new populations of “complex care” patients (or relatedly, “high utilizers”) that then receive additional attention, time and resources. The sorting of these patients into the population, crucially, draws from the EHR data to calculate a complexity score that is then used to rank severity of need and patient acuity. Consider the following:

The complexity score is basically a score that the EHR develops based on certain reviews of the record. If you have diabetes, you have a risk score of one. If you haven't been seen in our clinic in two years, that's another one added. That risk score is basically, it's a gauge if you will to tell you how complex this particular patient can be based on all these different present criteria. The score will be from 1 to 40 for the complex care score. It's an accumulation of different things that happened in the case, like if you're a smoker, one point, diabetic, two points. Haven't seen the clinic, three points.

That score allows us to identify more acute patients who need probably more wraparound services. What the outreach team does is it basically prioritizes patient outreach by that complexity score. It's kind of a scrub of the record. That's how all of that relates into it is we'll define specific populations based on their medical histories and such. They get put into these populations. Then they're managed basically on that score. (Interview, Data Analyst)

Here, populations represent classifications based a quantitative calculation of “complexity.”

Demonstrating the political nature of these new technologies sorting with data, more “complex” and “acute” patients then received specific resources, as the health system created a Vivitrol

program specifically for those “complex care” patients. Vivitrol (Naltrexone) is a monthly injection used to manage alcohol and drug dependency. But as an expensive medication, it is not made freely available to any patient suffering from alcohol or drug dependency: rather, it is targeted for “high-risk” patients based on the complexity score. I also observed visits between these “complex” patients and special nurses designated for their care: unlike the usual 15-minute encounter between one provider and the patient, these visits were an hour long and comprised of two nursing staff, with at least one check-in from a primary care provider through the course of the visit. In this case, the ordinal sorting of patients into a specific population redirects not only material resources in terms of expensive medications and treatments, but also the ordinal redistribution of time and attention from clinic staff.

However, not all of the population forms I outlined above appeared through local health system fieldwork. After synthesizing and constructing the population form ideal types presented in Table 4.1, I at first expected populations to also be created through *consumer* preferences, especially since there are several workgroups and regular meetings that take place monthly dedicated to the issue of patient experience. In an interview with a QI coordinator who was charged with responsibility for both hospital patient experiences measures and the complex care program, I asked him about these competing means of creating populations:

TMC: So you mentioned the importance of quantification for the complexity score. Is there an effort to quantify patient preferences, or is there a way that that’s captured?

Coordinator: We have what you call HCAHPS and CGCAHPS Scores... HCAHPS score is a standardized tool that the state of California and CMS are using. It’s a scoring system and based on the results and the standing of the hospital you can compare the standing of the hospital throughout the United States. So yeah, this is a way of quantifying that.

TMC: Does the health system use patient satisfaction surveys to create similar patient populations, similar to the complexity score? You mentioned earlier there are high-risk population, a low-risk population. Are there efforts to create populations based on preferences?

Coordinator: That’s a good question. I actually do not know the answer to that, but I don’t believe so. (laughs) (Interview, QI Coordinator)

Here the QI coordinator acknowledges that certain structures of difference are more likely to be employed than others. Patient preferences, mobilized at a national panel by the orthopedic surgeon to recognize “meaningful differences” among patients (detailed above), in this instance is not pursued at this health system. That is not to say that it is not mobilized elsewhere, at another health system or in some other clinic for example. Rather, *local decision-making determines what is to serve as the basis of populations within any given health system to shape the practice of care*. The creation and inclusion of the complexity score with the EHR, for example, was not mandated by the state or CMS, and yet it quickly became instituted within health system overall, crossing from the emergency room and hospital-based services and spilling into primary care within the health system’s outpatient clinics.

This forming of a population is not only critical for the redistribution of resources, as the case of the Vivitrol program and the assignment of extra nursing care shows, but also can shape the way that providers “see” patients on the record before they come in. For the nurses seeing “complex” patients, the EHR screen provides a colored circle with a number inside it next to the patient name – this is the complexity score, and it redirects clinician focus to particular kinds of information. Most of primary care providers I spoke with also showed how the complexity score was included right alongside patient name, sex, and date of birth in their daily appointment log. This directs provider gaze to focus on particular kinds of calculated information, in this instance, provide an ordered rank of patient illness severity through complexity.

The local decision-making that accompanies data collection, calculation, and population management is further complicated by the fragmented presentation and availability of this information to staff. Based on my interviews and fieldwork, I learned that each EHR screen is fully customizable to the end user suggesting that the medical gaze may be recreated in partial

ways across different providers. To that end, one nurse manager did not include the complexity score, but instead added other information, namely race and ethnicity. In contrast, most of the primary care providers I interacted with did not necessarily consult race and ethnicity information prior to interacting with patients. If this information is not customized to show up on each patient line in the daily appointment log – in similar fashion to complexity score, as reviewed above – then it is tucked away behind layers of tabs. This suggests that there is a great degree of *local customization* as to which kinds of “population data” are used to reconfigure how providers “see” the patients they treat. As data are used in different ways by different providers, even within the same outpatient clinic within a larger health system, data-driven care is deeply intertwined with an overarching technopolitics.

Furthermore, data that are collected does not necessarily result in further social action. While “data-driven care” claims to use data to inform decision-making, redirect resources, and drive actionable change, not all data are acted upon in equal manner. A primary example of this emerged with a recent disparity plan that the health system developed, particularly to reduce “disparities” for Latinos with diabetes. The state program mandated the development of disparity reduction plan making use of mandated race, ethnicity, and language data collection that was initiated a few years ago. The health system, in turn, decided to target Latinos with poorly controlled blood glucose levels. Based on my observations of meetings with health center managers and clinical leadership, there was a considerable amount of discomfort among most (but certainly not all) staff when this Latino diabetes disparity reduction plan was discussed. Even after steady data collection on this demographic information from patients, and despite treating high numbers of minority groups within the outpatient clinics they each manage or provide care in, most staff were reluctant to use the data as a means of redirecting particular

attention or focus for issues of categorical social difference. This suggested that once race and ethnicity data are collected, there is a significant amount of work that must go into ensuring that it is used for population-level disparity reduction, despite the widely recognized principles of PHM already being in place in the organization.

I followed up on this in an interview with the team data analyst, who confirmed that data *collection* does not necessarily result in data *use*, even when data-driven care and PHM principles call for the use of data to inform decision-making, redirect resources, and improve care delivery. This further demonstrates the new loci of decision-making, not just as to which data should be collected or how scores are to be calculated, but also whether data are accessed, analyzed, and acted upon. I asked the data analyst why this plan was so challenging to implement, even after the data had already been collected for some time now. He in turn related to me that data collection does not necessarily result in data use, and that the disparity plan was created solely because it was tied to financial reimbursement under the pay-for-performance program. Consider the following exchange:

TMC: So – correct me if I’m wrong – even if the data were collected, on social determinants of health, complexity score, race, sexual orientation, whatever the case may be... without making people do the reduction plan, there’s no guarantee they would use it?

Data Analyst: Oh yeah. If there was no reduction plan, the data would just sit in a silo and would be good to use, good to have. It would be more like a trophy on a shelf. It’s good to look at. You reference it. Maybe you use it in your quality review committee, maybe. Because we have, like I was saying, because we have all those pieces in there together, it allows us to create a bigger picture of all the potential little pieces.

TMC: All the providers have access to this information too, right?

Data Analyst: Mm-hmm (affirmative). If they want it. We gave it to them...

We tend to do more over- data collection than use. I think that’s probably everywhere... We end up doing a little more analysis that we do taking action per se. There’s been lots of times where we say, “Maybe we should take action based on this data.” We pull back and we do our due diligence and we learn, maybe it was probably smarter waiting to get all the details or wait for the next release or whatever the case may be. (Interview, Data Analyst)

The reduction plan, in this case, was created as a part of a larger pay-for-performance program: in other words, the use of data to “act” on disparity was primarily driven by financial reimbursement. Even though the data had been readily available, the data were not widely used to “act” upon populations at least as it pertains to the categorical form. Even after a significant amount of work has already taken place to collect the data – programmers created a new build within the EHR, quality staff provide trainings on how to collect the information, and health center managers and receptions oversee the survey administration to patients and then enter the data into the proper EHR field – there is no guarantee that the data are used to find “meaningful differences” among patients, or that particular population forms are mobilized to redirect resources or drive different care practices. This is another case in which the *technopolitics of data-driven care* plays a critical role and reflects the emerging consequences of quantification in health care, as new loci of localized decision-making appear within individual health systems and across the health care delivery landscape.

The use of data and technology within health care – in this case, PHM tools drawing from EHRs – does not eliminate social and political decision-making within care, but instead transforms it in new ways. The social and political nature of data and technology is not merely imposed from the outside (e.g., through policy “action at a distance” from Washington DC), but is found embedded within the new local techno-politics of data-driven care. Policymakers, national leaders, advocates, and vendors claim to use technology and data to make health care “better,” but as the case of Population Health Management shows, there is hardly agreement on how change should be brought to medicine. Through various population forms, patients are “known” and “acted upon” in new kinds of ways. But of utmost importance, *there is difference within these new forms of knowledge-action*, as the multiplicity of forms to be found within a

given arena exemplifies the impossibility of arriving at a single, straightforward answer to what kinds of data about what kinds of populations would be useful in making care “better.” Instead, an emergent politics of *form* is now relevant – as multiple forms are recognized, social and political decisions must be made to select one population form over another. Increased data collection and technical capabilities within the context of health care provision cannot alone specify the “right” population form, as there are various possible forms in acting on difference. Applying the concept of “population” to recognize “patient heterogeneity” does little to address, and in fact only underscores, a more fundamental question: who and what constitutes a population, under which kinds of circumstances, and for which particular aims (Krieger 2012; Shim 2014; Cruz 2017)? While data and technology reconfigure the medical gaze from viewing individual patients on a one-on-one basis, the shift to particular generalized *forms* is imbued with a socially significant *techno-politics*.

