

Actor Networks in Health Care:
Translating Values into Measures of Hospital Performance

MaryBeth A. Farquhar

Dissertation submitted to the faculty of Virginia Polytechnic Institute and State University in
partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

In

Public Administration and Public Affairs

Joseph V. Rees, Chairman

Karen M. Hult

Anne Khademian

William L. Murray

Sally J. Phillips

James F. Wolf

June 30, 2008

Alexandria, Virginia

Keywords: performance measurement, public policy, networks, actor-network theory, health care policy, governance, hospitals, the National Quality Forum

Actor Networks in Health Care:
Translating Values into Measures of Hospital Performance

MaryBeth A. Farquhar

ABSTRACT

The health care system within the United States is in a state of transition. The industry, confronted with a variety of new technologies, new ways of organizing, spiraling costs, diminishing service quality and new actors, is changing, almost on a daily basis. Reports issued by the Institute of Medicine raise quality issues such as avoidable errors and underuse/overuse of services; other studies document regional variation in care. Improvement in the quality of care, according to health care experts is accomplished through measuring and comparing performance, but there are a number of disparate actors involved in this endeavor. Through a network of both public and private actors, collaboration on the development of a set of national performance measures is underway. Organizations such as the National Quality Forum (NQF), the Agency for Healthcare Research and Quality (AHRQ), the Centers for Medicare & Medicaid Services (CMS) and other have formed networks to develop and standardize performance measurement systems that can distinguish between quality services and substandard ones. While there is some available research about the processes involved in performance measurement system design, there is little known about the factors that influence the development and work of the network, particularly the selection of hospital performance measures. This dissertation explored the development of a national performance measurement system for hospitals, using an institutional rational choice perspective and actor-network theory as frameworks for discussion. Through qualitative research methods such as direct observation, interviews, participant observations and document review, a theoretically informed case study of the NQF's Hospital Steering Committee was performed, to address the following questions: How is a national performance measurement system developed and what is the role of federal agencies (e.g., AHRQ and CMS) in the process?

ACKNOWLEDGEMENTS

Creating a dissertation, particularly one as complex and lengthy as this one, requires focus and perseverance. I could not have completed my research without the unwavering support of family, friends, and colleagues. This manuscript is dedicated to my family, to my husband, Bruce and my daughter, Nichole who provided me with loving encouragement, and strong, gentle support throughout this process. Both of you reminded me of what is important—to laugh, to persevere, and to believe in myself. But you also had the wisdom to know when to push me forward and when to be patient. For this, I thank you.

To Joe Rees, you have been my advisor, my teacher, my mentor, and my friend throughout this process. Thank you for challenging me, and for the opportunity to take risks and explore areas that I felt, would enhance my research. You have made me a better researcher than I would have been otherwise, and have fostered my love of learning and exploration, not always for an end product, but sometimes just because. I am truly richer for having known you. Special thanks also go to my dissertation committee members but especially Jim Wolf, Karen Hult, and Bill Murray who took the time to read and comment on the multiple drafts of this paper. I truly appreciate your insights and constructive suggestions that ultimately led to an improved product.

To my friends and colleagues at Virginia Tech and the Agency for Healthcare Research and Quality, who provided support and encouragement throughout this long process but especially Bryce Hoflund, Boyce Ginieczki, Jason Fichtner, Judy Sangl, Chuck Darby, Mamatha Pancholi, and Howard Holland. Your persistent, but friendly prodding kept me motivated and your quiet support encouraged me to continue on when there was no end in sight. Every doctoral student should be so fortunate to work and learn from people like you. Special thanks go to the staff of the Information Resource Center at AHRQ who patiently answered my many questions

about literature searches, finding obscure articles, and manuscript formatting, thus saving me endless hours of research in these areas.

And last but certainly not least, I have to thank Gordon Lightfoot—singer and songwriter, whose songs kept me company late into the night and into the early morning hours as I worked to complete my degree. Most recently it was my privilege to meet you in person; however, I don't think I articulated very well your importance to this effort. I want you to know that your beautiful baritone voice made the process more enjoyable and considerably less solitary.

TABLE OF CONTENTS

PART I: INTRODUCTION	1
CHAPTER 1: THE CURRENT STATE OF HEALTH CARE IN THE UNITED STATES	7
CHAPTER 2: THE HISTORY OF PERFORMANCE MEASUREMENT	40
CHAPTER 3: RESEARCH STUDY METHODOLOGY	63
CHAPTER 4: THE HUMAN ACTORS OF THE HOSPITAL QUALITY NETWORK.....	78
CHAPTER 5: THE NON-HUMAN ACTORS OF THE HOSPITAL QUALITY NETWORK	132
PART III: THE HOSPITAL QUALITY NETWORK AND THE POLITICS OF PERFORMANCE MEASUREMENT.....	180
CHAPTER 6: THE NQF HOSPITAL MEASURES STEERING COMMITTEE.....	189
CHAPTER 7: FOLLOWING THE “FRAGILE THREAD” OF PERFORMANCE MEASURES—THE POLITICS OF TRANSLATION.....	259
PART IV: A DISTINCTIVE VIEW OF HEALTH CARE PERFORMANCE.....	269
CHAPTER 8: CONCLUSIONS	269
REFERENCES	285
Appendix A: Map of Standard Measurement Sets	305
Appendix B: A Glossary of Actor-Network Theory Key Elements*	306
Appendix C: Comparison of Selected Versions of the NQF CDP	308
Appendix D: Environmental Assessment of Health Care in 2001	311
Appendix E: Hospital Steering Committee Work Groups.....	322

Appendix F: Milestones for the NQF Hospital Performance Measures Project.....	324
Appendix G: Participants of the Hospital Measures Steering Committee.....	326
Appendix H: Priority Topics Identified for Measures using HCUP	328
Appendix I: Measures Recommended by NQF Staff for Further Evaluation	329
Appendix J: Hospital Care National Performance Measures Workshop Participants	333
Appendix K: Informed Consent Form	335
Appendix L: Interview Questionnaire	338
Appendix M: Selected Characteristics of the Sample.....	341
Appendix N: Summary of the Main Questions of the Study	345

LIST OF TABLES

Table 1: Summary of Selected Performance Measurement Systems in Health Care	18
Table 2: Data Collection Methodology.....	68
Table 3: The National Quality Forum Board of Directors (2003)*	87
Table 4: Results of Member Council Elections in 2000	89
Table 5: The National Quality Forum Strategic Framework Board	90
Table 6: NQF Measure Evaluation Criteria for Hospital Performance	155
Table 7: Draft Recommendations of the SFB and NQF Staff Recommendations	205
Table 8: Measure Sources Identified by NQF Staff.....	208
Table 9:NQF Hospital Care National Performance Measures Ballot (2002)	229
Table 10: Hospital Performance Measures and Source Deferred to Work Group	231
Table 11: Hospital Steering Committee Workshop Member Differences by NQF Council	242
Table 12: Hospital Measures for Workshop Discussion.....	244
Table 13: Group 2 Hospital Care Performance Measures and their Source	247
Table 14: Measures Excluded from Group 2 Measure Ballot	251

LIST OF FIGURES

Figure 1: International Comparison of Spending on Health, 1980-2004	9
Figure 2: Infant Deaths Within One Year, per 1,000 Live Births	10
Figure 3: The U.S. Scores-Dimensions of a High Performance Health System	11
Figure 4: Hospital Network Tiering: Excluding NQF, CMS, and AHRQ	79
Figure 5: NQF Governance & Committee Structure (1999-2007)	85
Figure 6: Conceptual Framework for Quality Improvement	93
Figure 7: NQF Consensus Development Process-Version 1.7	137
Figure 8: The CMS/Premier HQI Project	165
Figure 9: Hospital Participation in the HQA Voluntary Initiative	240
Figure 10: Gender Profile	341
Figure 11: Age Of Key Informants	341
Figure 12: Education Of Key Informants	342
Figure 13: Clinician Status	342
Figure 14: Position Classification	343
Figure 15: NQF Classification	344
Figure 16: Tax Status	344

PART I: INTRODUCTION

“Everyone wants to measure, but no one wants to be measured.”

*Dennis O’Leary, M.D., former President
Joint Commission for the Accreditation of Healthcare Organizations*

I first learned of the quote above while I was conducting interviews as part of my research, and have often thought about the truth of the statement. As a nurse, I was always looking for ways to provide better care to my patients and doing that, as cited by my health care colleagues, require measurement and comparison. When I met Dennis O’Leary at a conference, I asked about his statement, and he indicated that what I had heard was correct, but incomplete. What he had really said was, “everyone wants to measure *the other guy*, but no one wants to be measured *themselves*.” It’s true; the thought of being compared to others sends a chill up our spines, and sets in motion feelings of anxiety, frustration, and worry. We are vulnerable. Our reputations are at stake and in some cases, our very livelihoods. In fact one physician friend of mine claimed that clinician acceptance of information derived from performance measures were similar to the five stages of grief developed by Elizabeth Kubler-Ross.¹ Yet we are told that it is “human to classify” (Bowker and Star 1999) as we constantly sort one thing from another—clean dishes from dirty dishes, good service from poor service and we do it as a matter of course.

So why do health care organizations and providers feel so much anxiety with regard to measurement? When it comes to performance measures and performance measurement, health care organizations have to deal with a proliferation of measures, many of which are conflicting,

¹ The Kubler-Ross stages are denial and isolation; anger; bargaining; depression and finally acceptance. The clinician version of performance measurement data acceptance follows this sequence: “you don’t have any data; your data stinks; your data is OK but it doesn’t apply to me; your data is OK, it does apply to me, but the reason the performance is poor is that you guys can’t get *your* act together; and finally, we need to work with you guys to optimize our performance” (D.Varga, personal communication, August 8, 2007).

and others have methodologies that are “black boxes” and not reproducible or transparent. The target is moving, not stationary due to the changing nature of the science that supports measures and measurement. Thus, classifying or categorizing is often a contentious, “political” activity among groups of stakeholders. Being classified in one category affords certain rewards and recognition, while being classified in another does not and in some cases may have negative consequences. Therefore, selection of the criteria or measures used to categorize is infused with “politics” at various levels, much of which may be invisible. Bower and Star argue that “each standard and each category valorizes some point of view and silences another” (1999: 5), and this is undeniably true of health care. Arguments and decisions about the “minutiae of classifying and standardizing” have to be made by people; consequently negotiation and, as a result, “politics”² become an issue.

The cost, whether in funds or resources, of measuring the quality of care can be substantial and may not yield much of a return on investment for health care organizations. And with the number of measures available and the number of organizations requesting data to complete report cards on quality, the popular buzzword frequently used in the field is “harmonization.” Harmonization, or alignment of measures, is designed to produce a parsimonious core set of measures that everyone can use, thus reducing the resources devoted to reporting on health care quality and allowing for comparisons among providers. Obviously, some measures will be selected for inclusion in this core set, and others will be eliminated. Measure developers such as the Joint Commission for the Accreditation of Healthcare Organizations, the Centers for Medicare & Medicaid Services, and others have much at stake.

² I use the term “politics” throughout this dissertation, which has a variety of meanings, however, in this manuscript; “politics” will be interpreted in terms of Actor-Network Theory. A more in-depth discussion will be provided later on, but in general, “politics” can be defined as the events that occur as a result of actor relationships and the interplay of those relationships within a network that are used to achieve actor values and goals.

The merits of one measure over another are debated in a public forum, specifically the National Quality Forum. Traditionally, health care performance measurement systems are designed by expert panels and by the strength of scientific evidence supporting them. Schuster and his colleagues (1997) compared the development of pediatric and adult performance measures using expert panel rankings and found that indicators based on rigorous studies, such as randomized controlled trials, were more likely to be retained by the panel than those indicators based on descriptive studies and expert opinion. The evidence problem becomes apparent when typically 50-90 % of medical procedures do not have a rigorous evidence base to support them (Smith, 1991; Brook, 1991; Ellis et al., 1995; Gill et al., 1996). Furthermore, clinical guidelines for these procedures, also based on the evidence obtained from clinical trials, vary substantially (Fahey and Peters, 1996). Even where published clinical evidence does exist, it may not be widely accepted. A study conducted by Nicollier-Fahrni et al. found there was complete agreement between experts and published evidence only 20-30 % of the time (2003: 21).

Therefore, other factors appear to be at work during the measure selection process. While terms such as evidence-base, feasibility, risk adjustment, confidence intervals, statistical significance and the like are bandied about and touted as the “gold standard,” there is “no natural law that says the best measure shall win” (Bowker and Star, 1999: 14), particularly when multiple stakeholders are involved. Thus, negotiation and politics come into play.

In this dissertation, I use primarily Actor-Network Theory (ANT) as a lens to look at the interactions of the actors within the Hospital Quality Network. According to ANT, all actors, both human and non-human, interact within a network to produce certain effects, and through their shifting associations and disassociations, the effects change. This theory is applicable here because the network I examined is not restricted to human actors; specifically non-human actors

include performance measures, legislative acts, and the like. Much like the human actors, the non-human actors interact and change over time. ANT addresses questions such as: How is it that it turned out this way? Who is influencing it? And why are actors behaving in a particular way? (Underwood, 1998) This political maneuvering typically involves two processes: arriving at the categories and/or standards, and determining what will be visible and invisible within the network (Bowker and Star, 1999).

This is the focus of my dissertation—to examine the politics of performance measurement within the network from an ANT perspective, i.e., the way that classification decisions have been made in health care. How influential are the politics during the process of standard selection, and how does a particular point of view become prominent over all the others? What factors influence the outcome? To find the answers to these questions, one must examine the field of performance measurement, its history, the organizations involved, and the environment in which it functions. To limit my research, I specifically focused on the Hospital Quality Network. Although it had existed for some time in health care, the Network emerged in a more active, collaborative role at the time the National Quality Forum announced its hospital project in 1999.

This manuscript is divided into several parts—both for ease of reading and to provide a full picture of this network. Part I provides an overview of the current state of the health care environment within the United States. It discusses the issues related to the quality of care, the concept of performance measurement, the role of measurement in quality improvement, and some of the problems and criticisms associated with measuring performance. Chapter 1 describes the Hospital Quality Network. To understand the politics of performance measurement in terms of relationships and interactions, it is essential understand the actors in the network,

their governance system and the institutional self-interest of the organizations involved. The term “actors” is used throughout; an in-depth definition is provided later on, but for the purposes of general discussion, an actor can be defined as an entity that modifies another through reasoned activity that occurs over a period of time. I have selected a combination of two theories to provide structure and also to help explain the interactions of the actors involved in selecting a core set of measures for hospital performance that can be used at the national level. Specifically, I use Institutional Rational Choice Theory exclusively to map and describe the actors of the Hospital Quality Network because this theory provides clues as to what motivates the actors and why they behave in a particular way. Then I use Actor-Network Theory as the major framework in which to analyze the actors and their interactions within the network as they construct their vision of what high-quality hospital performance should look like. An overview of each theory is provided at the end of the first chapter and is offered as a way to describe the events surrounding the selection of an initial set of hospital performance measures for the nation.

Chapter 2 provides a historical overview of performance measurement from the early 1800’s through the present, tracing its evolution and its changing uses in health care. Chapter 3 discusses the methodology used in this research, and includes data collection methods, a general description of the sample, an analysis, and some limitations of the study.

Part II will describe the specifics of the Hospital Quality Network and its actors—both human and non-human—to set the stage for the reader; it also discusses the notion of power and politics from the perspective of Actor Network Theory. The actors and their “biographies” are described using Institutional Rational Choice Theory and provide a window through which to view each actor’s motivations for participation in the network. Their positions in the network and the resistances they had to overcome are also described.

Part III will portray the events of a microcosm of the larger network specifically, the work of the National Quality Forum's Hospital Performance Measures Steering Committee and the events that surrounded its endorsement of a core set of hospital measures for national use. Actor-Network Theory provides the primary framework in which to recount and describe the actor relationships and the politics of a measure selection process that is based on consensus development. It provides an "on-stage" view as well as a "backstage" perspective that paints a richer, more complete picture of consensus development and the measure selection processes. In this section, I have selected one actor—performance measures—and follows it through the consensus development process, illustrating its varying roles throughout the discussions of the hospital project. Part IV briefly discusses the findings of the research, draws some conclusions from this study, and suggests ideas for future research in this area.

CHAPTER 1: THE CURRENT STATE OF HEALTH CARE IN THE UNITED STATES

“There are many obstacles to rapid progress in improving the quality of health care, but none exceeds the fact that the nation lacks a coherent, goal oriented, consistent and efficient system for assessing and reporting on the performance of the health care system.”

*Performance Measurement: Accelerating Improvement,
Institute of Medicine, 2005*

Across the nation, many Americans are provided with high quality services that either maintain or restore health. The care delivered by the health care facilities in the U.S. is the most scientifically advanced in the world (IOM, 2005). Large investments in the health care system have led to advances in knowledge, technology and pharmaceuticals that prolong life and diminish human suffering. For example, the number of individuals who live with cancer has increased from 3 million in 1971 to approximately 9.8 million in 2001 (CDC, 2004). A number of other breakthroughs have occurred in the areas of cell restoration, prosthetic devices, and rehabilitation that have restored functioning to many. Advances in the area of stem cell research and genomics have the potential to improve the health and longevity of all Americans.

Yet, at the same time, concern has surfaced about the quality, the safety, the effectiveness and the consistency of care being delivered by the health care industry.³ The President’s Advisory Commission on Consumer Protection and Quality in the Health Care Industry (President’s Advisory Commission) reported, “while most Americans receive high-quality health care, too many patients receive substandard care” (1998: 1). Evidence of quality problems such

³ For the purposes of this dissertation, the health care industry consists of diverse enterprises such as pharmaceutical companies, acute care facilities, nursing homes, insurance companies, and provider services. Most of the industry can be placed within a few broad categories: providers, purchasers, consumers, health plans, and research and improvement groups.

as avoidable errors (Brennan et al., 1991; Phillips et al. 1998), underutilization of services (Soumerai et al. 1997; Chassin, 1997), overuse of services (Bernstein et al., 1993), and variation in service (Wennberg and Cooper 1998) are well documented in the research literature.

Variations in both the quality of care and its cost exist in this country, but there is no evidence to support the notion that more expensive care is better care (Baicker and Chandra, 2004; Fisher et al., 2003a, b). In addition, racial and ethnic disparities proliferate within our health care system. Minorities consistently receive lower-quality care than whites, a situation that persists even when accounting for insurance status and income level of the patient (Ayanian et al., 1993, 1999; Barker-Cummings et al., 1995; Epstein et al., 2000; Gaylin et al., 1993; Hannan et al., 1999; Herholz et al., 1996; Johnson et al., 1993; Petersen et al., 2002; Williams et al., 1995).

The cost of health care continues to escalate and contributes to the increasing number of uninsured, who make up more than 15 percent of the population (IOM, 2002, 2004). Many others have only minimal insurance coverage, and limited funds available to pay for services out of pocket (Collins et al., 2004; Henry J. Kaiser Family Foundation, 2004b; Henry J. Kaiser Family Foundation/eHealthInsurance, 2004). While there are many reasons for the rise in the rates of uninsured, “there is little doubt that rapidly rising health care costs, driven in part by waste in the current health system, hamper efforts to expand coverage” (IOM, 2005: 19).

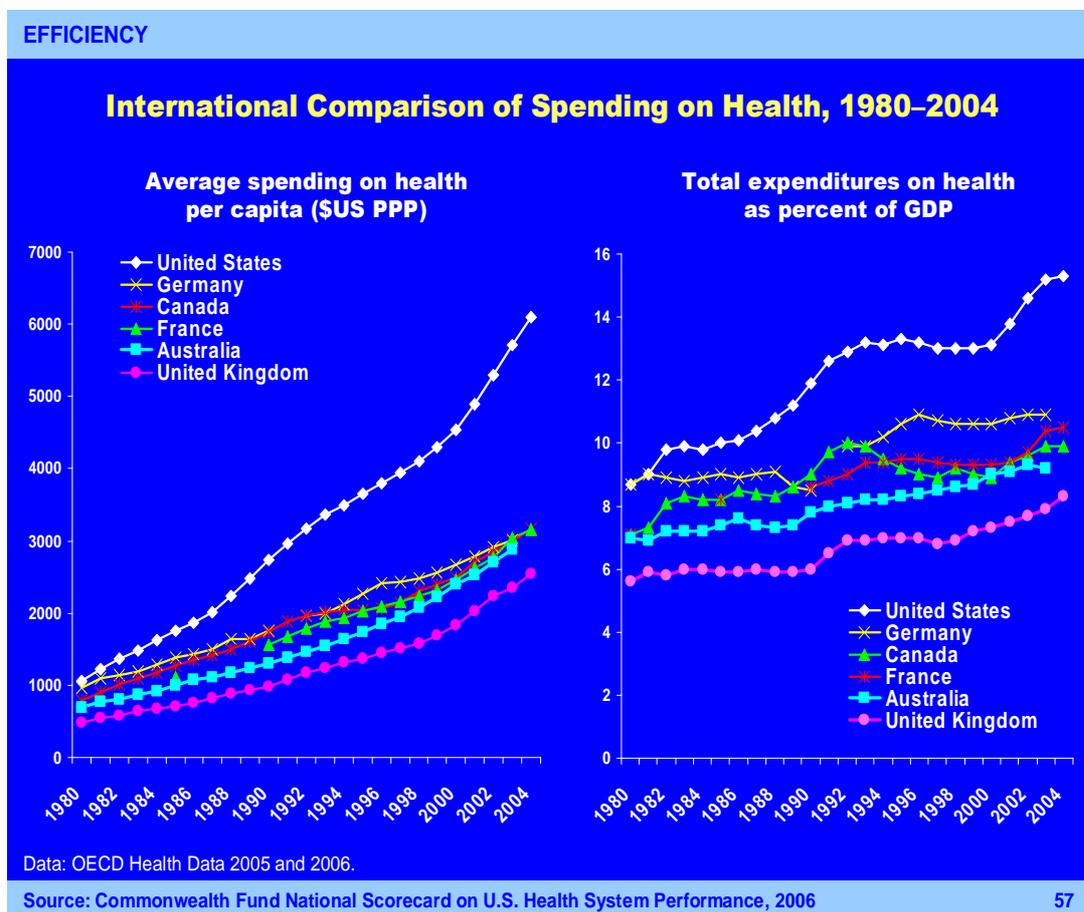
Americans are concerned not only about the rising cost of health care⁴ but also with the quality of the services they receive.⁵ The National Roundtable on Health Care Quality convened by the Institute of Medicine (IOM) reported, “serious and widespread quality problems exist

⁴ In a Kaiser Family Foundation survey, 46 percent of citizens polled in the U.S. indicated that they were worried about the overall costs of health care, while 21 percent stated that they were particularly concerned about the costs of prescription drugs (Henry J. Kaiser Family Foundation, 2004a).

⁵ From 2000 to 2004, a survey found that the number of people dissatisfied with the quality of their care rose by 10 percent. And during that same period, 40 percent of the respondents reported that the quality of their care had gotten worse (Henry J. Kaiser Family Foundation, et al., 2004).

throughout American medicine” (Chassin, Galvin, and the National Roundtable, 1998). A study by McGlynn et al. found that on average, adults receive only slightly more than half (54.9 percent) of the care that could provide a benefit (2003). And while per capita spending on health care in the United States vastly exceeds that of other countries,⁶ comparisons among industrialized nations reveal that other countries overall perform better than the U.S. (Hussey et al., 2004; Reinhardt et al., 2004). Figure 1 compares selected countries and their expenditures on health.

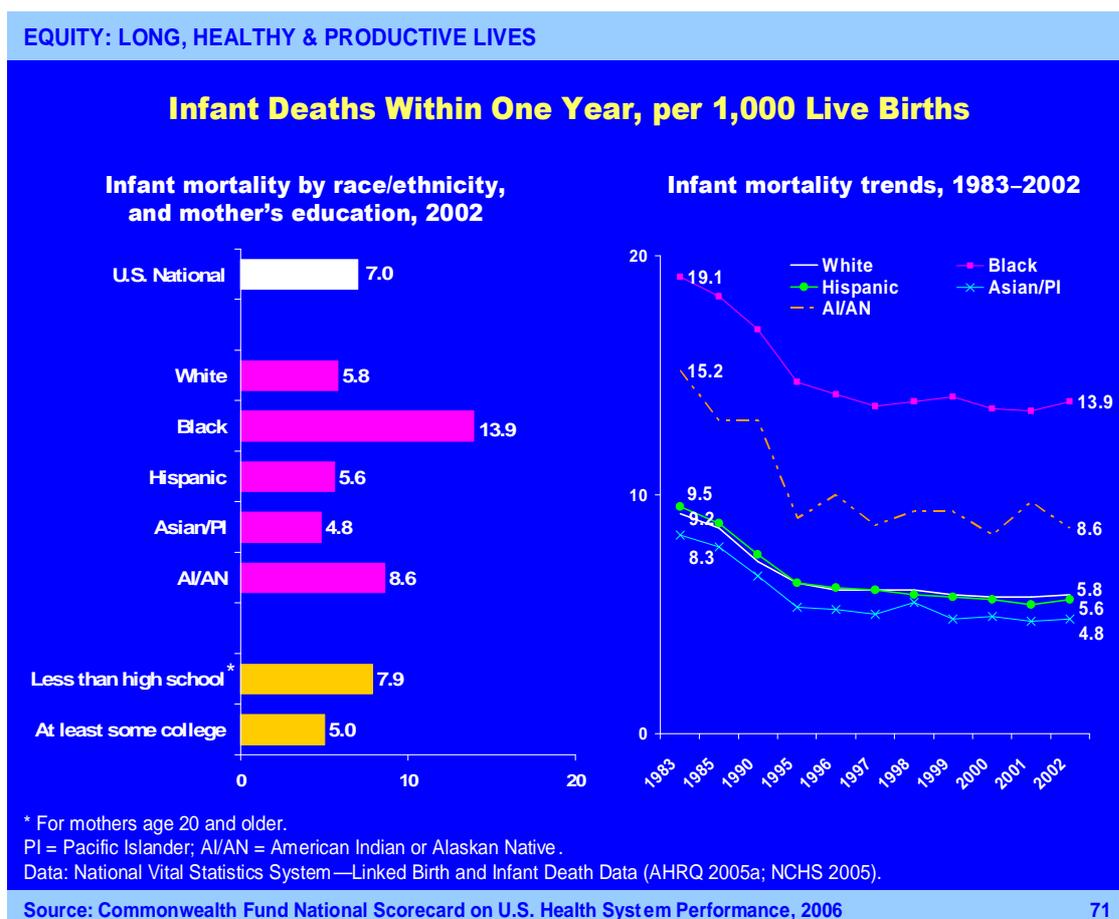
Figure 1: International Comparison of Spending on Health, 1980-2004



⁶ In 2001, the U.S. spent \$4,887 per person on health care, as compared to the \$2,792 Canada spent; the \$1,992 that the UK spent and the \$2,131 that Japan spent (Reinhardt et al., 2004).

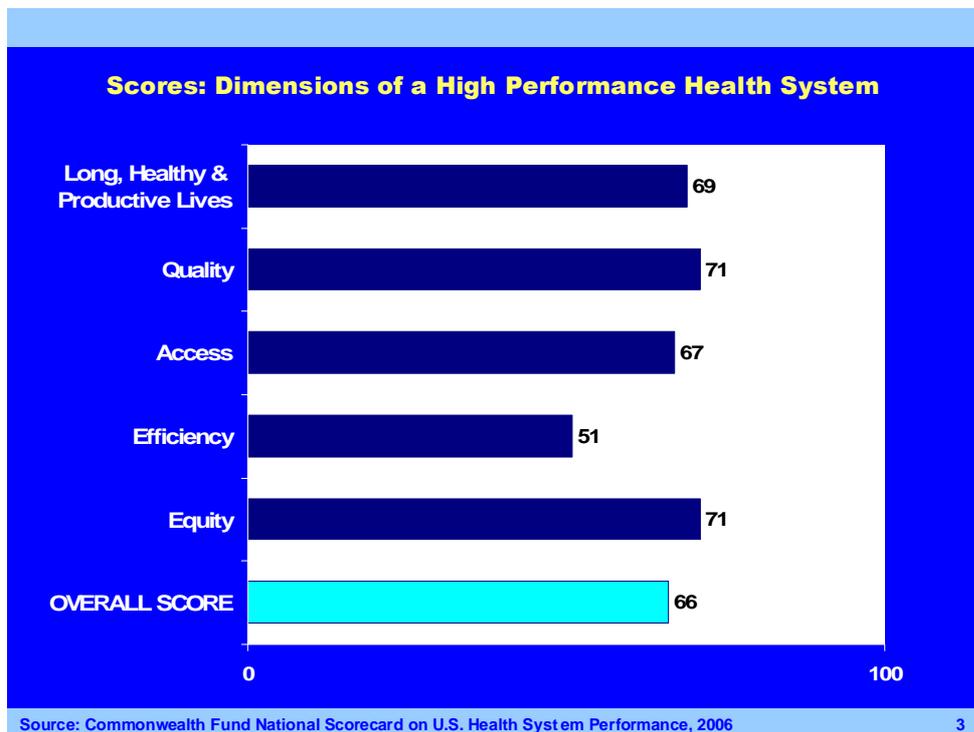
In fact, health care in the United States has been ranked 37th in the world for overall performance, behind countries such as France (first), Singapore (sixth), and Switzerland (20th) (World Health Organization, 2000). The United States performs poorly in terms of both life expectancy at birth and infant mortality, ranking in the bottom quartile of industrialized countries (Reinhardt et al., 2002). Furthermore, there is inequity with regard to health care as depicted in Figure 2. The national average for infant mortality is 7.0 deaths per 1,000 live births as compared to 2.7 for the top three countries (The Commonwealth Fund, September 2006). Noticeably higher are the infant mortality rates for African-Americans and American Indians/Alaskan Natives in comparison to the national average and to whites.

Figure 2: Infant Deaths within One Year, per 1,000 Live Births



A national scorecard developed by The Commonwealth Fund Commission rated U.S. health system performance on 37 key indicators across five domains. Figure 3 summarizes the average rates of health care performance for the country. Overall the U.S. scored only 66 out of a possible 100 points, with the scores on individual dimensions ranging from 51 to 71 (The Commonwealth Fund, September 2006).

Figure 3: The U.S. Scores-Dimensions of a High Performance Health System



Obviously, there are fundamental shortcomings affecting the quality of health care in America. “Quality of care is highly variable and delivered by a system that is too often poorly coordinated, driving up costs, and putting patients at risk” (The Commonwealth Fund, September 2006: 1). Improving the access to and the performance of our health care system is a matter of national urgency.

Some attribute these serious quality problems to a lack of knowledge on the part of clinicians (Wennberg et al., 1982; Wennberg, 1987), while others believe the causes of poor

quality are highly complex, and in the end difficult to resolve (Chassin, 1998). According to the Committee on Quality of Health Care in America, quality problems occur because of "fundamental shortcomings in the ways care is organized" (IOM, 2001: 26). An adequate health care delivery system, according to the IOM, must possess the capabilities to ensure services that are safe, effective, patient-centered, timely, efficient, and equitable (2001).

Yet, defining what quality is in health care is not easy. "Quality" is a complex, multidimensional concept that suggests different things to different people (Lee, 1996; McGlynn, 1997). Consequently, competing views of quality should be balanced among patients, purchasers, managers and health care professionals.⁷ Regardless of how quality is defined, according to experts in the field, the only way to know whether the quality of health care is improving is to measure the performance of those delivering it. Performance measures and performance measurement systems provide a tool to define what is meant by quality and to potentially determine if quality exists.

Performance Measurement and Quality of Care

While there are many responses to the challenges identified in the health care system, performance measurement, according to some experts, is central to quality improvement because it provides information on current and past performance that can help guide future improvement efforts. In particular, performance measures can distinguish between good and substandard performance. Accordingly, the development and application of performance measurement is essential to improving the quality of care as stated by I. Steven Udvarhelyi of Independence Blue

⁷ A widely used definition of quality in health care, which will serve as a working definition for this dissertation, is "the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge" (IOM, 1990).

Cross of Pennsylvania, in his testimony before the President's Advisory Commission (1998).

Measurement is one of the “first steps in the improvement process and involves the selection, definition, and application of performance indicators...” (Fine and Snyder, 1999: 24).

Performance measurement, while not the only influence, can act as a force to promote certain issues and agendas. Performance measurement conveys importance; what is important is measured while what is not measured is considered less important (Waggoner et al., 1999). By focusing people and resources on a particular aspect of an industry, performance measurement can be a driver of change and reform.

Performance measurement is a tool for continuous quality improvement. These measures provide data with which to potentially compare and assess differences in performance that in turn identify areas for service improvement. Improvement of services depends not only on detection and identification of problems, but also upon changing behavior. Effective strategies, such as those summarized by Bauchner and colleagues (2001), can influence changes in clinician behavior. But to determine which strategies are more (or less) effective, performance has to be measured and evaluated. Audits, in combination with reports involving specific recommendations, have resulted in behavior changes consistent with improvements in quality of care (Bauchner et al., 2001). Thus, performance measurement provides information identifying areas that may be targeted for improvement as well as information about progress toward established quality goals. Quality improvement initiatives cannot be realized without performance measures that are “coherent, robust, and integrated” into a system that is “purposeful, comprehensive, efficient, and transparent” (IOM, 2005: 4)

Performance measures serve a vital role in the many public reporting efforts that are currently taking place across the country. Performance measurement is a tool of accountability. Accountability in health care is the “need to demonstrate to others the quality and value of the

health care provided” (JCAHO, 1994: 28). Performance measurement systems are a tool of public accountability because they document the quality of care provided, as well as reveal those areas needing improvement. However, performance measurement, in and of itself, may not affect improvement in the health care system overall unless it is reported to the public.

According to the IOM, “performance measures can serve as the foundation for public reporting programs intended to promote accountability among providers and to aid consumers in making informed choices, serve as the basis for payment incentives that reward providers who deliver more effective and efficient care, and guide and inform clinicians and organizations in their quality improvement initiatives” (IOM, 2005: 2). Public disclosure of performance seems necessary to advance improvement via incremental changes in consumer, professional, and managerial behavior (Leatherman and McCarthy, 1999; Fowles, 2000; Hibbard et al., 2003, 2005). In its report, the President’s Advisory Commission concluded that the standardization of performance measurement and reporting efforts of both public and private organizations influence quality. “A key element of improving health care quality is the Nation’s ability to measure the quality of care and provide easily understood, comparable information on the performance of the industry” (President’s Advisory Commission, 1998:3). One study by Hibbard et al. (2005) reported that, even if consumers pay little attention to public reporting initiatives, clinicians do, and are inclined to improve in order to protect their reputations within the field.

One way to achieve the goals of quality improvement and public accountability is through the use of organizational report cards. This type of report card can be defined as “a regular effort by an organization to collect data on two or more other organizations, transform the data into information relevant to assessing performance, and transmit the information to some audience external to the organizations themselves” (Gormley and Weimer, 1999:3). The benefits

of report cards are that they reduce information asymmetries that affect market competition and consumer choice (Gormley, 1998; Gormley and Weimer, 1999), they encourage performance improvement (Leatherman and McCarthy, 1999; President’s Advisory Commission, 1998; Fowles, 2000; Barr et al., 2002; Hibbard et al., 2003, 2005), and they provide a mechanism to demonstrate accountability (Gormley and Weimer, 1999; Barr et al., 2002). Problems with report cards include assessment issues due to limitations of data and theory⁸ (GAO, 1994; Gormley and Weimer, 1999), organizational response and participation problems⁹ (GAO, 1994; Gormley and Weimer, 1999), and consumer receptivity issues¹⁰ (Gormley and Weimer, 1999). In 2001, the IOM’s Committee on the National Quality Report on Health Care Delivery published *Envisioning the National Health Care Quality Report*, which outlines a national approach to public reporting of health care system performance (IOM, 2001). While various efforts report the different aspects of health care such as health plans, nursing homes and home health agencies, there are few report cards that “compare data by hospital or that are designed for public dissemination, and still fewer that report patient satisfaction” (Barr et al., 2002: 51).

Problems and Criticisms of Performance Measurement and Performance Measurement Systems

While the current environment of health care clearly favors measuring performance to correct market asymmetry to improve the quality of care, there are others that are critical of

⁸ This means that all the relevant variables may not be able to be measured and that the theoretical links between variables that can be measured and those that are conceptually appropriate are weak (Gormley and Weimer, 1999: 7; GAO, 1994: 4).

⁹ Organizations can either use report cards for self-improvement or respond dysfunctionally. Some dysfunctional responses include non-participation, “cream skimming” or “cherry picking,” i.e., selecting only those cases that make a provider’s numbers look better), manipulating the numbers, and blaming the messenger (Gormley and Weimer, 1999: 13-15).

¹⁰ These problems include weak motivation to use report cards, cognitive limits of individuals, and information inequalities, which is the difference between those who process the information well, i.e., high income, well educated groups and those that do not, i.e., low income, poorly educated groups (Gormley and Weimer, 1999: 15-17).

performance measures and performance measure initiatives based upon the problems associated with the measures themselves, and/or their implementation. Historically, performance measurement efforts have not been effective, for several reasons. One important reason is the industry's inability to develop a consensus about what performance is, what constitutes "good" performance, and how to measure it. According to Meyer, "performance is not an easy subject. Clearly there is a need to study and rethink what is meant by performance and how to measure it" (Meyer, 2002a: 51). Performance is difficult to measure because either the parameters are constantly changing or they are only partially known (Meyer, 1994). Because performance is a moving target, problems arise with regard to the reliability and validity of measures. Inadequately developed and tested measures can lead to inaccuracies in determining quality.

A second problem is the existence of multiple and, at times, conflicting performance measures and performance measurement systems. Several organizations develop and market a variety of performance measures designed to report information about the quality of services provided by health care facilities and practitioners. The problem with these existing systems is that they are neither uniform nor standardized and as a result, no total picture of quality emerges from any single performance measurement system. Organizations such as the National Committee on Quality Assurance (NCQA), the Joint Commission for the Accreditation of Healthcare Organizations (the Joint Commission), the American Medical Association (AMA), the Foundation for Health Care Accountability (FACCT),¹¹ the Agency for Healthcare Research and Quality (AHRQ), the Centers for Medicare & Medicaid Services (CMS), and the Pacific Business Group on Health (PBGH) have developed a variety of performance measures, and

¹¹ In 2004, this organization ceased operations. As indicated in a letter from President, David Lansky, PhD, "We no longer feel we can make a distinctive contribution to health care reform. Other organizations have assumed some of our past roles..." (September 23, 2004). Despite its closure, organizations such as the IOM and others continue to refer to its work in performance measurement, specifically bringing ideas such as public reporting into the mainstream.

several have developed performance measurement systems, some of which have been thoroughly tested and validated while others have not. Measures that assess the same underlying health care construct, but have no correlation between results raise questions about the validity of the measures, which contributes to some of the skepticism surrounding measuring performance. Table 1 provides a detailed summary of some of the performance measurement systems developed by the organizations cited above. These are only a few of the performance measurement sets that are currently available to the health care industry. There are others that focus on specific aspects, procedures, or departments within health care organizations—all developed with the intent of improving health care quality. Appendix A provides a more extensive pictorial view of the various organizations involved in the health care performance measurement field and includes developers, provider organizations and vendors, to name a few.