CONCLUSION

As health care increasingly finds itself under pressure to take into account “the social” conditions of medical-social arrangements – whether these are national concerns over rising cost or political pressures to improve quality – data and technology are widely expected to this accounting work. Using the case of Population Health Management (PHM), I examined one major case of using data to take into account patient heterogeneity through the sorting of patients into populations. In many ways, this turn to recognize difference is often associated with the recognition of the “social nature of health” is concerned. At many of the key events I attended, the social determinants of health were widely discussed as an “emerging issue,” with many now reflecting on the role of medicine in addressing social inequality through the use of data. As one

executive representing a prominent safety net hospital suggested, information technology is rapidly becoming implicated in these concerns:

We need to transform data into actionable information, and use that data to stratify patients and match the intervention. We know that most of the stuff that happens to patients happens out there in their day-to-day life. We know food deserts, job stability, education, and social factors all affect health. But as clinicians, *this is what you can do* – you can match the patient to the intervention. (Field note, remarks from executive and board member representing safety net hospital)

Technopolitics also refers, then, to how data and technology are being used to know and act upon “the social.” This requires social and political decision-making that occurs at the local level, as is *distributed across actors*. Primary care providers, health center managers, local data analysts and coders, technology specialists, and organizational leadership each do infrastructural work (Bowker and Star 1999; Star 1999) to bring about this change, each with varying degrees of knowledge and appreciation for these issues.

Under this new paradigm of data-driven care, data will have to be used to drive further action, fueling additional political concerns of how and whether data are used for which kinds of purposes:

Quality [work in health care] has transformed, and in many ways that are good and bad, which is we’ve become obsessive with data. That’s a good thing in some sense, but only if you do something with the data. Our role here now is to figure out – and I really take the role seriously – how do we transform all of this to better meet the needs of the future? (Interview, Director of Quality)

As this quote and the field note above make abundantly clear, data and technology are being used to reshape health care. But this is not carried out in a deterministic manner: instead of closing off debates, struggles, and disagreements, here the availability of new forms of data creates new opportunities for deliberation as to how providers should “see” patients, where resources should be directed, and through which means actors should be held accountable for patient care and outcomes.

To return to the former AHRQ director’s definition of quality – “doing the right thing for the right patient at the right time, every time” – we must now consider what constitutes the “right data” under data-driven care initiatives. This will consist of sociotechnical decisions, as social and political processes are increasingly made technoscientific within matters of care. But this also means that social scientists must continue to contribute a wide array of insights – theoretical arguments, data analytic skills, coordination across distinct social groups of people – in order to create a more just world that takes health equity and the role of medicine seriously. Social scientists cannot merely attend to national level issues, but neither can we remain grounded in the local without a broader understanding of how this change is happening across medicine. As the accountability endeavor in health care continues to spread to new domains, implicating new kinds of actors – data scientists, technology vendors, administrators, and policymakers and experts – social scientists must find a way to have a seat at the table. We must engage to bring the change to medicine, bringing “the social” into health care yet again, simultaneously transforming medicine and society, side-by-side, one through the other, in the spirit of making things “better.”

CONCLUSION

The Sociotechnical Transformation of American Medicine

Taylor M. Cruz

Department of Social and Behavioral Sciences

University of California, San Francisco

CONCLUSION

The Sociotechnical Transformation of American Medicine

In the previous three chapters, I presented the origin, design, and consequences of quantification. In the second chapter, I showed how quantification emerged at the turn of the twentieth century out of multiple conflicting critiques of medical-social relations. These critiques were put forth by physician leaders, public health officials, health economists, and minority health advocates, spurring successive movements towards evidence-based medicine, managed care, and disparity monitoring. Quantification served as a means of continuing these programs of reform to the institution of medicine, as multiple social groups developed expectations surrounding the measurement of health care delivery. In the third chapter, I demonstrated that evaluative work (e.g., endorsement through the National Quality Forum) confronts a tension between simplicity and complexity that is ultimately flattened through the technical specifications mandated by quantification. But rather than settle controversy and stabilize reality, these metrics then generate new kinds of politicization within health systems around the country as on-the-ground actors struggle to make sense of specific means of measurement. In the fourth chapter, I considered the local decision-making that takes places through new data-driven care initiatives. Local health systems have considerable flexibility in the use of data and quantified knowledge within internal operations, creating new loci of interpretation and justification that creates novel forms of complexity across organizations. By examining the social and political dimensions of quantification, these three chapters address the historical development of “reform by numbers,” the move between national endorsement and health system implementation, and the changing jurisdiction of medicine in the era of accountable care.

In this conclusion, I now turn to consider the question of medical power in the twenty-first century. The topic of medical social power has a long history within social science

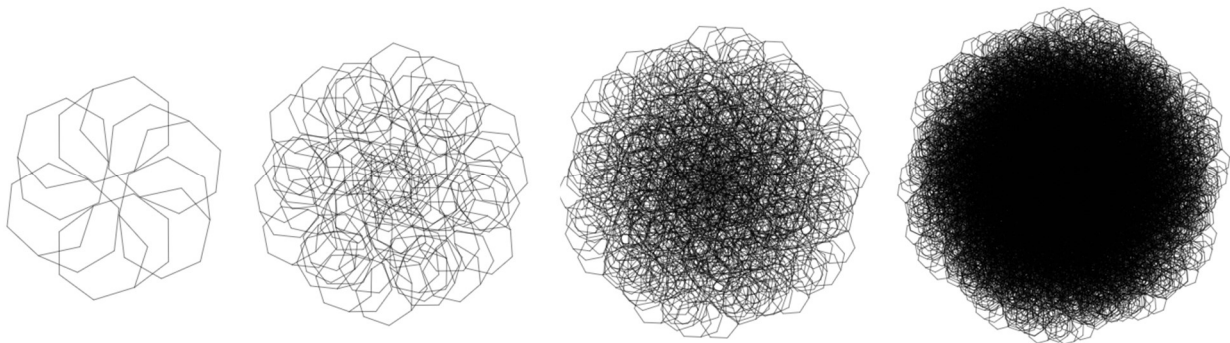
scholarship (Starr 1982; Timmermans and Oh 2010). Freidson's (1970) work seems most clearly aligned with the Golden Years of Doctoring, privileging the importance of the profession for evaluating itself in terms of licensing and credentialing. Marxist scholars considered the on-going crisis of medicine, looking to corporatization as a means of understanding medical-social relations in the 1970s and 1980s (Navarro 1986; McKinlay and Stoeckle 1988). Contemporary sociologists have instead privileged the importance of evidence-based guidelines in understanding the shifting nature of medical work (Hafferty and Light 1995; Timmermans and Berg 2003). Medicine has undergone clear transformation over the past few decades, as documented by this influential scholarship.

The key argument of this dissertation is that medical-social relations have *transformed* with the rise of quantified knowledge, data, and technology, in some ways continuing these decades-long developments but in others transforming them into a wholly new phenomenon. Following Boltanski and Chiapello's (2006) lead, we might argue that there is a "new spirit" of medicine: the Golden Years of Doctoring, once characterized by "autonomy," "trust," "professionalism," have evolved through times of "corporatization," "evidence," "management," onto our current era of "transparency," "accountability," and "value." Scholarship from STS scholars has convincingly demonstrated that science and technology refracture existing relations, creating new alliances, groupings, objects, and justifications (Bijker 1995; Latour 2005; Callon 1986). If this is the case, then quantified accountability has not just "continued" the transformation (Light 1993; Timmermans and Oh 2010), but accelerated it in new kinds of consequential ways.

As I showed in the previous three data chapters, pinning down "accountability" and "transparency" is quite the daunting sociotechnical task. For all of the focus on "data-driven"

decision-making, for all of the public ranking systems aiming to compare hospitals and providers, for all of the financial incentives and penalties distributed across competing entities, we cannot understand these things outside of thinking about social complexity (Law and Mol 2002). I draw the reader's attention to the successive fractals depicted in Figure 5.1.

Figure 5.1. The Complexity of Sociotechnical Transformation



In the first image, we see six overlapping shapes: let us consider each of them as six prominent stakeholder groups, say providers, payers, purchasers, patients, policy makers, and the general public. Some groups have closer affinities with others – for example, some purchasers (large employers) have relatively little regard for the well-being of the general public beyond company employees, and patients may have relatively limited direct access to influential policy makers – but overall, there are a number of groups vying for political power to control matters of care. In the second image, the shapes have multiplied in number, in varying sizes and continue to overlap each other. Compared to the first image, the second image is complex. It not only contains more shapes, but it has introduced a dimension of size as smaller shapes hover near the original shapes. This also results in the piling on of more and more layers, refracting the size of original shapes and cutting through already existing boundaries.

But as time goes on, complexity only increases in greater and greater magnitude. The third and fourth images are merely successive steps to these developments. Taken together, these images reflect the complexity of the *sociotechnical transformation of American medicine*. My argument is that with the introduction of data, technology, and quantified knowledge into the health care realm, health politics have become *refracted* in a way that adds more and more layers of complexity to all dimensions of the work of medicine. The National Quality Forum, for example, is one prominent layer, or node, that gathers different stakeholders and conducts needed evaluation work. The implementation of Electronic Health Records (EHRs) is another layer with its own sets of concerns and issues, with vendors, CMIOs, and technical consultants continuing this work. New measure advisory boards make decisions about which measures to select from the list of NQF-endorsed measures for individual accountability programs, quietly drawing together key actors to implement programs. Public reporting entities work towards “transparency” by “releasing the data” to the general public hoping to “empower consumers,” but the source of the data, its manipulation through data cleaning, risk adjustment, and practices of disaggregation are lost in the excitement and anxiety over the newly released scores. The era of accountable care has ushered in unprecedented change to medicine, bringing in a full force sociotechnical transformation.