Table 1: Summary of Selected Performance Measurement Systems in Health Care

Organization	Performance Measurement System	What's measured?
NCQA ¹²	HEDIS® ¹³	Performance of Health Plans ¹⁴
Joint Commission ¹⁵	ORYX ¹⁶	Performance of Health Care Organizations
AMA ¹⁷	AMAP ¹⁸	Performance of Individual Physicians ¹⁹
FACCT ²⁰	Consumer Information Framework ²¹	Consumer assessments of quality for various conditions ²²
AHRQ	a) Conquest ²³ b) CAHPS® ²⁴	a) Clinical Performance Measures b) Consumer Assessment of Health Plans
CMS ²⁵	QISMC ²⁶	Clinical and non-clinical care ²⁷
PBGH ²⁸	Quality improvement Fund/CCHRI ²⁹	Health plans, medical groups, and individual physicians.

¹² A non-profit organization that provides information to purchasers and consumers regarding the quality of managed care plans.

¹³ Health Plan Employer Data and Information Set, which includes 56 measures across 8 domains of care (McIntyre et al., 2001).

¹⁴ Allows for plan-to-plan comparisons across commercial, Medicaid and Medicare populations.

¹⁵ A non-profit organization that accredits a range of health care delivery organizations.

¹⁶ Uses an approved measurement system to report performance measures. In 2000, JCAHO approved 25 measures across 5 initial core measurement areas such as acute myocardial infarction, heart failure, pneumonia, pregnancy and related conditions, and surgical infection prevention.

¹⁷ A professional trade association of physicians.

¹⁸ Proposed as a voluntary program to measure individual physician performance in relation to national standards in credentials and qualifications, care environment, clinical process and patient outcomes.

¹⁹ In March 2000, the AMA ended the AMAP due to the “lack of interest from the market place, and a poor business plan” after an expenditure of \$12 million over a four-year period.

²⁰ A not-for-profit organization that collaborates with other organizations on quality assessment measures.

²¹ Organized into five areas based on how consumers think about their care: the basics, staying healthy, getting better, living with illness, and changing needs.

²² Conditions such as adult asthma, alcohol abuse, breast cancer, diabetes, health status under 65 years of age, and depressive disorders are some examples (FACCT, 2003).

²³ A quality improvement tool that links two databases; specifically clinical performance measures and medical conditions (AHRQ, 2001).

²⁴ A patient satisfaction survey that is applied across commercial, Medicare and Medicaid populations, and is required for NCQA accreditation and CMS reimbursement (AHRQ, 2001).

²⁵ Requires performance measurement activities for organizations that participate in the Medicare+Choice program, i.e., HEDIS and CAHPS data.

²⁶ Quality Improvement System for Managed Care is required by CMS for organizations that participate in the Medicare+Choice program.

²⁷ Organizations must develop at least two-quality improvement and performance improvement projects that demonstrate sustained improvement in clinical and non-clinical care.

²⁸ A non-profit business coalition of purchasers that seek to improve the quality of health care while moderating cost.

²⁹ California Cooperative Health Care Reporting Initiative that consists of 11 health plans (PBGH 2003).

A third problem with performance measures is that the development and maintenance of measures is resource-intensive. Depending upon the type of measure being developed and the number of measures under development, the cost could be anywhere from \$147,000 to \$1.1 million, (Booz, Allen and Hamilton, 2006:2) which includes testing, data collection, analysis and measure refinement. Further, the development process is very lengthy; typically taking anywhere from 14 to 36 months, (Booz, Allen and Hamilton, 2006: 2) again highly dependent upon the number of measures under development and the amount of testing to be completed. Further performance measure maintenance can be costly. For instance, measures that use the ICD-9-CM codes³⁰ in their definitions need to be updated each time codes are refined or added, which can be as frequent as every few months. If the measure is not maintained by the developer (or someone else), it could be perpetuating out-of-date practices that at best, could provide ineffective treatments, or at worst, could potentially injure patients.

The fourth problem centers on obtaining and providing information about performance. Unsurprisingly, with the increased demand for information, and with the multiple, and sometimes competing requests for performance data, the associated burden on health care organizations to provide such data can be enormous. For example, the data sources required to populate performance measures are often extracted from medical and billing records, which in today's health care system is still primarily a paper system. Thus a health care organization expends a number of resources such as time, personnel, and money gathering and collating the data that are required to satisfy the growing demand for performance information by purchasers, accreditors, health plans, government agencies and others.

³⁰This stands for the International Classification of Diseases, version 9, clinical modification. The ICD classification system is "used to translate diagnoses of diseases and other health problems from words into an alphanumeric code, which permits easy storage, retrieval and analysis of the data" (ICD-10, 1992b: 2).

Yet despite the increasing numbers of performance measures available, there is scant evidence of its usefulness, particularly in the area of performance outcomes, a fifth criticism of performance measures. Few studies show a relationship between performance measurement and improved patient outcomes (Kazandjian and Lied, 1999). Nevertheless, at least two empirical studies suggest this connection.³¹ Further, some experts believe that measuring performance does not improve the quality because to a certain extent, current measures in use only focus on one part of a system of care, which promotes a “studying for the test” mentality, which in turn, focuses attention on certain areas, often to the detriment of others. Thus, while it seems that the focused areas such as heart attack, congestive heart failure and pneumonia are improving, in actuality, the overall quality of care may be declining because resources are being drawn away from other areas to those areas identified as important by quality measurement initiatives.

In addition to the problems mentioned above, other impediments to performance measurement, particularly in health care, include professional autonomy and the attitudes of practitioners, a sixth problem. Traditionally, physicians and other health professionals have enjoyed a high degree of autonomy and public confidence. The complexities of the health care system, as well as its scientific, technology-driven underpinnings imply that only those within the health care profession are capable of assessing performance. Self-assessment, whereby providers monitor and discipline their own without external intervention, is considered a professional privilege as well as an obligation. However, professional autonomy only exists to the extent that society “allows it [professions] to maintain its prerogatives by reasons of confidence in its integrity and belief in its general beneficence” (Scarlett, 1991: 130). Thus, autonomy is related to public accountability, and when public confidence begins to wear away,

³¹ They are the mortality rate studies of patients after coronary artery bypass graft surgery (CABG) conducted in New York (Hannan et al., 1994), and the cesarean section cohort study of the Maryland Hospital Association Quality Indicator Project (Kazandjian and Lied, 1998).

as it has with the health care professions, there is an ensuing loss of professional autonomy. Performance measurement, as well as the public reporting of performance, poses a threat to professional autonomy and self-assessment in that performance is audited by “lay examining groups” (Lyons and Petrucelli, 1978: 8).

The attitude of health care professionals is another barrier to performance measurement initiatives. Performance measurement has been used to identify problems with individual performance so that those responsible could be embarrassed and punished. This “shame and blame” culture that currently exists within the industry must be positively changed to encourage a team approach focused on quality, safety and performance improvement. Another attitude held by some practitioners is that the problem of quality resides with other departments or disciplines. These professionals do not see themselves as either a part of the problem or a solution to it. The inability of practitioners to realize that they are part of a delivery system and that all departments and disciplines have a role to play in the improvement of organizational systems contributes to the challenges of performance measurement.

A seventh problem associated with performance measures includes measurement gaps in certain areas. Based on the IOM’s six aims for quality improvement—safety, effectiveness, patient-centeredness, timeliness, efficiency, and equity (IOM, 2001), the current measures used in health care are not comprehensive. Measures are lacking or absent in the areas of efficiency, equity, and patient-centeredness.

Moreover, the measures that are used in health care often provide “perverse incentives” or produce unintended consequences of being measured, an eighth criticism of performance measures. For example, the publication of cardiac surgeon’s mortality rates in New York (Chassin et al., 1996; Omoigui et al., 1996) has been linked to limiting access to care and/or delaying treatment to patients by encouraging surgeons to deny their services to the sickest

patients. What's more, if this type of behavior is not identified and actively discouraged, it could be rewarded through pay-for-performance initiatives, thus encouraging providers to act in their best interests and not in the interest of the patient.

Additionally, “up-coding” of medical events or recording an incident to make it look more complex than it actually is, can be encouraged by measuring performance. What this practice does is maximizes income and gives rise to advantages in risk-adjustment³² often associated with certain measures. However, even with risk-adjustment, measures are often considered unreliable because of issues such as small population or sample size. Hofer et al. found that the quality of diabetes measures was so poor that by removing three patients with the highest HbA1c³³ measures from a physician profile would have improved the performance of several of the outliers that were categorized as “poor performers” (1999). “Such large effects from comparatively small changes show both the problems of making judgments from an unreliable measure and why the temptation to game is so powerful” (Hamblin, 2007: 185). Hence, performance measures that inadvertently encourage perverse incentives loose credibility not only with consumers but providers as well.

Overall, it is apparent that the way we measure performance is not perfect and what we measure is not comprehensive. However, while not perfect, there are those that believe performance can be measured and measured accurately using well constructed, reliable and valid measures of quality that have been thoroughly tested and evaluated for unintended consequences and that these measures currently exist. Based on this idea, what has emerged from the health care industry is a number of performance measures and performance measurement systems that

³² Risk-adjustment is an algorithm that is applied to certain performance measures (i.e., mortality for example), that is designed to “level the playing field” or to not penalize those providers that take care of the sickest patients. It is an attempt to impede perverse incentives associated with measures.

³³ This is a measure of diabetic care, specifically a laboratory blood test that if maintained within a certain range, helps physicians monitor patient’s blood sugar levels in order to avoid long-term complications from the disease.

provide a highly fragmented, incomplete picture of care quality. Thus, there is no system currently in place that can measure, track, and comprehensively report the quality of the care delivered in the United States.

The availability and use of measures to demonstrate quality in health care is accelerating; however, there remain large gaps in the understanding of their relevance, reliability and meaning. Furthermore, performance measure development has been “uneven across different settings, populations, and health conditions” (IOM, 2005: xi). A contributor to the problem is that the health care industry is significantly behind other industries in using information. Desired goals in health care, such as conducting performance measurement and evaluating health systems, to name a few, are hindered by the need for an established, standardized information structure (IOM, 2001). The U.S. lags behind other nations in the use of the electronic health record (EHR); only 17 percent of the physicians in this country use an EHR as compared to the 80 percent of physicians in other countries (The Commonwealth Fund, September 2006). Data contained in an electronic format can significantly reduce the burden associated with performance measurement and improvement activities. Information can be quickly aggregated and analyzed so that information on quality can better inform decision-makers within health care.

It is clear that a comprehensive, industry-wide effort is needed to “focus the development of quality measures that would enhance the Nation’s ability to evaluate and improve health care” (President’s Advisory Commission, 1998: 3). As a result, what has emerged, with the help of the government, is a network of organizations focused on the issues of health care quality measurement and improvement.

The Hospital Quality Network

The entity that has emerged in response to a variety of environmental factors is a loosely structured Hospital Quality Network (HQN), which consists of a range of public and private stakeholders within the health care industry. The actors of the network consist of providers, health plans, consumers, research and quality improvement types, and purchasers of health care, to name a few. Each has a different perspective in what it considers to be quality health care, and a different set of priorities and values based upon each organization's mission, structure, governance, and culture. These organizations collaborate in an effort to improve the quality of care provided to Americans, to lower the costs of care, and to offer information about the quality of care to consumers so that they can make better health care decisions. To identify all the actors and study all their interactions within this network is beyond the scope of this dissertation; however it may be possible to study selected interactions of the main actors in a microcosm. Specifically, in March 2001, a Hospital Steering Committee was formed by the National Quality Forum (NQF), a not-for-profit, private membership organization consisting of the various health care constituencies, to discuss and establish an agreed-upon set of core performance measures for hospitals. This core measure set would become the national standard for evaluation of hospital quality; the intent was to publicly report hospital performance, in an effort to drive quality improvement and to provide information for greater consumer choice.

In summary, the quality of the health care system within this country is substandard, as compared to other industrialized nations. Performance measures can help identify those areas that require improvement and foster change, but there are problems associated with measures and measurement, and the most immediate being the plethora of performance measures available for a variety of quality initiatives that are neither uniform nor standardized. In an effort to consolidate the measures by which to judge the quality of health care, the National Quality

Forum was established to provide an opportunity for discussion and debate, in the hope that consensus could eventually be reached on a core set of measures that could be used for a variety of purposes by various stakeholders. Obviously, consensus development consists of negotiations, iterative processes and conflict. At issue are questions about how consensus is developed among a broad spectrum of stakeholders during a measure selection process for a national effort, specifically why one point of view is advanced while another is neglected. Other issues requiring investigation include: how do these negotiations take place, who develops the process by which they take place, and who determines the final outcome for measure selection?

The Research Questions

In their book, *Sorting Things Out: Classification and Its Consequences*, Bowker and Star contend that to “classify is human” and that there is a certain invisibility about it, due to the ordering of human interaction via classification systems. They argue that “each standard and each category valorizes some point of view and silences another” and test this hypothesis by examining widely adopted classification designs to uncover how classification decisions have been made (Bowker and Star, 1999: 5). Arguments and decisions about the “minutiae of classifying and standardizing” have to be made by people, and consequently negotiating and politics are very much a part of the landscape. Politics typically takes place around two processes: arriving at the categories and/or standards, and determining what will be visible and invisible within the system. In some instances, the latter is an ethical choice (Bowker and Star, 1999).

One example of the ethical issues involved in classification is the “decision of the U.S. Immigration and Naturalization Service to classify some races and classes as desirable for U.S. residents, and others as not, [which] resulted in a quota system that valued affluent people from

northern and western Europe over those (especially the poor) from Africa or South America” (Bowker and Star, 1999: 6). An example of what becomes visible and invisible as referred to by the authors is the classification of students according to standardized test scores that are designed to place more importance on some kinds of knowledge skills rather than others (Bowker and Star, 1999). By emphasizing these particular knowledge skills, certain privilege is extended to those meeting the valorized criteria, in the form of access to better institutions of higher learning, which subsequently may lead to better job placement (and a higher earning potential) after graduation.

Once a standard is established, the politics are forgotten—or if they are recorded, are buried deep within the meeting minutes or other documents. Once a standard is recognized it becomes embedded in the mechanics of the system, i.e., computer programs, data collection tools, etc. Because of this, it can be difficult to determine why a particular standard was selected and why other options were not. Given the increasing number of measures and increasing requests for performance data by the various stakeholders, how influential are politics during the consensus develop phase of measure selection? How does a particular point of view become more prominent over the other points of view, and what factors influence the outcome? Why do certain values trump others during this process? These are some of the broader questions concerning the political aspects of performance measure selection.

On a smaller scale, this dissertation will explore these questions and others related to the selection of standards—specifically performance measures and the politics associated with a consensus driven measure selection process in health care. Whose point of view prevails over the others, and why? What actors influence the measure selection outcome, and how do they make their values and goals more prominent over others? I will examine the selection of measures for a national hospital performance measurement system and the various forces that

can influence the outcomes in a consensus development process. The NQF Hospital Steering Committee will be used as a microcosm of the larger HQN. While there are many frameworks and criteria that can be used to design a performance measurement system, there is little written about the factors that affect the work of a network of organizations convened to design a national system of performance measures for hospitals, or the federal agencies' role in the process. And still less is written about the politics involved in the selection of performance measures for a national system that emerge from a consensus process. The study of the NQF's Hospital Steering Committee, as a microcosm of the larger hospital network is a unique opportunity to explore those factors that influence the selection of performance measures to be used in a national performance measurement system for hospitals.

For the purposes of this dissertation, I will combine certain aspects of two theories to address the different facets of the HQN. Specifically, I will use Institutional Rational Choice Theory (IRC) and Actor-Network Theory (ANT), forming a "hybrid" theory that can address the overall research question: How are hospital performance measures selected and what is the role of politics in a formal consensus development process? The specific research questions are:

1. What role does politics play in the selection of performance measures by the NQF Hospital Steering Committee?
 - a. How did the members of the Committee interact as representatives of the various actors of the HQN during the consensus development process?
 - b. What were their critical tasks during the measure selection process and who or what factors influenced these tasks?
 - c. Whose point of view prevailed and why?
2. What is the role of federal agencies, specifically AHRQ and CMS, in the measure selection process?
 - a. Do federal agencies become involved in the politics of measure selection and consensus development?
 - b. How influential are they in the measure selection process?

- c. What do they contribute to the process?
3. What role do clinicians play in the selection of performance measures?
 - a. How influential are they in the selection process and in developing consensus?
 - b. What influences do clinicians employed by the federal agencies, particularly AHRQ and CMS have in a consensus development process?
 - c. What points of view, if any, do they silence and why?
 4. How has the field of performance measurement and the performance measures themselves, influenced the selection of a core set of performance measures for hospitals?
 - a. What types of roles do performance measures play throughout the selection process?
 - b. Has the nature of performance measures directly or indirectly assisted in silencing certain values and viewpoints within health care?

Politics, according to Lasswell (1936), is largely a function of who gets what, when and how. In short, politics matters. Performance measurement, to a certain extent, determines the same thing, particularly in the current health care environment of pay-for-performance programs, public reporting and accountability. The selection of measures used to judge performance is significant because they categorize behaviors as “good, bad, or indifferent.” A provider’s rating often determines whether payment is received and in what amount, or whether accreditation is awarded, or in some cases whether a lawsuit is brought. There are various consequences associated with measure selection, including ethical, financial, autonomy gain or loss; and altered identities or status. It provides certain benefits and advantages for particular technologies; and also provides visibility for select providers (Kindleberger, 1983). According to Bowker and Starr, “for any individual group or situation, classifications and standards give advantage or they give suffering. Jobs are made and lost; some regions benefit at the expense of others” (1999: 6). And because the selection of performance measures for these and other programs is determined through a collaborative network with a variety of actors, the answers to

these and other questions about the network, its actors, and the consensus development process can provide more efficient and effective ways to participate in the network and in the network governance process.

Institutional Rational Choice Theory

In order to understand the measure selection process for a national hospital performance measurement system and its associated politics, a thorough understanding of the motivations and self-interests of the actors within the network is required. To accomplish this, I selected Institutional Rational Choice Theory (IRC), which provides a basic framework to describe the actors of the network identified by the key informants interviewed for this project; it helps explain individual action as part of a group working toward a common goal, which, in this case, is improving the quality of care provided by hospitals. Rational choice “refers to decision making based on an internally consistent ordering” which seeks the highest feasible level of subjective satisfaction for the actor (Caporaso and Levine 1992: 80). The theory is based on the assumption of the rational actor, or one that acts consistently with his or her attitudes and beliefs. Actions are taken only when they will advance the individual’s beliefs or attitudes (Azoulay, 2004). In general, the theory assumes 1.) “humans are intendedly rational; 2.) their behavior is strongly influenced by institutional rules; and 3.) they seek to influence institutional rules in order to alter others’ behavior” (Sabatier 1999: 263). Accordingly, institutions can be viewed as governance or rule systems in that they represent “rationally constructed edifices established by individuals seeking to promote or protect their interests” (Scott, 2001: 34). Thus, IRC provides a starting point to describe and map the actors of the network; it also provides clues as to why they behave as they do, particularly when developing consensus around a core set of measures for a national hospital reporting system. It also allows insight into actor’s organizational structures,

governance rules, and mission statements that constrain the interactions of individuals, and the choices available to them. In this study, individual actors, (often senior leaders in their establishments) are constrained by the mission, culture, history, organizational norms and established networks of their respective organizations. According to Goffman, “information about the individual actors help to define the situation, enabling others to know in advance what he will expect of them and what they may expect of him. Informed in these ways, the others will know how best to act in order to call forth a desired response from him” (1959: 1).

Actor-Network Theory

The second but primary theoretical lens used in this project is Actor-Network Theory (ANT), also known as the “sociology of translation.” It is applicable to this research because this theory specifically looks at the interactions of the actors within a network, how they change over time, and how these interactions order and structures the network and its end products. It is a relational and process-oriented theory that rejects the notion that only human action and meaning construct society and social interaction. It examines the shifting associations (and disassociations) between heterogeneous actors³⁴ within a network. It asks questions such as: How is it that it turned out this way? Who is influencing it? And why are actors behaving in a particular way? (Underwood, 1998) What differentiates this theory from others is the concept of the non-human actor, which influences and interacts with other actors within the network. Due

³⁴ Goffman defines an actor as “either male or female; is an individual that appears in the presence of others, in which there is some reason for him/her to mobilize their activity in order to convey an impression to others, which is in their interest to convey” (1959: 4). Latour defines an actor as “any entity, both humans and nonhumans, that modifies another entity in a trial; they act and their competence is deduced from their performances; the action, in turn, is always recorded in the course of a trial and by an experimental protocol, elementary or not” (2004: 237). I have combined elements of each into one definition, and thus for the purposes of this dissertation, an actor is an entity, either human or nonhuman, that modifies another entity through mobilizing activity for some reason over a period of time. The activity or performance the actor conveys is that performance which is in its best interest to convey to achieve its goals and, which in turn determines its competence within a particular situation.

to the very nature of non-human actors, they contribute to the interactions within the network, which affects the final order of the network, and thus, influences the network effects or products. And because they participate in the interaction, they become part of the politics that play a role in developing the network order.

At the heart of the theory is the idea of a heterogeneous network: the “identities, shapes, and forms of actors, be they human or non-human, are a product of the relations established within a network” (Goodman and Walsh, 2001: 198). ANT recognizes the interconnectedness of the elements of the network and assumes that each actor has agency, that is, that none is passive. Yet, each actor’s influence over and resistance to other actors varies considerably (Callon, 1993). A network, according to the theory, consists of various materials that are “patterned” or arranged together to achieve a variety of effects. It is this “social engineering” that, when individual resistances of each actor are overcome, results in a cohesive network that produces institutional and organizational effects or products (Law, 1992a). These effects can be in the “forms of power, hierarchy or market exchange” as well as “other forms of coordinating and governance mechanisms” (Thompson, 2003: 74). ANT can examine the means and materials by which actors establish their power, how they overcome resistances and stabilize their position, and how they maintain and promote the network ordering (institutionalizing) into large-scale phenomena, thus assuring their position within the network. Therefore, the bottom line from the ANT perspective is that the “social” is nothing more than “patterned actor-networks which are made up of heterogeneous materials” (Thompson, 2003: 73) with a focus on examining how actor-networks are formed, as well as how they are maintained, and/or collapse.

ANT consists of several essential concepts and terms that describe actor relationships and the network that emerges from these relationships. These key elements are described below, and

because they will be used throughout this manuscript, are included in a glossary, located in Appendix B.

Punctualisation involves “the substitution of a network by a point” (Law, 1992b: 385). This point or node represents a complex network of ordered actors and things that is often not readily visible. As John Law states:

“All phenomena are the effect or the product of heterogeneous networks. But in practice we do not cope with endless network ramification. Indeed, much of the time we are not even in a position to detect network complexities. So what is happening? The answer is that if a network acts as a single block, then it disappears, to be replaced by the action itself and the seemingly simple author of the action...So it is that something much simpler comes, for a time, to mask the networks that produce it” (1992a: 5).

For our purposes we can use the automobile as an example. An automobile is a complex system, consisting of various mechanical and computerized parts. All of these components are essentially invisible to the driver, who deals with one object—an automobile. Furthermore, according to ANT, when the actor-network begins to break down, the effect of punctualisation tends to disappear. In our example of the automobile, if the engine is not working properly, the driver becomes aware of the fact that an automobile is a collection of parts—alternators, fan belts, and electrical wiring, etc., and not just a single entity that was designed for transportation. But once the engine is fixed, the punctualisation effect is reestablished, and the network of parts that make up an automobile become invisible once again, and are replaced with the notion of a single entity. Much like the automobile, networks and network patterns that are widely used and

performed can be punctualised, thereby simplifying the social without having to draw on the infinite complexities associated with the network.³⁵

Inscription, according to Law is “a process of creating artifacts that would ensure the security of actors’ interests” (1992a). This idea has to do with the notion of durability. Some materials are more durable than others, e.g., the spoken word does not last very long, and text has a somewhat longer lifespan. Therefore the idea behind inscription is to maintain network relational patterns that are of interest to an actor for as long as possible; to do so, one must associate them with durable materials. For example, the office procedures of an organization, once decided, are incorporated into the office software, procedure manuals, and the like, which increases the likelihood of permanence.

Translation, a central tenet of ANT, is broadly defined as the struggle to create an ordered network “which generates ordering effects such as devices, agents, institutions, or organizations” (Law, 1992a). Translation is a process—a “multifaceted interaction in which actors 1) construct common definitions and meanings; 2) define representatives; and 3) co-opt each other in the pursuit of individual and collective objectives” (Bardini, 1997: 20). It shows how actors with differing interests become aligned. Callon, in his (1986) case study of scallops, fishermen, and scientists in St. Brieuc Bay, France, illustrated the stages of translation that include problematisation, interessement, enrolment and mobilisation. This seminal work documents the “moments of translation” by describing the efforts of three scientists who studied the cultivation of scallops in Japan and proposed these techniques to the fishermen of St. Brieuc Bay, as a solution to their problem of declining scallop harvests. *Problematisation* is the first moment of translation, in which a central actor identifies a real-world issue. Other actors are

³⁵ As noted by Law and others, punctualisation is a *process*, rather than something that can be achieved indefinitely; it is a relational effect that is recursive and reproduces itself (Law, 1992a).

identified that have similar interests in the issue or problem identified by the focal actor. In Callon's study, once the scientists returned to France, they set about determining whether the Japanese techniques of anchoring scallops were transferable to St. Brieuc Bay. As they investigated this question, the scientists came into contact with other actors within the network, i.e., their colleagues, the fishermen, and the scallops. The scientists tried to link the interests of the actors to the specific knowledge they acquired in Japan, thus establishing themselves as an *obligatory point of passage*³⁶ The line of reasoning is:

“...if the scallops want to survive (no matter what mechanisms explain this impulse), if their scientific colleagues hope to advance knowledge on the subject (whatever their motivations might be), if the fishermen hope to preserve their long term economic interests (whatever their reasons) then they must: 1) know the answer to the question: how do the scallops anchor?, and 2) recognize that their alliance around this question can benefit each of them” (Callon, 1986: 205-206).

The second moment of translation is *interessement*, which is the manner by which the focal actor tries to impose and stabilize other actors' identities. It is a process of convincing the other actors to accept the identities and issues as defined by the focal actor as well as reinforcing the resolve to move through the obligatory passage point to address the problem. In Callon's case study, the scientists had to define the roles of the actors in relation to the identified problem, i.e., what's in it for them. At the same time, they were trying to exclude voices of dissent both from outside of the network as well as within. The scallops' role included threatened extinction in the bay and the assumed ability that they could anchor to the collection vessels in a fashion

³⁶ An obligatory point of passage can be defined as an unavoidable point or narrative bottleneck through which actors that the focal actor wants to include in the network must pass, in order to continue to exist and develop and to articulate their identity and reason for being (Singleton and Michael, 1993: 229-230).

similar to the Japanese variety. The fishermen, in their role, were convinced through a series of meetings and eventually accepted the scientist's approach to their problem of scanty harvests. The scientific colleagues were persuaded of the scientist's methods for anchoring the scallops through a series of meetings, conferences and publications. Once the actors are convinced, they can proceed to the third moment of translation—*enrolment*. This phase involves the successful alignment of the interests of the actors of the network with the focal actor's (or enroler's) interests as well as solidifying the intent of the network.³⁷ It is at this stage where the actors commit resources such as time, money, expertise, knowledge, etc. "Enrolling actors into a network is a political process in which humans and non-humans, as individuals or groups persuade, coax, or entice others to become parts of their network; that it is in their best interest to do so, or that they share a common cause" (Atkinson, 2002: 2). In Callon's study, the scallops, if they were to be enrolled into the network, had to anchor themselves to the collection baskets. The scientists entered into "negotiations" with the scallops and through their experiments tried to encourage the scallops to anchor in sufficient quantities. The colleagues of the scientists were accepting of the anchorage theory as long as enough evidence (in the form of successful anchoring) was available. The fishermen, according to Callon "watched like amused spectators and wait[ed] for the final verdict" (1986: 213). Yet despite the agreement reached by all the actors involved, a network would emerge only after the fourth moment of the translation process—*mobilization*. It is the process by which the enrolled actors become mobilized to address the problem identified by the focal actor in the initial phase of translation. In this example, the focal actors, the scientists were able to mobilize their colleagues, the fishermen, and

³⁷ The enroler is considered "successful" if other actors in the network have been organized for the benefit of the enroler (Somerville, 1999).

the scallops as long as the images of these actors, as envisioned by the focal actor, persisted within the network. Specifically:

“The scientists were able to get some (representative) scallops to anchor but would the mass of scallops follow suit? Likewise, would the scientific community follow the scientists’ closest colleagues into believing the data? And would the fishermen follow the directions of their representatives? Affirmative answers to these questions would allow the scientist to position themselves at the head of three populations—the scallops, the scientific community, and the fishermen; through the intermediaries or representatives, each population would be successfully translated and tied into a network on the terms set by the scientists” (Murdoch, 1997: 739-740).

Thus, the four moments of translation, problematisation, interessement, enrolment and mobilisation “lead the enrolled actor or numbers of actors through a non-reversible obligatory passage point that transforms and consolidates (‘black-boxes’)³⁸ both the newly incorporated actor(s) and the network” (Atkinson, 2002: 1).

Another key idea of ANT is that of *immutable mobiles* (Latour, 1990). These are entities like maps, photographs, paintings, textual descriptions, and all kinds of images, but can also include devices such as ships, cloths, or scientific instruments (Thompson, 2003). These entities are ‘fixed’ in that they remain stable, but can also be ‘mobile’ in that they can be “rearranged and reconfigured through the network of places and agencies to which they are attached or through

³⁸ Stable ties can be used as packages or black boxes and include devices, texts, standardized sets of organizational relations, social technologies, and organizational forms (Law, 1992a). While black boxes continue to face resistance, according to ANT, they are never complete, as actors can defect at any time (Callon, 1986) however; there is an associated property of irreversibility about them. Irreversibility is “the extent to which it is subsequently impossible to go back to a point where that translation was only one amongst others; and the extent to which it shapes and determines subsequent translations” (Callon, 1991: 150), which suggests that translations are difficult to undo and hence constrain future possibilities.

which they operate” (Thompson, 2003: 73). Displacement or mobility of these entities is achieved through network transformations (Latour 2005: 223). Thus, immutable mobiles have the combined properties of mobility, stability, and combinability (Latour, 1987). An example of an immutable mobile is the ICD statistical classification system. This tool “translates diagnoses of diseases and other health problems from words into an alphanumeric code” (ICD-10, 1992b: 2). These codes remain relatively stable, are designed to be combinable and comparable, and are used in a variety of settings. This coding system has become “the international tool for standard diagnostic classification for all general epidemiological and many health management purposes” (ICD-10, 1992b: 2). Thus, the ICD statistical classification system is an immutable mobile; the codes remain stable, they are used in a variety of networks in which they can be combined with other sources of information and they can be used for comparison purposes across networks.

The concepts illustrated above all relate to describing the association and disassociation of actors within a network. Building networks, according to Law, is a process of overcoming the resistance of actors to achieve ordering, and to weave them together in a network, which, in turn, generates effects like organizations, power, standards, or other structures (1992a).

Because this manuscript specifically discusses the issue of politics in networks, it is essential to have a general understanding of “politics” from an ANT perspective. Specifically from this point-of-view, politics is viewed in terms of associations among actors within a network. Here politics is defined “as the entire set of tasks that allow the progressive composition of a common world” (Latour, 2004: 53). What’s more, Latour states, “that associations [within networks] are not enough, that they should be *composed* in order to *design* one common world.” (Latour, 2005: 259). This is where the notion of “power” enters into the equation. Networks are a design of forces that “emerge from and dissolve into the play of power” (Brown and Capdevila, 1999: 38). In other words, power is a network effect that comes

from the ordering of associations and disassociations of actors. Power shapes the identity of the network, connects it to the world around them, and is ultimately responsible for its collapse. Hence two concepts are included in the definition of politics from an ANT perspective—relationships and power. Part II will provide a more in-depth discussion of politics and power from the ANT perspective.

“A major advantage of actor-network theory is its capacity to encompass transformation and change” because it is based on the regrouping and the formation of new objects and new networks (Goodman and Walsh, 2001:3). A second advantage of this theory is that it offers a non-dualistic approach to technology. ANT recognizes that all actors, be they human or non-human have a role in social processes rather than “merely props for social action” (Prout, 1996: 199). I selected this theory because of its ability to encompass transformation and change, which makes it particularly useful in the turbulent health care environment. In addition, this theory recognizes all the actors (both human and non-human) within the network and focuses on actors by intent, not because they are human but because their role is significant to a particular network. In this dissertation, ANT is woven throughout, sometimes implicitly and at times explicitly, and where possible with examples to illustrate the elements of the theory. Thus, ANT provides a unique way to view my research, and was particularly helpful in organizing and reporting on the extensive amount of materials and interviews analyzed for this project.

In sum, my research will focus on the NQF Hospital Measures Steering Committee and the deliberations surrounding the selection of a core set of performance measures for hospitals. Initially IRC is used to describe the actors and then ANT provides an overall framework in which to view the consensus development process of measures selected for a national performance measurement system. Each theory contributes to understanding the actors in the network—their characteristics, their motivations and self-interests, however, ANT is used

exclusively to view the interactions among network actors as they are made and re-made throughout the discussions and events surrounding this NQF project.

CHAPTER 2: THE HISTORY OF PERFORMANCE MEASUREMENT

Performance measurement has been around for quite some time, however serious efforts to measure performance in terms of quality did not appear until the mid- 1950's, in both business and in health care. Both areas have unique challenges when reporting performance, and both have experienced what Neely calls “measurement madness” or society’s obsession with performance measurement (2002).

This chapter will trace the evolution of performance measurement—first briefly in business, and then more extensively in the health care industry. The business aspect provides some of the general concepts of performance measurement, which form the basis of health care measurement. However, the adoption of these concepts by the health care industry was slow, and more often than not, quality and performance measures development efforts arose sporadically, and often independently of each other, within the field.

Performance measures and their uses in both areas have changed over time. In business, performance measures were initially used for productivity management, then for quality improvement and benchmarking, and finally evolved into the multifaceted approaches used today to assess overall organizational performance from a variety of perspectives, such as financial, internal business, customer and innovation and learning.³⁹ In health care, measurement has evolved from its original purpose, i.e., exclusive use as an internal quality improvement tool, to include its current multiple uses—credentialing and accreditation, accountability, public reporting and pay-for-performance.

³⁹ These are linked elements of the “Balanced Scorecard” developed by Kaplan and Norton (see Kaplan and Norton, 1992).

To help frame the discussion, ANT can be used to follow non-human actors, i.e., performance measures and performance measurement throughout their interactions with other actors in the network. As stated previously, one of the advantages of ANT is its ability to encompass transformation and change. An actor can at any moment redefine its identity and relationships within a network when new elements (or relationships) are brought into the network (Goodman & Walsh, 2001). As actors, human and non-human interact; they become inscribed by all the elements of the network, and, through the inscription process, become a product of what has gone before. Yet, the actors of the network are in a state of constant flux and they continually group and regroup to form new practices and entities. The actors in the network, in the ANT tradition, can consist of a variety of entities: organizations, computers, algorithms and risk adjustment methodologies, reports, managers, statisticians, spreadsheets, and the like. For example, as a quality improvement nurse, I interact with other actors within the hospital network. I run a randomized program at my computer terminal to select medical records for auditing and data collection. I pull the records, review their text, extract the data of interest and put it into a spreadsheet. I run my performance algorithms and risk adjustment models, and develop trend lines, which I analyze. I develop a report, which I send to the hospital administrators, who will present it to the hospital board. The people on the board will never hear me, or see me, but my interaction with all the other actors, the algorithms and the computer, as well as the medical records personnel and the hospital statistician, help shape and re-shape my identity and the relationships in the network. It helps differentiate my role from that of the medical records staff and other actors. And with the introduction of other actors into the network, different relationships form and old ones may or may not re-form which in turn, may change my role or identity as a network actor. For example, if additional performance data demands are made, the personnel office may decide to hire an additional nurse for data

collection, thus changing my role and my interactions. Hence, all the actors within a particular network interact; this in turn, shapes the way they act. My network is organized in such a way as to define my role as well as to produce a “scientific” product, i.e., the hospital performance report.

Likewise, the network that encompasses measuring quality and performance of hospitals has a particular order: the reader will see that the actors of the network are organized in such a way as to produce a “heterogeneous scientific product,” which in this case is a core set of performance measures. The reader should also note that the network ordering also provides clues as to how performance measures are viewed and used within a particular industry. Therefore, while reading the history of performance measures, the reader should be attuned to the different actors and the nature of their interactions and the changes that occur over time; these are reflected in the changing role of performance measures and performance measurement.

Additionally, by following the history and evolution of performance measurement, it becomes clear that politics often play a role, especially in determining why certain movements took hold while others did not. Yet, very little is written in the scholarly literature about the politics of performance measurement, particularly from an ANT perspective. While there is anecdotal evidence of politics influencing how performance is measured, there has not been an in-depth study about this subject in health care.

The History of Performance Measurement—the Business Perspective

Age of Business Productivity

The first performance measures were developed over 100 years ago. In the early 1900s, information about a firm's performance was rarely disclosed, to prevent public intrusion into business operations.⁴⁰ At this time, performance measures were largely based on productivity management (Burgess 1990; Kendrick 1984; Sink 1985) and product enhancement research (Bicheno, 1989). Managers often relied on incomplete or single measures of productivity (Craig and Harris 1973; Mundel 1987) for decision-making. The use of productivity measures continues into the present day; however those within industry have largely dismissed the idea that a single measure can be used to determine overall organizational performance.

Age of Quality

Beginning in the early 1980s, managers across a broad range of industries became interested in the concept of quality (Scott and Cole, 2000), due to the competitive threat posed by Japanese firms. In response, both the United States and European nations sought to change their manufacturing practices and philosophies, which required modification of traditional performance measures (Dixon, et al, 1990). As a result, the first performance measurement crisis, “measurement myopia,”⁴¹ occurred. Based on Skinner's manufacturing performance research (1969), a stream of literature ensued that related quality, time, cost, and flexibility with organizational processes and performance (Garvin 1987; Stalk 1988; Gerwin 1987; Slack 1983). What emerged was the benchmarking concept that was popularized by Camp in 1989.

⁴⁰ The U.S. Steel Corporation broke the secrecy policy in an attempt to ward off anti-trust and restraint of trade allegations in 1905. To assuage public opinion and to depict the company as socially accountable, U.S. Steel made public their performance audit results. However, this disclosure was a rarity among firms at the time.

⁴¹ Measurement myopia can be defined as measuring the “wrong things” (Neely and Austin, 2002).

Benchmarking or the “comparison of organizations to the best of their kind or to standards established by knowledgeable professionals” (Ammons 1995: 16) remains popular today and is often used in combination with performance measurement for quality improvement purposes.

Age of Globalization

Today, “measurement madness” or society’s obsession with performance measurement shows no signs of abating (Neely and Austin, 2002; Meyer 1994). There is a proliferation of measures and measurement systems as well as exponential growth in companies that provide these services. As a result, there is confusion over what is of value to organizations, often resulting in misplaced priorities. Still, performance measurement is becoming more complicated due to new technologies (e.g., the Internet), new environmental arrangements (e.g., globalization), and new organizational structures (e.g., partnerships, alliances and networks). A single approach to performance measurement is not conducive to survival or profitability in an increasingly global marketplace.

The History of Performance Measurement—the Healthcare Perspective

While measurement of quality has been around for at least 250 years (Loeb, 2004), it was not widely used in health care facilities until the 1950s with the formation of the Joint Commission for the Accreditation of Hospitals (JCAH). This delay in measuring care quality can be attributed to several factors: the assumption of high quality of care, the implied insult to physicians, public discomfort with performance measurement, and the statistical lens through which performance is routinely examined (Eddy 1998: 8). The subject of measurement is never neutral. Health care, by its very nature, has a variety of stakeholders, making measurements complicated and often at cross-purposes. Questions of what to measure, which measures to

choose, and determining what the measures will be used for are a significant source of debate and negotiation, making measurement even more challenging. Furthermore, the demand for performance information about hospitals and health care providers has produced a “proliferation of measures” supported by a burgeoning performance measurement and quality improvement industry. Providers are under more and more scrutiny by a number of organizations requesting information on a variety of measures that are often mutually exclusive. Because the electronic health record is not yet a standard in the industry and is most likely years away, performance data is collected via a process called “abstraction,” which requires nursing personnel or other trained staff to gather the information manually from the patient’s paper medical record. The Joint Commission estimates that the time burden to extract the data required to report on the “core measures” currently reported on a government website is approximately 22-27 minutes per chart for conditions such as acute myocardial infarction, heart failure, and pneumonia⁴² and costs anywhere from \$25 to \$47 per reported measure per quarter (Booz, Allen and Hamilton, 2006:5). Beside abstraction costs, other factors associated with the burden of collecting and reporting performance include the costs associated with analysis of data quality and integrity, the application of risk adjustment, data and trend analysis and reporting of health care data. Over the years, performance measures in health care have evolved, becoming more sophisticated. Its uses have also changed as more and more, the demands for provider accountability by purchasers and regulators are heard by the political leadership on Capitol Hill.

What follows is a brief history of performance measurement in health care, its uses and other significant themes that emerged from my research. This history is not meant to be all-

⁴² JCAHO. (1999). *A Comprehensive Review of Development and Testing for National Implementation of Hospital Core Measures*. Last accessed on August 17, 2007.

inclusive, but to give the reader an idea of how and why performance measures have changed over time.

The Early Years (1750-1910):

Performance Measures as a Tool for Tracking Disease and Performance

Measuring the quality of health care can be traced back to the 1700s, with the establishment of hospitals in the colonies. In the United States, the development and history of hospitals essentially followed that of hospitals in Europe, but within a shorter time frame. Hospitals, founded in this country in the early 17th century, were originally almshouses that served the poor, the aged, orphaned, the insane, the ill and the debilitated (Starr, 1982). They only occasionally provided care for the sick. As hospitals in America became more widespread, they changed their focus from care-taker of the dependant to serving the sick, although limiting their services to the poor or lower classes. The earliest effort to collect information about the care provided in these early hospitals was the collection and categorization of patient outcomes by diagnosis at the Pennsylvania Hospital in 1754. In an effort to provide hospitals with a more attractive identity, hospital managers and physicians began excluding those with contagious illness, incurable patients, and those with chronic illness, thus only accepting patients that could be cured. This practice kept the patients admitted to a reasonable level, as well as reduced the mortality rate, and in effect served to “combat the traditional image of the institution as a house of death” (Starr, 1982: 151).⁴³

Another early effort at tracking and measuring health care quality was initiated by Florence Nightingale in the 1860s. She had tracked the mortality rates in British military

⁴³ Excluding incurable patients is a form of risk adjustment, which is used today to “level the playing field” among hospitals in an effort not to penalize those hospitals that treat the sickest patients. Selecting patients that can be cured is a form of “cherry picking” or “cream skimming”, which is a gaming strategy and serves to make the hospital look better when compared to other hospitals.

hospitals in Scutari during the Crimean War (1854-1856), and through her efforts of improved nutrition, sanitation, and infection control, mortality rates were reduced from 40% to only 2% (Starr, 1982: 154). As the Civil War began, military hospital authorities took note of what Nightingale had learned in Scutari, and applied it to a system of over 130,000 beds in military hospitals. The application of the principles of cleanliness and ventilation resulted in a mortality rate of only 8% in Union hospitals, which treated over 1 million soldiers in one year (Starr, 1982).

As if to foreshadow the consumer involvement of the 20th century, a movement to reform hospitals came from a group of upper-class women in New York in 1872. These women founded the State Charities Aid Association, and formed a committee to “monitor the conduct of public hospitals...” (Starr, 1982: 155). The committee consisting mostly of women, found deplorable conditions in hospitals and appealed to men of their own class to rectify the situation. Thus, the quality of care did not originate with physicians but with those “outsiders” who saw a need and sought to reform the hospitals.

In the early days of hospitals, performance measurement was not a very high priority, as demonstrated by the limited efforts cited here. Concerns centered on getting the needed care to those who required it, and the quality of care, while important, was not a main concern of hospital personnel. Any efforts to improve hospital quality were often from outside the health care industry.

The Modern Era and New Delivery Systems (1910-1950):

The Development of Structural Performance Measures as Tools for Quality Assurance

Over the next 30 years, advances were made in both medical knowledge and hospital organization; however performance measurement did not emerge as a practical tool to gauge quality of care until 1910. The earliest significant effort to evaluate the quality of care in the United States is attributed to Ernest A. Codman of Boston's Massachusetts General Hospital, who developed the "end result system of hospital standardization" which can be classified as outcome measures. This system tracked hospital patients after discharge to determine treatment effectiveness (McIntyre et al, 2001; JCAHO, 2003). At about the same time, the Flexner Report was made public. This report evaluated the state of medical schools and focused on the quality of medical education provided. Flexner used "common sense" indicators that included availability of books, lab equipment, and specimens (Colton, 2000) to determine the adequacy of physician education. As a result of the findings of this report, many medical schools were closed and there was a call for the development of standards with regard to physician education and preparation. The effects of this report reached far beyond physician education; it also stimulated standards development in other areas of health care, specifically hospitals.

In 1912, the Clinical Congress of Surgeons of North America passed a resolution at its annual meeting that "some system of standardization of hospital equipment and hospital work should be developed." And in 1913, the American College of Surgeons (ACS) resolved to develop standards of hospital construction, administration and equipment (also known as structural measures), as well as establish a "structured examination of surgical practice." While Codman's system became a "stated objective" of the ACS, it was not the basis for hospital standardization. Instead, Franklin Martin, the founder of the ACS along with John Bowman, developed a new set of objectives which eventually became the "Minimum Standard for

Hospitals” that was used for site evaluations beginning in 1918 (JCAHO, 2003). These standards,⁴⁴ compiled on a single page, focused more on the quality of individual physicians, controlling for things such as recruitment, certification and interaction (Brennan and Berwick, 1996) and on the “functional performance measures,” such as structure of records and procedures (Berki 1972; Smith and Kaluzny 1986) rather than on system-wide outcomes. The first accreditation survey results were reported at an ACS conference at New York’s Waldorf Astoria Hotel on October 24, 1919 by Director J.G. Bowman. As reported, of the 692 hospitals surveyed, only 89 met the ACS Minimum Standard. Although the statistics were reported in aggregate publicly, the names of the hospitals were not disclosed. In an effort to assure that the identities of the non-compliant hospitals, which included some of the leading facilities of the time, would not be disclosed, the ACS leaders met in the furnace room of the hotel, and burned the original reports (Roberts et al., 1987). Although the reports were incinerated, the ACS did not retreat from its commitment. Its members continued to work on improving the quality of hospital care, albeit in more discreet ways. According to Brennan and Berwick, the efforts of the ACS represented an important attempt to “link intellectually grounded forms of standard setting and audit with the actual conduct of day-to-day care” (1996: 100).

Another significant event that occurred in this period was the passage of the Hospital Survey and Construction Act (PL 79-725), commonly known as the Hill-Burton Act of 1946. This law provided national support for the development of rural and community hospitals, and provided standards for construction and regional planning. This not only expanded the number of hospitals in the United States but also contained a provision for a “community service

⁴⁴ These standards called for organizing hospital medical staffs; limiting staff membership to well-educated, competent, and licensed physicians and surgeons; framing rules and regulations to ensure regular staff meetings and clinical review; keeping medical records that included the history, physician examination, and laboratory results; and establishing supervised diagnostic and treatment facilities such as clinical laboratories and radiology departments (Roberts, et al., 1987: 936-40).

obligation.” As hospitals began charging for their services, it became clear that the poor and dependent would not be able to pay for their hospital care. In an effort to provide for the disenfranchised, the government intervened and provided funding for hospital construction with the stipulation that a certain percentage of acute hospital care be provided to the poor at no charge. Faced with the increasing number of hospitals, the ACS was confronted with a significant increase in the number of institutions requiring review and evaluation. And, based on these resource demands, the ACS, in 1951, sought to form a non-profit collaborative to continue the hospital reviews.

During this era, performance measurement, albeit only structural measures, began to emerge as a tool for recognizing quality, or the lack thereof, in the hospital setting. The Flexner Report and the efforts of the ACS to publicly disclose inadequate physician preparation and hospital performance highlighted the need for standardizing certain aspects of care. With the development of structural measures, the role of performance measurement was strengthened, but remained limited with regard to health care quality improvement and it only played a marginal role in the day-to-day operations of hospitals.

The Golden Years (1950-1980): Performance Measures as Tools for Accountability

In the 1950's, Paul A. Lembcke developed a medical auditing system that used “scientific methods.” It emphasized the need for “explicit and objective measures of quality” that used state-of-the-art methods (Ostrow: 1983: 24). With the publication of Lembcke's ideas, a change occurred within the hospital accrediting organization—a collaborative was formed that consisted of the American Hospital Association, the American Medical Association, the American College of Physicians, and the Canadian Medical Association, which eventually formed the JCAH in 1951. The JCAH became the predominant performance measurement organization for hospitals, and in 1953 began offering accreditation based on its Standards for Hospitals using the Minimum

Standards developed by the ACS. The program offered was private and voluntary, and was basically free to hospitals prior to 1964, when JCAH began charging accreditation fees. And with the passage of the Social Security Amendments in 1965, JCAH's dominance of the performance measurement field was assured, due to a provision requiring hospital accreditation for participation in Medicare and Medicaid programs. The JCAH program, partly because it was voluntary and private, became the preferred mechanism to demonstrate compliance with quality standards required by the federal government. Hospitals, after meeting the standards, could apply for reimbursement from the government for treating Medicare, and Medicaid beneficiaries.

Also as a result of the 1965 legislation, in an effort to exert some control over expenditures, all hospitals were required to establish and maintain a utilization review board. These boards were composed of physicians, and were responsible for assuring that the services rendered by the hospital were reasonable and necessary.⁴⁵ These committees had no formal evaluation criteria for decision making, could not deny payment, and essentially had no incentive to be effective (Starr, 1982).

In 1966, the JCAH changed its approach to the accreditation of hospitals, and adopted the "optimal achievable standards" methodology. This was a result, in part, of more sophisticated techniques used to assess quality, as well as Medicare implementing more rigorous guidelines as part of its conditions of participation.⁴⁶ These refinements were largely based on Avedis Donabedian's seminal 1966 article entitled, "Evaluating the Quality of Medical Care." In this work, Donabedian suggests that quality be measured in the areas of structure, process and

⁴⁵ The development of the utilization review boards was intended to both detect and deter misuse of Medicare funds. However, despite the existence of these boards, the incidence of Medicare fund abuse grew over the next 10 years, leading to additional legislation in 1972.

⁴⁶ This change in JCAH policy occurred primarily for three reasons: 1) most American hospitals were already meeting the Minimum Standards; 2) Medicare set more rigorous guidelines, creating an obligation to respond; and 3) the techniques used to assess and improve quality had grown more and more sophisticated (Luce et al., 1994: 264).

outcomes.⁴⁷ Ultimately the JCAH, as well as other organizations, adopted Donabedian's model, which is currently in use today. In 1969, the JCAH made further refinements to its accreditation program to stay abreast of the changing field of health care. The refinements included the creation of four councils that were charged with developing standards and survey accreditation procedures, which were later published in the *Accreditation Manual for Hospitals*. And for the first time in the history of the JCAH, registered nurses and hospital administrators joined physicians in conducting accreditation surveys of hospitals.

In 1972, Congress passed legislation replacing the utilization review boards of hospitals with the Professional Standards Review Organizations (PSRO), which were made up of physicians only (but could not be state medical societies). These entities were established as a type of performance measurement system. Their charge was to monitor the need and quality of care provided to those enrolled in federal health programs. The PSROs had the power to deny Medicare payment, if the services rendered were determined to be unnecessary. As expected, the AMA strongly objected to these review boards, and ultimately succeeded in modifying their charge. As a result of their political maneuvering, the federal government could not own the data collected in these reviews, the scope of the PSRO was limited to inpatient services, preadmission certification for elective surgery would become voluntary instead of mandatory, and only physicians would be involved in the decision making of the PSRO (Starr, 1982). The AMA objected to these organizations because they considered the PSRO a government intrusion into medical practice, while political liberals objected to the exclusion of consumer participation in the program. This dissent surrounding the development and operations of the PSROs ultimately led to the discarding of national norms (Starr, 1982).

⁴⁷ Structure refers to the physical and personnel staffing uniqueness of patient care; process refers to care delivery methods; outcomes are the results of the care provided.

The PSRO legislation also required the Secretary of the Department of Health and Human Services to validate the surveys conducted by the JCAH for those hospitals participating in Medicare, conduct surveys based on alleged complaints of noncompliance with Medicare standards by accredited hospitals, and establish more rigorous standards than JCAH if the standards endorsed by the accrediting organization were deficient. The law also required an annual review and evaluation of the JCAH, which was to be included in the Secretary's Departmental report sent yearly to Congress.

At the close of the decade, the JCAH made some additional changes, including acquiring a fifth corporate member, the American Dental Association. They also abandoned the Accreditation Councils established previously, replacing them with several Professional and Technical Advisory Committees, one for each accreditation program offered by the JCAH.

The Age of Information and Consumerism (1980-2000):

Performance Measures as Tools for Quality Improvement and Public Reporting

Beginning in the 1980s and continuing into the present day, the quality of goods and services provided within the United States became a major issue in a broad number of fields across the country. Health care was no exception. With the passage of the Tax Equity and Fiscal Responsibility Act (TEFRA) in 1982, the controversial PSRO program was replaced with utilization and quality control peer review organizations (PROs).⁴⁸ These organizations were responsible for the efficient and effective delivery of quality health care services to Medicare participants. Specifically, they were responsible for “validating assignments to diagnosis-related groups (DRGs),⁴⁹ reviewing readmissions, reducing unnecessary hospital admission and

⁴⁸ Peer Review Improvement Act, title I, subtitle C of the Tax Equity and Fiscal Responsibility Act of 1982 [P.L. 97-248].

⁴⁹ This is a system designed to classify hospital cases into one of approximately 500 groups. These groups are expected to have similar resource use, and were developed as part of the Medicare prospective payment system. Currently Medicare uses this system to pay for inpatient hospital care. The groupings are based on diagnoses,

operations, and lowering death and complication rates” (Luce et al., 1994: 265). A random selection of medical records was evaluated, based on six criteria:

- The adequacy of discharge planning,
- Medical stability at discharge,
- Unexpected deaths,
- Nosocomial⁵⁰ infections,
- Unscheduled returns to surgery, and
- Trauma suffered in the hospital (Luce et al., 1994: 265).

Problems identified by PROs could be remedied by “formally notifying an institution or practitioner, requiring continuing medical education, preadmission or more thorough retrospective reviews, referral to medical staff committees, informing licensing and accrediting

procedures, age, gender, and the presence of complications or comorbidities (AHRQ website, accessed September 30, 2006).

⁵⁰ Nosocomial infections are those infections acquired during a hospital stay.

bodies, and sanctions, imposed only by the Inspector General, such as loss of Medicare billing privileges” (Luce et al., 1994: 265). However, increasing problems and dissatisfaction with the program prompted Congress to request a study into the effectiveness of the PROs. In the reports to Congress, the IOM and the Health Care Financing Administration (HCFA)⁵¹ indicated that the PROs were ineffective because of limited scope, and called for their restructuring. By the close of the decade, a new program, in addition to the existing PRO program, was implemented—the Medicare Program to Assure Quality. This program updated the conditions of participation, encouraging the modernization of quality improvement methods used by the JCAH, while the existing PROs would look specifically at outcomes related to facilities and clinician performance.

In the mid-1980s, the concept of total quality management (TQM) was shown to be applicable to the health care industry (Berwick, et al., 1990). TQM could improve health care processes, which would potentially lead to better outcomes and reduce costs. However, hospitals have implemented these types of programs very slowly, due to the feeling that “outsiders are telling hospital employees how to do their jobs” (Brashier et al., 1996: 31). Hospitals seem only to accept TQM techniques either “when they perceive a need for a competitive advantage, or as they realize that they must change to meet future demands” (Brennan and Berwick, 1996: 309-310). The application of TQM to the health care industry signaled a move away from retrospective review of process and outcomes, sometimes known as quality assurance, to a more proactive analysis, known as quality improvement (Colton, 2000: 8). However, health care organizations that used this technique for the first time often applied it only to administrative services, with the intent to eliminate unprofitable services, increase efficiency, improve outcomes and, as a result, increase profits.

⁵¹ This agency is now known as the Center for Medicare & Medicaid Services (CMS).

The strongest impetus for applying TQM, i.e., adopting the quality improvement paradigm, was the JCAH, response to what was called a “crisis in confidence”. The *Wall Street Journal* ran a series of articles that were highly critical of the JCAH’s survey process. Due to both external and internal pressures in 1987, the JCAH changed its name to the Joint Commission for the Accreditation of Healthcare Organizations (JCAHO), to better reflect the scope of organizations that were now accredited. It also announced a new project—the *Agenda for Change*. This project heralded a new design to modernize the accreditation processes. The design included three aspects or “planks” that included creating the standards by which hospitals would be accredited, creating a new survey process that was more patient-centered and more reproducible and with less inter-rater variability, and linking performance measurement to a contemporary accreditation process. “The new standards introduced for improving organizational performance were essentially a manifesto for the industrial model of quality improvement” (Colton, 2000: 35). At about the same time, JCAHO field-tested several sets of standardized measures in a variety of areas, and announced its intent to incorporate them into their accreditation program. The idea was to require accredited hospitals to collect and submit performance data to JCAHO to maintain their accreditation.

The Department of Health and Human Services (DHHS) has also played an important role in standardizing quality measurement and quality reporting. In 1986, HCFA issued report cards comparing hospital mortality rates. However, these early efforts were discarded, due to the strong opposition by the hospital industry and others.

Other significant events at the close of the decade included the signing into law of the Omnibus Budget Reconciliation Act of 1989 (PL 101-239). This law addresses the quality issue, in that it recommended that Medicare’s payment schedule, based on “customary, prevailing, and reasonable” charges, be replaced with one that is based on resource costs related to a “relative

value scale.” The thought was that changing the payment schedule affects practice patterns which, in turn, influence not only costs but quality as well. In addition, the law also provided federal support for health services research activities. Under its auspices, Congress established a new federal agency—the Agency for Health Care Policy and Research (AHCPR)⁵² to enhance the quality, application and efficiency of health care services, as well as to improve access to health care services. In effect, establishing the AHCPR highlighted the federal government’s interest in improving outcomes of care and treatments by supporting research in the health services and supporting the development and use of practice guidelines.

The 1990s saw multiple efforts to measure the quality of health care. Data from these efforts provide information for purchasers and consumers to help evaluate the value of the care received. Further, as quality measures began being used for purposes other than internal quality improvement, quality measurement became more and more sophisticated. In the 1990s, quality measures began developing and incorporating methods to adjust for severity of illness in order to provide meaningful comparisons of the quality of health care facilities. One measure set that uses these methods is the Health Plan Employer Data and Information Set (HEDIS). Originally released in 1989 by a coalition of health plans and large employers, it was refined and adopted by the National Committee on Quality Assurance (NCQA), which incorporated it into its accreditation program. Any health plan that sought accreditation through the NCQA was required to report on the HEDIS measures. By mid-1990 the public purchasers were requiring the use of these measures to participate in the Medicare program, and the information gleaned from them was subsequently made available to the public. An important area included as a compliment to the HEDIS measures was the addition of consumer satisfaction measures. Known originally as the Consumer Assessment of Health Plans Survey (CAHPS), these measures were

⁵² This agency is now known as the Agency for Healthcare Research and Quality (AHRQ).

designed to capture the consumer's experiences of the care they received, and were incorporated into the NCQA accreditation program in 1997. The CAHPS surveys evolved from measuring patients' satisfaction with their health plan to measuring their experiences in various care settings, such as hospitals, nursing homes, and hemodialysis treatment centers.⁵³

In the hospital arena, HCFA modified its conditions of participation that had closely tracked the JCAHO standards instituted under the *Agenda for Change* program, signaling the need to begin measure alignment or harmonization in 1994. At the close of the decade, JCAHO initiated a second effort to standardize performance measures and collect performance data for public reporting. It instituted a system that allowed organizations seeking accreditation to select measures whose specifications were standardized, to satisfy accreditation requirements. This was a prelude to the introduction of a standardized set of core measures, known as ORYX, in 2002.

Throughout the 1990s, there were various efforts to inform policy decisions regarding the quality of health care across the nation. In 1996, President Clinton established the Advisory Commission on Consumer Protection and Quality in the Health Care Industry, to "advise the President on changes occurring in the health care system and recommend such measures as may be necessary to promote and assure health care quality and value, and protect consumers and workers in the health care system" (President's Advisory Commission, 1998). In its final report entitled *Quality First: Better Health Care for All Americans*, released in March, 1998, fifty recommendations⁵⁴ were included that were crafted to enhance the quality of health care within the country. Based on this report, a planning committee was convened in June 1998 to create

⁵³ This family of surveys is now known as Consumer Assessment of Health Providers and Services, and retains the same acronym.

⁵⁴ The recommendations included a call for public-private partnerships to improve the quality of care provided in the US, establishing national aims for improvement, establishing a core set of quality measures that could provide easily understood comparable information on industry performance, strengthening the market to improve quality, building capacity to improve quality, and addressing the problem of the uninsured.

what would become known as the National Quality Forum (NQF), the private sector entity that was called for by the President’s Advisory Commission. At about the same time, the IOM issued a statement about the quality of care provided by the health care industry:

“Serious and widespread quality problems exist throughout American medicine. These problems, which may be classified as underuse, overuse, or misuse, occur in small and large communities alike, in all parts of the country, and with approximately equal frequency in managed care and fee-for-service systems of care. Very large numbers of Americans are harmed as a result. Quality of care is the problem, not managed care” (Chassin, Galvin, and the National Roundtable, 1998).

In 1999, the IOM issued the report entitled *To Err is Human: Building a Safer Health System*, which indicated that anywhere from 44,000 to 98,000 people die each year due to medical errors. At the close of the decade, the NQF was established and an NQF Strategic Framework Board was convened to design a national strategy for quality measurement and reporting. This strategy incorporated guiding principles and priorities for a national system, including the roles of key players, as well as identifying barriers to strategy implementation and potential solutions thereof. The IOM report set the stage for the first projects that would be undertaken by the NQF—the Never Events Project and the Hospital Performance Measures Project.

Back to the Future (2000-Present):

Performance Measures as a Tool for Quality Improvement, Accountability, Public Reporting, and Pay-for-Performance

As indicated above, JCAHO established its ORYX system—a core set of standardized measures that would be required of hospitals seeking accreditation, beginning in 2002.

Organizations would have to report and submit data on three of five standardized measures.⁵⁵ Over the next few years, JCAHO intends to require standardized measures for other types of organizations and conditions (JCAHO, 2005).

In 2002, with the emergence of the Hospital Quality Alliance (HQA),⁵⁶ a public-private collaboration of the various stakeholders in health care, hospitals were voluntarily reporting performance data on what was known as a “starter set” of performance measures. Despite the low number of hospitals participating in the initiative, hospital performance data on three conditions—heart attack, heart failure and pneumonia were reported on the DHHS website in October 2003. Based in part on the slow recruitment and reluctance of hospitals to participate in public reporting, Congress included in the Medicare Prescription Drug, Improvement and Modernization Act of 2003 a provision that tied Medicare payment updates to public reporting data on hospital performance from FY 2005 through 2007. Hospitals that did not report the ten measures would receive a deduction of 0.4 percentage points from their Medicare update payments. Other pay-for-performance projects, most notably the CMS⁵⁷-Premier Hospital Quality Improvement Demonstration project, were developed at about the same time. This project, initiated in March 2003, was designed to test the concept of linking payment to the provision of high quality care. With over 270 hospitals participating, this project initially shows that linking quality and payment does result in overall improvement in the care provided. Based on the initial results of public reporting and pay-for-performance initiatives, it appears that quality improvement requires certain incentives—either in the form of recognition or payment.

⁵⁵ The five measures are in the areas of acute myocardial infarction, heart failure, pneumonia, pregnancy and related conditions, and surgical infection prevention (JCAHO, 2005).

⁵⁶ For more information on the HQA, see Chapter 3: Identifying and Describing the Human Actors of the Hospital Quality Network in this dissertation.

⁵⁷ In 2001, the Secretary of Health and Human Services changed the name of this agency. Previously known as HCFA, the name was changed to the Centers for Medicare & Medicaid Services (CMS).

As we close our review of the history of performance measurement, it appears that we have come full circle in the areas of quality measurement and improvement. As stated by Ernest A. Codman in 1917, “for hospitals to improve, they must 1.) find out what their results are; 2) analyze their results to find their strong and weak points; 3) compare their results with those of other hospitals; and 4) welcome publicity not only for their successes, but for their errors...Such opinions will not be eccentric a few years hence.”

The Changing Nature of Performance Measures: Applying Actor-Network Theory

Tracing the history of performance measurement and the various quality initiatives illustrates how the idea of quality has changed over time, as well as how measures themselves and their uses have changed. Interest in tracking outcomes was not a widespread phenomenon in health care, even with evidence provided by Florence Nightingale in the form of lower mortality rates. Measures at that time involved trending and were not risk-adjusted for severity of illness. Further, if a hospital did track its mortality rates, it was considered the hospital’s own business and was not shared with outsiders. Overall, if measurement did occur, it was primarily for trending and not widely used for quality improvement. At this stage, performance measures were relatively few. However, as measures became more and more useful within settings of care, more and more measures were developed for these purposes.

From there, performance measures became more precise. These same measures used for quality improvement, although slightly refined, were then used for credentialing and accreditation. They were used by such organizations as the NCQA and the Joint Commission, as part of the requirements for accreditation or certification. Performance data was often used only as one criterion of many to award accreditation and certification, such as when the NCQA and the Joint Commission accreditation programs were recognized by CMS as part of the conditions

of participation for Medicare and Medicaid payment for providing care to those beneficiaries of the program. At this stage, performance measures were frequently based on structure and processes that needed to exist to provide care, not on the outcomes of care.

In the late 1990's, with the formation of the NQF, a new use for performance measures was instituted on a larger scale. In addition to the uses for quality improvement and credentialing, the NQF brought the use of performance measures for accountability and public reporting into the mainstream. Their first projects addressed the patient safety aspect of care and the quality of hospital care delivered in the U.S. Thus, performance measures had an additional use—comparing hospital quality and publicly reporting hospital performance.

More recently, performance measures have been used to link payment to the provision of high quality of care. Otherwise known as pay-for-performance (P4P), these initiatives demonstrated the various attempts by the federal government to provide incentives to health care facilities to render high quality care. The assumption was that high quality care provides value to patients and that it costs less because it is delivered in an efficient manner. Letting the market forces work, according to some, will eliminate providers that do not practice efficiently or effectively, and thereby improve the overall quality of the health care system.

As illustrated by the history recounted here, performance measures and performance measurement have various uses which translate into multiple identities—as a scientific, evidence-based tool to judge performance; as tool for payment; as a threat to reputations and livelihood; as a tool for quality reporting; as an incentive to provide better care; and as a tool to maintain the status quo. Thus, as measures evolved over time, their interactions with other actors as well as the changing actors and their positions within the network help shape and re-shape their identities.

CHAPTER 3: RESEARCH STUDY METHODOLOGY

For this dissertation I considered several methods of inquiry within the qualitative framework. Qualitative research can be broadly defined as research that “is multi-method in focus (Brewer and Hunter, 1989) and involves an interpretive, naturalistic approach to its subject matter” (Denzin and Lincoln, 1994: 2). Thus, qualitative researchers attempt to study phenomena in a natural setting, without manipulation of variables (Caudle, 1994). Furthermore, qualitative researchers focus on the socially constructed nature of reality, the relationship between researcher and subject, and the situational context that directs the inquiry (Denzin and Lincoln, 1994). Other assumptions include a philosophical belief in the phenomenological paradigm, i.e., an appreciation of naturalistic inquiry, qualitative methods, inductive analysis, purposeful sampling, and holistic thinking (Patton, 1998).

Research questions, particularly those that ask *how* and *what* rather than *why* suggest that qualitative methods should be considered. *How* and *what* questions provide a description of what is going on, whereas *why* questions tend to suggest causal relationships among phenomena (Creswell, 1998). A qualitative design should be selected if variables are not easily identifiable (Marshall and Rossman, 1999), theories do not explain participant behavior or need to be developed, if a detailed view of the issue is required, or if the nature of the question requires studying individuals in their natural setting (Creswell, 1998). For this dissertation, I chose to study the NQF Hospital Measures Steering Committee, which I consider as a microcosm of the larger Hospital Quality Network (HQN). The HQN and the actions surrounding the endorsement of the nation’s first core set of performance measures for hospitals are large and complex; key variables were difficult to identify and required observation of the individual participants in their natural setting.

The advantages to conducting qualitative research are substantial, especially in this context. The flexibility of the research design provides the opportunity to gather additional information and sometimes different kinds of information, because the researcher is not constrained by pre-determined research categories (Patton, 1987). This approach allows a certain richness of examination that may not be captured in quantitative research designs. For these reasons, I selected a qualitative research design to study the NQF's Hospital Steering Committee and the network that emerged around the design of a national hospital performance measurement system; I was particularly interested in the role of federal agencies, specifically AHRQ and CMS, in this process. My research questions ask *how* and *what*, seeking to provide a description of what the network looks like, how the actors relate to one another, and how the Hospital Steering Committee operates, by tracing the selection and endorsement process within the changing health care environment.

The Case Study Method of Inquiry

A case study contributes to the in-depth understanding of the individuals, organizations and other social phenomena that occur in a complex environment. It permits the “holistic and meaningful characteristics of real-life events” (Yin, 1989: 14). Stake (1995) identifies three different types of case studies: the intrinsic case study,⁵⁸ the instrumental case study,⁵⁹ and the collective case study.⁶⁰ To study a case, data should be gathered on the nature of the case, its historical background, the physical setting (as well as other contexts including economic,

⁵⁸ When the case itself is of primary interest to the research, that is, the case is a “given” (Stake 1995).

⁵⁹ Involves research of a case to gain understanding of something else, such as an issue or theory refinement (Stake, 1998, 1995).

⁶⁰ Involves examining several cases of a particular project (Stake, 1995).

political, legal, and aesthetic), other cases through which the case is recognized, and those key informants through whom the case can be known (Stake, 1998: 88).

Based on the nature of my research questions, I selected as the research design an instrumental case study that is theoretically informed, because issues in performance measurement system design need refinement at the national level, e.g., how performance measures are selected for NQF endorsement and the federal government's role in this process. Through the use of the instrumental case study, reasonable generalizations can be drawn from my research, such as recommendations about designing performance measurement systems for other health care facilities within the United States, and potentially for hospital performance measurement systems located in other countries. This case study is theoretically informed, using institutional rational choice theory to identify the actors within the network, and to suggest their reasons for participation in the network. Actor-network theory is employed to follow a particular actor or set of actors as they interact and react within the network. In this case, I follow both human and non-human actors, specifically performance measures undergoing the consensus development process (CDP) for endorsement, the CDP itself, and the NQF Hospital Steering Committee over the course of its deliberations. Using this theory, I examine the actors at the beginning of the process, and document changes as the consensus building and measure selection proceed to the end product, a performance measurement system that can be used at a national level.

Data Collection Methods

I employed a variety of data collection strategies for this study. Data collection included in-depth, open-ended interviews with key informants⁶¹ from the performance measurement network described further below, the examination of pertinent organizational and committee documents and archival records, direct observation, and participant observation.

I conducted 30 in-depth interviews of key informants⁶² including: NQF staff, the Hospital Steering Committee members, and representatives from other public and private agencies that participated in the selection of performance measures for hospitals such as AHRQ, CMS, AHA and the Joint Commission. Individuals who were interviewed include directors and project managers, executive directors, advisory board members for the project, individuals from the AHA, and other stakeholders, such as representatives of consumer groups and purchasers. Individuals were selected based on their participation in the NQF Hospital Steering Committee, interviewee availability, and NQF Council designation. Additional contacts were identified by using the snowball or chain sampling technique, in which interviewees provided the names of others who could give additional insight into processes and issues (Caudle, 1994).

The interviews were audio taped, and were completed in approximately 60 minutes. At the time of the interviews, I also requested additional interviewee time for follow-up questions or clarifications, which on average took about 30-45 minutes, and took place after the interview. After the completion of each interview, I recorded my impressions and observations in my field notes, which were included in my analysis. The interviews were transcribed, and the

⁶¹ Key informants were provided with a description of the research project and an informed consent document that advised them of their right to refuse participation and their ability to withdraw from the study at any time. A copy of the consent form is included in Appendix K. IRB Expedited Approval was sought from the VA Tech University Institutional Review Board and was granted on October 13, 2004 (IRB# 04-461).

⁶² A copy of the questionnaire used to interview the key informants of this study is in Appendix L.

interviewees were given an opportunity to review their transcripts. I then identified key issues and themes and coded them using N6 software for qualitative data analysis.

Documents from both organizations and committee meetings were obtained throughout the interview process. Specifically, I collected and analyzed documents related to NQF's Hospital Performance Measures Project, including working papers, minutes from Committee meetings, logs, announcements, formal policy statements, briefing materials, letters and memoranda. In addition, I examined documents located on the NQF website, as well as publications, newsletters and press releases distributed by the NQF and other agencies and organizations involved in the project. I was able to review additional documents that were housed at AHRQ during my intern rotation at the agency. These documents included project summaries, statements of work, voting and comment documents, letters and memos, and notes. I also examined documents from other sources such as journal and news articles, commentaries, and government reports, such as those issued by the GAO.

I was able to attend several NQF Hospital Committee meetings during my tenure at AHRQ, and was able to obtain meeting materials which I included in my document examination. I used direct observation of the Hospital Steering Committee and its members and recorded my observations in my field notes. Further, I was able to attend several related meetings, such as meetings of the QuiC and the HQA. Again, I was able to record my impressions and observations of the meeting participants in my field notes. All these observations were coded and included in my analysis.

I also had the opportunity to act as a participant observer in the NQF CDP. On behalf of AHRQ, I acted as a liaison to two steering committees and four technical expert panels charged with evaluating additional hospital performance measures. After each meeting or conference call, I recorded my impressions and observations in my field notes. I also made notes about my

role as a federal agency representative for each meeting, and noted how representatives from other federal agencies behaved in relation to the committee members. I had the same opportunity to act as a participant observer in several of the Principals Meetings and several measurement subgroups of the HQA. I employed the same methods that I used for participation in the NQF processes for the HQA and recorded my observations and impressions. Table 2 summaries the data collection methods that were used for each research question.

Table 2: Data Collection Methodology

Questions	Participant Observation	Interviews	Direct Observation	Document/ Archival Records	Artifacts
Role of the Federal Government	X	X	X	X	X
NQF Critical Task		X		X	X
Physician Role		X	X	X	X
NQF interaction with industry	X	X	X	X	X
Influence of the field of performance measurement		X	X	X	X

Sample Description

The population that I studied for this research included all those that participate in the hospital quality enterprise at the national level. From this broader population, I selected 32 potential key informants, based on my selection criteria, and contacted them by e-mail to request their participation. I also included an abstract of my proposed research to inform their decision to participate. Of the 32 individuals contacted, 30 responded to my request to be interviewed, and two did not respond.

Overall, the sample consisted of mostly women, between the ages of 50-59, who had been working in the health care industry or a related industry for the majority of their careers.

The majority possessed Masters degrees, were not clinicians, and held executive positions within their organizations. More often than not, they worked for non-profit organizations or federal agencies.

Appendix M provides graphs and charts of the characteristics of the sample interviewed for this study. These characteristics were collected during the interview process and offer a broad description of the sample, and may provide additional insights into the behavior of individuals with similar traits. For example, are the responses from physicians about the NQF CDP different from nurse responses? Do those with more education respond differently from those with clinical training? Do individual with different NQF Council designations respond differently to the CDP and measure selection? These are a few of the questions that these data elements can address.