If more and more actors are being brought into the fold of medicine through critique (Chapter 2), and quantification work not only does not result in closure (Chapter 3) and instead creates new sites of decisions-making (Chapter 4), then medicine has both expanded and become increasingly politicized. It is a deep irony of the quantified accountability endeavor: while external entities aim to “democratize” medical power and hold the profession “accountable,” the account-building ushered in has created much administrative complexity that *obscures* as much

as it reveals (consider Figure 5.1). It is for this reason that the entire social landscape is characterized by both “trust” and “distrust” – while on the one hand, everyone needs to use data and technology to carry on their own work, on the other hand, individuals and entities are extremely skeptical towards others when it comes to these issues. Because decision-making is buried beneath layers of complexity, it is difficult to assess what on-the-ground actors can “trust.” Issues of data “accuracy,” the configuration of specific technological forms, and the building of reports all emerge throughout the field, in some ways making their effects felt and in others leaving room for customization. But because of this push towards standardization and hard-fast technical rules, in many instances without full explanation, there is widespread distrust that these developments were made properly. These generates a new politics *through* technology within health care today.

THEORETICAL CONTRIBUTIONS

This project’s engagement with Latour’s actor-network theory and Boltanski’s sociology of critique sought to not only explore these theoretical directions through empirical social science research, but to build upon these programs and modify them. Drawing from Boltanski (2011), I showed how quantification emerged out of long-standing critiques of medical-social relations. Directly challenging dominant “critical” approaches in medical sociology – such as the work by Navarro, McKinlay, and Waitzkin, for example (Navarro 1985; Navarro 1986; McKinlay and Stoeckle 1988; Waitzkin 1978) – which develop a strong “all-knowing” critique of the capitalist basis of health care, I demonstrate that such an approach fails to take into account the multiple sources of critique within the social world. Indeed, many of the critiques I outlined in Chapter 2 were developed during the same time period that the Marxist approach to medical sociology achieved dominance. Critiques of inefficiency and issues of cost control led to

payment reform in the 1980s and 1990s; critiques of equity led to new political work on health disparities; and critiques of effectiveness created the conditions for the emergence of clinical guidelines under evidence-based medicine. And yet, these critical sociologists oftentimes only recognize these “other” critiques in a partial and disconnected manner – or even worse, they do not consider them at all. By taking the work that all actors do, from health economists to public health advocates to physician leaders, to “render reality unacceptable” so as to facilitate future social change and considering it as part of the everyday work of making things “better,” sociologists are able to understand the dynamics of reform in a new light.

However, Boltanski’s turn to consider all actors and their critical capacities lacks serious attendance to power differentials in a stratified society. On the one hand, his approach is admirable for considering how legitimacy and worth are co-constructed in everyday work, particularly between parties engaged in public dispute (Boltanski and Thévenot 2006). This allows us to consider the significant work carried out on the part of everyday people, without immediately resorting to “power” as a full-force explanatory variable (as is common within critical sociology). On the other hand, taking all actor’s accounts seriously, while a noble endeavor, ignores that certain accounts are almost always afforded greater authority over others. Thus, my analysis did not strictly consider the everyday work of critique of “on-the-ground actors” per se as suggested by Boltanski (2011), but rather focused in particular on actors and sources with access to institutional legitimacy for the purposes of public policymaking. I believe this is a much-needed modification to the sociology of critique, as there is the risk of losing sight of the uneven distribution of legitimacy and credibility in society when the sociologist “democratizes” critique for “all.”

Latour's actor-network theory offers a full and engaging approach with which to make sense of science, technology, and society, and by extension, the use of quantified knowledge, data, and technical accounts in medical-social relations. His call to "reassemble the social" as a series of associations through which the social is constituted, rather than a separate domain that stands alone, proved particularly useful in this dissertation (Latour 2005). I demonstrated "the social" is actively made and constituted through "the technical," resulting in the *sociotechnical*. Decisions regarding "social factor" risk adjustment for readmissions, the labors and difficulties of documentation, and the use of forms in Population Health Management all lend empirical support to the bold claim that the social and technical no longer reside in separate domains.

At the same time, Latour's approach, especially in its earlier versions (Latour 1987; Latour 1988), requires significant modification in studying the political use of numbers. The "immutable mobiles" in this project (quantified quality measures) were both immutable (fixed technical specifications of measures) and mobile (implemented at sites beyond the point of original design and endorsement), but they frequently did not result in "closure" as I showed in Chapter 3 (Latour 1987). Controversies, it turns out, do not always disappear after standard technical devices are produced: they may obscure the complex administrative work it takes to build them, including disagreements over design, but these disputes may emerge anew when the devices are used in practice. If this is the case, then there is a much bigger question of what numbers, technology, and devices really "stand in for": if technical work stands in and *explains* multiple things at once, as Latour (1988) contends, then what is it that they explain? Furthermore, account-building from quantified knowledge, data, and technology is not a straightforward process: there are multiple ways that the technical may be mobilized, as I showed in Chapter 4 with the use of population forms under Population Health Management.

Thus, Latour's work requires significant modification for empirical examination of the social dynamics of science and technology, especially when these issues are intertwined with politics and policy; I have outlined in this dissertation a few starting directions for this work in sociology. Nonetheless, his call to reconceptualize the social is a welcome voice within the overall project of social science, especially as our world becomes increasingly more and more quantified, technical, and complex within contemporary society.

POLICY IMPLICATIONS

Policy makers are charged with extraordinarily difficult tasks: in making policy, they oversee the stewardship of limited resources, negotiate across political party lines, manage large agencies, and adhere to extremely detailed legal orders and laws. In many ways, this project has only deepened my respect for the everyday work of those working in Washington DC and Sacramento making enormously consequential decisions that affect entire political jurisdictions. As policy makers are often extremely short on time, it is unlikely that they will read the full draft of this or any future work. Still, I would like to offer some key points to keep in mind as this work continues.

As quality measures and technical forms are used within accountability programs today, the sheer complexity of this work is often underappreciated. Indeed, while "accountability" attracts interest and support across the aisle, there is a tremendous amount of work that goes into the implementation of these programs. At CMS, staff must design or contract with outside entities to design and test measures, submit them for NQF endorsement, convene committees to select measures for use in individual programs, and provide on-going support to health systems that report with the measure. Within individual health systems and practices, implementing Electronic Health Records (EHRs) involves tremendous effort: in addition to adopting the

technical system, programmers must create customized builds, staff must transfer paper records into a digital format readable by the EHR system, providers must undergo extensive training in the use of the technical system, and consultants and data scientists must remain on-site to troubleshoot on-going issues. Even with the EHR infrastructure in place, there is still an incredible amount of work needed to create reportable data. Providers and staff must enter in encounter information, technical coders must program specific pulls to draw from different fields within the EHR, quality staff validate entries to ensure data accuracy, and teams present results back to frontline providers as a part of local performance feedback.

And yet, all of this work is done within the context of limited budgets, changing financial incentive structures, and heightened external pressure. As new requirements are published each year, there is often little guidance on how to carry out this work as health systems grapple with moving between normal system operations and real system change. Few people probably anticipated how truly difficult it would be to bring about sociotechnical transformation: however, given the topic's central prominence on the health care reform agenda, I believe that we must address how this complexity impacts the use of resources within health care today. Modernizing health care through EHRs, quality measures, and patient portals in and of itself might be a worthwhile pursuit, but it must be pursued with sustained investment if it is to be successful. As a result, these initiatives may unintentionally contribute to overall further expense increase, as public agencies, health systems, and consulting organizations shift in size and number to accommodate these changes.

In most instances, however, data and technology are viewed as automatic cost-cutting solutions, and there are often enormous expectations placed onto technical fixes to what are usually deeply social and political problems. While there certainly are situations in which

technological adoption has resulted in cost savings (although even these initiatives are not without their own set of consequences), data and technology alone cannot address questions of *worth*. For all of the activity around “value-based” care and “transparency,” an underlying question remains: what is the *value* of health care *in* society, and *to* society? In policy circles, year-to-year rising medical expenditures are typically presented as wasted dollars that could be better spent on K-12 education, social welfare programs, the military and national defense, and basic scientific and applied research. And yet, if health care continues to be valued more highly than these other arenas, and the public would like a modernized health system, then there may be at least some justification for this use of resources. However, to pursue “value-based” care without clear and transparent alignment around the value of medicine in society today will only result in waste, empty promises, and unmet expectations.

FUTURE DIRECTIONS

In the future I will expand upon this work by comparing my fieldwork experiences at a large public safety-net health system with a large private health system; this will allow me to assess the full range of social heterogeneity across health systems, as well as whether and how quantification renders such heterogeneity knowable. As I have detailed throughout the dissertation, the stakes for what is made “knowable” and ultimately “actionable” are quite high, as data and quantified knowledge inform influential decision-making within public agencies and local health systems. Indeed, uncovering the boundary between the quantifiable and the non-quantifiable is not merely a philosophical thought-exercise or an epistemological puzzle, but a deeply social, political, and moral project. As things are made quantifiable, they draw attention to certain features for new kinds of audiences, informing particular understandings of the world and driving successive social action. By drawing upon comparative fieldwork, I will pursue this

line of inquiry to consider what quantified metrics reveal as health systems report their “performance” to the outside world and what they fail to capture (particularly as it pertains to social heterogeneity).

While I have sought to outline the contours of the emerging sociotechnical transformation of American medicine, there is much room for future scholarship to continue and build upon this work. Because the effects of the uptake of quantified knowledge, data, and technology are acutely felt across health systems across the country, and because of the sheer complexity of these developments, social scientists have much to offer in our era of health reform. Because new nodes of power and additional layers of decision-making are becoming incorporated as part of care and policy, sociologists should attend to these nodes in greater detail than I cover here (e.g., beyond the National Quality Forum). I believe that medical sociology must embrace this changing direction: as both medicine and society transform, sociology must find new ways to engage with the ever-changing and complex social world. While I have drawn from a few theoretical directions to make sense of this change, additional work should explore what this changing version of sociology might look like for both medical sociology and general social science audiences. Furthermore, just as these dynamics have also unfolded in the field of education (Espeland and Sauder 2016), social scientists might study how these forces have developed in other social worlds, such as the criminal justice system, child and social welfare, and environmental protection policy. Quantified knowledge, technical records, and novel forms of accountability have proliferated across multiple sectors of society, and comparative research would lead to general sociological contributions to the role of numbers, data, and technology in society today.

CLOSING THOUGHTS

This leads me to the final note of the conclusion: the future of American medicine. I will offer a few preliminary thoughts on this point based on my work this far, while acknowledging the ever-changing landscape of health care. And just as I opened with three relatively recent occurrences, I will close with a few more suggestive events:

January 2017: Joint Letter to Trump Administration and Congress Supporting Value-based Care

The new administration and composition of Congress brought in unparalleled uncertainty as to the future of the Affordable Care Act. The ACA was heavily attacked during the elections, creating an air of political uncertainty surrounding health care reform. However, a joint letter representing various stakeholders urged Congress to continue value-based health care specifically. The letter is signed by providers (American Medical Association, American Academy of Family Physicians, American Academy of Nursing, American College of Physicians, American College of Surgeons), commercial insurers (Aetna, Anthem, Blue Shield of California), employers and purchasers (Pacific Business Group on Health, National Alliance for Health Care Purchasers, Silicon Valley Employers Forum), health care organizations (American Hospital Association, America's Essential Hospitals), as well as measurement-based entities (National Committee for Quality Assurance, Press Ganey Associates, Vizient). There is a "new consensus in health care," as one NCQA staff member related to me during an interview, evidenced by this letter suggesting new political alliance: "The move towards value-based care is succeeding, measurably improving healthcare quality and contributing to historically low costs. Now is not the time for policymakers to signal a shift away from value-based care, either through action or inaction." This kind of alliance is deeply symbolic of newly organized relations across providers, purchasers, payers, patients, and policy advocates within health care.

March 2018: Value-Based Payment and Pay-for-Performance Summit, San Francisco, CA

At a recent public gathering on the topic of measurement-based delivery system reform, a series of presentations by two physician leaders reflect markedly different political orientations. A first keynote address, presented by the director of Clinical Standards and Quality and Chief Medical Officer at CMS, outlines the path forward with the new administration. She suggests that the momentum behind the value-based care movement will continue forward, with new care models and quality measures currently in development. Appointed to her director position in 2012, she embodies a true "physician leader": trained as an MD, her work has clearly brought about significant change to American medicine. The second presenter acts as a markedly different "physician leader": she serves as the president-elect of the American Medical Association, and she presents her work in advocating for providers across the country in the implementation of measurement-based reform. "CMS, frankly, *does not listen*. They do whatever they want now!" she relates, frustrated and looking expectantly at the CMS Medical Officer. Her slides contain charts and graphs created from her own data generated as part of her private practice's operations. She speaks in terms of risk adjustment, R-squared values, and performance benchmarks, leveraging her data to argue against CMS current administration of value-based care. The tension in the air is palpable, and the medical director of CMS looks away from the presentation, tight-lipped and uncomfortable. Each represent clearly different "paths" within medicine, despite receiving training and socialization within the medical profession.

April 2018: CMS Data-Driven Patient Care Strategy and Public Release of MA, CHIP Data

The current administrator of CMS, appointed under the Trump administration, announces a new overarching health care strategy. The strategy “will empower patients with the information they need as consumers of health care to enable them to make informed decisions about the care they need. Ultimately, the cornerstone of patient-centered system is data – quality data, cost data, and patient’s own data.” In addition to retooling existing work on interoperability to drive this new form of data-sharing, CMS also announces the first public release of encounter data from Medicare Advantage and the Children’s Health Insurance Program (CHIP) for researchers seeking to “drive health care innovation.” Data’s significance for public decision-making has increased not only in scope (the *kind* of data shared, expanding beyond quality to include information on cost), but also in span (the *representation* of data from different programs) and audience (the *users* and *use* of publicly released data). The strategy “ensures that CMS will support industry innovation in unleashing the power of data to drive system transformation – enhancing efficiency, improving quality, and reducing cost.” The health care data agenda, as described by the chief administration back in November 2014, is here to stay, continuing to transform care across the partisan divide.

These events reflect the refracturing depicted in Figure 5.1. These developments have made evaluating health care, and tracking questions of political power, extremely difficult. Consider the joint letter, with seeming alliance across stakeholders in pursuing the measurement-based transition from “volume to value.” Most of the prominent provider organizations, commercial insurers, large employers engaged in the business of health care, and measurement entities appear to be on the same page, with widespread agreement on the turn to quantified accountability. But as I reviewed in the third chapter, there are strenuous disagreements as to the technical details of holding entities “accountable.” Similarly, the CMS use and release of health care data seems to attract interest across political orientations: even if different administrations seek to pursue the data agenda in alternate ways, neither contest the basic premise that data be used to make care “better.” And yet the strong spilt between physician leaders representing the medical profession’s interests as advocate versus the public’s interest as public official reveals deep divisions over how data might be used. In this instance, the profession itself is fractured suggesting the diffusion of political power across different entities.

It is often socially and politically difficult to oppose quantification and evaluation – one can only propose that the measurement or the assessment be done in a different way, thus

shifting disagreements to technical details of evaluation (Dahler-Larsen 2012). Indeed, the use of quantified knowledge, data, and technology has created new conflicts organized around the development of “new” and “better” technical apparatuses. Stakeholders that disagree with the technical specifications of a measure become involved in the evaluation process, and design alternative metrics that better suit their goals and objectives (as in the case of the blood pressure metric in Chapter 3). Local databases can cause internal conflict within health systems, as administrators and medical staff argue over who has the better data. Clinician frustration towards EHRs is met with requests of feedback and suggestions for improvement, and staff grapple with questions of “the right data” (as in the case of population definition under PHM in Chapter 4). When there is distrust of the reporting practices of certain organizations, another coalition emerges with the goal of improving “transparency” and “data accuracy” this time around. As a result, the sheer number of quality measures, accountability programs, rating and ranking systems has skyrocketed, even within a relatively short time period.

And yet, few consider abandoning this particular health care reform agenda. Physician leaders now couch their arguments in statistical terms; health systems have invested billions in EHR infrastructure; and changing administrations direct data *strategy* without considering the overall value of *data*. The consequences of the widespread institutionalization of quantified knowledge, data, and technology are far-reaching and expansive. But on their own, they do not solve the larger question over controlling matters of care, as they merely shift these conflicts to new domains. In a powerful review of quantification as accountability in law, the sociologist of quantification Wendy Espeland (Espeland and Vannebo 2007:39) forcefully argues that the drive to replace judgement with rational calculation at once displaces power while obscuring its true location, buried beneath complex administrative work:

The locus and shape of discretion may change, but it does not disappear, sometimes moving to locations that make it harder to observe. Tracking discretion can be difficult, as assumptions, uncertainty, and ambiguity are buried in layers of small decisions, the traces of which are hard to recover. In such situations, responsibility becomes diffuse and abstract than when it is exercised by known individuals or groups whose decisions must be defended publicly. It may be harder to indict procedure than persons.

Taking this final note in relation to the turn to quantification in health care, the question of holding the historically powerful institution of medicine “accountable” is reconfigured across complex arrangements between the new quality measures, accountability programs, actors and experts, and technological developments that form each year. As I have argued throughout this dissertation, the use of measurement in health care has both expanded the number of people implicated in matters of care *and* drawn them to the project of medicine and society in an increasingly partial way. That is to say, accountability is today more *spread* than it is solid, *diffuse* across layers of complex administrative work. If this the case, then sociology must find ways to attend to these newly complex, digitized and quantified phenomena to track how power operates through data and technology. And it is through that work that sociology may continue to deliver upon its promise in the ever-changing social world we live in.

REFERENCES

- Alonso, William and Paul Starr. 1987. *The Politics of Numbers*. New York: Russell Sage Foundation.
- American Hospital Association. 2015. *TrendWatch: Rethinking the Hospital Readmissions Reduction Program*. Retrieved (<https://www.aha.org/system/files/research/reports/tw/15mar-tw-readmissions.pdf>).
- Anderson, Benedict. 1983. *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. London: Verso.
- Anderson, Margo J. 1988. *The American Census: A Social History*. New Haven, CT: Yale University Press.
- Arrow, Kenneth J. 1963. "Uncertainty and the Welfare Economics of Medical Care." *American Economic Review* 53(5):941–73.
- Ball, Robert M. 1995. "Perspectives on Medicare: What Medicare's Architects Had In Mind." *Health Affairs* 14(4):62–72.
- Beckert, Jens and Christine Musselin. 2013. "Introduction." Pp. 1–28 in *Constructing Quality: The Classification of Goods in Markets*, edited by J. Beckert and C. Musselin. Oxford: Oxford University Press.
- Benjamin, Walter. 2008 [1936]. *The Work of Art in the Age of Its Technological Reproducibility and Other Writings on Media*. Cambridge, MA: Harvard University Press.
- Berg, Marc. 1997. *Rationalizing Medical Work: Decision-Support Techniques and Medical Practices*. Cambridge, MA: MIT Press.
- Bijker, Wiebe. 1995. *Of Bicycles, Bakelites, and Bulbs: Towards a Theory of Sociotechnical Change*. Cambridge, MA: MIT Press.