Data Analysis

According to Marshall and Rossman “data analysis is the process of bringing order, structure, and interpretation to the mass of collected data” in a research project (1999: 150). While generating categories can help focus the study, tightly structured analysis plans may filter out the “unusual or serendipitous” (Marshall and Rossman, 1999: 150-1). Crabtree and Miller (1992) offer a continuum of analysis strategies that can be used to focus the research study. At one end of the continuum are the “technical, scientific and standardized” strategies, which rely on an objectivist perspective of inquiry and pre-configured categories of meaning. At the other end are the “immersion strategies” which rely on the researcher’s interpretative and intuitive

ability. Between these two extremes lie other analysis strategies, including “template” and “editing” strategies⁶³ (1992: 17-20).

On this continuum, my analysis strategy lay somewhere between the template and editing strategies. After a literature review, and from my preliminary knowledge of the hospital industry, I was able to develop broad categories to begin to organize my results, such as information about the NQF, general information about hospitals, and quality. I gathered several public documents that I found in the literature, on the NQF website, and through my attendance at various hospital quality meetings. I imported the text of these documents into the N6 qualitative software program, and started coding text segments into categories or nodes. For example, based on the documentation I examined, I was able to categorize the information into several free nodes or non-related categories, and tree nodes, or others that could be divided into several related sub-categories within a node. Based on my evaluation of the documents, I refined several of the nodes and added sub-categories to many. Then I conducted five interviews, transcribed them and used the N6 program to begin organizing the text into nodes. Based on these preliminary interviews, I was able to further refine many of the nodes, and eventually expanded them to include seven free nodes⁶⁴ and 56 tree nodes.⁶⁵ I anticipated that other nodes would emerge throughout the data collection and analysis process of my research, and I wanted to retain some flexibility to add categories, to combine categories, or to further sub-divide nodes; thus the process was iterative. After each interview I re-visited the node structure to assess whether or not the structure adequately captured what I was finding in the field. Through this

⁶³ The template strategy relies on code groups or templates that may be revised as the analysis proceeds. The editing strategy is less pre-configured and relies on the researcher’s ability to examine text in order to produce categories of meaning (Crabtree and Miller, 1992: 17-20).

⁶⁴ The free nodes included broad categories such as ANT, Hospital Quality Alliance, Background, Ken Kiser, Politics, Health Care, and Networks.

⁶⁵ The tree node categories included the Role of the Federal Government, NQF Steering Committee, NQF and Hospital Industry Interactions, Performance Measure Selection, Base Data including gender, age group, etc., CMS Role, AHRQ Role, and NQF Role, including subcategories such as Strategic Framework Board, CDP, and Issues.

process, I was able to identify themes, recurring ideas or language, patterns of beliefs, and actor linkages, which led to revisions of categories throughout the proposed study.

I used triangulation⁶⁶ procedures to improve the accuracy and credibility of the study. For example, data from in-depth interviews, document analysis, and observation were used for corroboration and further elaboration of themes and concepts derived from my field research. Further, interviews that went beyond the “usual” sources helped verify data and themes I identified. Other sources of interviews include directors from consumer groups such as the AFL-CIO, ARRP, and the National Partnership for Women and Families, purchasers other than CMS (such as Ford Motor Company and TRICARE), quality improvement organizations (such as the California Institute for Health Systems Performance and the Institute for Healthcare Improvement), and provider organizations such as the American Nurses Association and the American Medical Association. I also used alternative sources of data such as documents produced by other parts of the federal government, for instance, the Office of Management and Budget, the General Accountability Office, Congress, and the courts. These documents provided not only alternative viewpoints, but also support for and confirmation of the data I had already collected. I continually reviewed the data to determine if there were alternative explanations, which also helped in verifying my findings.

Once data collection was complete, and all the information was coded, I was able to compare nodes to determine if any patterns emerged. For example, I compared responses from representatives of the NQF Councils to assess similarities and differences. I also did node searches to find intersections of text within different nodes and overlap between nodes. These techniques, built into the N6 software, gave me a richer view of my research and highlighted the similarities and differences within the sample. I also did cross-tabulations to examine attitudes

⁶⁶ Involves “working to substantiate an interpretation or to clarify its different meanings” (Stake, 1995: 173).

and values of the different groups (e.g., whether physicians had different opinions about the NQF CDP), to determine if a pattern or theme was emerging.

Limitations of the Research

While there are several advantages to using a qualitative research design for this study, there are also some concerns. Usually, qualitative research findings are not generalizable to the larger population, due to limitations in the cases sampled for observation, problems of temporal sampling, and limitations on selectivity of interviewees and document sampling. Yin states that case studies are generalizable “to theoretical propositions and not to populations or universes” (1994: 10). The goal, according to Yin, is “analytic generalization” rather than “statistical generalization.” The former expands and generalizes theories and the latter specifies frequencies (Yin, 1994: 10).

In this study, based on Yin’s analytic generalization argument, the findings can be used to expand performance measurement propositions to include the political aspects of performance measurement within networks of actors. I identified the parameters of the network using Institutional Rational Choice Theory, and from the NQF member and meeting attendee lists. I further defined the network using information obtained from interviews conducted for this study and attendance at other meetings, such as the Hospital Quality Alliance (HQA) meetings. The problem of generalizability to other populations or settings that utilize networks or that can be characterized as collaborative networks can be addressed through the use of multiple sources of data. For this project, I used multiple research techniques and different data sources, such as direct and participant observation, document analysis and in-depth interviews with key informants. Furthermore, key informants provided additional examples from their performance

measurement work with different populations and settings, which were also germane to hospitals. Therefore, certain themes emerged, regardless of the data source.

Other problems associated with qualitative research include reactions of panel members to the presence of the researcher, investigator change, and concerns regarding researcher bias and subjectivity (Patton, 1998). Reactivity problems associated with researcher presence are well documented in the anthropological literature. For my research, the meetings that I attended were open to the public; therefore, I believe that reactivity problems (if any) could be related to public attendance rather than researcher presence. However, such was highly unlikely, as those participating in panel discussions were seasoned, senior level executives who are able to function in these types of public settings.

Investigator changes, or when the observer “goes native” and is absorbed into the local culture or organization, and are of concern in field research, particularly when participant observation is used as a research method. During the course of my study, I had an opportunity to participate in the AHRQ Internship program, and later became an employee of the Agency. Although my relationship with AHRQ gave me access to different meetings (all of which were open to the public with the exception of the QuiC meetings), I did have to pay extra attention to my reactions and impressions during data collection. To counteract these concerns, I was attentive to and recorded my reactions to interviews, informants, and documents in the research record, and when possible used triangulation with information from key informants to validate my impressions.

The predispositions and biases of the researcher also may affect analysis and interpretation. On the one hand, rigorous procedures are aimed at data validation, yet, due to the social construction and interpretative nature, data inevitably represent some degree of personal

perspective rather than absolute truth. To address this issue, the idea of “emphatic neutrality”⁶⁷ suggests that researchers can be seen as impartial within the research process. Again, I paid close attention to my reactions throughout this research project. While I am a clinician, I found during the course of my research that I did not automatically have a provider perspective about performance measurement. I discovered that I, as well as other nurses that I interviewed and observed in meetings, often share the consumer point-of-view when it comes to measures and measurement. Therefore, it was important that I present a balanced view of all perspectives of measurement. In order to accomplish this, I recorded my impressions, indicating whether they were more favorable to one group over another, and why. I tried to validate the arguments through discussions with other meeting attendees, to ascertain whether my impressions were correct or the product of bias. I also recorded these discussions and any counterarguments in the field notes that were coded and incorporated into my analysis.

In summary, the research design contributed to identifying the actors of the Hospital Quality Network, as well as provided information about their interactions within the smaller network of the NQF Hospital Steering Committee. However, to begin to understand each actor’s motivations and the reasoning behind their interactions, a more thorough exploration is required. Chapters 4 and 5 in Part II provide descriptions of both the human and the non-human actors of the network. These descriptions not only provide background information about the actors, but also provide some insight as to why certain actors behave as they do; this will assist in understanding how politics factor into the selection of performance measures for the nation.

⁶⁷ According to Patton, emphatic neutrality is a “stance in which the researcher is perceived as caring about the people under study, but neutral about the findings” (1998: 21).

PART II: MAPPING THE HOSPITAL QUALITY NETWORK

“The coexistence within the U.S. health care system of a wide variety of providers, organizational forms, and funding sources has been viewed by many as a positive attribute that contributes to the rapid diffusion of new technology, the enhancement of quality of care, and the capacity of the system to innovate and adapt to change.”

*National Center for Health Services Research, 1977
(in Litman & Robins, 1997, p. 25).*

Due to the complex, transboundary nature of the challenges facing the health care industry, traditional government-initiated programs are being replaced with newer innovative approaches, such as collaborative networks consisting of the traditional players in health care, as well as the non-traditional ones such as the AFL-CIO and the Chamber of Commerce. Alliances of public and private organizations through network structures are one approach to innovative problem solving. Organizations come together to focus on solving problems, and through an iterative process, come to redefine their relationships which, in turn, adds new elements to the mix that may provide new insights and breakthroughs to complex issues. To describe and map the network that has emerged for improving hospital performance, I used Institutional Rational Choice Theory (IRC) as a starting point to investigate the organizational actors of the network that were identified by interviews with key informants from within the field. I also used Actor-Network Theory (ANT) to classify actors into human and non-human categories. Recall the definition of an actor: an entity, either human or nonhuman, that modifies another entity through mobilizing its activity for some reason over a period of time. The activity or performance the actor conveys is that performance which is in its best interest to convey and in turn determines its competence within a particular situation.

In ANT, actors can be either human or non-human, and their importance is determined by their significance to the maintenance of the network. The actors' relationships are important, since they are consistently being made and remade as different elements of the network come into contact. As actors form connections, the network becomes patterned or organized in a certain way, which results in certain end products that remain stable until new elements are introduced and the order within the network is shuffled again. For example, Law's description of "knowledge" and "science", based on the heterogeneous ordering of materials, is germane to our discussion:

"'Knowledge', then is embodied in a variety of material forms. But where does it come from? The actor-network answer is that it is the end product of a lot of hard work in which heterogeneous bits and pieces—test tubes, reagents, organisms, skilled hands, scanning electron microscopes, radiation monitors, other scientists, articles, computer terminals, and all the rest—that would like to make off on their own are juxtaposed into a patterned network which overcomes their [individual] resistance.⁶⁸ In short, it is a material matter but also a matter of organizing and ordering those materials. So this is the actor-network diagnosis of science: that it is a process of 'heterogeneous engineering' in which bits and pieces from the social, the technical, the conceptual and the textual are fitted together, and so converted (or 'translated') into a set of equally heterogeneous scientific products" (Law 1992a: 2).

In other words, networks constitute an order of heterogeneous materials whose individual resistance to the proposed order has been overcome, or as Law puts it, "in this view the task of sociology is to characterize these networks in their heterogeneity, and explore how it is that they come to be patterned to generate effects like organizations, inequality and power" (Law, 1992a:

⁶⁸ Resistance can be thought of as an actor's self interests or preferences and "overcoming resistance" involves convincing an actor that it is in its best interests to change preferences or inclinations (translation) into a punctualised actor-network.

2). Thus, social life can be viewed as an ordered network of different entities, consisting of human and non-human actors. As a result, we can use the ANT lens to analyze the Hospital Quality Network, or more specifically the microcosm of the NQF Hospital Steering Committee, to provide clues as to why certain standards were favored over others for a national system.

In Chapter 4, the human actors (or in this case, the organizational actors) of the Hospital Quality Network are described in some detail. The Hospital Quality Network consists of several organizations that have come together to improve the quality of care delivered in our nation's hospitals. Chapter 5 continues the mapping of this network by describing the non-human actors, which are considered to be performance measures, public reporting and report cards, and the NQF consensus development process.

CHAPTER 4:

THE HUMAN ACTORS OF THE HOSPITAL QUALITY NETWORK

The network consists of several human actors that were identified through field interviews that were completed for this dissertation. The organizations were identified and placed into one of three tiers, based on their involvement within the network. Their “involvement” and their tier was determined by the number of times an organization was mentioned during interviews with key informants and the frequency with which the organizations’ name appeared in the documents examined for this research (Figure 4).

Organizations in the first tier include: the National Quality Forum, the Centers for Medicare & Medicaid Services, the Agency for Healthcare Research and Quality, the Joint Commission for the Accreditation of Healthcare Organizations, the National Committee on Quality Assurance, and the American Hospital Association. The second tier consists of the American Medical Association, and the Institute of Medicine. The third and final tier consists of the Consumer-Purchaser Disclosure Project, the AFL-CIO, and the Hospital Quality Alliance.⁶⁹ The remaining organizations, the Federation of American Hospitals (FAH), AARP and the American Nurses Association (ANA), while significant actors in the network, were not cited as often as the ones that will be described in this chapter.

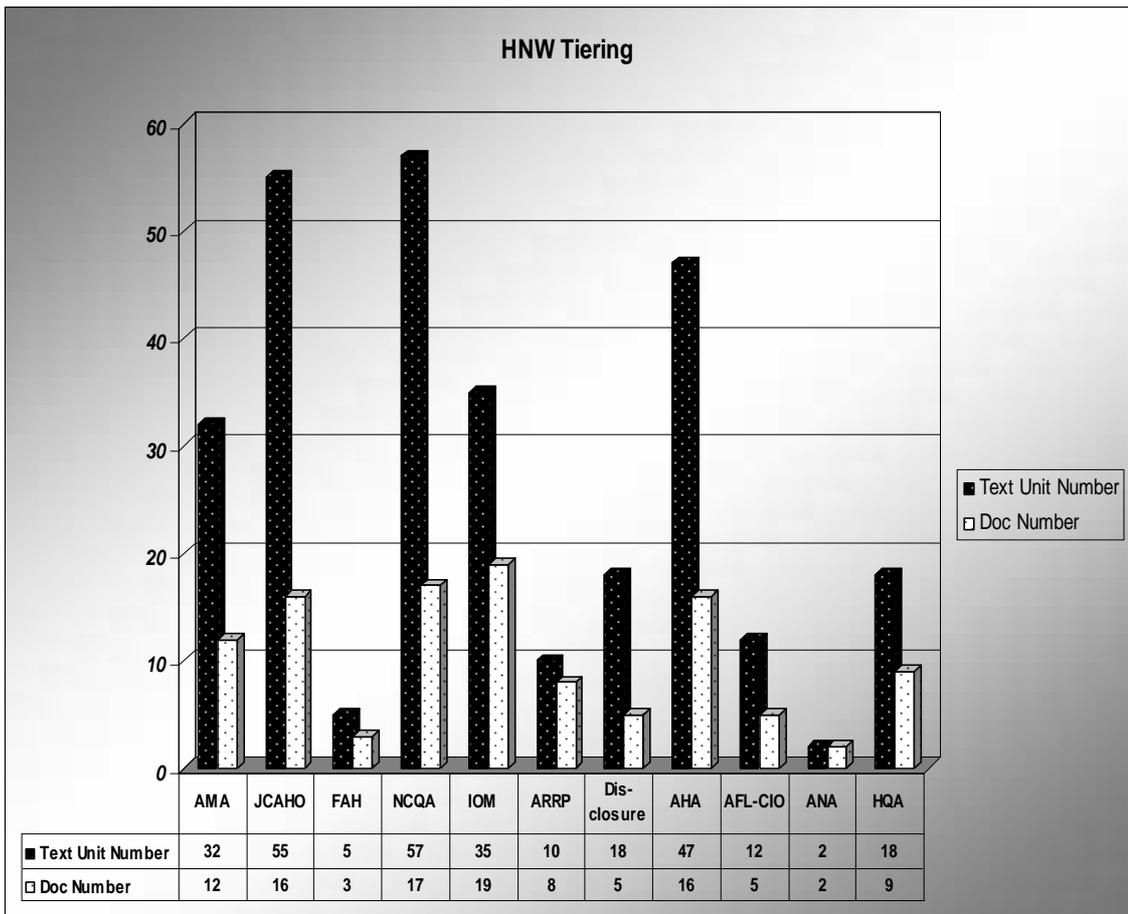
After identifying and tiering the major actors in the network,⁷⁰ Institutional Rational Choice Theory (IRC) was used as a basic framework to describe the participating organizations.

⁶⁹ Previously known as the National Voluntary Hospital Reporting Initiative, the name changed in 2005 to the Hospital Quality Alliance.

⁷⁰ I specifically chose organizations rather than individuals as actors within the network due to the nature of the health care performance measurement field. It is primarily a very small community and in order to maintain anonymity, I have opted to describe the organizations that occupy this space. The individuals interviewed were selected to provide a “punctualised” effect in that they are considered senior leaders and many have long-time associations within their respective organizations, and thus have the ability to influence the direction of the organization, commit resources, and in short, speak for and fully represent their establishments within the network.

Their origins, mission and purpose, governance structure, and the resources available to each organization were explored and documented, which provide some insight into their motivations for participation in the network.

Figure 4: Hospital Network Tiering: Excluding NQF, CMS, and AHRQ⁷¹



Organizational Actors-Tier 1

The National Quality Forum

The National Quality Forum (NQF) is a private, not-for profit, open membership organization that was established to promote consensus around the alignment of performance

⁷¹ NQF, CMS and AHRQ were excluded because they were explicitly incorporated into the study. They were also excluded from the figure for aesthetic reasons and ease of interpretation.

measures, reporting mechanisms, as well as a national strategy for health quality improvement. The organization, incorporated as the National Forum for Health Care Quality Measurement and Reporting, pursuant to section 501(c) (3) of the Internal Revenue Code in May 1999. The original mission that appeared in the NQF's Inaugural Newsletter, issued in Summer 2000, was to "improve healthcare quality by developing a clear, coordinated and coherent approach to measuring and reporting health care quality; by identifying national goals for quality improvement; and by promoting the uniform collection and public dissemination of comparative data" as quoted by Kenneth W. Kizer, M.D., the first President and CEO of the organization.

The origins of the NQF stemmed from the President's Advisory Commission on Consumer Protection and Quality in the Health Care Industry that was established during the Clinton Administration and convened in the late 1990s. Established by Executive Order 13017 on September 5, 1996, and later amended with Executive Order 13040 on March 26, 1997, the Commission was created to focus on the concerns that surfaced about the quality of care provided within the United States Health Care System.⁷² What was needed was a national commitment to quality improvement, not only from the President, but also from Congress. "The purpose of the health care system must be to continuously reduce the impact and burden of illness, injury and disability, and to improve the health and functioning of the people of the United States" (President's Advisory Commission Executive Summary, 1998: 1). The Commission, co-chaired by Donna E. Shalala, the Secretary of Health and Human Services, and Alexis M. Herman, the Secretary of Labor, consisted of 34 prominent individuals who represented the various and diverse constituencies of the health care industry.⁷³ The Commission

⁷² The President's Commission was governed by Public Law 92-463 as amended, the provisions of which set forth the standards for the formation and use of advisory committees.

⁷³ Representation on the Commission was from consumers, business, labor, health care providers, health plans, State and local governments, and health care quality experts.

was charged to “advise the President on changes occurring in the health care system and, where appropriate, to make recommendations on how best to promote and assure consumer protection and health care quality” (President’s Advisory Commission Preface, 1998: 1). More specifically, the Commission was to: review available data regarding consumer information and protections for those enrolled in health care plans and to make recommendations for improvement, review work that defines, measures, and promotes quality of health care to develop consensus on approaches to the delivery of quality health care in a changing environment, and to collect and evaluate data on availability of treatment and services and make recommendations where required (President’s Advisory Commission Preface, 1998: 1). In total, the Commission made more than 50 recommendations to the President to advance the quality of health care in the country. The recommendations can be summarized under six major categories: 1) Public-Private Partnerships, 2) National Aims for Improvement, 3) Quality Measurement and Reporting, 4) Strengthening the Market to Improve Quality, 5) Building the Capacity to Improve Quality, and 6) Addressing the Uninsured.

The Commission unanimously proposed, under the Public-Private Partnerships category, “the creation of two entities, one public and the other private intended to provide ongoing national leadership in health care quality improvement” (President’s Commission Executive Summary, 1998: 2) rather than require government regulation or oversight. The public entity, the Advisory Council for Health Care Quality, was to identify national aims for quality improvement, specify the objectives and goals for quality measurement, and track the progress toward those aims and objectives. It was envisioned that this entity would monitor improvements in and the effectiveness of the health care delivery system, and report directly to Congress. To date, neither Congress nor the Administration, either present or past, ever realized this public council, possibly due to the lack of interest and motivation.

The private entity, the Quality Forum, sought to develop a national strategy for “quality measurement and reporting to advance national aims for improvement” (Miller and Leatherman, 1999: 233). The NQF has attempted to fill the void by pursuing the functions originally envisioned for the council, i.e., national goal setting for health care quality improvement.⁷⁴ On June 17, 1998, funded with Foundation grants and in-kind support from the United Hospital Fund of New York, then-Vice President Gore convened the Planning Committee that would establish the private entity envisioned by the President’s Commission. The Forum for Health Care Quality Measurement and Reporting, dubbed the “Quality Forum,” was established with the aim of “implementing a comprehensive plan for measuring health care quality and reporting the results of such measures to the public” (President’s Commission Executive Summary, 1998: 2). Accordingly, this private organization ideally should consist of public and private stakeholders, including public and private purchasers of health care, consumers, health plans, health care providers, quality improvement organizations and others, in an effort to create a core measure set for reporting as well as to focus the development of new measures for future improvement activities. To achieve these activities, several steps were recommended by the Commission. They were: 1) to identify a core set of quality measures that were applicable to each sector of health care (e.g., hospitals, nursing homes, and health plans) that could be used for standardized reporting, 2) to support the focused development of additional quality measures that would improve the nation’s health care and 3) to ensure that comparative, valid, reliable, and comprehensive information is available for use by consumers, purchasers, practitioners, quality organizations, and others and that this comparative information will reside in the public domain

⁷⁴ There are those in the health care industry that believe the Quality Interagency Coordination Task Force (QuIC), a committee consisting of the various federal partners in health care, replaced the need for the Council. While the QuIC did serve as a focal point for organizing the federal partner programs for quality improvement, it did not include in its charter the overall goals originally envisioned by the President’s Commission, such as tracking national progress in quality improvement. To date, the QuIC has not reconvened since 2001.

(President's Commission Executive Summary, 1998: 3). Thus attention was focused on identifying gaps in performance measurement, reducing the burden of multiple reporting requirements, and encouraging the sharing of best practices across the health care industry. By providing for broad participation, this private organization could be attuned to and respond with greater flexibility to the changing health care environment. An additional provision specified substantial purchaser and consumer representation involved in its governance, with the hope that market forces would advance this initiative and stimulate competition on the basis of quality.

Creation of the Quality Forum required the resolution of issues such as governance, organizational structure and financial support from a neutral convener. In the six-month planning phase, the groundwork for operations including the solicitation of funding⁷⁵ and the development of an organizational structure was completed. The NQF, formally incorporated in the District of Columbia in May 1999, became operational in February 2000, with a founding membership of 57 organizations, as of July 2000.⁷⁶

The original mission of the Quality Forum as cited by the Planning Committee consisted of four goals: "1) the creation of an intellectual framework for quality measurement and reporting; 2) standardization of quality measures; 3) public access to valid, comparable data; and 4) use of the data to facilitate improvement by health care providers and plans, inform consumer and purchaser choice, and stimulate market demand for quality improvement" (Office of the Vice President Press Release, 1998). As a result, the NQF will: develop a framework to guide a national strategy for quality measurement and reporting; ensure system-wide capability to measure and report on quality; marshal market demand for quality; foster and inform consumer

⁷⁵ Founding grants for the NQF came from the Robert Wood Johnson Foundation, the California HealthCare Foundation, the Horace W. Goldsmith Foundation, the United Hospital Fund of New York, and the Commonwealth Fund.

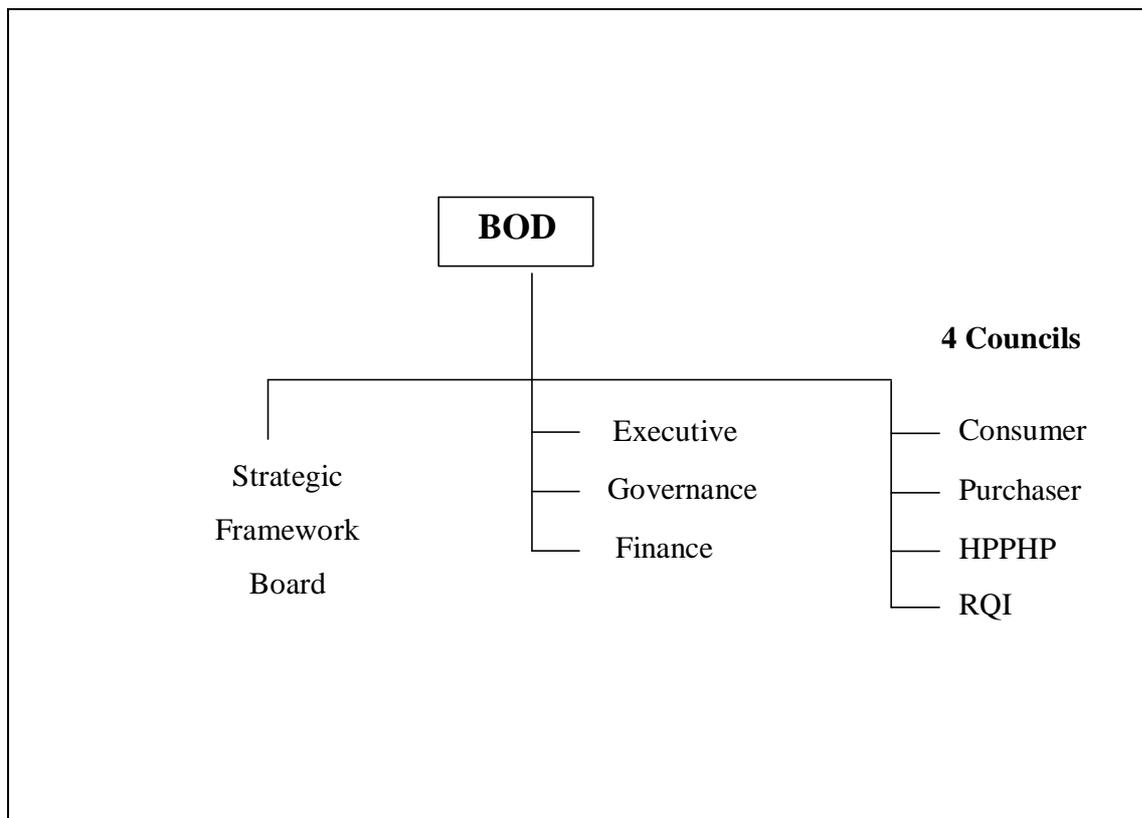
⁷⁶ As of March 2008, the membership consists of 375 organizations.

choice and use of quality information; promote collection and dissemination of data that providers need to improve quality; and reduce the burden on providers and health plans of measuring quality by promoting standardization of quality measurement and reporting (National Quality Forum, 1999).

According to Kenneth W. Kizer, M.D., President and CEO of the NQF, the uniqueness of this member organization of organizations is that “we have the largest public and the largest private purchasers sitting next to each other with the idea that they will be pursuing the same agenda and they will be able to leverage both their arenas towards the same end” (National Institute of Health Policy, 2002: 3). However, the NQF’s strength, e.g., a diverse group of stakeholders that encompasses all the actors of the industry, is also its weakness. The conflicts and disagreements that exist in health care manifested themselves as the NQF became operational. For example, one of the first challenges to the NQF came from a group of organizations representing physicians, who indicated that they wanted their own council, instead of being included in the provider-health plan council. The issue was ultimately heard by the NQF Board, which voted to have physicians serve on the provider council with the other providers of health care. Another issue that emerged was why consumers had a vote equal to providers; which is not surprising given the history of the health care system in the U.S. Other issues included getting stakeholders on the “same page” to move the quality agenda forward, competing public reporting efforts that use measures that are inconsistent with NQF-endorsed measures, state laws requiring reporting of adverse events, data collection issues in the field, complete lack of measures in certain domains of health care, whether NQF should endorse measures solely for public reporting or whether they should also endorse measures for quality improvement, and intellectual property issues which delay projects, to name a few.

The NQF is unique to the health care industry in the United States in both its governance structure and operation. Classified as a voluntary consensus standards-setting organization, it has implications for public policy. Its structure and operation has been designed to achieve its mission as well as to promote collaboration among the diverse member organizations and constituencies to move the quality agenda ahead. Figure 5 shows the NQF governance structure.

Figure 5: NQF Governance and Committee Structure (1999-2007)



Source: NQF Board of Directors Meeting Materials, 2007

In 1999, the NQF was governed by a 23-member Board of Directors (BOD), all drawn from the four major constituencies of health care: purchasers, consumers, provider/health plans (HPPHP) and quality improvement/research (RQI). There were 15 “at large” members, with the majority of votes allocated to purchasers and consumers per the NQF bylaws. Out of the 15 “at large” seats, three are held by representatives of the federal government that include AHRQ,

CMS, and the Office of Personnel Management (OPM).⁷⁷ There are also five non-voting liaison members, including representatives of the National Committee on Quality Assurance (NCQA), the Joint Commission for the Accreditation of Healthcare Organizations (The Joint Commission), the American Medical Association (AMA),⁷⁸ the National Institutes of Health (NIH), and the Institute of Medicine (IOM). These groups were included as members of the BOD based on the significant knowledge and experience they possessed in quality measurement. The remaining four seats were designated for each chairperson of the constituency member councils of the NQF. The initial BOD members and their organizational affiliation are shown in Table 3.⁷⁹

The NQF membership consists of the various stakeholders in the health care delivery system, including purchasers, insurers, providers, government agencies, and consumers. A major benefit of membership is a “seat at the table” in national decision- making. Members of the NQF participate in the national dialogue about health care quality measurement and reporting via four councils:

Consumer Council—this council consists of consumer organizations as well as labor unions at the national, state, regional, and local levels. Its purpose is to “develop a shared vision of consumer needs for quality information, and how the NQF can best meet those needs” (National Quality Forum, 1999).

⁷⁷ At the NQF Annual Meeting held in October 2005, it was announced that the Veterans Administration would replace OPM on the Board.

⁷⁸ The AMA representative was included due to the AMAP program that was to accredit individual physicians.

⁷⁹ In addition to the BOD, there are currently three standing committees (i.e., Executive Committee, Governance Committee and the Finance Committee) that advise and govern the NQF. The Technical Advisory Committee was intended to a) direct the analysis of existing measures, b) establish technical criteria for evaluating measures and c) assess measurement sets. However, these functions were performed by the various ad hoc committees and project-specific Steering Committees (NQF BOD Meeting Materials, 2007).

*Table 3: The National Quality Forum Board of Directors (2003)**

<p>Gail L. Warden (Chair) President and Chief Executive Officer Henry Ford Health System Detroit, Michigan</p>	<p>Christopher J. Queram Chief Executive Officer Employer Health Care Alliance Cooperative Madison, Wisconsin</p>
<p>William L. Roper, M.D., MPH (Vice-Chair) Dean, School of Public Health University of North Carolina Chapel Hill, North Carolina</p>	<p>John C. Rother Director of Policy and Strategy AARP Washington, DC</p>
<p>Bruce E. Bradley Director, Managed Care Plans General Motors Corporation Detroit, Michigan</p>	<p>Nancy Ann Min DeParle Administrator Health Care Financing Administration Baltimore, MD</p>
<p>John M. Eisenberg, M.D., M.B.A. Director Agency for Healthcare Research and Quality Rockville, Maryland</p>	<p>Gerald M. Shea Assistant to the President for Government Affairs AFL-CIO Washington, DC</p>
<p>William A. Gillespie, M.D. Executive Vice President Kaiser Permanente Oakland, California</p>	<p>Michael A. Stocker, M.D., M.P.H. President and Chief Executive Officer Empire Blue Cross and Blue Shield New York, New York</p>
<p>William E. Golden, M.D. Immediate Past President American Health Quality Association Washington, DC</p>	<p>Marina L. Weiss, Ph.D. Senior Vice President for Public Policy and Government Affairs, March of Dimes Washington, DC</p>
<p>Lisa I. Iezzoni, M.D. Professor of Medicine Harvard Medical School Boston, Massachusetts</p>	<p>Linda K. Wertz State Medicaid Director Texas Health and Human Services Commission Austin, Texas</p>
<p>Janice R. Lachance Director Office of Personnel Management Washington, DC, representing the QuIC</p>	<p>Liaison Members Margaret E. O’Kane President National Committee for Quality Assurance Washington, DC</p>

Kenneth W. Kizer, M.D., M.P.H. President and Chief Executive Officer The National Quality Forum Washington, DC	Dennis S. O’Leary, M.D. President, Joint Commission on Accreditation of Healthcare Organizations Oakbrook Terrace, Illinois
Judith L. Lichtman President National Partnership for Women and Families Washington, DC	Kenneth I. Shine, M.D. President Institute of Medicine Washington, DC
John R. Lumpkin, M.D., M.P.H. Director Illinois Department of Public Health Springfield, Illinois	Randolph D. Smoak, Jr, M.D. Chair, Governing Body, American Medical Accreditation Program, American Medical Association Chicago, Illinois

* Bolded names were members of the initial BOD in July 2000.
Source: National Quality Forum, 2003.

Purchaser Council—this council consists of employers, private corporations, regional purchasing coalitions, business groups and government agencies that have an interest in reducing the costs of care, improving quality, and enhancing productivity of the workforce through improved health. The purpose is to “provide a venue for public and private purchasers to build demand and capacity for quality improvement” (National Quality Forum, 1999).

Provider and Health Plan Council—this council consists of physician and nurse organizations, health plans, health systems, hospitals, ambulatory care facilities, home care agencies, long-term care and other residential facilities, and groups of health care practitioners. Its purpose is to “promote a shared agenda on quality among providers and collaboration with purchasers” in an effort to “improve quality and reduce the cost and burden of reporting duplicative measures” (National Quality Forum, 1999).

Research and Quality Improvement Council—consists of organizations that conduct research, education, or projects to improve health care quality measuring and reporting, such as accrediting bodies, professional schools, policy or quality centers, federal, state, and local government agencies, and supporting industries such as medical suppliers. Its purpose is to

“build on advances made by quality improvement organizations and experts to promote use of quality measures by purchasers, consumers, providers, and policymakers” (National Quality Forum, 1999).

In late 2000, the NQF held its first elections for leadership of each Member Council. A Chair and a Vice-chair were selected by an electoral process. The purpose of the Chair and Vice-chair of a council is to encourage discussion among members about current projects of the NQF and to present that information to the BOD. The results of this election are provided in Table 4.

Table 4: Results of Member Council Elections in 2000

<p><u>Consumer Council</u></p> <p>Judith L. Lichtman—Chair National Partnership for Women & Families Washington, DC</p> <p>Brian W. Lindberg—Vice Chair Consumer Coalition for Quality Health Care Washington, DC</p>	<p><u>Research & Quality Improvement Council</u></p> <p>William E. Golden, M.D.—Chair American Health Quality Association and University of Arkansas for Medical Sciences Little Rock, Arkansas</p> <p>Deborah M. Nadzam, Ph.D., F.A.A.N.—Vice Chair The Cleveland Clinic Foundation Cleveland, Ohio</p>
<p><u>Provider & Health Plan Council</u></p> <p>William A. Gillespie, M.D.—Chair Kaiser Foundation Health Plan, Inc. Oakland, California</p> <p>Thomas R. Reardon, M.D.—Vice Chair American Medical Association Chicago, Illinois</p>	<p><u>Purchaser Council</u></p> <p>Christopher Queram—Chair Employer Health Care Alliance Cooperative Madison, Wisconsin</p> <p>Lee Partridge—Vice Chair National Association of State Medicaid Directors Washington, DC</p>

Source: National Quality Forum, 2000.

In December 1999, in order to realize the goals charged to the NQF, a nine-member Strategic Framework Board (SFB) was established to develop recommendations and a conceptual framework for a national strategy for health care quality measurement and reporting, one of NQF's earliest initiatives. The members of the SFB (see Table 5) consisted of experts from areas such as health care quality measurement, quality reporting, research, health care purchasing, accreditation and certification, education, information technology and health care delivery. The effort cost over \$1 million and lasted approximately 18 months. While the BOD was briefed at various intervals throughout the length of the project, the draft recommendations were presented at the June 2001 Board meeting. The draft report was made available to the membership for comment, and revisions were made by the SFB. The final report, with membership input, was subsequently published in October 2001 and included 17 recommendations for the NQF to act upon. Based on additional input from the public, the NQF staff revised and reordered the recommendations which, subsequently the NQF membership voted to approve.

Table 5: The National Quality Forum Strategic Framework Board

Donald M. Berwick, M.D., M.P.P. President and Chief Executive Officer, Institute for Healthcare Improvement Boston, MA
Christine K. Cassell, M.D. Dean, Oregon Health and Science University Portland, OR
Molly J. Coye, M.D., M.P.H. President and Chief Executive Officer, The Health Technology Center Menlo Park, CA
Robert S. Galvin, M.D. Director, Global Healthcare, General Electric Fairfield, CT
Judith H. Hibbard, Dr.PH. Professor, University of Oregon Eugene, OR
Brent C. James, M.D., M.Stat. Vice President, Medical Research and Continuing Education, Intermountain Health Care Salt Lake City, UT

Sheila Leatherman Adjunct Professor, University of North Carolina School of Public Health Chapel Hill, NC
Elizabeth A. McGlynn, Ph.D. Director, Center for Research on Quality in Healthcare, RAND Santa Monica, CA
Helen L. Smits, M.D. Visiting Scholar, New York University New York, NY
Project Staff Kenneth W. Kizer, M.D., M.P.H. President and Chief Executive Officer
Robyn Y. Nishimi, Ph.D. Chief Operating Officer

Source: NQF: A National Framework for Healthcare Quality Measurement and Reporting, 2002

Board approval was obtained in May 2002. The final recommendations were categorized into several areas: four quality standard principles, five strategic priority areas of action, and eight policy statements. Interestingly enough for a consensus-driven organization, and prior to NQF membership vote and approval, the SFB requested permission from the NQF BOD to further disseminate the draft report presented in June 2001 to a wider audience. The Board unanimously approved the SFB request, and a compilation of their deliberations was published in a special issue of *Medical Care* in 2003.