- Boltanski, Luc. 2011. *On Critique: A Sociology of Emancipation*. Malden, MA: Polity Press.
- Boltanski, Luc. 2016. "Alain Desrosières's Reflexive Numbers." Pp. 33–42 in *The Social Sciences of Quantification: From the Politics of Large Numbers to Target-Driven Policies*, edited by I. Bruno, F. Jany-Catrice, and B. Touchelay. Switzerland: Springer International Publishing.
- Boltanski, Luc and Eve Chiapello. 2006. *The New Spirit of Capitalism*. New York: Verso.
- Boltanski, Luc and Laurent Thévenot. 2006. *On Justification: Economies of Worth*. Princeton, NJ: Princeton University Press.
- Bourdieu, Pierre. 1984. *Distinction: A Social Critique of the Judgement of Taste*. Cambridge, MA: Harvard University Press.
- Bowker, Geoffrey C. and Susan Leigh Star. 2000. *Sorting Things Out: Classification and Its Consequences*. Cambridge, MA: MIT Press.
- Bruno, Isabelle and Emmanuel Didier. 2013. *Benchmarking: L'État Sous Pression Statistique*. Paris: La Découverte.
- Burstin, Helen R. et al. 1999. "Benchmarking and Quality Improvement: The Harvard Emergency Department Quality Study." *American Journal of Medicine* 107(5):437–49.
- Callon, Michel. 1986. "Some Elements in the Sociology of Translation: Domestication of the Scallops and the Fishermen of St. Brieuc Bay." Pp. 196–223 in *Power, Action, and Belief: A New Sociology of Knowledge*, edited by J. Law. London: Routledge & Kegan Paul.
- Callon, Michel, Cécile Méadel, and Vololona Rabearisoa. 2002. "The Economy of Qualities." *Economy and Society* 31(2):194–217.
- Catlin, Aaron C. and Cathy A. Cowan. 2015. *History of Health Spending in the United States, 1960-2013*. Retrieved (<https://www.cms.gov/Research-Statistics-Data-and-Systems/>)

Statistics-Trends-and-Reports/NationalHealthExpendData/Downloads/HistoricalNHE
Paper.pdf).

Charmaz, Kathy. 2007. *Constructing Grounded Theory: A Practical Guide through Qualitative Analysis*. Thousand Oaks, CA: Sage.

Clancy, Carolyn M. 2009. "What Is Health Care Quality and Who Decides?" *Agency for Health Care Research and Quality*. Retrieved March 6, 2016
(<http://www.hhs.gov/asl/testify/2009/03/t20090318b.html>).

Clarke, Adele E. 2005. *Situational Analysis: Grounded Theory After the Postmodern Turn*.
Thousand Oaks, CA: Sage.

Clarke, Adele E., Janet K. Shim, Laura Mamo, Jennifer Ruth Fosket, and Jennifer R. Fishman.
2003. "Biomedicalization: Technoscientific Transformations of Health, Illness, and U.S.
Biomedicine." *American Sociological Review* 68(2):161–94.

Colgrove, James. 2002. "The McKeown Thesis: A Historical Controversy and Its Enduring
Influence." *American Journal of Public Health* 92(5):725–29.

Council on Ethical and Judicial Affairs. 1990. "Black-White Disparities in Health Care." *Journal
of the American Medical Association* 263(17):2344–46.

Cruz, Taylor M. 2017. "The Making of a Population: Challenges, Implications, and
Consequences of the Quantification of Social Difference." *Social Science & Medicine*
174:79–85.

Dahler-Larsen, Peter. 2012. *The Evaluation Society*. Stanford, CA: Stanford University Press.

Daston, Lorraine. 1992. "Objectivity and the Escape from Perspective." *Social Studies of Science*
22:597–618.

- Desrosières, Alain. 1998. *The Politics of Large Numbers: A History of Statistical Reasoning*. Cambridge, MA: Harvard University Press.
- Dranove, David. 1988. "Demand Inducement and the Physician/Patient Relationship." *Economic Inquiry* 26(2):281–98.
- Durkheim, Emile. 1982 [1895]. *The Rules of Sociological Method*. New York: Free Press.
- Espeland, Wendy Nelson and Michael Sauder. 2016. *Engines of Anxiety: Academic Rankings, Reputation, and Accountability*. New York: Russell Sage Foundation.
- Espeland, Wendy Nelson and Mitchell L. Stevens. 2008. "A Sociology of Quantification." *European Journal of Sociology* 49:401–36.
- Espeland, Wendy Nelson and Mitchell L. Stevens. 1998. "Commensuration as a Social Process." *Annual Review of Sociology* 24:313–43.
- Espeland, Wendy Nelson and Berit Irene Vannebo. 2007. "Accountability, Quantification, and Law." *Annual Review of Law and Social Science* 3(1):21–43.
- Feinstein, Alvan R. 1967. *Clinical Judgment*. Huntington, NY: Krieger Publishing Company.
- Feldstein, Paul J. 2007. *Healthy Policy Issues: An Economic Perspective*. Chicago, IL: Foundation of the American College of Healthcare Executives.
- Fourcade, Marion. 2011. "Cents and Sensibility: Economic Valuation and the Nature of 'Nature.'" *American Journal of Sociology* 116(6):1721–77.
- Fourcade, Marion. 2016. "The Ordinalization of Society." *Sociological Theory* 34(3):175–95.
- Fourcade, Marion and Kieran Healy. 2013. "Classification Situations: Life-Chances in the Neoliberal Era." *Accounting, Organizations and Society* 38:559–72.
- Freidson, Eliot. 1970. *Profession of Medicine: A Study of the Sociology of Applied Knowledge*. Chicago, IL: University of Chicago Press.

- Fuchs, Victor R. 1978. "The Supply of Surgeons and the Demand for Operations." *Journal of Human Resources* 13 Suppl:35–56.
- Fuchs, Victor R. 1974. *Who Shall Live? Health, Economics, and Social Choice*. New York: Basic Books.
- Gamble, Vanessa N. and Deborah Stone. 2006. "U.S. Policy on Health Inequities: The Interplay of Politics and Research." *Journal of Health Politics, Policy and Law* 31(1):93–126.
- Grol, Richard. 2001. "Successes and Failures in the Implementation of Evidence-Based Guidelines for Clinical Practice." *Medical Care* 39(8):II46-II54.
- Grundy, Emily. 2005. "The McKeown Debate: Time for Burial." *International Journal of Epidemiology* 34(3):529–33.
- Habermas, Jürgen. 1981. *The Theory of Communicative Action, Volume I*. Boston, MA: Beacon Press.
- Hafferty, Fred W. and Donald W. Light. 1995. "Professional Dynamics and the Changing Nature of Medical Work." *Journal of Health and Social Behavior* 35(Extra Issue):132–53.
- Hanks, William F. and Carlo Severi. 2014. "Translating Worlds: The Epistemological Space of Translation." *HAU: Journal of Ethnographic Theory* 4(2):1–16.
- Haug, Marie. 1988. "A Re-Examination of the Hypothesis of Physician Deprofessionalization." *The Milbank Quarterly* 66(Suppl 2):48–56.
- Hay, Joel and Michael J. Leahy. 1982. "Physician-Induced Demand: An Empirical Analysis of the Consumer Information Gap." *Journal of Health Economics* 1(3):231–44.
- Hoffman, Beatrix. 2012. *Health Care for Some: Rights and Rationing in the United States since 1930*. Chicago: University of Chicago Press.
- Horkheimer, Max. 2002 [1972]. *Critical Theory: Selected Essays*. New York: Continuum.

- Horkheimer, Max and Theodor W. Adorno. 2007 [1944]. *Dialectic of Enlightenment*. Stanford, CA: Stanford University Press.
- Illich, Ivan. 1976. *Medical Nemesis: The Expropriation of Health*. New York: Random House.
- James, Paul A. et al. 2014. “2014 Evidence-Based Guidelines for the Management of High Blood Pressure in Adults: Report from the Panel Members Appointed to the Eighth National Joint National Committee (JNC 8).” *Journal of the American Medical Association* 311(5):507–20.
- Karpik, Lucien. 2010. *Valuing the Unique: The Economics of Singularities*. Princeton, NJ: Princeton University Press.
- Kingdon, John. 1984. *Agendas, Alternatives, and Public Policies*. Boston: Little, Brown.
- Komaroff, Anthony L. 1982. “Algorithms and the ‘Art’ of Medicine.” *American Journal of Public Health* 72:10–12.
- Kotchen, Theodore A. 2014. “Developing Hypertension Guidelines: An Evolving Process.” *American Journal of Hypertension* 27(6):765–72.
- Krieger, N. 2012. “Who and What Is a ‘Population’? Historical Debates, Current Controversies, and Implications for Understanding ‘Population Health’ and Rectifying Health Inequities.” *The Milbank Quarterly* 90(4):634–81.
- Lamont, Michèle. 2012. “Toward a Comparative Sociology of Valuation and Evaluation.” *Annual Review of Sociology* 38:201–21.
- Lamont, Michèle. 2009. *How Professors Think: Inside the Curious World of Academic Judgment*. Cambridge, MA: Harvard University Press.
- Latour, Bruno. 1987. *Science in Action: How to Follow Scientists and Engineers through Society*. Cambridge, MA: Harvard University Press.