The SFB developed a visual conceptual framework and subsequently crafted a purpose statement for a quality measurement and reporting system. Areas included in the final recommendations are “setting national standards, local adoption, measurement and reporting, evidence based standards, quality measurement and reporting, public awareness, and the development of a reporting strategy” (National Business Coalition on Health, 2001: 3). Its purpose statement indicated that a national quality and measurement reporting system should:

“Evaluate the degree to which the U.S. health care system is providing safe, effective, timely and patient-centered care; Assess whether the

delivery of the high quality care is efficient and equitable; Enable substantial progress to be made toward achieving established national goals; Provide easily accessible information on quality to a variety of audiences, including consumers, purchasers, and providers, to facilitate individual and collective decision making; provide information that regulators, purchasers and providers can use to support continued improvement and achievement of goals” (National Business Coalition on Health, 2003: I-3).

One part of the NQF’s charge to the SFB included setting national standards because the public sector Advisory Council that was directed to do that task was never created. At the same time, while developing recommendations for the NQF strategic vision, the SFB also considered several areas to include in its proposed national goals for quality. The considerations included:

- What type of data elements should be collected on a national level?
- Under what circumstances and of what time along the process of care would data be collected?
- What types of data sharing arrangements are necessary to reduce burden?
- What display issues should be considered?
- When does information need to be available to providers of care?
- When does information need to be available to consumers?
- How can you integrate systems that are useful to consumers and providers alike?

(National Business Coalition on Health, 2001: 5).

During its deliberations, the SFB selected one diagnostic area, and then applied its assumptions to see how intended and unintended consequences of a measurement and reporting system for quality affected the health care system, which led to the conceptual framework that was ultimately produced (see Figure 6).

Figure 6: Conceptual Framework for Quality Improvement



Source: NQF: A National Framework for Healthcare Quality Measurement and Reporting, 2002: B-3.

Membership dues are the principal source of funding for the NQF. Dues are determined via a sliding scale based on what category an organization is in and its operating budget. However for Federal, State or Local Government members, operating budget is not considered, but size of population served is. Membership dues range from \$1,000 to \$25,000 per year. These dues support the operating budget of NQF. The NQF staff consists of approximately 16 individuals with varying educational backgrounds and experience. Some staff members are also clinicians. The first president and chief executive officer was Kenneth W. Kizer, M.D., M.P.H., whose tenure lasted from 1999 until 2005. In addition, the NQF also receives funding from both the public and private sector. These funds tend to be earmarked for particular projects, such as the consensus standards project for the reporting of healthcare-associated infections sponsored by the Texas Medical Institute of Technology, or the project on best practices for the treatment of substance use disorders sponsored by the Robert Wood Johnson Foundation.

Centers for Medicare & Medicaid Services

The Centers for Medicare & Medicaid Services (CMS) is charged with the administration of the Medicare and Medicaid programs (Title XVIII and Title XIX of the Social Security Act) that were signed into law on July 30, 1965 (PL 89-97). This law extended health care insurance to those 65 years old or older, and provided health care services to “low-income children deprived of parental support, their caretaker relatives, the elderly, the blind and individuals with disabilities” (CMS 2005). CMS is one agency of the U.S. Department of Health and Human Services, and provides a variety of services to the public. Not only does CMS manage Medicare, the federal portions of the Medicaid program and the State Children’s Health Insurance Program (SCHIP), it also provides survey and certification services, and produces original research. In 2005, it financed \$661 billion in health care services, or one-third of the nation’s health care spending and approximately 75 percent of all public spending on health care (Hoffman et al., 2007). It provides administrative services to nearly 44 million beneficiaries, with federal entitlement outlays of approximately \$402 billion⁸⁰ (Hoffman et al., 2007). Its mission is to “ensure effective, up-to-date health care coverage and to promote quality care for beneficiaries” (CMS, 2006).

To accomplish its mission, CMS developed a Strategic Action Plan, which can be outlined though five key objectives. They are:

- Skilled, Committed, and Highly-Motivated Workforce
- Accurate and Predictable Payments
- High-Value Health Care
- Confident, Informed Consumers

⁸⁰ In 2006, Part A benefit payments were \$189.0 billion, Part B totaled \$165.9 billion, and Part D benefits were \$47.1 billion.

- Collaborative Partnerships (CMS, 2006).

The agency was originally formed in 1977 and was then known as the Health Care Financing Administration (HCFA).⁸¹ Prior to its formation, the Social Security Administration was responsible for the administration of Medicare and federal aid to State Medicaid Programs was managed by the Social and Rehabilitation Services in the Department of Health, Education, and Welfare (HEW)⁸². HCFA was created to more effectively coordinate the Medicare and Medicaid programs. Since 1965, congressional leaders and others have tried to establish a coordinated quality assurance program for beneficiaries of these programs as evidenced by the 1972 amendment to the Social Security Act (SSA) that established the professional standards review organizations (PSROs). These entities were seen as being “enforcers” by HCFA, but despite this perception, in reality, the PSROs could only make recommendations and were not permitted to impose sanctions. Overall, the program was not effective and was replaced by the Medicare Utilization and Quality Control Peer Review Organizations when the prospective payment system was adopted in 1983.⁸³ Known as “Peer Review Organizations” (PROs), they were intended to be a more efficient mechanism to review “medical necessity, reasonableness, and quality of care and the appropriateness of care setting” (Sprague, 2002: 4). In 1986, one of the provisions of the Omnibus Budget Reconciliation Act was passed by Congress required the PROs to focus on quality monitoring. Congress also tasked the secretary of DHHS to evaluate the PROs and the quality of their work in Medicare. IOM conducted the study and found the effectiveness of the PROs mixed—they were laudable in certain aspects but limited in others. The IOM concluded that the structure of the PROs should be revised and should have as its main

⁸¹ In 2001, this Agency was renamed as the Centers for Medicare & Medicaid Services (CMS) by Secretary Tommy Thompson.

⁸² HEW was divided into the Department of Education and the Department of Health and Human Services (DHHS) in 1980.

⁸³ The impetus for the PROs was the Tax Equity and Fiscal Responsibility Act of 1982, which contained a provision among others that expanded HCFA’s quality oversight responsibilities through these local bodies.

foci quality of care, patient outcomes, provider communications, and care administered outside the hospital.

In 1992, HCFA announced that it was revamping the structure of the PROs to be more in line with the goals envisioned by the IOM study. The PROs role would be one of technical assistance to providers with the goal of quality improvement. New data systems and methods of analysis were incorporated into this system, and training was available to local providers that assisted in the development of local quality improvement projects. In November 2001, DHHS Secretary Tommy G. Thompson announced the Quality Initiative, which demonstrated the Bush Administration's commitment to "assure quality health care for all Americans through published consumer information coupled with health care quality improvement support through Medicare's Quality Improvement Organizations (QIOs)" (CMS, 2004).⁸⁴ In mid-2002, CMS released the performance results of the 20 quality indicators used in the project. Overall, the results showed an improvement in the quality of care delivered by those states participating. With the documentation of improvement in a number of clinical areas, the procedural requirements were relaxed by CMS, and the PROs⁸⁵ were left to focus on working with local partners in ways that were most effective for them.

Other significant changes to the Medicare and Medicaid programs that affected CMS operations in 1996-97 included changes in welfare reform, the Health Insurance Portability and Accountability Act of 1996 (HIPAA), and the Balanced Budget Act of 1997. In the area of welfare reform, the entitlement program Aid to Families with Dependent Children (AFDC) was replaced by the Temporary Assistance for Needy Families (TANF). The welfare link to

⁸⁴ This initiative originally was launched in 2002 as the Nursing Home Quality Initiative, and in 2003 was expanded to include the Home Health Quality Initiative and the Hospital Quality Initiative. These were part of the focus on quality improvement within the Department and included other projects such as the Doctor's Office Quality and the End-Stage Renal Disease projects.

⁸⁵ The name PROs was formerly replaced in 1998 with the Quality Improvement Organization (QIO) by Secretary Tommy Thompson of the Department of Health and Human Services (DHHS).

Medicaid was removed and enrollment or termination of Medicaid benefits was not automatic with receipt or loss of welfare cash assistance. The HIPAA regulations had several provisions that affected CMS operations. HIPAA amended the Public Health Service Act, the Employee Retirement Income Security Act of 1974 (ERISA) and the Internal Revenue Code of 1986 in an effort to provide Federal regulations for improving continuity of health care coverage in health insurance markets. In addition, the Medicare Integrity Program was created that permitted CMS to competitively contract for program integrity work, and which formed the national administrative simplification standards for electronic health care transactions. As a final provision, if Congress failed to enact substantive privacy legislation, DHHS was required to issue the regulations. In 1997, the Balanced Budget Act contained a number of provisions relating to Medicaid and Medicare. Some of the provisions were establishing a State Children's Health Insurance Program (SCHIP) and new Medicaid managed care options and requirements for states. On the Medicare side, provisions included making education and information available to beneficiaries to inform their health care choices; requiring CMS to develop and implement additional prospective payment systems in certain areas;⁸⁶ slowing the rate of Medicare spending; providing a range of beneficiary protections; expanding preventive benefits; and through research and demonstration projects, testing other approaches to payment and service delivery.

More recent legislation that has affected CMS operations is the Medicare Prescription Drug, Improvement, and Modernization Act of 2003 (MMA)⁸⁷. This Act created a discount prescription drug card program on a voluntary, limited cost basis to enrollees, and provided limited financial assistance for purchasing these prescription cards to low-income beneficiaries

⁸⁶ These include inpatient rehabilitation services, hospital outpatient departments, skilled nursing facility services, home health services, and outpatient rehabilitation services.

⁸⁷ Public Law 108-173.

until 2006.⁸⁸ At that time, the new voluntary Part D outpatient prescription drug benefit would be available from private drug plans and the Medicare Advantage plans.⁸⁹ “Part D provides subsidized access to prescription drug insurance coverage on a voluntary basis, upon payment of premium, for all beneficiaries, with premium and cost-sharing subsidies for low-income enrollees” (Hoffman, et al., 2007: 5). For the first time in its history, CMS will consider beneficiary income in relation to Medicare benefits. Beneficiaries with lower incomes will be eligible for subsidies for the Part D prescription drug program, and those with higher incomes were required to pay a greater share of the Medicare Part B premium beginning in 2007.

The organizational structure of CMS is typical of a bureaucracy consisting of a variety of centers and offices; it employs approximately 4,100 individuals located at its headquarters in Baltimore. It also has ten regional offices throughout the United States. In FY 2005, CMS requested appropriations in the amount of \$482.1 billion (Centers for Medicare & Medicaid Services, 2005). Activities such as survey and certification account for a portion of the requested budget (\$270 million), but 68 percent of the FY 2005 budget request financed activities such as claims processing, appeals, inquiries and provider assistance. CMS processed 1.1 billion claims and respond to approximately 51 million inquiries. The unit costs to process Part A claims remained constant from FY 2004-FY 2005 (\$0.87 per claim) while the Part B processing costs decreased (from \$0.65 in FY 2004 to \$0.63 in 2005) (Centers for Medicare & Medicaid Services, 2005). With Medicare, Medicaid, and SCHIP, accounting for approximately one-third of the nation’s health spending, CMS is the largest public purchaser of health care services in the country.

⁸⁸ Lower incomes are defined as incomes less than 150 percent of the federal poverty limit.

⁸⁹ Employers that provide prescription drug benefits to their retirees that is comparable to Medicare’s drug plan will be eligible for a federal subsidy.

Agency for Healthcare Research and Quality

The Agency for Healthcare Research and Quality (AHRQ) is the lead Federal agency on quality research and patient safety. As one of 12 agencies of the U.S. Department of Health and Human Services, AHRQ has a broad range of programs aimed at improving the quality of health care, reducing its cost, and broadening access to essential health care services. According to Agency Director Carolyn M. Clancy, M.D., “Our [AHRQ’s] focus is on getting research results into the hand of those who can put it to practical use as rapidly as possible” (AHRQ, 2005:iii). The main functions of the Agency include sponsoring as well as conducting health services research “that provides evidence-based information on health care outcomes; quality; and cost, use, and access” (AHRQ, 2001). The results of the research help inform decision makers across the spectrum of health care, such as patients and clinicians, health care system leaders, purchasers, and policy-makers. Approximately 80 percent of the AHRQ budget is awarded as grants and contracts to researchers and research institutions within the U.S.

Created by Congress in 1989 as the Agency for Health Care Policy and Research (AHCPR) from the Omnibus Budget Reconciliation Act (OBRA) of 1989,⁹⁰ it has had a turbulent history. AHCPR replaced the National Center for Health Services Research and Health Care Technology Assessment, thus ending the Center’s long-standing struggle for funding and respect on Capitol Hill.

The new Agency’s charge was to “carry out research, demonstrations, guideline development, training, and dissemination activities with respect to health care services and systems of information regarding the following areas: the effectiveness, efficiency, quality, and outcomes of health services; clinical practice including primary care; health care technologies, facilities and equipment; health care costs, productivity, and market forces; health promotion and

⁹⁰ Created as Title IX of the Public Health Service Act by Section 6103 of OBRA 1989.

disease prevention; health statistics and epidemiology; and medical liability” (Gray, 1992: 40). As per provisions in the Act, a 17-member National Advisory Council for Health Care Policy, Research, and Evaluation was established to provide direction with regard to agenda setting and priorities for the Agency. Funding for AHCPR was to come from two streams— the Public Health Services (PHS) Act and the Social Security Act.⁹¹ With a modest budget of \$53 million in 1989, the Agency’s funding continued to grow in the next several years, reaching \$162,386,000 in FY 1995.

To carry out its mission regarding outcomes research and practice guideline development, the Agency funded its Patient Outcomes Research Teams (PORTs). These teams were multi-disciplinary centers “that would focus on particular medical problems and review and synthesize available research, analyze practice variations and patient outcomes using administrative data augmented by primary data collection, disseminate the results, and evaluate the effects of dissemination” (Gray et al, 2003: W3-289). AHCPR had two assets that could be useful to both the first Bush and the Clinton Administration in terms of health system reform and policymaking. The first included an intramural research program that could tap the National Medical Expenditure Survey (NMES), which could be used to analyze policy options under consideration at the time. The second was that the Agency employed a substantial number of researchers, who were familiar with data as well as the general policy issues in health care.

In 1994, with a change in leadership, the Agency sought to expand its customer base from health services researchers to include “those who would make use of the products of its work and that the research community should be seen as being the Agency’s partners in meeting

⁹¹ This was in addition to the usual appropriations for the Agency. From the PHS Act, the Agency was to receive 40 percent of the evaluation money, which was approximately one percent of the appropriated funds of the Act. The amendment to the Social Security Act, under the Medicare trust fund, provided some budgetary authority to support the outcomes research program in President George H. Bush’s budget, which amounted to approximately \$53 million in 1989 (Gray et al, 2003: W3-287).

the needs of the customers on whose goodwill the Agency's support would depend" (Gray et al, 2003: W3-293). To meet this goal, the Agency was reorganized from the "extramural and intramural" dichotomy into centers that reflected the substantive work being accomplished such as the Center for Outcomes Effectiveness. Furthermore, the Agency's NMES database was enhanced to make it more comprehensive and up to date, and re-named the Medical Expenditure Panel Survey (MEPS). Other changes included a move away from the PORT grant program, cited by the leadership as being ineffective and expensive to administer.

When the Republicans gained control of both the House and the Senate after the 1994 elections, the Agency, as well as other federal entities, found themselves in an atmosphere that favored a reduced welfare state, as well as reduced federal regulation.⁹² The Agency's name appeared on a list of discretionary programs that were to be eliminated. In the end, under a joint resolution of the House-Senate, the AHCPH was to have a reduction in funding by 75 percent with a provision supporting the MEPS project only.⁹³ A hearing of the House Ways and Means Committee cited the Agency for "wastefulness and unwarranted interference with the practice of medicine."⁹⁴ And while the House Appropriations Committee approved a budget for \$125 million, when it went to the floor, an amendment was introduced to eliminate the Agency's funding. In the end, the amendment was withdrawn, and the Agency was given \$65.5 million by the House. The AHCPH fared somewhat better in the Senate, where the Appropriations Committee passed a \$125,310,000 budget for the Agency. After reconciliation of the Senate and House appropriations, the bill finally passed in April 1996 and funded the Agency at

⁹² Known as the Contract with America and led by Speaker Newt Gingrich (R-GA), a contentious battle with the Democratic administration and the Congress ensued and ultimately led to government shutdowns and delays in passing the Appropriations bill until well into the fiscal year.

⁹³ A FY 1995 "rescission" also took back \$3 million from the Agency's budget, which reduced the budget to \$159,386,000 at the 8th month of the FY (Gray et al, 2003: W3-294).

⁹⁴ This was borne out by reports issued from several organizations, including the U.S. General Accounting Office (GAO), a Physician Payment Review Commission (PPRC), and the Office of Technology Assessment (OTA) in 1995.

\$125,169,000, which was 21 percent less what had been appropriated in FY 1995 (Gray et al, 2003).

The core argument for the AHCPR's continued existence included its creation, which had bipartisan support, with key sponsorship by Republicans in both the House and Senate. Furthermore, the reason for the Agency's creation remained—the outcomes/effectiveness problem and the high cost of care. In addition, the Agency had developed several allies in the private sector that supported and used its work. Organizations such as the American Association of Health Plans (AAHP), the American Hospital Association (AHA), the American Medical Association (AMA), the Health Insurance Association of America (HIAA), the Association of American Medical Colleges (AAMC), the American Nurses Association (ANA), the Paralyzed Veterans of America, and other professional organizations supported continued funding of the Agency. Many provided support in terms of testimony and tracking of breaking issues, but also by lending lobbying resources to prevent the Agency's closure. In the end, it was argued by Congressional supporters that “it is essential to have a federal agency that works with the private sector to provide consumers with information to make informed choices, measure and improve the quality of care and improve the cost and effectiveness of our health care system” (Gray et al, 2003: W3-300).

The Agency survived efforts to “zero out” its budget, but it was reduced by approximately 21 percent as a result of the debates on Capitol Hill. In 1997, amidst budgetary struggles, the Agency changed leadership. The new director, John Eisenberg brought considerable strengths as a health services researcher, was well-connected within the health care industry and was politically savvy, with “trusting and friendly relationships with key staff on both sides of the aisle” (Gray et al, 2003: W3-301). Under Eisenberg's direction, the Agency made several changes, such as the reorganization of research activities into topical centers with a

focus on consumers and the initiation of the Quality Interagency Coordination Task Force (QuIC).⁹⁵ Additionally, the AHCPR placed more value on building relationships with the industry in an effort to learn more about what its customers valued and needed, as well as establishing an active dissemination program that promoted the Agency's activities to its various stakeholders and constituencies. One other important change included the replacement of guidelines development and dissemination with "evidence-based practice centers (EPCs)." The twelve EPCs, which were external organizations, would be responsible for the accumulation and aggregation of data that would be disseminated via reports; these could be used by other organizations (mostly private-sector) to develop practice guidelines for the health care industry.

In December 1999, under the Healthcare Research and Quality Act, the AHCPR reauthorizing legislation passed in both the House and the Senate, and included a name change to the Agency for Healthcare Research and Quality (AHRQ). The legislation eliminated from the Agency's purview the development of clinical practice guidelines. The reauthorization legislation also stipulated that the AHRQ is a "science partner" that works collaboratively with the public and private sector organizations to improve the quality and safety of patient care. Specifically, AHRQ would:

- Meet the information needs of its customers—patients and clinicians, health system leaders, and policymakers—so that they can make more informed health care decisions
- Build the evidence base for what works and doesn't work in health care and develop the information, tools, and strategies that decision makers can use to make good decisions and provide high-quality health care based on evidence

⁹⁵ The QuIC consisted of several federal agencies including HHS, AHRQ, CMS (formerly known as HCFA), the Departments of Defense, Veterans Affairs, Labor, and Commerce, OPM, OMB, the U.S. Coast Guard, the Federal Bureau of Prisons, the National Highway Transportation and Safety Administration, and the Federal Trade Commission. The goal is to "ensure that all federal Agencies that purchase, provide, study, or regulate health care services are working in a coordinated way toward the common goal of improving the quality of care" (QuIC, 2001: 1).

- Develop scientific knowledge in these areas, but will not mandate guidelines or standards for measuring quality (AHRQ, 1999: 2).

AHRQ's mission is to "support research designed to improve the quality, safety, efficiency, and effectiveness of health care for all Americans" (AHRQ, 2001: 1). To support its mission, the Agency is comprised of four offices and five centers that share a budget of \$318 million for FY 2007. The offices and centers are supported by a multidisciplinary staff of approximately 300, including physicians, nurses, pharmacists, researchers, statisticians and others. The Agency takes direction from the Secretary of the Department of Health and Human Services. In addition, the National Advisory Council for Healthcare Research and Quality, a twenty-one member panel, provides advice and recommendations to the Director of AHRQ and the Secretary of DHHS with regard to priorities for a national agenda in health services research. This panel is composed of a number of public and private sector experts who are appointed by the DHHS Secretary to serve three-year terms. The panel's purpose is to provide varied perspectives on the health care system in the U.S., and prioritize the research areas that AHRQ should address to promote quality improvements, better outcomes, and cost-effectiveness of clinical practice (AHRQ, 2006: 1). The Director of the Agency also has an in-house panel, consisting of center and office directors as well as some senior advisors, to provide input and recommendations related to daily operations of projects as well as strategic planning for the Agency. Some of the projects that AHRQ has established include outcomes research (PORTS), diabetes treatment (Diabetes Quality Improvement Project), research on ethnic and racial disparities in health care, and reducing medical errors and promoting patient safety. Other projects and quality improvement efforts include the Centers for Education and Research on

Therapeutics (CERTs),⁹⁶ Consumer Assessment of Health Providers and Services (CAHPS®),⁹⁷ Healthcare Cost and Utilization Project (HCUP),⁹⁸ Medical Expenditure Panel Survey (MEPS),⁹⁹ the National Guideline Clearinghouse™ (NGC),¹⁰⁰ Evidence-based Practice Centers (EPCs),¹⁰¹ the U.S. Preventive Services Task Force (USPSTF),¹⁰² and the Quality Interagency Coordination (QuIC) Task Force.⁹⁵

Joint Commission for the Accreditation of Healthcare Organizations

The Joint Commission for the Accreditation of Healthcare Organizations (Joint Commission) is a private, not-for profit organization that evaluates and accredits health care organizations/programs across the nation. It was established in 1951, and is the oldest and largest standards-setting and accrediting body in health care. According to the President of the Joint Commission, Dennis S. O’Leary, M.D., “The Joint Commission is widely recognized for its cutting-edge leadership role in developing standards and performance measures, and for the adaptability of its rigorous evaluation processes to emerging new forms of health care delivery

⁹⁶ Established as a permanent program under the 1999 Healthcare Research and Quality Act, its purpose is “to help reduce adverse drug events and promote the safe and effective use of pharmaceuticals by conducting state-of-the-art research that increases awareness of the uses and risks of new drugs and drug combinations, biological products, and devices as well as mechanisms to improve their safe and effective use” (AHRQ, 2001: 6).

⁹⁷ CAHPS is a survey and reporting kit that incorporates the perspectives of consumers and others and “provides reliable and valid information to help consumers and purchasers assess and choose among health plans, hospitals, and other entities.”

⁹⁸ HCUP is a “Federal-State-industry partnership to build a multi-state health care data system for research, policy analysis, and quality measurement and improvement.” It comprises several databases, web products and quality measurement and improvement tools to identify, track, analyze, and compare trends in hospital care.

⁹⁹ MEPS is a “national survey of health care use, expenditures, sources of payment, and insurance coverage for the U.S. civilian non-institutionalized population as well as a national survey of nursing home residents.” This database consists of information from approximately 10,000 households and approximately 24,000 individuals.

¹⁰⁰ The NGC, developed in partnership with the American Medical Association and the American Association of Health Plans, is a web-based resource that provides information on evidence-based clinical practice guidelines. The NGC “helps health care professionals and health system leaders select appropriate treatment recommendations by providing full text or an abstract of recommendations, comparing and evaluating different recommendations, and describing how they were developed.”

¹⁰¹ EPCs “conduct systematic, comprehensive analyses and syntheses of the scientific literature to develop evidence reports and technology assessments on clinical topics that are common, expensive, and present challenges to decision-makers.” This information is provided in partnerships with specialty societies and health systems and the like, who will use the findings to develop tools and other materials for quality improvement.

¹⁰² USPSTF is an “independent panel of preventative health experts, convened by AHRQ, who are charged with evaluating the scientific evidence for the effectiveness of a range of clinical preventive services and producing age-specific and risk factor-specific recommendations for these services.”

organizations” (JCAHO, 2006). Through evaluation of organizations based on compliance with established standards, the Joint Commission has granted accreditation to more than 15,000 health care facilities across the country.¹⁰³ For over 75 years, the Joint Commission and its predecessor organization monitored and evaluated the performance of health care organizations.

The Joint Commission has its origins in the American College of Surgeons (ACS), which developed the Hospital Standardization Program, with a resolution that “some system of standardization of hospital equipment and hospital work should be developed” (JCAHO, 1990). In 1917, a single page document, the *Minimum Standard for Hospitals*, was developed and adopted for use in the Hospital Standardization Program (Roberts, Coale, and Redman, 1987). The ACS began on-site hospital inspections in 1928 and subsequently released the results of its findings. Of the 692 hospitals inspected, only 89 met the minimum standard, which garnered immediate support of the hospital program by the medical community. Due to the expense associated with the program, more than \$2 million had been spent by 1950; the ACS could no longer be the sole source of support for the program and sought the participation of other organizations. In 1951, the Joint Commission on Accreditation of Hospitals (JCAH) was created and included in addition to the ACS, the American College of Physicians, the American Hospital Association, the American Medical Association, and the Canadian Medical Association. In 1959, the Canadian Medical Association withdrew its corporate membership to develop its own accrediting organization in Canada, and was eventually replaced by the American Dental Association in 1979.

In 1965, Social Security Amendments, specifically the Medicare Act (PL 89-97), contained a provision that hospitals that are JCAH accredited were “deemed” to be in

¹⁰³ The Joint Commission has an affiliate organization, the Joint Commission Resources, Inc. It is a not-for-profit organization that provides health care consulting services as well as worldwide accreditation services.

compliance with the Medicare Conditions of Participation for Hospitals.¹⁰⁴ By being in compliance, hospitals were able to participate in the Medicare and Medicaid programs and receive medical payments for services provided to these populations.¹⁰⁵

Shortly after the passage of the Medicare Act, the JCAH Board decided to review and revise the minimum standard with a focus on achieving optimal achievable levels of care, rather than the minimal essential level of care. In 1970, the four-year hospital standards revision project was complete, and the JCAH Board approved the new standard, which was subsequently published in the *Accreditation Manual for Hospitals*.

In 1971, the Social Security Act was amended to require the Secretary of the U.S. Department of Health and Human Services to validate JCAH survey results. The law also indicated that the Secretary was to conduct surveys of hospitals, based on complaints of noncompliance with Medicare standards, and to establish standards “higher than those of the Joint Commission if it were found that those of the Joint Commission were deficient” (Brooks, 1995: 147).

Over the next several years, the JCAH convened several councils to establish a variety of accreditation programs in areas such as long term care, psychiatric facilities, substance abuse programs, ambulatory health care facilities, hospice care, and home care. And in 1987, the organization changed its name to the Joint Commission for the Accreditation of Healthcare Organizations, to better reflect the scope of its activities.

In 1986, Dennis O’Leary took the helm of the Joint Commission and announced an “Agenda for Change” that was designed to enhance the visibility and authority of the

¹⁰⁴ The Medicare Conditions of Participation for Hospitals were issued by the federal Health Care Financing Administration (HCFA) in the *Code of Federal Regulations*, 1965).

¹⁰⁵ Hospitals have a choice of being surveyed by either the JCAHO for a fee or by state agencies that are acting as agents of CMS (formerly HCFA) and do not charge for inspections.

organization and establishing it as the national standard-bearer of quality of health care. As part of this initiative, a series of projects were designed to emphasize organizational performance, specifically they were designed to 1) reformulate the standards, 2) redesign the survey process, and 3) develop performance measures and a related reference database (Brooks, 1995: 147). As an extension of its basic mission, the Joint Commission developed an accreditation process for managed care plans or health maintenance organizations (HMOs). However, this ill-fated project eventually was folded into an ambulatory care accreditation program, primarily due to lack of interest by the managed care industry. Instead, the Joint Commission refocused its attention on health care organizations and developed, in collaboration with a national advisory group of executives, a Health Care Network Accreditation Program that was launched in 1994. The goal was to use the accreditation process as a driver of continuous quality improvement in health care organization performance, thus emphasizing performance goals and outcomes instead of process specifications. In 1998, in response to the refocused efforts on performance outcomes, the ORYX¹⁰⁶ system was integrated into the accreditation process.

In 1999, the Joint Commission began a concentrated effort to focus on patient safety in health care organizations. It revised its mission statement to reflect this effort.¹⁰⁷ It also introduced other initiatives such as the Shared Visions—New Pathways¹⁰⁸ initiative and the National Patient Safety Goals.¹⁰⁹ In recent years, the Joint Commission has focused on various

¹⁰⁶ ORYX: The Next Evolution in Accreditation is intended to support the accreditation process as well as quality improvement efforts in health care organizations. One component of this initiative is to use “standardized” measures that currently exist rather than develop new ones. Hospitals began collecting data on these “core” measures in July 2002.

¹⁰⁷ The original mission statement was similar to the revised statement, but the debate centered on whether to add the explicit reference to the term “patient safety.” Some on the Board thought the term “quality” included patient safety while others did not. In the end, the Board voted to include the term “patient safety” in the revised mission statement (J. Loeb, personal communication, April 24, 2006).

¹⁰⁸ This initiative, announced in 2002, was to focus the accreditation process on care systems that are believed to be critical to patient safety and the quality of care delivered by health care organizations. It became effective in 2004.

¹⁰⁹ These goals and related accreditation requirements, established in 2002, were intended to improve patient safety within health care organizations, and became effective in 2003.

areas that have emerged within the health care environment. Some include patient safety and health information technology, nursing shortage issues, further alignment of measures with other measure developers, and collaboration with organizations such as the Occupational Safety and Health Administration (OSHA) and the American Society for Healthcare Engineering (ASHE) to improve patient safety.

The mission of the Joint Commission is “to continuously improve the safety and quality of care provided to the public through the provision of health care accreditation and related services that support performance improvement in health care organizations” (JCAHO, 2005). Today, Joint Commission accreditation is recognized as “a symbol of quality that reflects an organization’s commitment to meeting certain performance standards” (JCAHO, 2005). Accreditation lasts for three years, at which time the organization is required to undergo an on-site survey by an inspection team. Laboratories are required to be re-accredited every two years.

The Joint Commission is governed by a 29-member Board of Commissioners, whose members have diverse backgrounds and experience in health care, business, and public policy. They include physicians, administrators, nurses, employers, a labor representative, health plan leaders, quality experts, ethicists, a former health insurance executive, a consumer advocate and educators. The commissioners serve renewable three-year terms and their charge is to provide oversight of the Joint Commission in relation to its mission. Board meetings occur three times a year, with an additional retreat to focus on the strategic direction of the organization.

Five standing Board Committees also provide oversight on specific areas that the organization is involved in, including the Executive, Accreditation, Standards and Survey Procedures, Finance and Audit, and Governance committees. Other Board Committees include the Board Appeal Review Committee, Human Resources and Compensation Committee, and Strategic Issues Workgroups. These workgroups address issues such as the future value of

accreditation, information technology in health care, performance measurement, public policy, and alignment of quality and payment.

The Joint Commission relies on a number of Advisory Councils that are composed of organizational representatives accredited in specific programs. They provide advice on 1) improving the value of the accreditation program, 2) new products, 3) standards or survey process changes, and 4) environmental influences (JCAHO, 2005). The advisory councils include councils on ambulatory care, behavioral health home care, hospitals, laboratories, long term care/assisted living, performance measures, business/employer priorities, health care safety, core measures, communications with health care professionals, and nurse staffing/shortages.

In addition, seven Professional and Technical Advisory Committees (PTACs) provide guidance for each accreditation program on standard and survey procedure changes. Members include representatives of professional organizations that are nominated by their organization, as well as advocates. There are two associated workgroups with the PTACs. They include a Public Advisory Group that provides guidance on current and evolving health care issues, and the Sentinel Event Advisory Group that reviews and provides guidance on patient safety goals to the Joint Commission.

The Joint Commission headquarters is located in Oakbrook Terrace, Illinois with a satellite office in Washington, DC. In total, the organization employs over 1,000 individuals and has an approximate annual budget of \$138 million. In 2007, Dennis O'Leary announced his retirement, and a new president was named. Mark Chassin, M.D. began his tenure as President of the Joint Commission on January 1, 2008. The activities of the organization are supported by a combination of accreditation fees, as well as the sale of products such as manuals, books, and software, consulting services, grants and contracts.

National Committee for Quality Assurance

Originally formed in 1979, the National Committee for Quality Assurance (NCQA) was designed to “fend off federal monitoring of health plans” (Bodenheimer, 1999: 489). It was reconstructed by a coalition of HMOs and some large employers; the NCQA was transformed from an advocate for HMOs to an organization that focused on quality and quality improvement. Established in 1990 as an independent organization, with support from the Robert Wood Johnson Foundation, the employers and the main trade association for the managed care industry, its stated vision is to “transform health care quality through measurement, transparency and accountability” and its broad, overall mission is “to improve the quality of health care.” It is a private, 501(c) (3) not-for-profit organization, which has played a central role in driving quality improvement within health care. The NCQA currently manages a number of voluntary programs that include accreditation of managed care organizations, preferred provider organizations and disease management, certification programs in such areas as physician organization and utilization management and credentialing, and physician recognition programs that include conditions such as diabetes, heart/stroke, and back pain. It also is charged with the management and evolution of the Health Plan Employer Data and Information Set (HEDIS[®]), which is a performance measurement tool originally developed by The HMO Group. It measures the performance of health plans on a variety of important dimensions of care and service and allows for objective side-by-side health plan comparisons.¹¹⁰

In accordance with its mission and core values, the NCQA began conducting accreditation surveys health plans in 1991, and with employer mandates these quickly became the national standard, with nearly 90 percent of all health plans in the country collecting some

¹¹⁰ Some of the measures included in the set are: advising smokers to quit, antidepressant medication management, beta blocker treatment after a heart attack, breast cancer screening, cervical cancer screening, comprehensive diabetes care, controlling high blood pressure, as well as prenatal and postpartum care.

HEDIS[®] data elements. In 1994, the NCQA released its first Accreditation Status List, which publicly reported the status of all health plans undergoing the accreditation process and in 1996 released its first Accreditation Summary Reports, which provided detailed information about accredited health plans.¹¹¹ To improve the validity and reliability of the reports, the HEDIS Compliance Audit Program was launched in 1997. This program included a standardized methodology for verification of the HEDIS collection and calculation processes. This program paved the way for the NCQA to include HEDIS[®] as a requirement in its accreditation program, which occurred in 1999. Throughout this decade, the NCQA sought to expand its products and services to the health care industry. In 1997, it launched its accreditation program for managed behavioral health care organizations as well as its certification program for physician organizations. In 1999, NCQA announced its plans to accredit preferred provider organizations (PPO), and in 2001 accredited its first PPO.

From 2001 to the present, NCQA has launched a number of initiatives aimed at improving the quality of health care within the U.S. These include new accreditation programs, such as the VA Human Research Protection Accreditation Program, which was developed in partnership with the Department of Veterans Affairs, new websites for consumers, such as Healthchoices (www.Healthchoices.org), and new tools, such as the Quality Dividend Calculator, which measures the financial impact of selecting an accredited health plan.

Under the direction of President Margaret E. O’Kane, NCQA continues to refine and enhance its programs. Health plans that want to receive NCQA’s seal of approval will have to comply with a more rigorous set of standards and must report performance in more than 40

¹¹¹ According to Bodenheimer, a health plan has the option to refuse to publicly disclose its HEDIS results. For example, in 1996 a total of 51 percent (329 HMOs) permitted the disclosure of the data to the public, while in 1997, only 45 percent of the HMOs (292) did. As noted by NCQA, those plans that did not disclose their data had “significantly lower scores than the plans that permit publication” (1999: 489).

areas. According to the NCQA website, “these standards will promote the adoption of strategies that we believe will improve care, enhance service and reduce costs, such as paying providers based on performance, leveraging the Web to give consumers more information, disease management and physician-level measurement” (NCQA, 2007).

Located in Northwest Washington, D.C., the NCQA employs approximately 170 individuals and currently offers a wide variety of programs that are designed to improve the quality of health care. It is advised by a Board of Directors whose members represent a variety of stakeholders. Since its inception, the NCQA has worked with broad coalitions that provide insight and support to its many oversight programs designed to assure that the programs offered are “relevant, scientifically sound, and feasible” (NCQA, 2007).

Its reported revenue for 2006 was approximately \$23 million. NCQA is supported financially through accreditation fees, which make up about 40 percent of its operating budget, and by many other organizations through grants and other contributions. It accepts contributions from corporations and from corporate and philanthropic foundations and provides a complete list thereof on its website. Sponsors range from foundations and government agencies, such as AHRQ and the Commonwealth Fund, to the corporate sponsors such as Baxter and Bristol-Meyers Squibb. Other sources of revenue include contracts, educational programs and publications.

While not a requirement of all employers, “today many Fortune 100 companies will only do business with NCQA Accredited health plans,” and “about three quarters of the nation’s largest employers use HEDIS[®] data to evaluate the plans that serve their employees” (NCQA 2007).

American Hospital Association

The American Hospital Association (AHA) is the national organization that represents and serves all types of hospitals and health care networks, as well as their patients and communities. “Close to 5,000 hospitals, health care systems, networks, other providers of care and 37,000 individual members come together to form the AHA” (AHA 2006). Its mission is “to advance the health of individuals and communities. The AHA leads, represents and serves hospitals, health systems and other related organizations that are accountable to the community and committed to health improvement” (AHA 2006). Originally founded in 1899 as the Association of Hospital Superintendents of the United States and Canada by a small group of hospital administrators, it served as an exclusive club for hospital superintendents to discuss common concerns and interests. Their goal was “to facilitate the interchange of ideas, comparing and contrasting methods of management, the discussion of hospital economics, the inspection of hospitals, suggestions of better plans for operating them, and such other matters as may affect the general interest of the membership” (AHA 2005). However, in 1906, the membership restrictions changed to include associate memberships (without voting privileges), which were awarded to those “next in line to the superintendent.” At the same time, the organization officially changed its name to the American Hospital Association, which remains in effect today and shortened its goal to “the promotion of economy and efficiency in hospital management.” These changes reflected the broader vision of the organization and by 1913, the membership included trustees, medical staffs, and superintendents of nursing. Individual membership was the norm until 1918, when the AHA adopted an institutional membership structure. These changes were an effort by the organization to broaden its mission and to reflect the social and health care environment of the times. In fact, throughout its history, the AHA

periodically revises its mission to reflect the changes in health care and American society, in an effort to better serve its membership.

Much like its goals and mission, the governance structure of the AHA evolved over the years, and included a number of entities that provided direction in membership and policy issues. The House of Delegates was established in 1938 to provide overall direction and governance to the organization. Delegates for each state were, and continue to be, apportioned in the House of Delegates based on the total institutional dues paid in a particular state. In addition, six councils also were established to provide direction to the AHA in the areas of policy and policy development. In 1968, in an effort to improve communication and relationships with the State Hospital Associations, the AHA established Regional Advisory Boards. This structure provided the opportunity for additional meetings, between House meetings, to discuss policy issues. This governance structure remained more or less intact until 1987, when the AHA made significant changes in organizational structure and process through the New Directions and Initiatives Project. This restructuring eliminated the use of standing councils and *ad hoc* committees for policy development. It also replaced the Regional Advisory Boards with the Regional Policy Boards, for the purposes of policy debate and discussion of policy options. And, finally, it designated the Board of Trustees as the final approval authority for policy decisions, rather than the House of Delegates.

The AHA headquarters are located in Chicago, with an additional office located in Washington, D.C. The Washington office, principally responsible for lobbying for the Association,¹¹² consists of the Office of the President as well as staff working in the areas of

¹¹² According to the Associated Press, the AHA spent \$7.3 million in the first six months of 2007 to lobby the federal government including Congress, the CMS, the DHHS, the OMB, the IRS and other agencies on “issues related to Medicare and Medicaid reimbursement, rural hospitals, access to inpatient care and coverage of uninsured patients” (Associated Press Business News-MSN Money, August 17, 2007).

AHA policy, communication, and national advocacy. Under the direction of President Richard J. Umbdenstock, this private, not for profit organization employs approximately 450 individuals and has an estimated budget of \$79 million per year. The activities of the AHA are supported by a variety of revenue sources, including membership dues ranging anywhere from \$500-\$10,000 annually (depending upon the type and size of the member organization), and the sale of products, as well as various grants and contracts.

As envisioned by its leadership, the AHA has three principal functions: standards and research development, education, and advocacy (Lesparre et al. 1998). These functions are reflected in the variety of activities that AHA offers to its members, such as assisting hospitals and other providers to form networks for patient care, conducting research on the structuring and deliver of health care, and producing educational materials. Through its advocacy activities, the AHA “ensures that members’ perspectives and needs are heard and addressed in national health policy development, legislative and regulatory debates, and judicial matters. Our advocacy efforts include the legislative and executive branches and include the legislative and regulatory arenas” (AHA 2006).

Organizational Actors-Tier 2

Actors that appear in this tier are based upon their involvement in the Hospital Quality Network, and although they are important to the industry, according to the key informants interviewed for this study, their involvement is not as active as those that appear in the first tier.

American Medical Association

The American Medical Association (AMA) is the largest private trade association of physicians and medical students in the country. Its mission is “to promote the art and science of medicine and the betterment of public health” (AMA 2007a). Its vision is to become “an

essential part of the professional life of every physician” (AMA 2007a) through advocacy, education and the promotion of ethical standards. The AMA was established in 1847 by Nathan Smith Davis and others at a meeting in Philadelphia. The purpose of this founding meeting, attended by approximately 250 delegates from 28 States, was to raise the standard of medical education. Additional goals of the newly formed association included scientific advancement, standards for medical education, establishment of a code of medical ethics, and improvement of the health of the public. It created, in 1858, a Committee on Ethics that drafted a code of ethics for American physicians. The Association was eventually incorporated in 1897. In 1901, the AMA reorganized and created the House of Delegates, which is the main policy-making body of the Association. The House of Delegates consists of a number of elected representatives and others, as designated in the Bylaws of the AMA. The role of the House of Delegates is to “transact all business of the Association not otherwise specifically provided for in the Constitution and Bylaws, and shall elect the general officers except as otherwise provided in the Bylaws” (AMA, 2007b). Throughout the early 20th century, the AMA continued to play an influential role in establishing standards for physicians and physician practice, including their education, and training. It established the Liaison Committee on Medical Education in 1942 “to maintain standards for medical undergraduate programs and to accredit medical schools in the United States and Canada” (Microsoft®Encarta® Online Encyclopedia, 2007). It also was one of four other medical organizations that established the Joint Commission for the Accreditation of Hospitals, which would eventually become the largest accreditor of health care organizations in the country. Throughout the 1960s, AMA membership continued to grow and consisted of approximately 70 percent of all physicians in practice at that time. As the membership grew, the Association became more involved in political processes on Capitol Hill and formed the American Medical Political Action Committee in 1961 to more effectively represent its interests.

The Association used this vehicle to lobby against the creation of a national health insurance plan for those over 65 years of age. Although not successful in blocking the Medicare and Medicaid legislation, which was signed into law in 1965, it was able to limit the proposed law by limiting the costs covered by it. During the 1970s, the AMA would sponsor a number of health campaigns, including its long-running campaign against tobacco use. By mid-decade however, membership had dropped by approximately 20 percent, as physicians increasingly sought to join organizations that were more relevant to their medical specialties than the AMA.

During the 1980s, the AMA continued its efforts to promote the interests of physicians and their patients, through the adoption of various resolutions related to discrimination against AIDS patients and lobbying against legislation that would affect the confidential doctor-patient relationship. Another issue that emerged during the 1990s had to do with the increasing number of medical malpractice lawsuits filed against individual physicians. The AMA lobbied legislators at both the national and state levels in an effort to try to regulate and limit malpractice awards. However, this effort was ultimately unsuccessful, as the American Bar Association argued that to regulate and/or limit malpractice claims and awards would “take certain legal rights away from individuals” (Microsoft®Encarta® Online Encyclopedia, 2007). Other activities of the AMA during this time included a successful lobbying effort against the Clinton Health Care Reform Plan, campaigns against domestic violence, child and elder abuse, and sexual assault, and support for a women’s right to obtain an abortion.

However, the 1990s also brought controversies and national attention to the AMA. One controversy included the AMA’s endorsement of Sunbeam Corporation health care products, for which the Association was to be paid millions of dollars. After receiving protests from the membership about the organization’s commercial interests and ethics, the AMA was forced to rescind the agreement with Sunbeam. A second controversy involved the dismissal of George

Lundberg, the longtime editor-in-chief of the Association's successful *Journal of the American Medical Association (JAMA)*¹¹³ by an executive vice president of the AMA.¹¹⁴ Most of the membership disapproved of Lundberg's dismissal, and this incident led to debates about AMA leadership and the "journalistic independence" of *JAMA*.

Also in 1999, the AMA, after intense debate within the House of Delegates, voted to establish an affiliated national labor organization for physicians. The Physicians for Responsible Negotiations was formed in response to concerns that "government regulations and health insurers unfairly dominated the health care industry" (Microsoft® Encarta® Online Encyclopedia, 2007). This union is expected to give physicians more negotiating power in their dealings with managed care companies, as well as to provide a unified voice that can better represent physicians' interests during health policy discussions at the federal and state levels.

Today, the AMA is headquartered in Chicago, Illinois, with a satellite office located in Washington, D.C. It has a membership of approximately 240,000 physicians and medical students (approximately 33 percent of the 850,000 practicing physicians).¹¹⁵ Income for the AMA is derived from member dues that range from a \$20 annual student membership to \$420 annually for a full membership, and accounted for approximately \$47 million in 2006 (AMA 2006 Annual Report). Additional income is generated through other products, such as medical books and journals, educational programs, and medical malpractice insurance sales. However, a significant amount of revenue is generated from the sale of physician prescribing data to

¹¹³ This weekly peer-reviewed journal covers general medicine topics and includes other features such as employment opportunity listings, book reviews, calendar of events, case reports, and obituaries. Its circulation is estimated at 361,000 and a subscription is included in the AMA membership dues.

¹¹⁴ Lundberg had been fired over his decision to run an article in *JAMA* about the definition of sex during the impeachment trial of President Bill Clinton in 1999.

¹¹⁵ Some indicate that membership would be even lower if only those physicians who paid full dues were counted. Discounted dues are offered to students, residents, interns, retired or semi-retired physicians and newly graduated physicians who are establishing their practices.

pharmaceutical companies, which accounted for approximately \$33 million of the Association's income for 2005.¹¹⁶

The AMA is governed by several entities, including the Board of Trustees, which is responsible for the vision and direction of the AMA, the House of Delegates, which is primarily focused on policy-making for the organization, and the AMA Councils, which have a key role in the development of health policy statements. There are six AMA Councils, including the Council on Constitution and Bylaws, the Council on Ethical and Judicial Affairs, the Council on Long Range Planning & Development, the Council on Medical Education, the Council on Medical Service, and the Council on Science and Public Health.

Institute of Medicine

The Institute of Medicine (IOM) is a private, not-for-profit organization that was chartered by Congress in 1970 as part of the National Academies.¹¹⁷ It functions in an advisory role as a research organization and its membership is by invitation only. Its mission is “to serve as adviser to the nation to improve health” (IOM 2007). It acts as an independent source to supply policymakers, professionals, and the public with unbiased, evidence-based information on health and health science.

Its mission is to “provide advice that is unbiased, based on evidence, and grounded in science” with an overall aim “to increase the impact of what we do.” The IOM's goals include: to “Wisely select the topics and the portfolio of projects to undertake; conduct work with the highest possible quality, in a timely and efficient manner; communicate strategically and effectively; evaluate what we do from the vantage point of our sponsors, constituents, and

¹¹⁶ Individual physicians can “opt out” of the program through the Physician Data Restriction Program.

¹¹⁷ The National Academies were originally given their charter by President Abraham Lincoln. Besides the IOM, they consist of the National Academy of Sciences, the National Academy of Engineering, and the National Research Council.

audiences; and mobilize the human, physical and financial resources needed for the task (IOM, 2003: 13-15). According to Harvey V. Fineberg, the President of the IOM, “In a world full of complex and conflicting health information, the Institute of Medicine is a beacon to those seeking objective, evidence-based guidance” (IOM, 2007: ii).

The IOM projects are usually complex, and are selected by a five-member Executive Committee. Selected projects rely on the expertise of unpaid professionals to gather information and current research, to participate in the deliberations, and to develop a final report. Individual members are selected from a slate of candidate nominees with the required expertise and perspectives necessary to complete the project. Biases and potential conflicts of interest may disqualify individuals from participation in the committee. Potential committee members must be reviewed by IOM leadership, and receive approval from the IOM President, who then formally requests their appointment from the National Research Council (NRC) Chairman.

The product, usually a report, is developed through a consensus process. It must go through the NRC’s institutional process, which is a structured, formal peer review process, conducted by the NRC Report Review Committee (RRC). This process requires that the findings and recommendations of the project be based on scientific evidence when evidence exists, and then on expert opinion if there is no evidence available. For the most part, meetings are open to the public; however, this is not a requirement, and some committees choose to conduct their deliberations in private. Reviewers for this peer review process are solicited by the RRC staff, and are approved by the NRC. Once selected, the reviewers, who remain anonymous until the study is published, evaluate the report, and provide input, which is then presented to the study staff. The study staff must respond to each recommendation and either make the change requested or provide a rationale for not incorporating the request.

Project sponsors are not permitted to serve on committees that they fund, either in whole or in part. Furthermore, to avoid the appearance of sponsor influence on the findings and recommendations of the committee, sponsors can view the report for up to 10 days before its publication, but are not permitted to comment on it. All IOM reports are released to the public, not to any individual organization.

The IOM conducts studies on a broad and diverse range of topics. They include: mental health, child health, food and nutrition, aging, women's health, education, public policy, health care and quality, diseases, global health, workplace, military and veterans, health sciences, environment treatment, public health and prevention, and minority health.

The IOM is also an invited membership organization. Its members are selected based upon their expertise and accomplishments, as well as the capacity for public service. One fourth of the membership is drawn from fields other than health care and is organized into 12 sections, prearranged by discipline. Members are expected to engage in the activities and work of the IOM, and provide their expertise in IOM studies, workshops, and other healthcare-related activities.

Located in Washington, D.C., under the direction of Harvey V. Fineberg, the President and Chair of the Council of the Institute of Medicine, this organization consists of approximately 1,543 members and is staffed by approximately 157 individuals operating on a budget of \$14,000,000. The IOM receives contributions from both public and private sponsors as well as from its membership, but the majority of the funds are contributed by the federal government through individual contracts. The IOM does not receive direct appropriations from Congress, but will undertake projects funded out of federal agency appropriated funds. Some revenue is generated from the sale of its reports and other educational materials.

It is organized into nine programs: Population Health and Public Health Practice, Health Sciences Policy, Health Care Services, Global Health, Food and Nutrition, Children, Youth, and Families, African Science Academy Development, Military and Veterans Health and Health Policy, Educational Programs and Fellowships. The IOM is governed by a Governing Council that consists of the IOM president and 20 members elected by the membership; all serve three-year terms. This Council provides guidance on policy issues and reviews and approves the program plan and budget.

Organizational Actors-Tier 3

The actors that appear in the final tier are not viewed as being heavily involved in the Hospital Quality Network. This perception may be attributed to their relative “newness” to the network as compared to the actors in the other tiers. Nevertheless, the actors that appear in this tier are an important component of the network, and are becoming more and more prominent in health care.

Consumer-Purchaser Disclosure Project

The Consumer-Purchaser Disclosure Project (the Disclosure Project) is a group of health care purchasing coalitions that represent the purchaser, consumer, and labor organizations perspectives on health care policy issues. Established in 2001, the original group consisted of The National Business Coalition on Health, the Employer Health Care Alliance Cooperative in Madison, Wisconsin, the Pacific Business Group on Health, and the Midwest Business Group on Health. The project’s stated vision is that, with comparative information on performance, “Americans will be able to select hospitals, physicians, and treatments based on national standardized measures for clinical quality, consumer experience, equity and efficiency” (The Consumer-Purchaser Disclosure Project, 2001). The original purchaser groups, realizing that

natural allies such as consumers and labor organizations were likely to have similar views on health policy issues, were sought after and eventually included in the membership of the Disclosure Project, in an effort to advance its agenda of improving the transparency of health care performance.

According to the Disclosure Project, transparency will improve both quality and affordability. Quality improvement, in its view, can be accomplished by “1) consumers using valid performance information to choose providers and treatments, 2) purchasers building performance expectations into their contracts and benefit designs, and 3) providers acting on their desire to improve, supported with better information” (The Consumer-Purchaser Disclosure Project, 2001). The Disclosure Project focuses its efforts on four priority areas. They are: “encouraging development of quality measures relevant to consumers and purchasers; promoting the endorsement of a robust set of performance measures through the NQF; encouraging the use of NQF-endorsed measures supplemented by other qualified measures to fill gaps in NQF measurement sets, to promote improvement; and enhance the availability of data to support public reporting” (The Consumer-Purchaser Disclosure Project, 2001).

The Disclosure Project is an informal coalition consisting of approximately 30 organizations representing purchasers, consumers, and labor and is managed by two co-chairs: Peter Lee, Chief Executive Officer of the Pacific Business Group on Health, and Debra Ness, President of the National Partnership for Women & Families. It is governed by a leadership team consisting of seven individuals representing employers, consumers, and quality and policy experts. The Disclosure Project is largely funded by participating organizations that donate both time and other resources to accomplish the group’s stated goals. In February, 2003, the group moved to Washington, D.C. and was housed in the headquarters of the National Partnership for Women & Families; it received financial support from the Robert Wood Johnson Foundation.

Since its inception, the Disclosure Project has established itself as a leading voice for advancing the consumer and purchaser interests of transparency in the health care arena. Since 2003, the Disclosure Project has produced a number of publications advocating and publicizing the consumer-purchaser aspect of key issues in health care, such as measurement, affordable health care, and the use of electronic data in measuring performance. It's most recent accomplishment involved an agreement between the New York Attorney General and Cigna, relating to the health plan's physician ranking program.¹¹⁸ With the strong advocacy of the Disclosure Project and other consumer and purchaser groups, what could have potentially been a step backward for the transparency movement, was turned into an advance by establishing standards for accuracy, disclosure, transparency and fairness in physician comparisons. The agreement between the Attorney General's Office and Cigna, announced on October 29, 2007, incorporates the principles and guidelines developed by the Disclosure Project, which include disclosure of the how the program is designed, criteria for ranking physicians, and a process to appeal incorrect ratings. According to Peter Lee, co-chair of the Consumer-Purchaser Disclosure Project, "Today's agreement is a great model for how health plans should conduct physician measurement and reporting programs. Employers across the country can look to this agreement as a framework for how to assure that their health plans are meeting the needs of their employees" (Consumer-Purchaser Disclosure Project, 2007).

American Federation of Labor and Congress of Industrial Organizations

The American Federation of Labor and Congress of Industrial Organizations (AFL-CIO) is a national trade union center, and is the largest voluntary association of unions in the country. It consists of approximately 55 national and international labor unions and represents about 10

¹¹⁸ Attorney General of New York Andrew Cuomo began an investigation into physician ranking programs and practices of payers, citing concerns about "potentially deceptive programs" that were "designed to steer patients to the cheapest, but not necessarily the best doctors, letting profits trump quality" (Porter, 2007).

million people in the workforce, such as teachers, engineers, bakers, farm laborers, pilots, doctors and nurses, to name a few. Created in 1955 by the merger of the American Federation of Labor and the Congress of Industrial Organizations, its stated mission is “to bring social and economic justice to our nation by enabling working people to have a voice on the job, in government, in a changing global economy and their communities” (AFL-CIO, 2008). To achieve its mission, the AFL-CIO works to accomplish four goals: To build a broad movement of American workers by organizing workers into unions, to build a strong political voice for workers, to change unions to provide a new voice to workers in a changing economy, and to change the labor movement by creating a new voice for workers in the community (AFL-CIO, 2008).

The AFL-CIO is governed by delegates elected by the members of the federation. The delegates meet every four years to set policy and goals for the labor movement as well as to elect officers and vice presidents to govern the daily work of the federation. The current elected officers, each with a term of four years, include John J. Sweeney, president; Richard Trumka, secretary-treasurer; and Arlene Holt Baker, executive vice president.¹¹⁹ These officers, and 44 vice presidents elected by the membership, form the Executive Council, which is required to meet twice a year. The duty of the Council is to provide overall guidance to the officers with respect to the general daily work of the federation, but specifically this group oversees the legislative program for the organization. The Executive Committee, formed after a constitutional amendment in 2005, consists of the president, the vice presidents from the ten largest unions, nine other vice presidents selected by the Executive Council, and two ex-officio members. This Committee governs the organization between Executive Council sessions and approves the budget and issues charters. The AFL-CIO also has a General Board that assists with the

¹¹⁹ The executive vice president office was established in 1995 by a federation constitutional amendment.

governance of the organization. It consists of the Executive Council, the Chief Executive Officer of each member union, the President of each Constitutional Department of the AFL-CIO,¹²⁰ and four regional delegates elected by the AFL-CIO state federations. Their duties are limited, and they only take up issues that are referred by the Executive Council; for example, the General Board provides the AFL-CIO endorsement for candidates running for the Office of the President and Vice-President of the United States.

At the state and local level, the AFL-CIO charters and organizes state federations and central labor councils. The 51 state federations typically focus on state economic issues and state legislative lobby efforts, while the over 500 central labor councils tend to focus on county or city issues related to elections, zoning and the like. Each organization has its own charter that determines its jurisdiction, governance structure, mission, etc. These bodies are permitted representation and possess voting rights at the quadrennial convention. They generally work to organize and mobilize their members around collective bargaining campaigns, lobbying, strikes, picketing, and boycotts.

Located in Washington, D.C., the AFL-CIO headquarters is largely funded by union membership dues. Its partisan political activities, however, are supported by voluntary donations from the membership. The organization has become more active in health care as issues such as cost and quality have come to the forefront in negotiations among union members and employers.

¹²⁰ These are departments mandated by the AFL-CIO constitution and are governed by Article XII. The departments are autonomous in that they may have their own constitution, membership, officers, governance structure, dues and organizational structure, as long as they conform to the AFL-CIO policies and constitution. Currently, there are six departments: Building and Construction Trades Department, Maritime Trades Department, Metal Trades Department, Department for Professional Employees, Transportation Trades Department, and Union Label Department.

Hospital Quality Alliance

The Hospital Quality Alliance (HQA), formerly known as the National Hospital Voluntary Reporting Initiative, is a public-private collaboration consisting of a number of organizations, including hospitals, purchasers, government agencies, health plans, quality groups, and consumers. The purpose of the HQA is to “encourage hospitals to voluntarily collect and publicly report quality performance information in a consistent, standardized way” (HQA 2002). This effort is intended to make hospital performance information available to the public and to advance the quality improvement agenda for hospitals.

On December 12, 2002, the leaders of the AHA, AAMC and the Federation of American Hospitals (FAH) announced a collaborative designed to partner with government, the Joint Commission, NQF and others, in an effort to align their performance measurement and improvement efforts. Charter organizations of the HQA included the AHA, AAMC, FAH, CMS, AHRQ, the Joint Commission, AMA, AARP, AFL-CIO and NQF. Through a collaborative process, the HQA envisions that “more efficient and equitable care at all levels of a patient’s experience” (HQA, 2007). Its mission is to “facilitate continuous improvement in patient care through implementing measures that portray the quality, cost and value of hospital care; develop and use measure reporting in the nation’s hospitals; and to share useful hospital performance information with the public” (HQA 2007). Its goals are to 1) “provide meaningful information to the public on hospital quality by working together; 2) align quality measurement sets e.g., give hospitals predictability and streamline data collection; and 3) to support quality and safety improvement” (N.Foster, personal communication, May 2006).

The structure of the HQA consists of a collaborative network that uses consensus to make decisions with regard to issues that affect the reporting of hospital performance in a public forum. In light of its selected structure, the HQA did not establish any bylaws, or require the

financial support of its membership, although resources such as time, materials, staff expertise, and funding were donated by several of its members. This core group, known as “the Principals” and chaired by the AHA president, met on a periodic basis to discuss relevant issues to the industry, such as hospital recruitment for participation in public reporting projects sponsored by CMS and to plan the rollout of hospital performance data to the public.

An important aspect of the HQA is to report credible and user-friendly information about the quality of care delivered by the nation’s hospitals. Originally, hospitals were asked to submit performance data on three common conditions—heart attack, heart failure and pneumonia, and in October 2003, these data obtained for a starter set of 10 measures were displayed for the first time on a CMS web site intended for clinicians and providers.¹²¹ However a more consumer-friendly version of the CMS website was eventually developed and *Hospital Compare* (www.hospitalcompare.hhs.gov) debuted on April 1, 2005. Since that time, the HQA has worked collaboratively with its partners to add measures in other areas such as surgical care, mortality, and patient experiences of care i.e., HCAHPS.¹²²

HQA has also expanded the Principals to include other organizations such as the American Nurses Association, the National Association of Children’s Hospitals and Related Institutions, the Consumer-Purchaser Disclosure Project, the U.S. Chamber of Commerce, the National Business Coalition on Health, America’s Health Insurance Plans, the National Association of Public Hospitals, and the Blue Cross Blue Shield Association. The HQA is considering additional partners that may represent areas that currently do not have a voice in the

¹²¹ For the first data display on the CMS website, the number of hospitals that volunteered for this initiative was approximately 450 out of the 4,000 hospitals eligible to report. However, when Congress passed both the Medicare Modernization Act and the Deficit Reduction Act, which tied hospital market basket updates to quality reporting, the number of hospital and the number of measures they were reporting vastly increased, and today essentially all hospitals provide some performance data for public reporting on the website.

¹²² This is a consumer survey that assesses patient experiences in certain areas such as doctor communication, hospital cleanliness, pain control, communication about medicines, and discharge information. The results of this survey are projected to be published for participating hospitals on the *Hospital Compare* website in 2008.

collaboration, such as the Society for Critical Care Medicine. Further changes in this collaboration have included the development of a dues structure for membership to support the activities of the HQA. Funding from the collection of dues will be used to support a staff person who will work exclusively on HQA issues and projects. An additional development includes opening the HQA quarterly principals' meetings to the public, as well as inclusion of a public comment period during which individuals can provide input into the decision-making of the collaborative.

Over the past few years, a number of Quality Alliances has been established using the HQA as a model.¹²³ HQA also has begun some transformational activities, including commissioning a white paper entitled “Envisioning the Roadmap for National Hospital Quality Reporting” that reported the opinions of the health care leaders with respect to hospital quality reporting and developing a roadmap for a single reporting system. The paper indicated that there is a shared vision about the need to access meaningful information to inform decision-making and that the most formidable obstacle is the lack of consensus around a model for quality measurement. The report recommends that a time-bound roadmap for each segment of the health care industry be developed and executed and maintains that the commitment and participation of all stakeholders in health care is crucial: “Understanding that no single stakeholder should control the entire hospital quality reporting system, and that consensus among the stakeholders is key to simplifying the nationwide need for quality reporting, industry leaders must step up their commitment and participation in a nationwide public-private partnership to gain and then sustain momentum for change” (Booz, Allen, and Hamilton, 2006: 1).

¹²³ These alliances include the AQA, formerly known as the Ambulatory Quality Alliance; the Pharmacy Quality Alliance; the Alliance for Pediatric Quality; the Cancer Quality Alliance; the Surgical Quality Alliance; and the Kidney Care Quality Alliance.

Summary

This chapter has mapped the human actors of the Hospital Quality Network (HQA), describing those that, through my research, have emerged as the first or primary tier of actors within the field of performance measurement, and the less-mentioned actors described in the second and third tiers of the network. The next chapter will discuss the non-human actors and their relevance to the network.

CHAPTER 5:

THE NON-HUMAN ACTORS OF THE HOSPITAL QUALITY NETWORK

Also found within the Hospital Quality Network are non-human actors that influence all the other actors in the network, thus providing input that can change each actor's ideas about how to construct the world of performance measurement in health care.

John Law provides an example:

I am standing on a stage. The students face me, behind serried ranks of desks, with paper and pens. They are writing notes. They can see me, and they can hear me. But they can also see the transparencies that I put in the overhead projector. So the projector, like the shape of the room, participates in the shaping of our interaction. It mediates our communication and it does this asymmetrically, amplifying what I say without giving students much of a chance to answer back (Thompson, 1990). In another world it might, of course, be different. The students might storm the podium and take control of the overhead projector. Or they might, as they do if I lecture badly, simply ignore me. But they don't, and while they don't the projector participates in our social relations: it helps to define the lecturer-student relationship. It is a part of the social. It operates on them to influence the way in which they act (Law, 1992a : 2).

What differentiates ANT from other theories is the issue of the non-human actor, which influences and interacts with other actors in the network. Non-human actors contribute to the interactions in the network, which in turn influences the final order of the network, and thus, may change the network effect. And because they participate in the interactions, they become part of the politics that play a role in developing the network order.

I have identified three non-human actors within this network—the NQF consensus development process, public reporting and report cards, and performance measures themselves. These are not the only three non-human actors within this network; I chose these three actors based on the themes that emerged as I conducted interviews with the industry’s key informants for this study.

The NQF Consensus Development Process

Origins

With the publication of the President’s Advisory Commission Report to the Clinton White House, the planning of what was to become the National Quality Forum (NQF) was underway. From this report, a recommendation emerged that specifically called for the creation of an entity to “develop and implement effective, efficient, and coordinated strategies for ensuring the widespread public availability of valid and reliable information on quality” (AHRQ, 1998), thus implying standardization of measures via agreement of stakeholders and the coordination of quality improvement efforts at the national level.

With this in mind, the Planning Committee determined the main objective of the NQF would be to “develop consensus among health care’s diverse stakeholders about standardized indicators that can be used to measure and report on health care quality” (NQF, 2000a). From the beginning the consensus development process (CDP) was, by design, to be reviewed and redesigned as new knowledge and experience was acquired. To date, the formal CDP has been revised eight times, the last of which was in fall 2007.¹²⁴ This process is specifically aimed at activities that will ultimately lead to an NQF-endorsed measure or sets of measures that, in the

¹²⁴ Prior to 2007, the last revision of the CDP (version 1.7) was in August 2004.

end, will promote standardization; this will allow for comparisons among health care entities as well as decrease the burden of reporting for health care organizations. Other products that are expected from this process include explanatory text and supporting documentation, such as guidelines from reporting indicators and standards.

The National Technology Transfer and Advancement Act

The NQF CDP is a voluntary consensus process that has its roots in the National Technology Transfer and Advancement Act (NTTAA)¹²⁵ of 1995 and the Office of Management and Budget (OMB) Circular A-119-Federal Register (63 FR 8545).¹²⁶ Specifically, the NTTAA defines voluntary consensus standards-setting organizations¹²⁷ and provides guidance to the National Institute of Standards and Technology (NIST), which is charged with implementation of the Act. In short, the Act directs NIST to coordinate with other governmental agencies, including federal, state, and local governments, to increase the reliance on standards developed by voluntary consensus and to promote the use of private sector standards. The OMB Circular A-119 outlines detailed requirements for federal agencies under the NTTAA. These documents direct federal agencies to use consensus technical standards developed by voluntary consensus standards bodies such as the NQF, in lieu of government-unique standards. If a federal agency chooses not to use standards created by a voluntary consensus process, it must explain its reasoning to the OMB. Reasons an agency may choose not to use a voluntary consensus standard are that use of the voluntary consensus standards is inconsistent with law or is otherwise

¹²⁵ Pub L No 104-113, Stat 15, USC 3701.

¹²⁶ OMB Circular A-119 (1998) is based on 31 U.S.C. 111, which gives OMB broad authority to establish policies for the improved management of the Executive Branch.

¹²⁷ Voluntary consensus standards bodies are domestic or international organizations that plan, develop, establish, or coordinate voluntary consensus standards using agreed-upon procedures. “A voluntary consensus standards body is defined by the following attributes: openness; balance of interest; due process; an appeals process; and 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 objections and the reasons why, and the consensus body members are given an opportunity to change their votes after reviewing the comments” (OMB Circular A-119, revised February 10, 1998).

impractical. These documents also provide guidance to federal agencies that choose to participate in voluntary consensus standards body deliberations and describe procedures for reporting on government-unique standards when used in lieu of voluntary consensus standards. Voluntary consensus standards can be identified through a series of databases maintained by NIST. If no voluntary consensus standard exists, however, the agency may use a government-unique standard and is not required to file a report.

The NTTAA and OMB Circular A-119 are intended to achieve the following goals:

- Eliminate the cost of developing government standards and decrease the cost of goods procured and the burden of complying with regulation;
- Provide incentives and opportunities to establish standards that serve national needs;
- Encourage long-term growth for U.S. enterprises and promote efficiency and economic competition through harmonization of standards; and
- Further the policy of reliance upon the private sector to supply government needs for goods and services (OMB Circular A-119, 1998).

As also outlined in these documents, federal agencies are encouraged to participate in voluntary consensus standards bodies' discussions and to assist in the development of standards when it is in the public interest to consult and when it is compatible with their missions, authorities, priorities and budget resources. The purpose of participation is to eliminate the need for development and maintenance of government-unique standards, and to further the national goals of increased use of a standardized metric, to use environmentally sound and energy efficient products and practices, and to improve public health and safety (OMB Circular A-119, 1998). The agency may provide staff to participate in meetings and workshops as well as provide direct support in areas such as financial, administrative and/or technical support, and assistance in the identification and development of standards where none currently exist.

Limitations for federal participation focus on maintaining the independence of voluntary consensus standards bodies. Federal agencies should refrain from becoming involved in internal management decisions; they should not dominate the proceedings of such bodies; and agency personnel “must avoid the practice or the appearance of undue influence relating to their agency representation and activities” in relation to voluntary consensus standards bodies (OMB Circular A-119, 1998).

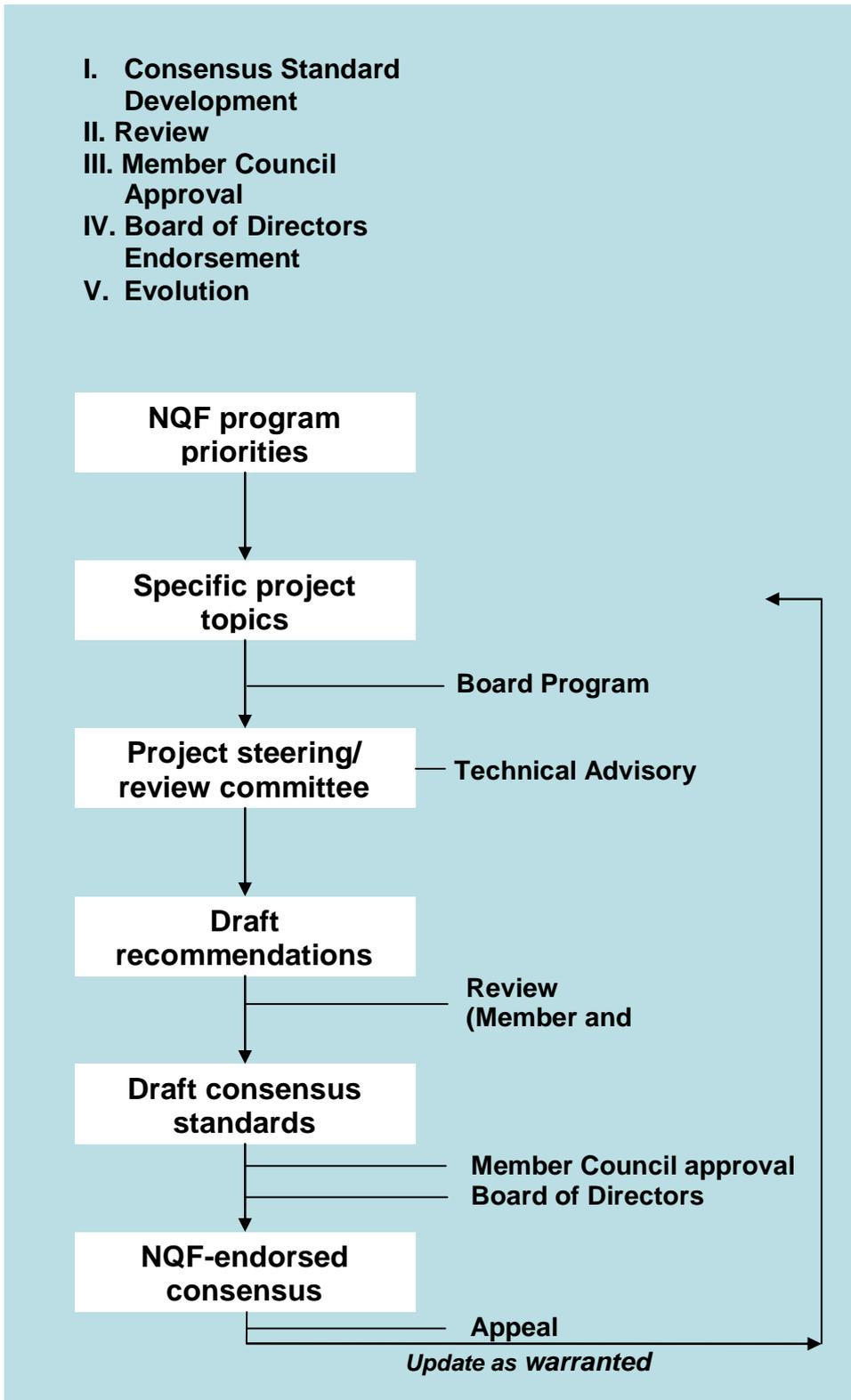
The NQF Consensus Development Process-Version 1.7

The NQF Consensus Development Process (CDP) rests on five principal steps that have been refined as the organization gains more experience and knowledge with the consensus process for standards setting. (For the differences between the earlier versions and the current 1.8 version, see Appendix C). The five steps are: consensus standard development, review, member council approval, board of directors’ endorsement, and evaluation. An overview of the NQF Consensus Development Process is presented in Figure 7. As indicated by the figure, both public and member involvement is a key component of the process. All NQF products undergo at least three levels of review and approval, including review by the projects’ Steering Committee, by NQF Council members and the general public, and by the NQF Board of Directors (NQF, 2004: 3). Additional reviews may be required from the Strategic Advisory Council as well as other organizations, based upon the specific issues or needs of a project.

This process speaks to the five elements that characterize a voluntary consensus process as designated in the NTTAA. These elements are openness, balance, due process, consensus,¹²⁸ and an appeals mechanism (NQF, 2004).

¹²⁸ In OMB Circular A-119, consensus is defined as “general agreement, but not necessarily unanimity” and also includes a process aimed at the resolution of objections made by interested parties.

Figure 7: NQF Consensus Development Process-Version 1.7



Source: NQF Consensus Development Process-Version 1.7 (2004).

Consensus Standard Development

To begin the process, each project is assigned to an NQF staff person, who is the principal contact for the project. Each project is guided by a Steering Committee (SC),¹²⁹ composed of representatives from the four diverse Member Councils, whose general purpose is to “develop specific project plans, provide advice about the subject, ensure input from relevant stakeholders, and review draft products” (NQF, 2004: 5)

Representation from outside the membership will be sought if the needed perspective or expertise is not available from within the Member Councils. Selection of members is accomplished through NQF staff review of nominee qualifications, with the final composition being selected by NQF management, which gives preference to individuals whose organizations are members of NQF. Technical Advisory Panels (TAPs) will also be involved in a project, particularly to provide technical expertise and experience on a topic. They review and evaluate the evidence supporting the measures presented for endorsement, and advise the Steering Committee as to the soundness of the scientific evidence base underlying the measures. In addition to these committees, when indicated and when tied to the NQF’s quality measurement framework, the advice of the Strategic Advisory Council (SAC) may be sought. Once all input is assimilated and consensus is obtained, a draft report, including recommendations, will be approved by the SC prior to becoming available to the membership or general public. A “Commentary” section is included in the report so that SC members may express any dissenting views or perspectives that were not covered in the report, or may emphasize the importance of certain points in the report itself. Project information, such as a list of representatives on the SCs

¹²⁹ The NQF also uses Review Committees in primarily the same capacity as Steering Committees, with the difference being that a Review Committee oversees minor projects or the updating activities related to endorsement.

or TAPs of projects, a schedule of meeting dates, and draft products for public review is made available on the NQF website.