- Latour, Bruno. 1988. "The Politics of Explanation: An Alternative." Pp. 155–76 in *Knowledge and Reflexivity: New Frontiers in the Sociology of Knowledge*, edited by S. Woolgar. Newbury Park, CA: Sage Publications.
- Latour, Bruno. 2005. *Reassembling the Social: An Introduction to Actor-Network-Theory*. New York: Oxford University Press.
- Law, John and Annemarie Mol. 2002. *Complexities: Social Studies of Knowledge Practices*. Durham, NC: Duke University Press.
- Light, Donald W. 1993. "Countervailing Power: The Changing Character of the Medical Profession in the United States." in *The Changing Medical Profession: An International Perspective*, edited by F. W. Hafferty and J. B. McKinlay. New York: Oxford University Press.
- Link, Bruce G. and Jo C. Phelan. 2002. "McKeown and the Idea That Social Conditions Are Fundamental Causes of Disease." *American Journal of Public Health* 92(5):730–32.
- Lynch, Thomas. 2015. "A Critique of Health System Performance Measurement." *International Journal of Health Services* 45(4):743–61.
- Marcuse, Herbert. 2008 [1964]. *One-Dimensional Man: Studies in the Ideology of Advanced Industrial Society*. London: Routledge.
- Marmor, Theodore R. 2000. *The Politics of Medicare*. New York: Aldine de Gruyter.
- Marx, Karl. 1998 [1932]. *The German Ideology: Including Theses on Feuerbach and Introduction to the Critique of Political Economy*. Amherst, NY: Prometheus Books.
- Mayes, Rick and Robert A. Berenson. 2006. *Medicare Prospective Payment and the Shaping of U.S. Health Care*. Baltimore, MD: John Hopkins University Press.

- McKeown, Thomas. 1976. *The Role of Medicine: Dream, Mirage, or Nemesis?* London: Nuffield Provincial Hospitals Trust.
- McKeown, Thomas. 1976. *The Modern Rise of Population*. New York: Academic Press.
- McKinlay, John B. and John D. Stoeckle. 1988. "Corporatization and the Social Transformation of Doctoring." *International Journal of Health Services* 18(2):191–205.
- Merry, Sally Engle. 2016. *The Seductions of Quantification: Measuring Human Rights, Gender Violence, and Sex Trafficking*. Chicago: University of Chicago Press.
- Messerli, Franz H. 2014. "Hyperten Guidelines: Political Correctness Trumping Expertise?" *Hypertension* 63:e81.
- Mol, Annemarie. 2002. *The Body Multiple: Ontology in Medical Practice*. Durham, NC: Duke University Press.
- Mol, Annemarie. 1999. "Ontological Politics. A Word and Some Questions." *The Sociological Review* 47(S1):74–89.
- National Quality Forum. 2012. *2012 NQF Report to Congress: Changing Healthcare by the Numbers*. Retrieved (http://www.qualityforum.org/Publications/2012/03/2012_NQF_Report_to_Congress.aspx).
- Navarro, Vicente. 1985. "U.S. Marxist Scholarship in the Analysis of Health and Medicine." *International Journal of Health Services I* 15(4):525–45.
- Navarro, Vicente. 1994. "Professional Dominance or Proletarianization? Neither." Pp. 273–83 in *Health Policy and Nursing: Crisis and Reform in the U.S. Health Care Delivery System*, edited by C. Harrington and C. L. Estes. Boston, MA: Jones and Bartlett.
- Navarro, Vicente. 1986. *Crisis, Health, and Medicine*. New York: Tavistock.

- Navarro, Vicente. 2001. "The New Conventional Wisdom: An Evaluation of the WHO Report Health Systems: Improving Performance." *International Journal of Health Services* 31(1):23–33.
- Oberlander, Jonathan. 2003. *The Political Life of Medicare*. Chicago: University of Chicago Press.
- O'Brien, Eoin. 2017. "Who is an Expert?" *Hypertension* 35(11):2326–27.
- Porter, Theodore M. 1995. *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton, NJ: Princeton University Press.
- President's Advisory Commission on Consumer Protection and Quality in the Health Care Industry. 1998. *Quality First: Better Health Care for All Americans*.
- Quadagno, Jill. 2000. "Promoting Civil Rights through the Welfare State: How Medicare Integrated Southern Hospitals." *Social Problems* 47(1):68–89.
- Quadagno, Jill S. 2005. *One Nation, Uninsured: Why the U.S. Has No National Health Insurance*. Oxford: Oxford University Press.
- Radin, Beryl A. 2006. *Challenging the Performance Movement: Accountability, Complexity, and Democratic Values*. Washington, D.C.: Georgetown University Press.
- Robson, Keith. 1992. "Accounting Numbers as 'Inscription': Action at a Distance and the Development of Accounting." *Accounting, Organizations and Society* 17(7):685–708.
- Rose, Nikolas. 1991. "Governing by Numbers: Figuring out Democracy." *Accounting, Organizations and Society* 16(7):673–92.
- Rose, Nikolas and Peter Miller. 1992. "Political Power beyond the State: Problematics of Government." *British Journal of Sociology* 43(2):173–205.

- Rosenthal, Meredith B. 2015. "Can Federal Provider Payment Reform Produce Better, More Affordable Healthcare?" Pp. 221–46 in *The Future of Healthcare Reform in the United States*, edited by A. Malani and M. H. Schill. Chicago: Chicago University Press.
- Rueschemeyer, Dietrich and Theda Skocpol. 1996. *States, Social Knowledge, and the Origins of Modern Social Policies*. Princeton, NJ: Princeton University Press.
- Scott, James C. 1998. *Seeing like a State: How Certain Schemes to Improve the Human Condition Have Failed*. New Haven, CT: Yale University Press.
- Scott, Susan V and Wanda J. Orlikowski. 2012. "Reconfiguring Relations of Accountability: Materialization of Social Media in the Travel Sector." *Accounting, Organizations and Society* 37(1):26–40.
- Shim, Janet. 2014. *Heart-Sick: The Politics of Risk, Inequality, and Heart Disease*. New York: New York University Press.
- Skocpol, Theda. 1997. *Boomerang: Health Care Reform and the Turn Against Government*. New York: W. W. Norton & Company, Inc.
- Star, Susan Leigh. 1999. "The Ethnography of Infrastructure." *American Behavioral Scientist* 43(3):377–91.
- Stark, David. 2009. *A Sense of Dissonance: Accounts of Worth in Economic Life*. Princeton, NJ: Princeton University Press.
- Starr, Paul. 2011. *Remedy and Reaction: The Peculiar American Struggle over Health Care Reform*. New Haven, CT: Yale University Press.
- Starr, Paul. 1982. *The Social Transformation of American Medicine*. New York: Basic Books.
- Stone, Deborah. 2002. *Policy Paradox: The Art of Political Decision-Making*. New York: Norton.

- Stone, Deborah A. 1989. "Causal Stories and the Formation of Policy Agendas." *Political Science Quarterly* 104(2):281–300.
- Thomas, Stephen B., Georges C. Benjamin, Donna Almaro, and Monica J. Lathan. 2006. "Historical and Current Policy Efforts to Eliminate Racial and Ethnic Health Disparities in the United States: Future Opportunities for Public Health Education Research." *Health Promotion Practice* 7(3):324–30.
- Timmermans, Stefan and Marc Berg. 2003. *The Gold Standard: The Challenge of Evidence-Based Medicine and Standardization in Health Care*. Philadelphia, PA: Temple University Press.
- Timmermans, Stefan and Hyeyoung Oh. 2010. "The Continued Social Transformation of the Medical Profession." *Journal of Health and Social Behavior* 51(1 suppl):S94–106.
- U.S. Department of Health, Education, and Welfare. 1979. *Healthy People: The Surgeon General's Report on Health Promotion and Disease Prevention*.
- U. S. Department of Health and Human Services. 2000. *Healthy People 2010: Objectives for Improving Health*.
- U. S. Department of Health and Human Services. 2005. *Methodological Issues in Measuring Health Disparities*.
- U. S. Department of Health and Human Services. 1985. *Report of the Secretary's Task Force on Black & Minority Health*.
- U. S. Department of Health and Human Services. 1990. *Healthy People 2000: National Health Promotion and Disease Prevention Objectives*.
- U. S. Department of Health and Human Services. 1980. *Healthy People 1990: Promoting Health/Preventing Disease: Objectives for the Nation*.

- Vatin, Francois. 2013. "Valuation as Evaluating and Valorizing." *Valuation Studies* 1(1):31–50.
- Waitzkin, Howard. 1978. "A Marxist View of Medical Care." *Annals of Internal Medicine* 89(2):264–78.
- Waring, Justin, Davina Allen, Jeffrey Braithwaite, and Jane Sandall. 2015. "Healthcare Quality and Safety: A Review of Policy, Practice and Research." *Sociology of Health & Illness* 38(2).
- Weber, Max. 1968 [1922]. *Economy and Society: An Outline of Interpretive Sociology*, edited by G. Roth and C. Wittich. New York: Bedminster Press.
- Weisz, George et al. 2007. "The Emergence of Clinical Practice Guidelines." *Milbank Memorial Fund Quarterly* 85(4):691–727.
- Welsh, Valerie. 1998. *Target Setting: How Low or High Do We Aim?*
- Wiener, Carolyn L. 2000. *The Elusive Quest: Accountability in Hospitals*. Hawthorne, NY: Aldine de Gruyter.
- Wright, Jackson T., Lawrence J. Fine, Daniel T. Lackland, Gbenga Ogedegbe, and Cheryl Dennison-Himmelfarb. 2014. "Evidence Supporting a Systolic Blood Pressure Goal of Less than 150 Mm Hg in Patients Aged 60 Years or Older: The Minority View." *Annals of Internal Medicine* 160(7):499–504.

APPENDICES

Data Sources

Tables and Figures

Taylor M. Cruz

Department of Social and Behavioral Sciences

University of California, San Francisco

APPENDIX A

Data Sources

PHASE I: TECHNICAL DOCUMENTS, POLICIES, AND REPORTS

Centers for Medicare and Medicaid Services

1. 2012, Mar. National Impact Assessment of Medicare Quality Measures.
2. 2015, Mar. National Impact Assessment of the Centers for Medicare and Medicaid Services (CMS) Quality Measures Report.
3. 2015, Aug. A Blueprint for the CMS Measures Management System.
4. 2016, May 2. CMS Quality Measure Development Plan: Supporting the Transition to the Merit-based Incentive Payment System (MIPS) and Alternative Payment Models (APMs).
5. 2016, Nov 4. Final Rule, Medicare Access and CHIP Reauthorization Act of 2015.
6. 2015, Sept. The CMS Equity Plan for Improving Quality in Medicare.

Department of Health and Human Services

7. 2011, Apr. HHS Action Plan to Reduce Racial and Ethnic Health Disparities.
8. 2011, Dec. Measures Under Consideration for Calendar Year 2012 Rulemaking.
9. 2012, Dec. Measures Under Consideration for Calendar Year 2013 Rulemaking.
10. 2013, Dec. Measures Under Consideration for Calendar Year 2014 Rulemaking.
11. 2014, Dec. Measures Under Consideration for Calendar Year 2015 Rulemaking.

Legislation

12. 2008, Jul. 15. Medicare Improvements for Patients and Providers Act of 2008 (MIPPA). PL 110-275.
13. 2010, Mar. 23. Patient Protection and Affordable Care Act of 2010 (PPACA). PL 111-148.
14. 2013, Jan. 2. American Taxpayer Relief Act of 2012. PL 112-240.
15. 2013, Dec. 26. Bipartisan Budget Act of 2013. PL 113-67.
16. 2014, Apr. 1. Protecting Access to Medicare Act of 2014. PL 113-93.
17. 2015, Apr. 6. Medicare Access and CHIP Reauthorization Act of 2015 (MACRA). PL 114-10.

Congressional Research Service

18. 2009, Aug. 5. Measuring Health Care Quality: Measure Development, Endorsement, and Implementation.
19. 2010, Mar. 24. Public Health Workforce, Quality, and Related Provisions in the Patient Protection and Affordable Care Act (PL 111-148).

20. 2013, May 17. Public Health, Workforce, Quality, and Related Provisions in ACA: Summary and Timeline.
21. 2017, Jan. 6. Increasing Choice, Access, and Quality in Health Care for Americans Act (Division C of PL 114-255).

National Quality Forum

22. 2010, Mar. 1. Report to the Congress and the Secretary of the U.S. Department of Health and Human Services 2009.
23. 2011, Mar. 1. Report to the Congress and the Secretary of the U.S. Department of Health and Human Services 2010.
24. 2012, Mar. 1. Report to the Congress and the Secretary of the U.S. Department of Health and Human Services 2011.
25. 2013, Mar. 1. Report to the Congress and the Secretary of the U.S. Department of Health and Human Services 2012.
26. 2014, Mar. 1. Report to the Congress and the Secretary of the U.S. Department of Health and Human Services 2013.
27. 2015, Mar. 1. Report to the Congress and the Secretary of the U.S. Department of Health and Human Services 2014.
28. 2016, Mar. 1. Report to the Congress and the Secretary of the U.S. Department of Health and Human Services 2015.
29. 2012, Mar. Cardiovascular Endorsement Maintenance 2010: Technical Report.
30. 2014, Nov. 14. NQF-Endorsed Measures for Cardiovascular Conditions 2014: Technical Report.
31. 2015, Aug. 31. NQF-Endorsed Measures for Cardiovascular Conditions 2014-2015 Phase 2: Technical Report.
32. 2016, May 3. NQF-Endorsed Measures for Cardiovascular Conditions 2015-2016: Final Report.

PHASE II: NATIONAL POLICY EVENT FIELDWORK AND EXPERT INTERVIEWS

National Policy Event Fieldwork

1. 2016, Jan. 7. UCSF Value Forum. San Francisco, CA.
2. 2016, Feb. 7-9. Value-based Payment Pay for Performance Summit. San Francisco, CA.
3. 2016, June 9-10. Accountable Care Organization Summit. Washington D.C. (webcast)
4. 2016, July 27. California Health Care Foundation Regional Atlas Event. Sacramento, CA.
5. 2016, Oct. 6. New England Journal of Medicine Catalyst: New Marketplace. (webcast)
6. 2016, Oct. 20. CMS Physician Compare Informational Session. (webcast)
7. 2016, Nov. 30-Dec. 2. MACRA Summit. Washington D.C.
8. 2017, Jan. 19. New England Journal of Medicine: Relentless Redesign. (webcast)
9. 2017, Jan. 30-31. AcademyHealth National Health Policy Conference. Washington D.C.
10. 2017, Mar. 8-10. Value-based Payment and Pay for Performance Summit. San Francisco, CA.
11. 2017, Mar. 22. Health Affairs Delivery System Innovation. Sacramento, CA.

12. 2017, June 6. Health Affairs Health Equity Event. Washington D.C. (webcast).
13. 2017, June 28-29. Accountable Care Organization Summit. Washington D.C. (webcast).
14. 2017, Dec 6-7. National Health Care Ratings Summit. Washington D.C. (webcast).
15. 2018, Feb. 28-Mar. 2. Value-based Payment and Pay for Performance Summit. San Francisco, CA.

Expert and Stakeholder Interviews

1. 2016, Nov. 30. Physician leader and medical profession society board member, Health system (redacted).
2. 2017, Jan. 26. Measurement advocate, Leapfrog Group.
3. 2017, Jan. 26. Policy expert and physician, Urban Institute.
4. 2017, Jan. 27. Measurement advocate, National Committee for Quality Assurance.
5. 2017, Jan. 27. Patient advocate and policy expert, FamiliesUSA.
6. 2017, Feb. 10. Measurement expert and staff member, National Quality Forum.
7. 2017, Mar. 2. Actuary and health care consultant, Milliman.
8. 2017, Mar. 2. Policy expert, AcademyHealth.
9. 2017, Mar. 6. Researcher and health care consultant, Value Institute.
10. 2017, July 11. Evaluator and staff member, Centers for Disease Control and Prevention.
11. 2017, July 13. Hospital advocate, America's Essential Hospitals.
12. 2017, Oct 20. Measurement consultant, Health care consulting firm (redacted).
13. 2017, Nov 22. Policy expert, Safety Net Institute.

PHASE III: LOCAL HEALTH SYSTEM FIELDWORK AND INFORMANT INTERVIEWS

Health System Fieldwork

I conducted fieldwork at a local health system between September 2017 and March 2018, with additional follow-up through May and June 2018. I conducted observations at the health system by attending two full work days (Tuesdays and Fridays, 9-5pm) over the course of nine months, resulting in approximately 400 hours of observation. The organization is a large public safety-net system that primarily provides services to Medicaid (Medi-Cal) patients and the underserved.

As a part of the fieldwork, I observed regular weekly and monthly meetings with quality team members, health center managers and clinic leadership, chief executives and administrators, and frontline clinicians and support staff. I also attended monthly outpatient clinic visits with the quality team, and recorded the general characteristics of clinics, organizational activity, and people (both patients and staff). Towards the end of the fieldwork, I shadowed select providers to gain a better understanding of clinic proceedings.

I also attended two program-specific statewide events with the health system's quality team:

1. 2017, Oct. 12. Data Summit. Oakland, CA.
2. 2017, Nov. 15. Learning Collaborative. Sacramento, CA.

Informant Interviews

1. 2018, Jan. 30. Programmer and data analyst.
2. 2018, Feb. 3. Quality coordinator and registered nurse.
3. 2018, Feb. 3. Program manager and analyst.
4. 2018, Feb. 9. Quality coordinator and registered nurse.
5. 2018, Feb. 20. Program manager and technical consultant.
6. 2018, Feb. 23. Quality coordinator and registered nurse.
7. 2018, Feb. 27. Health center manager.
8. 2018, Feb. 27. Programmer and data analyst.
9. 2018, Feb. 27. Programmer and data analyst.
10. 2018, Mar. 9. Complex care registered nurse.
11. 2018, Mar. 13. Director of quality.
12. 2018, Mar. 16. Chief Medical Informatics Officer (CMIO; former).
13. 2018, Mar. 20. Physician champion and primary care provider.
14. 2018, Mar. 20. Complex care registered nurse.
15. 2018, Mar. 20. Complex care registered nurse.
16. 2018, Mar. 23. Medical director and primary care provider.
17. 2018, Mar. 23. Assistant nurse manager.
18. 2018, Mar. 23. Program analyst.
19. 2018, May 22. Programmer and data analyst.
20. 2018, May 25. Licensed vocational nurse.
21. 2018, May 25. Licensed vocational nurse.

Follow-up Interviews

1. 2018, Feb. 6. Registered nurse and quality coordinator.
2. 2018, Feb. 9. Registered nurse and quality coordinator.
3. 2018, Feb. 23. Programmer and data analyst.
4. 2018, May 11. Programmer and data analyst.
5. 2018, May 25. Programmer and data analyst.

APPENDIX B

Tables and Figures

Overview

1. Table 3.1. Three Quality Measures.
2. Figure 3.2. Controlling High Blood Pressure (NQF 0018).
3. Figure 3.3. Ischemic Vascular Disease (IVD): Use of Aspirin or Another Antithrombotic (NQF 0068).
4. Figure 3.4. Hospital 30-Day, All-Cause, Risk Standardized Readmissions Rate (RSRR) following Heart Failure (HF) Hospitalization (NQF 0330).
5. Figure 3.5. Final List of Risk Adjustment Variables for NQF 0330 (NQF CV10).
6. Table 4.1. Four Forms of Difference in Population Health Management (PHM).
7. Figure 5.1. The Complexity of Sociotechnical Transformation.