Review of Draft Products

The review process begins with a pre-voting member review, with written comments forwarded directly to NQF. Comments received will be posted on the website and “formally considered before voting commences” (NQF, 2004: 7). Membership Councils are expected to consolidate their position on candidate measures, and the leadership of each Member Council represents that position to the Board of Directors. Once the draft report or product is made available to the membership, it becomes available to the general public for review and comment. Comments from organizations that are not members of NQF are forwarded to the appropriate Member Council for consideration in their discussions. Comments received are reviewed by staff who may, based on the comments received, revise the standard and re-circulate it to the Member Councils and SC. The revised drafts often appear with a summary of the changes as well as the source of the comments that warranted the revisions.

Member Council Approval

All members in good standing with the NQF are afforded the opportunity to vote on any consensus projects. Ballots are sent via confirmed delivery to the designated organization’s representative. The voting cycle for the first round usually is 30 days. Ballots have a specified deadline for return. Votes are tallied, the majority determining the outcome vote of a Member Council. Votes may be submitted with proposed modifications or conditions. For example, an organization may vote “yay” if the following conditions are met, otherwise, the vote will become “nay.” If all four Member Councils approve a standard without revisions after the one round of voting, then it will be forwarded to the Board of Directors for consideration of approval. If one or more Member Councils is unable to obtain agreement, then NQF staff work to resolve the

issues and submit the revised draft product to all Member Councils for a second round of voting, which is at least 14 days in duration but can not exceed 21 calendar days. After a second round of voting, only two Member Councils out of four need to approve the product for it to move forward to the Board of Directors for consideration.

Board of Directors Endorsement

Products approved by the membership are forwarded to the Board of Directors for either endorsement or re-consideration. “NQF endorsement of voluntary consensus standards will not be considered to have been achieved until the candidate standards/draft products have been approved by consensus of the NQF membership and endorsed by the Board of Directors” (NQF, 2004: 10). Decisions of the Board of Directors may be appealed in writing within 30 days of a decision. Written documentation must include “information clearly demonstrating that the appellant has interests that are directly and materially affected by the NQF-endorsed voluntary consensus standard(s), and the NQF decision had (or will have) an adverse effect on those interests” (NQF, 2004: 10). Appeals are handled by NQF staff in consultation with the project’s SC or TAPs, or other entities as deemed appropriate. Staff forwards the appeals package, as well as their recommendation to the Board, which the Board acts upon, and the decision is then made available to the membership and the general public.

Evaluation

Once the NQF-endorsed Voluntary Consensus Standards are promulgated, the NQF anticipated that issues may arise after widespread implementation. NQF has developed a mechanism to report implementation issues that arise. This information is forwarded to the measure developer/sponsor of the standard for consideration and response. This is part of the ongoing effort to assure that NQF-endorsed voluntary consensus standards are continuously evaluated and refreshed. Another mechanism to ensure standards maintenance involves the

Standing Working Groups (SWG), also known as Maintenance Committees. These SWG propose recommendations to the Board of Directors with regard to changes to standards that either do or do not require re-consideration through the CDP. Other recommendations involve user-recommended changes to an endorsed standard that a measure developer declines to make, thereby making the measure a potential candidate for endorsement withdrawal. In any event, the NQF staff, by design undertakes a “formal review of voluntary consensus standard-specific implementation issues for NQF endorsed products within 12 to 24 months of endorsement” (NQF, 2004: 12).

Key Issues of the Consensus Development Process

Several themes emerged during discussions with the interviewees about the NQF Consensus Development Process (CDP). But while they made specific criticisms of the CDP, the majority of those interviewed indicated that “if the NQF were to go away tomorrow, a similar organization would emerge to take its place.” Essentially, the CDP provides opportunities for multiple stakeholders, as well as the general public, to participate and provide input into the process. It does not mean that the multiple voices heard during the process will necessarily be incorporated into the final product or that just because there is an opportunity to be heard, that the constituency has felt that it was heard. The NQF has described its CDP as a “living process,” which has been reviewed and revised periodically to reflect the current needs of the industry as well as to increase its effectiveness. Nevertheless, certain issues have been raised with regard to this process, which often is used in lieu of federal rulemaking.

Identification of Project Topics

Although the Strategic Framework Board laid out a framework to guide the NQF with regard to project selection, some feel that it is not specific enough to guide prioritization. Projects that are selected should reflect the endorsed NQF priorities for measurement and

reporting which is often not the case. One reason is that the government entity that was suggested by the President's Commission was never established, probably due to lack of interest by Congress and the Administration. This governmental body's main purpose was to act as a coordinator among quality improvement organizations and to establish national goals for health care quality improvement. For this reason, as well as the lack of a steady stream of funding for the NQF, project initiation, selection and prioritization is often opportunistic, relying on funds from various public and private organizations with specific agendas and goals,¹³⁰ rather than on what is strategically pertinent to the industry and the nation. According to a recent IOM report, NQF's current funding structure may undermine its effectiveness: "NQF's projects tend to reflect the priorities of available funding sources rather than addressing a discrete set of national goals. As a result, a comprehensive strategy for the development and implementation of health care performance measures does not yet exist" (IOM, 2005: 40). What are needed are national goals that can be agreed upon by the various stakeholders, and that can be prioritized to help advance the health care quality agenda. NQF may act to fill this void; however coming to consensus on a set of national goals for health care, particularly with the diverse interests of the stakeholders in play, may not occur quickly enough to be beneficial.

Development of Consensus Standards

For each project initiated by the NQF, a Steering Committee is selected and an NQF staff person is assigned. The Steering Committee consists of at least one member of each of the four Member Councils, as well as others solicited through a call for nominations. The SC sets the direction of the project and makes the ultimate decision to forward measures to the membership

¹³⁰ NQF receives funds from the private and public sectors. It currently has an \$8 million dollar budget, of which 40 percent is derived from member dues; three percent is from report sales, registration fees, and other sources; and the balance is from grants and contracts (seven percent from federal sources, while 50 percent is from private foundations) that are linked to a specific project. More recently, NQF received \$3 million from the Robert Wood Johnson Foundation to begin an endowment.

and the Board. It is advised on technical issues by a TAP as appropriate. Some interviewees suggested that the evidence base for some projects was not well-developed, particularly with respect to the measures, the criteria for inclusion, and the rationale for recommendations.

Although the most recent practice of conferring with TAPs on technical issues, thus assisting with better informed decisions on measures, the SC decision-making is less structured. The interviewees suggest that this is in part due to the SC's make-up, which is designed to reflect the Member Council balance. In addition, the conflict of interest issue has been of concern, due to NQF's practice of including measure developers on SCs and TAPs. Along these same lines, some concern was expressed about how members of the SCs and TAPs are selected. Typically, there is a call for nominations, and preference is given to member organizations first. If there are no "qualified" members that represent a particular point of view, the NQF staff will consider nominations outside the membership. Staff respondents indicate that they strive for balanced representation from all Member Councils as well as gender, geographic, and racial and ethnic representation. The ultimate decision as to who will be on the SCs and the TAPs is the President and CEO's.

Public Review

As part of the consensus development process, proposed consensus standards are made available for public review and comment. The comment periods typically last for 30 days. After that time, staff collect, analyze and respond to comments, making revisions in consultation with the SC. Any revised drafts of the standards or recommendations include a summary of significant comments, and NQF responses to the comments relating to any action taken. One concern raised by the interviewees was that multiple drafts and review processes are inefficient, and they often expressed the view that the NQF CDP is very lengthy and time consuming, and that multiple reviews and revisions add significant time to the process without producing much

value in return. Another suggested that NQF staff have too much power and control over the process. The staff researches the measures or standards, then writes them up and presents them to the SC. They also participate in the SC and TAP discussions. Staffers also have the ability to determine which public comments to act on or not to act on, without providing justification for their decisions, and in some cases, do not follow up with the Steering Committee for their approval.

Member Voting

Members of the NQF in good standing may vote on proposed consensus standards or measures. A quorum within each Member Council is not required, and a simple majority approval in each Council may send a proposed consensus standard to the Board. If all four Member Councils fail to approve a proposed item, staff work with members and measure developers to modify the standard, if appropriate, and a second ballot is forwarded to the membership for a vote. In this case, if two of the four Member Councils approve the revised consensus standard, the item is forwarded to the Board for a vote. One issue that has been raised by interviewees suggests that there is unequal representation among the Councils. Many providers believe that the provider/health plan council is too large, and that physicians should have more representation as a Council itself, thus paving the way for representation on the Board, and on SCs and TAPs. It should be noted that not all members agree with this concern. For instance, the Purchaser Council recently voiced concern that the SC and TAPs are overly representative of providers. By design, the NQF was formed to give an equal voice to all stakeholders in health care, but specifically to include consumers and purchasers that were often not invited into discussions relating to health care policy.

Board Endorsement

Once a proposed consensus standard is approved by the membership, it advances to the Board for either endorsement or is returned to the membership for reconsideration. For an item to be endorsed as a consensus standard by NQF, a quorum of Board members is required for all Board actions. All Board actions are made public via various mechanisms, such as the NQF website and member updates. There is a window of opportunity for the public and NQF members to appeal decisions of the Board. Once an appeal is received, the Board must act on it within 60 days. To date, the NQF Board has received appeals from the various stakeholders, but has not overturned any of its decisions.

Concerns have surfaced about the composition and the size of the Board. Some interviewees believe that it is too large to make decisions and be fully effective. The Board currently consists of 25 voting members, seven liaison or non-voting members, and three officers (the NQF President and CEO, a treasurer, and secretary). The make-up of the Board has also raised concerns among stakeholders. Once again, the physician community voiced concerns that there is inadequate representation of physicians within the process. To clarify, it is not so much physicians *per se*, as the Board currently has 11 members that are M.D.s, but more to the point, key physician organizations and specialty representation. Currently, physician representatives of organizations reside either in the Research and Quality Improvement Council or in the Health Professional, Provider and Health Plan Council, which tends to defuse their power somewhat. The American Medical Association occupies a liaison seat on the Board, as a non-voting member. On the other hand, some think relatively few consumers and purchasers representatives are involved throughout the CDP. To date, the composition of the Board consists of five members representing the consumer perspective; five members representing the purchaser perspective; four members representing the provider/health plan perspectives; three members

representing the research and quality improvement aspect; and seven members of the Board that are not members of the NQF and are not classified in the Member Council structure.

Evaluation

NQF solicits feedback about consensus standards and their implementation, which it then provides to the measure developer for consideration. These themes consistently emerged throughout my interviews with key informants, as issues relating to the NQF consensus development process.

Recently, the NQF has established Consensus Standards Maintenance Committees charged with maintaining and evaluating endorsed standards and measures. Of concern is the process by which measures are refreshed and by which feedback from users is evaluated and incorporated into the process. Some suggest that the process needs to be more developed and uniform criteria should be established. In this view, user feedback should be analyzed across projects to determine trends and cross-cutting issues that relate to the measurement sets. One respondent recommended that substantial issues should be communicated to those charged with goal setting for measurement at the national level for evaluation.

Performance Measures and Performance Measurement

The second non-human actor identified within the Hospital Quality Network is performance measures and performance measurement. Measuring performance is the first step toward improving the quality of care delivered in a hospital. In essence, what gets measured gets done. The findings of a Veterans Administration study conducted by senior researcher Elizabeth McGlynn of the RAND Corporation showed that the VA scored well on what was measured, but not so well on what was not. In other words, providers tend to perform to meet the measures, to

the detriment of other initiatives, which is why it's important to carefully select what is measured. Measures should be aligned with national goals for quality improvement, if national goals are available. Currently, no national body sets the goals for improvement in the health care system in the United States, although there have been calls to establish one.¹³¹ And as a result, performance measures have been developed in an *ad hoc* fashion. "Progress in measurement has been uneven across different settings, populations, and health conditions. The task now is to develop a system of performance measurement that is more complete and reliable, that fills gaps in difficult-to-measure areas and for hard-to-reach populations" (IOM, 2005: xi).

In the larger world of performance measurement, organizational performance has been a central theme in research for quite some time (Rousseau 1997). This body of literature is growing exponentially, due to increased competition in the global marketplace as well as advances in the performance measurement field. Other reasons for this increased interest include the changing nature of work, increased competition among firms, specific improvement initiatives, national and international awards, changing organizational roles, changing external demands, and the power of information technology (Neely, 1999: 210).

Also emerging from the performance measurement field is a newer, smaller body of literature—an institutional perspective that draws upon sociological institutional theory to examine performance. "Performance assessment can be analyzed as a social invention or as an institution" (Meyer, 1994: 559). There are three themes relating to performance and institutional theory. The first is that rationality and efficiency are exceptions rather than the norm,

¹³¹ The President's Advisory Commission report that established the NQF also called for the creation of a federal entity to set national standards. The IOM report, published in 2005, is the first in a series of reports that also called for the formation of a national body to identify and set national goals for health care system improvement.

particularly in institutionalized organizations.¹³² Second, in those organizations that are institutionalized, performance is assessed based on the “logic of confidence and good faith” (Meyer, 1994: 560). In this case, rules and procedures are more important than their outcomes. Third, legitimacy replaces efficiency outcomes in institutionalized organizations. The concept of legitimacy pertains to “societal evaluations of organizational goals” (Scott, 1991: 169). In other words, the organizational values must be consistent with society’s values to receive legitimacy. Legitimacy, in turn, provides the organization a claim to resources necessary for survival. These propositions about performance remain vague, because there is little consensus regarding the definition of the key concepts of institutional theory (Tolbert and Zucker, 1999; Meyer, 1994).

Performance measurement’s entry into industries and professions that were previously impervious to evaluation can be explained by institutional theory. The emergence of performance measurement can be linked to certain environmental events, such as regulations requiring the use of performance measures, to imitation of industry leaders, or to organizations seeking legitimacy. But the social mechanisms by which performance measures grow and diffuse are relatively indistinct. Formal diffusion of measures through networks of experts and rule-making bodies can be assessed; however, influence of informal mechanisms is often hard to follow.

Categorizations of organizations also play a role in performance measurement. The category of “schools,” for example, is understood by organizational characteristics. Schools that have expected characteristics, such as processes and organizational design, have greater legitimacy than ones that do not (Meyer and Zucker, 1989). However, there is a critical

¹³² According to Scott and Meyer, “institutional sectors are characterized by the elaboration of rules and requirements to which individual organizations must conform if they are to receive support and legitimacy from the environment” (1983: 140). In contrast, “technical sectors are those within which a product or service is exchanged in a market such that organizations are rewarded for effective and efficient control of the work process” (1983: 140).

connection between organizational categorization and performance measurement, according to Meyer. Measurement results are not comparable unless the measures and the units being measured are comparable (Meyer, 1994). Categorization identifies and groups organizations for comparison as well as suggests which types of measures should be used for evaluation. In this sense “processes akin to institutionalization are preconditions to performance measurement” (Meyer, 1994: 564).

Performance measurement in health care is in a relatively early stage of development (Lee, 1996; Eddy, 1998), and is particularly difficult to develop, due to the multiple and conflicting goals of health care organizations (Meyer, 1994). Despite its importance, there is no good way to measure how the industry is doing. Most of the health care quality performance measurement literature is based upon the structure, process and outcome model developed by Avedis Donabedian. While other scholars such as Brook¹³³ (1973) sought to improve this model, Donabedian’s model remains the most useful and influential in assessing and measuring health care quality.

The model draws attention to three dimensions or levels in assessing health care interventions: structure, process, and outcome (Donabedian, 1966). “Structure” refers to health system characteristics and capacity. It includes facility and equipment adequacy, medical staff qualifications and organization, program structure and operation, and fiscal organization (Donabedian, 1966, 1980). It also can include patient characteristics, such as illness profile, and community characteristics like environmental risk (Schuster et al., 1998). In general, these measures determine the capacity for quality of an organization rather than the quality of performance. The best structural measures are those that have an influence over the provision of care and on patient health. In general, structural measures have not been linked to processes of

¹³³ Brook described the specific operation procedures that measure quality care processes.

care or health outcomes (Brook et al., 1990), there are exceptions, however; specifically, the volume of care or procedures provided have been associated with quality (Hannan et al., 1989, 1995; Phibbs et al., 1996; Stone et al., 1992).

Accreditation organizations such as JCAHO have relied upon structural measures in the past as a key component of certification; however, they have recently included some process and outcome measures in their accreditation programs. “Process” is what practitioners do to provide care, and consists of technical processes, interpersonal aspects of care (Schuster et al., 1998), and coordination and continuity of care (Donabedian, 1966, 1980). Technical process of care assess whether the right choices were made for diagnosis and treatment. These measures often are used to assess overuse and underuse of health care via appropriateness criteria, as well as to assess effective provision of care by comparison with professional standards. Interpersonal processes determine whether care was provided in a humane way and in accordance with the preferences of the patient. The best measures of process are those that are evidence-based¹³⁴ and provide an association between better processes and better outcomes (Schuster et al., 1998). For the most part, the literature on quality of care focuses on process measures. One limitation of these measures is the lack of conclusive scientific evidence that supports the standards and guidelines, which in turn limits consensus emerging around a particular set of performance measures (Lee, 1996). In such cases, expert consensus as to which measures are important is relied upon, but when no consensus is reached, the measure simply is not used to assess quality.

“Outcomes” are the results of health care interventions on patients’ health, such as recovery, function restoration and survival, as well as the effects on the health of a population (Donabedian, 1966, 1980). There are three general types of outcome measures: biological or

¹³⁴ Evidence-based means “proof” based upon “a foundation of solid science, especially using research that has applied rigorous scientific methods of epidemiology and has been published in peer-reviewed journals” (Eisenberg, 2001: 369).

clinical status, functional status, and consumer satisfaction (Schuster et al., 1998). Biological or clinical status indicates how the body is functioning, and is used to determine the success or failure of a particular intervention. Functional status assesses a person's ability to participate in the activities of daily living, whether physical, mental, or social. Consumer satisfaction is how patients feel about the care received (Schuster et al., 1998). The best measures of outcome contain important characteristics such as adequate risk adjustment,¹³⁵ relate to processes of care, account for different perspectives, designate sample size, and correlate to the care delivery system. While many performance measurement systems depend heavily upon process measures, outcome measures are becoming more popular, because they directly evaluate the health of a population (Schuster et al., 1998).

More recently, Total Quality Management (TQM), more popularly known as Continuous Quality Improvement¹³⁶ (CQI) in the health care industry, has emerged as a viable model for quality improvement (Brennan and Berwick, 1996). TQM, rooted in the manufacturing industry, addresses the nature and causes of variation. Traditional methods of performance measurement in health care, such as quality assurance and utilization review, describe the results of care processes, e.g., "after the fact" errors of clinicians. These methods do not address the process or system failures in health care organizations, whereas TQM focuses on the understanding and improvement of underlying systems of care. TQM has three basic principles: customer focus, continuous improvement, and teamwork (Dean and Bowen, 1994). Each principle is implemented by a number of practices, and each practice is sustained by a variety of techniques. All the principles relate closely to one another in the following way: continuous improvement

¹³⁵ The GAO defines risk adjustments as "methods for determining whether patient characteristics or poor quality are responsible for undesirable patient outcomes" (1994: 14).

¹³⁶ Some argue that TQM and CQI is essentially the same thing, and they are often used interchangeably in the literature. In fact, TQM is an approach to quality improvement that consists of principles, practices and techniques. CQI is a principle of TQM (Dean and Bowen, 1994).

helps identify and provide for customer needs and thereby produces customer satisfaction. Teamwork is essential to this process, due to the trans-boundary nature of continuous improvement processes. CQI, according to Batalden and Stoltz, has several elements. They are 1) new knowledge development, 2) leadership policy that fosters a shared purpose and promotes organizational learning, 3) tools and methods that accelerate work improvement, and 4) systematic strategies for applying knowledge to daily work processes (1993: 57-8). Formally, some health care organizations have adopted these principles but have adjusted them for a “better fit” for their use. For example, JCAHO describes TQM as a set of “strategies that stress the importance of leadership, external and internal customers’ needs, goal-driven design of new products and services, broad deployment of measurement systems, data-driven performance assessment, and systematic redesign of organizational processes and functions” (1994: 6).

Motwani et al. (1996) divide the TQM/CQI literature related to health into several themes. One theme is the description and synopsis of TQM, as applied from the manufacturing industry to the health care industry (Frist, 1992; Mayer, 1992; Matherly and Lasater, 1992; Mueller, 1992). Another describes the importance of implementation of TQM in health care organizations and its functional areas (Johnson, 1991; Berger and Sudman, 1991; Norlund, 1991; Godfrey et al., 1992). A third theme of the literature encompasses the development of conceptual models for assessing and implementing quality strategies in health care (Everett, 1993; Reitz, 1993; Luft and Hunt, 1986). The last theme involves the current practices of TQM implementation in health care organizations (Boyd and Haraway, 1991; Materna and Rothe, 1992; Burney, 1994). More recent research evaluates the effectiveness of TQM in terms of improvements in patient outcomes (Irvine et al., 2002; Irvine-Doran et al., 2002).

Characteristics of Performance Measures in Health Care

Lacking a national body to set goals and a vision for a national performance measurement system within the United States, the IOM has articulated a well-developed vision of the characteristics of performance measurement, and a performance measurement system that should be incorporated into any system developed for improvement of health care across the nation. The principles suggested are based upon the six quality aims identified in the *Quality Chasm* (IOM 2001) report: safety,¹³⁷ effectiveness,¹³⁸ patient-centeredness,¹³⁹ timeliness,¹⁴⁰ efficiency¹⁴¹ and equity.¹⁴² The ten design principles cited in the more recent IOM report on performance measurement are:

1. Comprehensive Measurement—performance measures and a performance measurement system should advance the purpose of the health care system and foster improvement in the six quality aims specified by the IOM.
2. Evidence-Based Goals and Measures—goals and measures should be based on the most recent scientific evidence, as reported in peer-reviewed journals. Measures should correspond to national goals and should include all aspect of the care of patients across the lifespan.¹⁴³
3. Longitudinal Measurement—measures should reflect health and health care of patients, both within settings and across settings over time.

¹³⁷ “Avoiding injuries to patients from the care that is intended to help them” (IOM 2001).

¹³⁸ “Providing services based on scientific knowledge to all who could benefit and refraining from providing services to those not likely to benefit (avoiding underuse and overuse, respectively)” (IOM, 2001).

¹³⁹ “Providing care this is respectful of and responsive to individual patient preferences, needs and values and ensuring that patient values guide all clinical decisions” (IOM, 2001).

¹⁴⁰ “Reducing waits and sometimes harmful delays for both those who receive and those who give care” (IOM, 2001).

¹⁴¹ “Avoiding waste, including waste of equipment, supplies, ideas and energy” (IOM, 2001).

¹⁴² “Providing care that does not vary in quality because of personal characteristics such as gender, ethnicity, geographic location, and socioeconomic status” (IOM, 2001).

¹⁴³ FACCT defines the continuum of care as “staying healthy, getting better, living with chronic illness, and coping with end of life” (1997).

4. Supportive of Multiple Users and Stakeholders—performance measures should support various uses such as quality improvement,¹⁴⁴ public reporting and accountability,¹⁴⁵ pay for performance initiatives and population health initiatives.¹⁴⁶
5. Measures Intrinsic to Care—collection of data to calculate measures should be byproducts of patient care processes, and therefore should reside in an electronic health record system.¹⁴⁷
6. A Central Role for the Patient’s Voice—patients should participate in the selection of measures for public reporting. Further, measures should include ratings and reports from patients and family caregivers.
7. Individual, Population, and System-Based Measurement—measures and a national measurement system should include measurement at the individual provider level, at the population level and at the health care system level. Measures should assess the various aspects of health care, including quality of care, unmet health care needs, efficiency, racial disparities, etc.
8. Shared Accountability—performance measurement in health care should not be constrained by the absence of a current, identifiable, single responsible agent. Measurement should reflect the patient and community priorities, and when no responsible agent can be identified, shared accountability among providers within the health care system should be assumed.

¹⁴⁴ Information used to improve the quality of health care delivered in health care organizations.

¹⁴⁵ Information derived from measures that assist in selection of providers for care, for inclusion in health insurance networks, and for decision-making for accreditation and certification purposes.

¹⁴⁶ Information used by decision makers to assess access to services, to address racial and ethnic disparities, and to contribute to decisions regarding disease surveillance and health protection.

¹⁴⁷ “A repository of electronically maintained information about an individual’s health care and corresponding clinical information management tools that provide alerts and reminders, linkages with external health knowledge sources and tools for data analysis” (Shortliffe et al., 2001).

9. A Learning System—performance measures and performance measurement systems should promote active and continual evaluation and learning, to improve performance and to advance the performance measurement methods and knowledge. This includes how to motivate performance improvement, redesigning effective and efficient care processes, and assessing progress in achieving the IOM six quality aims.
10. Independent and Sustainable—performance measurement systems should be continually assessed and enhanced and financed in such a way as to promote independence and sustainability (IOM, 2005: 143-145).

What has been presented above is a broad vision of the characteristics that performance measures and a performance measurement system should possess. At a more practical level, performance measures should be based on the current scientific evidence available, and address issues of importance, scientific acceptability, usability and feasibility. Table 6 defines each of these issues and provides additional definitions and examples for further clarification as specified by the NQF in its 2003 report, *A Comprehensive Framework for Hospital Care Performance Evaluation: A Consensus Report*.

Table 6: NQF Measure Evaluation Criteria for Hospital Performance

<p>Importance: Reflects the extent to which a measure quantifies a variation in quality, low levels of overall performance, and the extent to which it captures key aspect of the flow of care.</p>
<ul style="list-style-type: none"> i. The measure addresses one or more key leverage points for improving quality ii. Considerable variation in the quality of care exists iii. Performance in the area (e.g., setting, procedure, condition) is suboptimal, suggesting that barriers to improvement or best practice may exist.
<p>Scientific Acceptability: The degree to which a measure produces a consistent and credible result when implemented.</p>
<ul style="list-style-type: none"> i. The measure is well defined and precisely specified. Measures must be specified sufficiently to be distinguishable from other measures, and they must be implemented consistently across

<p>institutions. Measure specifications should provide detail about cohort definition, as well as the denominator and numerator for rate-based measures and categories for range-based measures.</p> <ul style="list-style-type: none"> ii. The measure is reliable, producing the same results a high proportion of the time when assessed in the same population. iii. The measure is valid, accurately representing the concept being evaluated. iv. The measure is precise, adequately discriminating between real differences in provider performance. v. The measure is adaptable to patient preferences and a variety of contexts of settings. Adaptability depends on the extent to which the measure and its specifications account for the variety of patient choices, including refusal of treatment and clinical exceptions. vi. An adequate and specified risk-adjustment strategy exists, where applicable. vii. Consistent evidence is available linking process measures to patient outcomes.
<p>Usability: Reflects the extent to which intended audiences can understand the results of the measure and are likely to find them useful for decision making.</p>
<p>The measure can be used by the stakeholder to make decisions.</p> <ul style="list-style-type: none"> i. The differences in performance levels are statistically meaningful. ii. The differences in performance are practically and clinically meaningful. iii. Risk stratification, risk adjustment, and other forms of recommended analyses can be applied appropriately. iv. Effective presentation and dissemination strategies exist (e.g., transparency, ability to draw conclusions, information available when needed to make decisions). v. Information produced by the measure can/will be used by at least one healthcare stakeholder audience (e.g., public/consumers, purchasers, clinicians and providers, policymakers, accreditors/regulators) to make a decision or take an action. vi. Information about specific conditions for which the measure is appropriate has been given. vii. Methods for aggregating the measure with other related measures (e.g., to create a composite measure) are defined, if those related measures are determined to be more understandable and more useful in decision-making. Risks of such aggregation, including misrepresentation, have been evaluated.
<p>Feasibility: The extent to which the measure is based on the way in which data can be obtained within the normal flow of clinical care and the extent to which an implementation plan can be achieved.</p>

- | |
|---|
| <ol style="list-style-type: none">i. The point of data collection is tied to care delivery, when feasible.ii. The timing and frequency of measure collection are specified.iii. The benefit of measurement is evaluated against the financial and administrative burden of implementation and maintenance of the measure set.iv. An auditing strategy is designed and can be implemented.v. Confidentiality concerns are addressed. |
|---|

Source: NQF (2003). A Comprehensive Framework for Hospital Care Performance Evaluation: A Consensus Report, pp. 8-9.

In addition, performance measures should possess the following traits for credibility among providers and the public:

- Explicit numerator
- Explicit denominator
- Explicit inclusions
- Explicit exclusions
- Data source
- Data element definition, including allowable sources and allowable values
- Data analysis logic and method
- Risk-adjustment method, if used, and associated data elements needed
- Cohort definition and sampling method, if applicable
- Format in which the measure should be reported (NQF, 2003: 12).

Regardless of these principles and preferred traits, performance measures and performance measurement will never be perfect. Analytic methods and risk adjustment will continue to be refined, and the process of arriving at consensus among stakeholders is continuous. It has taken several years for the health care industry to come to consensus on the validity of scientific and clinical concerns on which hospital performance measures are based,

even when the basic principles are accepted and supported by the hospital provider community. Imperfect measures can be improved through access to high quality data sets, or the adoption of an electronic health record system on a wide scale. Other issues, such as small sample size, are inherent problems with performance measurement overall, and cannot be corrected. The question is: how good is good enough? Decision makers will have to evaluate and decide when measures that are less than perfect should be acted upon and adjust their policy accordingly.

Uses of Performance Measurement in Health Care

Throughout history, performance measurement has had many roles. It began as a need identified by those outside the industry and, with the help of government, was eventually designated as a cornerstone of health care quality improvement. As the multiple identities of performance measurement evolved, its uses also have changed. Based on the research conducted for this dissertation, I have laid out the uses of performance measurement, and its evolution from internal quality improvement to current initiatives such as pay-for-performance (P4P). By evolution, I am referring to the changing nature of how performance measures and performance measurement are used in health care. These uses are not mutually exclusive, as performance measures and performance measurement truly have multiple identities, and instead of replacing one identity when a new use is developed, the identities are accumulated. How and which performance measures are used depends upon the goals of the organization and the context in which the measure is used. One key informant emphasized that, while the uses of performance measures have changed over time, the measures themselves have not. This means that a measure specifically developed for internal quality improvement and trending is now being used for public reporting and P4P, without any revision, refinement or risk adjustment, thus unfairly comparing and penalizing provider organizations.

Quality Improvement: Changing Behavior.

Originally, performance measurement and performance measures were developed as tools for internal quality improvement in hospitals and other health care organizations. Brought about by increasing concerns about the high mortality rates of hospitals, and initiated by those within the profession, these measures provided data for trend analysis. With the advent of new medical advances and the establishment of the nursing profession, more attention was given to quality improvement activities. Today, these measures provide data with which to compare and assess differences in performance within a hospital that, in turn, offer guidance for service improvement. These measures are not risk-adjusted and are used for trending and improvement activities. Accordingly, hospitals considered this information confidential and often used it in initiating programs to enhance the services they provide.

The “repair” of poor services often depends upon changing behavior. Effective strategies, such as those summarized by Bauchner et al. (2001) can influence changes in physician behavior. But to determine which strategies are more (or less) effective, performance has to be measured and evaluated. Audits, in combination with reports involving specific recommendations, have resulted in behavior changes consistent with improvements in quality of care (Bauchner et al., 2001). Thus, performance measurement provides information that identifies areas that may be targeted for improvement, as well as information about progress toward established quality goals within a facility.

Accreditation and Certification.

The next use of performance measures and performance measurement came about when CMS established “conditions of participation” when the Medicare and Medicaid programs were signed into law in 1965. Compliance with the conditions of participation is required for any health care facility that wishes to participate in the Medicare and Medicaid programs, thereby

influencing facility revenues. In order to participate in these federal programs, facilities must be accredited. Providing a voluntary accreditation process, the Joint Commission for the Accreditation of Healthcare Organizations (JCAHO) functions as an independent accrediting body, and its program is the most influential and pervasive, accrediting more than 15,000 health care facilities in the nation. Originally, the JCAHO focused its efforts on a structural approach to quality, basing accreditation on elements such as organizational frameworks of hospitals, medical record adequacy, safety standards and the like. More recently, the focus has been on quality assurance and efforts to coordinate quality monitoring activities among services within hospitals. For instance, in order to attain reaccreditation, hospitals must show ongoing performance improvement activities, as well as documenting problems and the specific actions taken to address them. Further, hospitals are required to submit performance data, using a subset of standardized performance measures (ORYX system) to attain and maintain accreditation. While this is only one piece of the accreditation process, it is a significant one.

An additional use of performance measures is certification of individual clinicians. While licensure¹⁴⁸ for physicians is the oldest and most pervasive form of quality assurance, the second is voluntary professional certification. Professional certification goes one step beyond licensure in that physicians must undergo testing via specialty boards to obtain certification. Testing typically consists of a paper/pencil test and some sort of skill testing, such as reading and interpreting x-rays. Thus, specialty boards set the minimum standard for training and knowledge within a particular subject area. Obviously, performance measures are used to classify which clinicians will receive certification and which will not. For a physician to be board certified is

¹⁴⁸ Licensure granted by states, sets a minimum standard for quality. It assures that the physician or nurse has a basic level of training and education relevant to the profession.

more prestigious and affords more opportunities for advancement and income than one who is not certified. Thus performance measurement plays a role at the individual clinician level.

Agenda Setting and Prioritizing Issues: The Case of Medical Errors and Patient Safety.

The third use of performance measures and performance measurement is agenda setting. As noted above, what gets measured gets done; thus, what is measured has resources devoted to improving it. One example is the emergence of the patient safety issue, which prior to 1998 was not part of the conversation within health care. With the release of the report *To Err is Human: Building a Safer Health System*, the IOM cited the prevalence of medical errors in health care and called attention to the issue of patient safety. This report provided the momentum for the various changes initiated by several actors, in both the public and private sectors.

As a result of this widely publicized report, President Clinton directed the Quality Interagency Coordination Task Force (QuIC)⁹⁵ to evaluate and develop a strategy for the federal government to improve patient safety and to detect and reduce errors in the health care system. In addition, Congress appropriated \$20 million¹⁴⁹ for the Agency for Healthcare Research and Quality (AHRQ) specifically earmarked for patient safety research (Eisenberg et al., 2000). Taking the lead from the public sector, the Joint Commission (the organization that certifies health care facilities), although concerned about patient safety since 1993, began requiring compliance with its National Patient Safety Goals for organizations seeking accreditation after January 1, 2003.

In performance measurement, priorities are reflected in the weight assigned to the measures and the areas selected for measure development (Jones and Duncan, 1999). Thus, these goals ensure a greater focus on the patient safety initiative, because they are part of the

¹⁴⁹ AHRQ actually received \$50 million in FY 2001 for medical error research (B. Ginieczki, personal communication, April 14, 2004).

accreditation process. Failure to comply may adversely affect the organization's accreditation status, which may have financial consequences in terms of reimbursement. Compliance with these goals will be reported in the organization-specific performance reports, as well as being made available on the JCAHO website (JCAHO, 2004). These are only a few of the organized efforts to track and report these mistakes. Many more organizations have different initiatives and resources directed toward patient safety.

Performance measurement, while not the only influence, can act as a force to promote certain issues and agendas. Performance measurement conveys the message of importance. Specifically, what is important is measured, while what is not measured is considered less important (Waggoner et al., 1999). By focusing people and resources on a particular aspect of an industry, performance measurement can be a driver of change and/or reform.

Accountability: The Instance of Performance Report Cards.

With the establishment of the NQF, performance measurement became a tool of accountability. Accountability in health care is the "need to demonstrate to others the quality and value of the health care provided" (JCAHO, 1994: 28). Performance measurement systems are a tool of public accountability because they document the quality of care provided, as well as revealing those areas needing improvement. Performance measurement, however, cannot affect improvement in the health care system unless it is reported to the public.

Public disclosure of performance seems to be necessary to advance improvement via incremental changes in consumer, professional, and managerial behavior (Leatherman and McCarthy, 1999; Fowles, 2000). In its report, the President's Advisory Commission (1998) suggested that the standardization of performance measurement and reporting efforts of both public and private organizations influence quality. "A key element of improving health care quality is the Nation's ability to measure the quality of care and provide easily understood,

comparable information on the performance of the industry” (President’s Advisory Commission, 1998: 3). One way to achieve the goals of quality improvement and public accountability is through the use of organizational report cards. This type of report card can be defined as “a regular effort by an organization to collect data on two or more other organizations, transform the data into information relevant to assessing performance, and transmit the information to some audience external to the organizations themselves” (Gormley and Weimer, 1999: 3). The benefits of report cards are that they reduce information asymmetries that affect market competition and consumer choice (Gormley, 1998; Gormley and Weimer, 1999), encourage performance improvement (Leatherman and McCarthy, 1999; President’s Advisory Commission, 1998; Barr et al., 2002), and provide a mechanism to demonstrate accountability (Gormley and Weimer, 1999; Barr et al., 2002). Problems with report cards include assessment issues due to limitations of data and theory¹⁵⁰ (GAO, 1994; Gormley and Weimer, 1999), organizational response and participation problems¹⁵¹ (GAO, 1994; Gormley and Weimer, 1999), and consumer reception issues¹⁵² (Gormley and Weimer, 1999). In 2001, the IOM published *Envisioning the National Health Care Quality Report*, which outlines a national approach to public reporting of health care system performance. Although various efforts report the different aspects of health care, such as health plans, nursing homes and home health agencies, there are few report cards that “compare data by hospital or that are designed for public dissemination, and still fewer that report patient satisfaction” (Barr et al., 2002: 51).

¹⁵⁰ This means that all the relevant variables may not be able to be measured and that the theoretical links between variables that can be measured and those that are conceptually appropriate are weak (Gormley and Weimer, 1999: 7; GAO, 1994: 4).

¹⁵¹ Organizations can either use report cards for self-improvement or respond dysfunctionally. Some dysfunctional responses include non-participation, “cream skimming,” manipulating the numbers, and blaming the messenger (Gormley and Weimer, 1999: 13-15).

¹⁵² These problems include weak motivation to use report cards, cognitive limits of individuals, and information inequalities, which is the difference between those who process the information well, e.g., high income, well educated groups and those that do not, e.g., low income, poorly educated groups (Gormley and Weimer, 1999: 15-17).