Chapter 3: Politics and the State Calculus

Table 3.1. Three Quality Measures

Name	Controlling High Blood Pressure	Ischemic Vascular Disease (IVD): Use of Aspirin or Another Antithrombotic	Hospital 30-Day, All-Cause, Risk Standardized Readmissions Rate (RSRR) following Heart Failure (HF) Hospitalization
NQF No.	NQF 0018	NQF 0068	NQF 0330
Steward	National Committee for Quality Assurance (NCQA)	National Committee for Quality Assurance (NCQA)	Centers for Medicare and Medicaid Services (CMS)
Type	Outcome	Process	Administrative
Politics	<i>Evidence</i>	<i>Record</i>	<i>Boundaries</i>

Figure 3.2. Controlling High Blood Pressure (NQF 0018)

<p>Measure Steward: National Committee for Quality Assurance</p> <p>Description: The percentage of patients 18 to 85 years of age who had a diagnosis of hypertension (HTN) and whose blood pressure (BP) was adequately controlled (<140/90) during the measurement year.</p> <p>Numerator: The number of patients in the denominator whose most recent BP is adequately controlled during the measurement year. For a patient's BP to be controlled, both the systolic and diastolic BP must be <140/90 (adequate control). To determine if a patient's BP is adequately controlled, the representative BP must be identified.</p> <p>Denominator: Patients 18 to 85 years of age by the end of the measurement year who had at least one outpatient encounter with a diagnosis of hypertension (HTN) during the first six months of the measurement year.</p> <p>Exclusions: Exclude all patients with evidence of end-stage renal disease (ESRD) on or prior to the end of the measurement year. Documentation in the medical record must include a related note indicating evidence of ESRD. Documentation of dialysis or renal transplant also meets the criteria for evidence of ESRD. Exclude all patients with a diagnosis of pregnancy during the measurement year. Exclude all patients who had an admission to a nonacute inpatient setting during the measurement year.</p> <p>Risk Adjustment: No</p>

Figure 3.3. Ischemic Vascular Disease (IVD): Use of Aspirin or Another Antithrombotic (NQF 0068)

<p>Measure Steward: National Committee for Quality Assurance</p> <p>Description: The percentage of patients 18 years of age and older who were discharged from an inpatient setting with an acute myocardial infarction (AMI), coronary artery bypass graft (CABG) or percutaneous coronary intervention (PCI) during the 12 months prior to the measurement year, or who had a diagnosis of ischemic vascular disease (IVD) during the measurement year and the year prior to the measurement year and who had documentation of routine use of aspirin or another antiplatelet during the measurement year.</p> <p>Numerator: Patients who had documentation of routine use of aspirin or another antiplatelet during the measurement year.</p> <p>Denominator: Patients 18 years or older by the end of the measurement year discharged from an inpatient setting with an AMI, CABG, or PCI during the 12 months prior to the measurement year or who had a diagnosis of IVD during both the measurement year and the year prior to the measurement year.</p> <p>Exclusions: Patients who had documentation of use of anticoagulant medications during the measurement year. Exclude patients using hospice services any time during the measurement period.</p> <p>Risk Adjustment: No</p>

Figure 3.4. Hospital 30-Day, All-Cause, Risk Standardized Readmissions Rate (RSRR) following Heart Failure (HF) Hospitalization (NQF 0330)

Measure Steward: Centers for Medicare and Medicaid Services

Description: This measure estimates a hospital-level, 30-day RSRR for patients discharged from the hospital with a principal diagnosis of HF. Readmission is defined as unplanned readmission for any cause within 30 days of the discharge date for the index admission. A specified set of planned readmissions do not count as readmissions. The Centers for Medicare & Medicaid Services (CMS) annually reports the measure for patients who are 65 years and older and are Medicare fee-for-service (FFS) beneficiaries hospitalized in non-federal hospitals or Veterans Health Administration (VA) hospitals.

Numerator: The outcome for this measure is 30-day readmission. We define readmission as an inpatient admission for any cause, with the exception of certain planned readmissions, within 30 days from the date of discharge from the index HF admission. If a patient has more than one unplanned admissions (for any reason) within 30 days after discharge from the index admission, only one is counted as a readmission. The measure looks for a dichotomous yes or no outcome of whether each admitted patient has an unplanned readmission within 30 days. However, if the first readmission after discharge is considered planned, any subsequent unplanned readmission is not counted as an outcome for that index admission, because the unplanned readmission could be related to care provided during the intervening planned readmission rather than during the index admission.

Denominator: This claims-based measure can be used in either of two patient cohorts: (1) patients aged 65 years or older or (2) patients aged 18 years or older. The cohort includes admissions for patients discharged from the hospital with a principal discharge diagnosis of HF and with a complete claims history for the 12 months prior to admission. The measure is currently publicly reported by CMS for those patients 65 years and older who are Medicare FFS beneficiaries admitted to non-federal hospitals. This measure can also be used for an all-payer population aged 18 years and older. We have explicitly tested the measure in both patients aged 18+ years and those aged 65+ years.

Exclusions: The HF readmission measure excludes index hospitalizations that meet any of the following exclusion criteria:

1. Admissions without at least 30 days of post-discharge enrollment in Medicare FFS;
2. Discharges against medical advice;
3. Admissions within 30 days of discharge from a prior HF index admission; and
4. With a procedure code for LVAD implantation or heart transplantation either during the index admission or in the 12 months prior to the index admission.


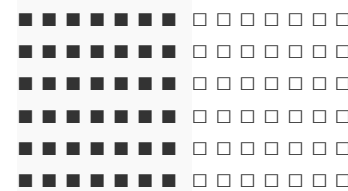
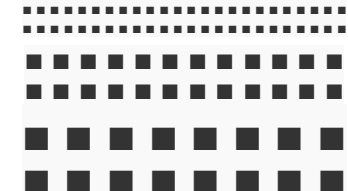
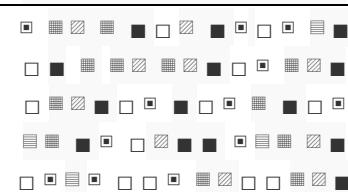
Risk Adjustment: Yes

Figure 3.5. Final List of Risk Adjustment Variables for NQF 0330 (NQF CV10)

<p>Demographic:</p> <ul style="list-style-type: none">• Age-65 (years above 65, continuous) for 65 and over cohorts; or Age (years, continuous) for 18 and over cohorts• Male <p>Cardiovascular:</p> <ul style="list-style-type: none">• History of CABG• Cardio-respiratory failure or shock• Congestive heart failure• Acute coronary syndrome• Coronary atherosclerosis or angina• Valvular or rheumatic heart disease• Specified arrhythmias• Other or unspecified heart disease• Vascular or circulatory disease	<p>Comorbidity:</p> <ul style="list-style-type: none">• Metastatic cancer or acute leukemia• Metastatic cancer or acute leukemia• Cancer• Diabetes or DM complications• Protein-calorie malnutrition• Disorders of fluid, electrolyte, acid-base• Liver or biliary disease• Peptic ulcer, hemorrhage, other specified gastrointestinal disorders• Other gastrointestinal disorders• Severe hematological disorders• Iron deficiency or other anemias and blood disease• Dementia or other specified brain disorders• Drug/alcohol abuse/dependence/psychosis• Major psychiatric disorders• Depression• Other psychiatric disorders• Hemiplegia, paraplegia, paralysis, functional disability• Stroke• Chronic obstructive pulmonary disease• Fibrosis of lung or other chronic lung disorders• Asthma• Pneumonia• End stage renal disease or dialysis• Renal failure• Nephritis• Other urinary tract disorders• Decubitus ulcer or chronic skin ulcer
--	--

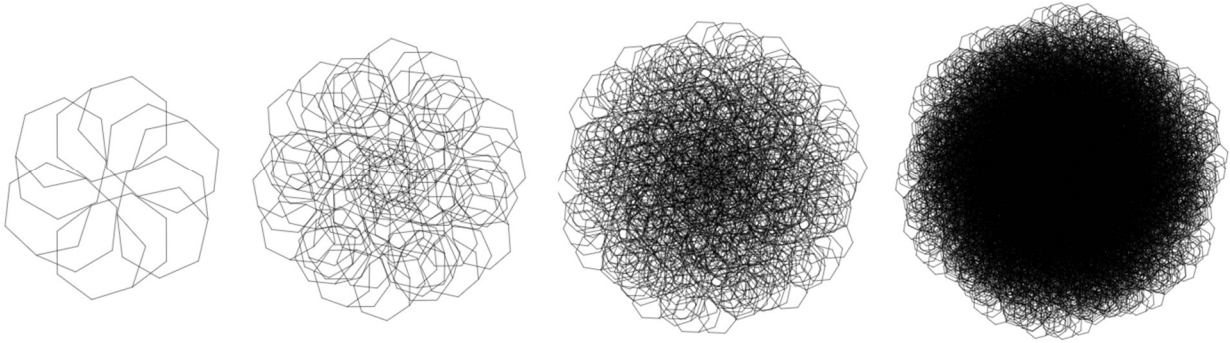
Chapter 4: Accounting for Difference

Table 4.1. Four Forms of Difference in Population Health Management (PHM)

	Structure of Difference	Paradigmatic Framework	Difference in Care	Visual Representation
Population as Citizens	Similarity; one single homogenous form	Standards (clinical guidelines, common workflows)	Remove difference; treat all patients similarly	
Population as Categories	Difference across, similarity within; multiple homogenous forms	Disparities, social determinants (different group outcomes)	Standardize difference; treat patients categorically	
Population as Classification	Similarity across, difference within; sorting on a single scale	Risk stratification (likelihood of becoming sick or costly)	Hierarchalize difference; treat patients based on ranking (e.g., groupings)	
Population as Consumers	Difference; Loose clusters of individual heterogeneity	Preferences (patient satisfaction)	Individualize difference; treat patients based on consumer requests	

Chapter 5: Conclusion

Figure 5.1. The Complexity of Sociotechnical Transformation



Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.

Author Signature  _____ Date May 5, 2018