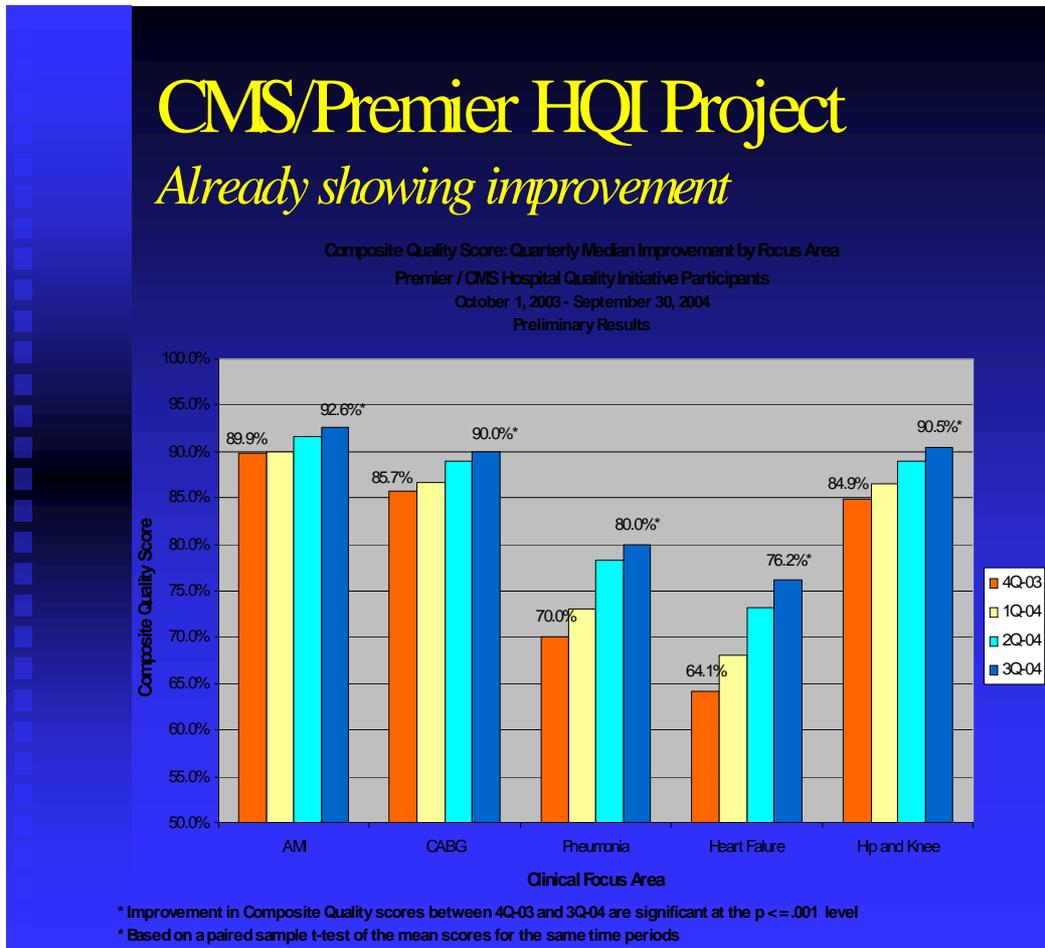
Pay for Performance.

The most current use of performance measures and performance measurement is Quality-based Purchasing, more commonly known as pay-for-performance (P4P). These types of programs depend upon market forces that presumably drive how health care is delivered. These purchaser initiatives were designed to derive the best value for health care expenditures and ensure that beneficiaries were receiving high quality care. In short, these programs reward providers that deliver high quality of care to patients, thus providing incentives to improve the quality of care delivered. In 2005, more than one-hundred P4P initiatives sought to align provider performance with payment (NQF, 2005).

With scant evidence to support the effectiveness of these initiatives, and with the increased popularity of this approach, CMS chose to test this concept in its Hospital Quality Improvement Project after there were several calls by the public for Medicare to provide payment for quality.¹⁵³ CMS partnered with Premier, a large hospital chain, to test new payment methods through a demonstration project linking payment to quality. The project began in 2003, with 278 hospitals within the Premier Chain participating. The goal of the project was to pilot-test the concept of P4P—using the power of purchasers by paying for high quality performance, processes and systems will be compelled to change at a more rapid rate, as compared to the incremental change processes currently in place in health care. To measure performance, there need to be standardized measures so that hospitals can be fairly compared to one another. CMS selected a number of performance measures and devised a weighting scheme that showed what each hospital would receive for improved performance. Preliminary results, as shown in Figure 8, suggest that improvement is already occurring.

¹⁵³ See IOM report 2002; *Health Affairs* article 2003; and the MedPAC report 2004.

Figure 8: The CMS/Premier HQI Project



Source: Centers for Medicare & Medicaid Services, S. Roman, M.D., personal communication, September 30, 2006).

This project is currently focused on health care facilities, particularly hospitals.

However, other demonstration projects sponsored by CMS that focus on clinician groups and individual clinicians are currently underway, and the results of these pilots should be available in the near future.

In sum, these are the current uses of performance measurement within health care. The health care community anticipates that there will be other initiatives that will utilize measures in new and innovative ways, all with the focus of improving the quality of care at the system, facility and/or individual clinician level within the country.

Key Issues of Performance Measures and Performance Measurement

Multiple Data Demands and Data Collection Burden.

Hospitals have been a focus in measuring performance and in quality improvement activities because they accounted for \$648.2 billion of the nation's health care expenditures in 2006 (Centers for Medicare & Medicare Services, 2006). Further, care delivered in hospitals is provided to individuals with serious illnesses, in which outcomes are highly sensitive to the quality of the treatments received. Because many measurement efforts have proceeded independently of each other, hospitals face multiple data reporting demands for various efforts, including accreditation and certification issues, public reporting, and quality improvement. In reality, many of the efforts use different measures or different specifications of the same measure requiring the collection of different data elements. This is due in part to the inherent tension among diverse stakeholder goals. Providers cite excessive burdens of data collection and the subsequent cost of collection and interpretation in relation to the various initiatives. Spending resources on collecting data for the different initiatives leaves little for quality improvement activities. At the same time, consumers and purchasers typically don't have enough information to make informed choices about their care. Their goal is to have a robust and wide-reaching set of performance measures that support the many goals of users. More recent efforts within the hospital community have focused on "harmonization" or alignment of measures so that there is agreement on one core set of measures that can be used by all stakeholders for a variety of purposes. These efforts are in their infancy however, and it may take some time for the measure developer organizations to come to consensus on which measures to include in a core set, the measure specifications, their data sources, and risk adjustment methodology, to cite some of the relevant and complex issues involved in these ongoing discussions.

Measure Relevancy, Accuracy and Timeliness.

As noted by the discussion above, there is an inherent tension among the goals of the various stakeholders for data use within health care. There also are tensions surrounding the relevancy of the measures selected for public reporting and other initiatives, their accuracy (i.e., how well they measure what they were designed to measure), and the “freshness” of the data, i.e., timeliness. What is currently measured in the hospital community and reported by the Hospital Quality Alliance often is described as a “hand full” of measures related to selected conditions.¹⁵⁴ Hospitals are required to submit data on ten measures, known as the “starter” measures, in order to get their full market basket update.¹⁵⁵ Additional measures are required by the Deficit Reduction Act (DRA) of 2005, raising the current 10 required measures to 21 in FY 2007, in FY2008, the number of measures required to be reported increased to 26. Even with reporting additional measures over the next few years, these measures do not begin to address the breadth of conditions and populations that are treated in the nation’s hospitals.

Further, these measures report on conditions that are considered life and death emergencies and in which a consumer has no choice of hospitals. For example, a patient having a heart attack is an emergency situation, and the paramedics will proceed to the closest hospital, in an effort to assure prompt and life-saving treatment. However, the situation is reversed in elective conditions or situations. Consider the instance where a woman with a normal pregnancy has time to investigate and select a hospital in which to deliver her baby, and could use

¹⁵⁴ As of December 2007, there are approximately 24 measures reported on the Hospital Compare Website. The conditions reported on include heart attack, heart failure, pneumonia, and surgical infections. Measures slated for public reporting in 2008 include HCAHPS (patient survey) and two 30-day mortality measures for AMI and heart failure (Medicare patients only). In 2009, pediatric asthma measures, selected measures from the Surgical Care Improvement Project (SCIP) and selected critical care measures will be added to the website.

¹⁵⁵ As required by Section 501(b) of the Medicare Prescription Drug, Improvement and Modernization Act of 2003 (MMA).

information about the obstetric care delivered at the area hospitals to make an informed decision. Yet, measures relating to obstetric care or elective conditions are not widely reported.

The accuracy of performance measures is also an issue with the hospital community and stakeholder groups. Measures typically have problems such as validity, reliability, sensitivity,¹⁵⁶ and specificity¹⁵⁷ which become relevant particularly when reporting performance in a public forum or when reimbursing providers for providing quality care. Validity is the extent to which a measure actually measures what it purports to measure (Testa and Simpson, 1994). In general, quality of care measures is defined quite narrowly in health care because of the various factors that may affect the outcome. For example, mortality is a poor indicator for quality of care for certain incurable conditions—the provider may do all the right things for the patient, but the patient still dies. Measures that assess whether a patient was given discharge instructions regarding the use of his or her prescribed medication are considered a valid measure of quality in health care.

Reliability is the extent to which repeated measurement yield consistent results (Testa and Simpson, 1994). Typically, reliability is poorest when results have to do with clinician-based data and judgments. For instance, reliability has been shown to be lacking anywhere from 5 to 40 percent when diagnosing a heart murmur for a patient, or whether a patient should undergo elective surgery. Although the reliability of clinical observations has not been sufficiently investigated, the methods used to assess quality in health care rely on clinical observation and assume their accuracy. Reliability is further complicated by the abstraction process used as a basis for many quality improvement programs. Abstraction involves reviewing the medical record, synthesizing the clinical information, and entering that information into a

¹⁵⁶ This refers to how small a variation a measure can detect in an attribute, and whether it can be reliably detected and measured.

¹⁵⁷ Specificity refers to how well a performance measure measures a specific attribute or condition.

database by medical record personnel. For example, the abstracted information may include discharged diagnoses (both primary and secondary), procedures conducted during an inpatient stay, and length of stay. This system assumes that all the care provided is documented in the medical record, which is not always the case. Obviously, interpretation of the medical record is also an issue. Medical abstractors may vary in the number of episodes of tonsillitis experienced by a patient, which were described in the medical record as “frequent” and “numerous.” Thus, the use of discharge data for quality assessment has raised concerns about reliability of the data, particularly when providers’ reputations and revenues are at stake.

Timeliness of the data that support quality initiatives is also an issue of concern to the health care community. Data collection through the abstraction process is tedious, slow and retrospective rather than prospective. Currently, there is not widespread adoption or use of electronic health records in this country. One key informant indicated that only about 15 percent of medical group practices have an electronic health record. Reasons often cited include financial burden, manpower costs, and overwhelming complexities of developing a comprehensive system that plans for contingencies. Electronic health records should be developed so that data to support performance measurement and reporting are by-products of the system, thus reducing the burden associated with these activities. However, without the widespread use of an electronic health record, data collection, collation and analysis can cause a time lag of at least three months to a year for national reporting efforts.

Needless to say, information on quality of care provided retrospectively does little to improve the quality of patient outcomes at the point of care. An ideal system would not only provide real-time information to better monitor the quality of care, but also would help predict and correct the trajectory of the outcome of care received at hospitals.

Measure Maintenance in a Dynamic Measure Environment.

As noted earlier, the health care industry environment is undergoing almost constant change. The performance measurement field is experiencing the same type of environment and is trying to keep up with the increasing demands of policy makers and others for more information on the quality and performance of the health care system. Because performance measures are based upon the scientific evidence, measure developers must routinely contend with new evidence that may have an effect on the specifications of their measures. Measures can become ineffective if they are not updated to reflect the advances in the knowledge base. Further, measures also depend upon other ever-changing factors. For example, performance measures typically consist of a numerator, or the population of interest, and a denominator, or the population at risk. ICD-9-CM codes are used to identify both numerator and denominator populations. These codes are routinely updated and must be incorporated into the performance measures to keep them effective. To maintain measures requires several activities that include conducting regular environmental scans, reviewing the evidence base, soliciting input from researchers, hospitals and other stakeholders, updating measures and data sources as needed, and communicating the changes to the industry. It is estimated that the cost range for measure maintenance is from \$26,000 to \$250,000 per measure, with an average cost of \$125,000 (Booz, Allen, Hamilton, 2006: 20). Based on the range of costs cited here, there appears to be some variability as to the quality of measure maintenance.

Retirement of measures is also a key feature in the lifecycle of a performance measurement program. As measurement goals are achieved consistently, or when measures do not effectively measure the event of interest, and are replaced by more accurate measures, they should be retired. Unfortunately, as many state legislatures are mandating public reporting of

hospitals, they are incorporating specific measures in the law, which in the long run will make maintenance and retirement of measures difficult.

HIPAA Concerns and the Issue of Data Ownership.

Other key issues related to the collection and reporting of performance data include concerns surrounding the Health Insurance Portability and Accountability Act of 1996¹⁵⁸ (HIPAA) and the issues of data ownership. Specifically, Title II of HIPAA,¹⁵⁹ also known as the Administrative Simplification provisions, requires the Department of Health and Human Services to establish standards for electronic health care transactions, develop a system of national identifiers for providers, health plans, and employers, as well as speak to the security and privacy of health data. The goal of this law is to protect individuals' health information while not hindering the "flow of information needed to provide high quality health care and to protect the public's health and well being" (OCR, 2003: 1).

The Privacy rule is significant to the public reporting of performance in that it establishes regulations for the use and disclosure of "protected health information." This type of information is considered to be "any information about health status, provision of health care, or payment for health care that can be linked to an individual" (45 C.F.R. 164.501). Interpreted in the broadest sense, it includes any part of an individual's medical record or medical payment history. Because performance information is gleaned from both of these sources, the HIPAA applies. De-identified data can be used and transmitted to reporting sponsors for the most part; however, in order to provide a "safe harbor," a number of data elements must be removed, including names, all geographic subdivisions smaller than a State, and all elements of dates

¹⁵⁸ Public Law 104-191 enacted on August 21, 1996.

¹⁵⁹ Title II is entitled "Preventing Health Care Fraud and Abuse; Administrative Simplification; Medical Liability Reform." The most significant provision is the Administrative Simplification, which consists of five rules promulgated by HHS: the Privacy Rule, the Transactions and Code Sets Rule, the Security Rule, and the Enforcement Rule.

(except year) for dates directly related to the individual (e.g., birth date, admission date, discharge date). Further, after all the identifying information is stripped out, if the remaining information can be used, alone or in combination with other information that can subsequently identify an individual, the covered entity can be held accountable if they had knowledge of the situation.¹⁶⁰

Many of the data elements that identify an individual are used to calculate and report performance measures. Fortunately, one of the permitted uses and disclosures includes health care operations. These include such functions as quality assessment and improvement activities, and competency assurance activities; however, these uses pertain to internal operations of the hospitals. Depending on who is doing the public reporting, for example (a government agency or a private entity), and/or who is interpreting the rule, there are other uses and disclosure provisions, including public interest and benefit activities such as health oversight activities, public health activities, and essential government functions.

Along these same lines is the question of data ownership, particularly when information is provided for public reporting. When large amounts of data are accumulated, researchers often try to take advantage of aggregated, readily available data sets. Very often, data use agreements that outline the uses and disclosure of data sets are signed by researchers, in order to gain access to these rich information sources. However, the question of who owns the data and who can have access to it has become a recent issue. Due to the complex legal issues and potentially severe penalties associated with HIPAA, the issue of data ownership is cause for concern. Hospitals can control the flow of information and protect personal health information of those they serve. Once the data leaves the facility however, and is transmitted to another entity, control over it by the originator is limited. According to the GAO, health care providers are

¹⁶⁰ 45 C.F.R. § 164.514(b).

“uncertain about their [legal] privacy responsibilities and often respond with an overly guarded approach to disclosing information...than necessary to ensure compliance with the Privacy rule” (Wilson, 2006: 313-316). Therefore, faced with fines for violators, most health care personnel tend to be overly guarded about releasing information. In public reporting initiatives that are voluntary, participation by hospitals is very limited.

Public Reporting and Report Cards of Health Care Performance

Comparative public reporting and the posting of quality report cards,¹⁶¹ while considered a fairly young field, has proliferated over the past several years. In 2003, IPRO, under contract to CMS, published the *Review of Hospital Quality Reports: for Health Care Consumers, Purchasers and Providers*. The purpose of this report was to inform CMS about the existing report cards, as well as their properties and the methods used for generating them. This report was a “review” rather than an inventory, because it was not considered comprehensive. Yet, the 36 entries reported on serve to illustrate the number of report cards available and their various sponsors. The types of sponsors, as well as the number of report cards for each type of sponsor, are listed below:

- Government (8)
- Proprietary/Commercial (9)
- Foreign (2)
- Purchaser or Business Coalition (4)
- Health Plan (2)
- Hospital Association (1)

¹⁶¹ A report card broadly refers to “a wide variety of information sources and tools that enable consumers to compare the quality and in some cases, other characteristics of health plans or providers” (AHRQ, 2006).

- Other Not-For-Profit (5)
- Individual Hospital or Hospital System (3)
- Specialty Society (2) (IPRO, 2003: iii)

These report cards give an account of the clinical quality of a hospital and do not include measures of patient satisfaction or report cards on health plans.

More recently AHRQ compiled a searchable Health Care Report Card Compendium that provides information on the available health care report cards, with comparative information on the quality of care in such areas as health plans, hospitals, medical groups, individual physicians, nursing homes, and other providers. As of November 2007, the number of entries in this resource was more than 211, the oldest of which dates back to 1996. All the entries demonstrate the range of methods currently used to report quality and performance data for health plans and providers.

Several objectives emerged in the public reporting and report card literature. They are:

- Providing consumers and purchasers with quality and performance information to inform health care decision-making (the consumer choice model)
- Promoting accountability to the public as well as informing providers that their performance is being monitored
- Providing feedback to providers and their organizations to assist in their quality improvement efforts (IPRO, 2003: v)

It should be noted that public reporting and report cards constitute only one strategy out of many approaches that may be incorporated into an overall quality improvement plan for a community. However, a study conducted by Hibbard et al. concluded that improvement

activities in hospitals that did not participate in any reporting initiative were less frequent than in those hospitals that did (2003), and that improvement was most dramatic among hospitals that publicly reported their performance (Hibbard et al., 2005). Thus, it appears that some sort of public disclosure for performance is necessary for quality improvement activities. Further, quality improvement efforts seem to be linked to report cards' public release. Hospitals that showed improved performance were much more likely to initiate a wide range of improvement activities directly after the performance report was released.

Yet, in the case of New York State's cardiac surgery report card, when consumers were provided with information on mortality rates in hospitals, consumer action (e.g., avoiding hospitals with high cardiac care mortality rates), did not provide the impetus toward improved care. In fact, consumers did not change their care-seeking patterns at all. It was found that the providers sought after and used the data reported to improve their practices, independent of any consumer activity (Chassin 2002). Earlier investigations indicated that hospitals were more concerned about how report cards affected their reputations, rather than their market share. A recent study supported this notion, finding that consumers who viewed a public report on hospital quality were more likely to have an accurate perception of their local hospital quality and were more likely to retain these perceptions for the greater part of two years (Hibbard et al., 2005). While professional pride is a motivator for quality improvement, these findings about consumers' memory of a hospital's reputation may affect the bottom line, particularly where fundraising and charitable donations are an important part of fiscal viability. The public reputation of hospitals may affect other areas, such as resource allocation by hospital boards, priority setting within a community as well as other unintended consequences that may affect the hospital's bottom line (Hibbard et al., 2005).

Characteristics of Public Reporting and Report Cards

Although a number of report cards are widely available, they all seem to have certain characteristics in common. While there are many publicly available report cards as evident by the discussion above, many others are available to a particular membership, such as members of a particular health plan and/or available on a fee-only basis. According to experts in the field, “content, format and access, are influenced by the type of sponsor agency or organization, the audience they are trying to reach, and their assumptions about how the report will be used” (S. Sofaer, personal communication, July 11, 2005).

In addition, access to comparative reporting on performance can be limited by factors such as who has resources to control data collection and distribution. For example, a health plan such as Blue Cross/Blue Shield may only be interested in collecting certain pieces of information that are relevant to its current initiatives and that pertain to its membership. Presumably, it would not go to the expense of gathering and reporting performance data on hospitals that are not part of its network. Hence, the reports it does provide pertain to a particular population and contain only those elements that are of interest to the senior leadership of the health plan. In recent years, however, Congress has enacted legislation to include mandates for more public reporting of health care performance. The most recent pieces of legislation include the Medicare Modernization Act of 2005 and the Deficit Reduction Act of 2006. In addition, pressure from consumer and purchaser coalitions is increasing, and as the costs of health care continue to rise,¹⁶² there are increasing calls for more public reporting and P4P initiatives.

¹⁶² Specifically, the Consumer-Purchaser Disclosure Project, a coalition of 60 employer, consumer, and labor organizations that collectively represent over 100 million beneficiaries in the United States, has called for a “public reporting of national standardized measures for clinical quality, consumer experience, equity, and efficiency” (Consumer-Purchaser Disclosure Project, 2004).

Other themes that have emerged with regard to the issuance of report cards and public reporting of performance include the diverse audiences that may use comparative report cards, as well as the effects of the format and content of report cards. Diverse audiences have distinct needs when using performance report cards. For example, consumers and purchasers may use the information contained in these reports to identify and utilize high quality providers, while providers use the information for quality improvement. In addition, how a report card is constructed influences consumer choice and change within the health care system. To fulfill their promise, report cards should have relevant content, and should be formatted for easy understanding of information; if the report appears on a website, ease of navigability also is an issue that should be addressed. Information contained in report cards should be “engaging, credible, easy to understand and actionable by the intended audience” (S. Sofaer, personnel communication, July 11, 2005).

Key Issues of Public Reporting and Report Cards

Comparison of providers using performance information can play an important role in health care reform. The “Theory of Action” of comparative report cards assumes consumers have the ability to compare and evaluate performance information, which will inform their ability to choose health care providers based on quality (Hibbard and Jewett, 1996). In other words, performance reports would influence choice, which then would influence market share, which in turn would create incentives for improvement. The literature is not clear on this aspect, however, and it offers a mixed picture on how consumers use performance information. While Americans believe that quality is more important than costs, benefits, and choice of physicians, various studies have shown that the impact of this information on decision outcomes is unclear. Those consumers who saw performance report cards were more likely to perceive that the categories reported were important to choice than those consumers who did not (Hibbard et al.,

2002). Yet in another study, Fowles and colleagues found that fewer than fifty percent of consumers read the report cards, and the perceived helpfulness was inversely related to an individual's level of health care experience (2000).

With increasing interest in the field of public reporting and report cards, researchers have developed a better understanding of public reporting and its desirable outcomes. Researchers suggest that multiple exposures of consumers to public reports are necessary, so that consumers will develop confidence and skill in their use. This may be due to the fact that consumers do not believe there is a problem with the quality of health care in the U.S., and if there is, it doesn't apply to them. Further, if report cards are used by consumers, they are only used by a small percentage of them—and only by those who perceive they have a choice and by those that feel they actively need to make a choice (rather than stay with their current health plan or provider). Part of the problem lies in how the reports are designed. Researchers typically design report cards, not consumers; therefore, these reports are considered “not responsive” to consumer needs and concerns by many experts in the field. Finding health care performance information also is an issue. Typically, one-third of those seeking information about health care quality experience difficulty finding it, and the difficulty increases in elderly and non-English-speaking populations (Berry et al., 2001).

In the end, collection and reporting performance data are costly and time-consuming. As one expert states, “poorly designed, unusable or unused quality reports wastes resources and reinforces a bias against the reporting, particularly public reporting of comparative quality data” (S. Sofaer, personal communication, July 11, 2005).

Summary

In closing, the chapters contained in Part II of this dissertation identified and described both the human and the non-human actors of the Hospital Quality Network. The human actors consist of organizations such as the NQF, CMS, AHRQ, IOM and others. A brief overview of each actor was provided and included its organizational mission and purpose, its origins, structure, and resources. Likewise, the non-human actors of the network were identified and described. A brief description of the actors included its purpose, characteristics, its uses in the network, and key issues associated with it.

The mapping of the network is important and was provided as background to the next section, which specifically describes the politics of performance measurement and the selection of measures through the NQF CDP. These descriptions of the actors provide clues as to how they behave and how they interact with other actors in the network in a political environment. What follows is a discussion that traces the activities and interactions of the NQF Hospital Steering Committee. It begins with a discussion about the nature of politics, and uses ANT to examine the role that it plays in the NQF measure selection and endorsement process.

PART III: THE HOSPITAL QUALITY NETWORK AND THE POLITICS OF
PERFORMANCE MEASUREMENT

“Science is the orderly arrangement of what,
at the moment, seem to be the facts”

Anonymous

At first glance, the terms “politics” and “performance measurement” in the same sentence seem to be in conflict. What does one have to do with the other? Politics is often viewed as that nebulous thing out there that happens behind closed doors with secret handshakes over agreed-upon courses of action to further one’s agenda. Performance measurement, as it is touted in health care, is founded on science and research, and uses such terms as precision, validity, reliability, and risk adjustment to name a few. Upon closer inspection, however, I found in the course of this research project that this is not necessarily the case. There is an intersection between these two concepts that is not apparent to the outsider, but seemed to be known and acknowledged by those who participate in the field of health care performance measurement.

There are many definitions of “politics.” In its usual context, politics refers to “the struggle and compromises between interests and human passions” (Latour, 2004: 247). Health care, however, is concerned with a broad array of issues, which suggests that a broader definition of politics is required. Politics with a “small p” can be defined simply as “who gets what, when, and how” (Lasswell, 1936). The “what” refers to anything of value that can be easily described and defined, e.g., money, power, influence, and status (Rhodes, 1992).

From these definitions and from one’s own experiences with the nature of politics, it may be surmised that the political process can be portrayed as somewhat “messy.” Some of its characteristics include 1) bargaining and compromise, 2) values-based and self-interests, 3) there is influence exerted (or not) and displays of power, either implicitly or explicitly throughout the process, 4) it involves the development of strategy and an argument to achieve established goals,

and 5) interpreting the relevant facts to support one's argument. And because at any one time multiple interests compete, compromise must be sought, albeit after much wrangling and quite often with displays of power. That is the essence of politics for many.

So, too, is the world of performance measurement, particularly with regard to the selection process of measures for a national performance measurement system for hospitals. The NQF plays a unique role in that it not only aligns and standardizes measures, but because its voluntary consensus standards-setting status, its products, (e.g., endorsed measure sets), carry the color of law for the federal agencies. Those measures endorsed by voluntary consensus bodies such as the NQF must be used by the federal government in its projects and initiatives, unless there is a practical reason not to, which requires justification to OMB. Politics enters into the process because there are multiple competing measures, and not all measures submitted to NQF receive endorsement; hence there are "winners and losers." Measure development demands extensive resources, which are sunk costs to major developers such as CMS and the Joint Commission. Those measures that are not selected for endorsement are both eventually phased out of use and abandoned by the developer, or developers may opt to further refine the measure and submit it at a later date for consideration by an NQF Maintenance Committee. Refinement of measures, if the developer chooses that option, involves additional resources of time and money, with no guarantee of future endorsement. Therefore, there is much at stake for measure developers, particularly with the amount of resources invested in their measures.

As with the democratic political process, which affords certain privileges to certain groups of people, so too does the politics of performance measurement. Overall, some inertia is associated with the field of health care performance measurement, because certain measures that are currently in use are embedded in accreditation and/or payment programs. Obviously, the measure developers who own and maintain these measures have an advantage over other

developers. While “tweaking around the margins” of an existing measure can be accomplished and is readily acceptable to the industry, major changes in measures is to be avoided, due to expense or “undue burden” on the hospitals of changing established systems to some extent to implement the new measure specifications. As suggested here, however, measures embedded in current programs are not stagnant, and tweaking around the margins often occurs during the process of consensus development. This is where bargaining and compromise and all the other aspects of politics come into play. As indicated by the mapping and tiering of the hospital quality network in the previous chapters, not all organizations enjoy the same status or position within the network. Since this is the case, more often than not an asymmetry occurs, where one organization exerts more influence within the network or initiates more action than another, in order to promote its self-interests. And as a result, sometimes measures are selected, not because of their validity or evidence base, but merely because of who is at the table when the decision is made. As we will see in this section, the process of measure selection or consensus development for NQF endorsement contains all the characteristics of a democratic political process, and can be just as chaotic and disorganized.

This plays out on a national stage within the Hospital Quality Network, but also can be seen on a smaller scale within the various NQF projects involving hospitals. Specifically, the Hospital Steering Committee appointed for the National Voluntary Consensus Standards for Hospital Care: An Initial Performance Measure Set is a microcosm of the larger Hospital Quality Network. Many of the members of the Steering Committee are also actors within the Hospital Quality Network. Since this was one of the first projects undertaken by the NQF, it was a prime opportunity for influencing the development of this organization. Clearly, this was an important opportunity for seasoned health care policy makers to shape and control this new organization that seemed to be so important to the future of health care.

Using the ANT lens, one can see that politics has a role to play in the ordering of the actors within a particular network. In essence, the actors create a network in a particular way to attain certain results, and how they do so has substantial political overtones. From this point-of-view, politics is seen in terms of associations among actors within a network. Networks consist of a series of social ties that are constantly reshuffled. Here politics is redefined “as the entire set of tasks that allow the progressive composition of a common world” (Latour, 2004: 53) and, “that associations¹⁶³ [within networks] should be *composed* in order to *design* one common world” (Latour, 2005: 259). And when networks are composed and designed, power becomes an important issue. Power is a network effect that emerges from the ordering of associations (and disassociations) of actors. Power shapes the identity of the network, connects it to the world around them, and is ultimately responsible for its collapse. According to ANT, power is not “stored up” but is produced through a process. It is a process whereby individuals are “enrolled rather than coerced, into the larger interests that are in play” (Munro, 1999: 430). Furthermore, power as viewed by ANT provides a certain amount of discretion to actors, specifically in terms of disassociation, i.e., the discretionary ability of defer or deny affirmations of membership (Munro, 1999:429). As theorized, power is exercised in networks by providing actors with the ability “to draw on and manipulate these effects to their (imagined) own ends” (Munro, 1999: 431). Yet, power relationships that rely on social ties typically do not last very long, hence “durability” of the network order often becomes a goal of actors in a network. Durability is a network’s ability to maintain its effects through a particular ordering of actors within the network. What makes networks durable is the presence of “immutable mobiles” or entities that remain “fixed” but can be “mobile” with in the network. However, the practicality of

¹⁶³ According to Latour, associations are considered “social,” and for ANT social “is the name of a type of momentary association which is characterized by the way it gathers together into new shapes” (2005: 65).

maintaining these associations, thus maintaining durability, has to be considered, specifically the constant investment in enrolling others and the cost paid for any extension of an interaction (Latour, 2005). Social ties, in this context means “something that has great trouble spreading in time and space, that has no inertia and is to be ceaselessly renegotiated” (Latour, 2005: 66). This is the precise reason that so much effort is directed toward replacing social ties with other types of links that maintain the network order.

One additional point should be made here, which thus far has been implicit in this discussion, specifically that networks, like the social, consist of asymmetries and inequalities. Not all actors are equal within the network which is largely due to the enrolment of actors to another’s cause and their ability to maintain the established order of actors within the network. ANT seeks to explain the source of asymmetries. In fact, if inequities exist, according to ANT, it provides proof that other actors are coming into play, thus generating the effect.

Overall, ANT and its view of politics provide a useful framework for this study. In the next chapter, I will describe the activities of the NQF Hospital Steering Committee in terms of a theatrical performance¹⁶⁴—where there is a stage, actors or performers, and a backstage often, but not necessarily, involving the same performers. The stage in this case is the setting of the Hospital Steering Committee meetings, conference calls and other venues; the actors are those individuals who participated in the discussions of the committee; and backstage can be viewed as those activities that support the activity on the stage, but that never came to light as part of the formal record of the committee discussions.

¹⁶⁴ The framework used here was developed and reported by Erving Goffman in *The Presentation of Self in Everyday Life* (1959). Goffman illustrates the ways in which individuals present themselves to others in work situations, how they control the impressions of others, and the kinds of things that they do (or not do) to sustain their “performance” (Goffman, 1959: xi).

The imagery of front and backstage is particularly appropriate when applying ANT. Consider what the audience sees in a theatrical performance. It is the final product, i.e., the play that is shown on stage, not the scurrying about behind the stage or the months of practice that went into finalizing the production. In a way, the backstage political maneuvering is put into context in that ANT focuses on actors and their relationships and network effects. ANT can be used to analyze the “forces and mechanisms of political power,” because the theory argues that power is not a property of an actor, but one of relationships established within the network that emerges only after translation is complete and the network of actors is stabilized (Latour, 1986). The last point is one difference from backstage political maneuvering, which usually involves a self-selected group of actors in which power is already established, and in which the actors often are proactive. ANT takes a broader perspective, and includes not only human actors but also non-human actors that may or may not have a significant influence on the relationships within the network. Thus, the network effect can be very different, depending upon several variables such as who is included in the network and what types of relationships are established.

In Chapters 4 and 5, the actors within this performance, both human and non-human, are described. They were selected based on interviews with key informants within the industry. These descriptions are germane to the deliberations described in these chapters because they shed some light on the relatively strong agendas of the actors involved. As one interviewee indicated, the actors’ comments and interactions during committee deliberations were “pretty much representing the organization that sent them to the meetings,” rather than reflecting their profession or personality. Finally, the backstage activities that I will describe should not be construed as anything diabolical but are presented to provide additional insight and information regarding the activities related to the Hospital Steering Committee and the selection of

performance measures for NQF endorsement, offering a “how to” primer, on how business is conducted in relation to developing consensus around measures of performance in health care.

Chapter 6 is organized into four acts that discuss both the on-stage activities (the formal record) and the backstage activities (the informal record). The formal record is derived from an examination of artifacts such as meeting summaries, journal articles, and press releases that reside in the public domain and were accessed through numerous searches of the literature and of the Internet, including archived documents on the NQF website. The informal record is gleaned primarily from interviews with key informants, but I also have included information from my direct observations and impressions and from my interactions with the actors as a member of the network. Piecing these elements together represents one interpretation of the activities that are based on the collective recollections of the key informants. To provide the reader with information about the context in which the Committee deliberations took place, an environmental scan of the state of health care is included in the appendix. Throughout the chapter, I have identified incidents that illustrate particular elements of ANT—punctualisation, enrolment, translation, inscription, and obligatory points of passage, which may help crystallize for the reader why certain measures are selected for endorsement and why some prevail over others.

I have suggested that throughout the deliberations of the Committee, the identity of performance measures changes based upon the making and remaking of the relationships among the actors of the network. In order to develop a richer understanding of how this occurs within this type of environment, one must suspend the notion that only humans have agency and can be designated as “actor.” Objects such as performance measures, according to ANT, can also be considered actors within the network. They influence and in turn are influenced by others, and thus are active participants in reconfiguring the relationships of the network. Because the

relationships within the network change, the face of performance measures also changes, which will be recounted in each of the acts of the CDP. Further, my purpose here is to account for how decisions regarding the selection of standardized measures are made. According to Latour and Woolgar, facts “only appear at the end of a long construction chain...” (1986[1979]). One does not see the controversies or how they are resolved or what compromises were reached. We only know the final product. This is true with regard to the NQF Hospital Steering Committee, its work and its final products. What is seen are the final products, which in this case involved the initial set of hospital performance measures and a framework for evaluating them, and not all the work, practical politics, and negotiating that were going on in the background.

For ease of reading and to focus the discussion, each act will begin with a sort of soliloquy by an imaginary narrator that will describe the element or elements of ANT that are pertinent to the act that will follow. Within each act, I have included both on-stage and backstage activities, some explicit and others implicit within the storyline. Finally, I describe what the identity of performance measures look like as the curtain falls in preparation for the next act.¹⁶⁵

In Chapter 7, I have chosen to follow “the fragile thread” of one actor, specifically performance measures, throughout the consensus development process. In typical ANT fashion, performance measures, non-human actors, play an active role within the network. Through their constant interactions with the various other actors of the network, performance measures (and all the other actors) are inscribed with all the interests and deliberations of the network that have gone before. Performance measures as objects can influence other actors, but can also be

¹⁶⁵ The reader should also be aware that there are a number of quotations throughout the various chapters in this section that do not have an associated reference. These are statements made by key informants that serve to illustrate the scenario, and to maintain confidentiality, their names have been withheld.

thought of as an archive of interests and events that were constantly being made and remade through time. Thus, this chapter will explore the multiple identities of performance measures and will trace their transformation as the events of the Hospital Quality Network unfolded.

CHAPTER 6:
THE NQF HOSPITAL MEASURES STEERING COMMITTEE

Act I: Setting the Stage for the Hospital Project (1999-2001)

The Narrator

This first act basically sets the stage for the NQF's Hospital Project and reflects the ANT concept of punctualisation. This concept "holds that in any specification an entity entails the substitution of a network by a point" (Law 1992b: 385). Entities, as a matter of course, are treated as a single block or point; yet beyond each point exists a complex network of people and things, ordered in a particular way that makes social life possible. It can be thought of as "a consolidation of the network" (Law, 1992a). Law uses the example of a television to make the point. "Most of the time a television is a single and coherent object with relatively few apparent parts. On the other hand when it breaks down, for that same user—and still more for the repair person—it rapidly turns into a network of electronic components and human interventions" (1992a: 4). Normally, these networks are invisible, but occasionally they become visible. Alan Prout provides an excellent example of punctualisation when the network behind the unified point becomes visible:

...normally I treat the computer on which I am writing this paper as a single block. When, however, it recently broke down and I had to call on the supplier's guarantee, part of the network (made up of people, things and their organization) that stands invisibly behind the machine were revealed. Similarly, when as social scientists we want to know how a device came into being and what goes into its existence, the unpicking (at least partially) of the punctualisations which efface the networks behind it becomes a key task (Prout 1996: 201-2).

Packaging networks as a single entity not only makes social life possible, but also provides some stability and durability. So, a single entity (human or non-human) “can be seen as packaging a network and extending it though time and space; it can ‘delegate’ a network, standing in for it, repeating it and performing its work in times and places remote from its organization” (Latour 1991: 261).

In the upcoming act, the announcement made by the CEO and President of NQF about the hospital project is an example of punctualisation. As the reader will see, the on-stage performance of NQF “packaged” the network, consisting of multiple actors that were working on the issue of burden and measure proliferation in prior years. Thus, the NQF punctualised the complex network that already existed.

On-stage: Announcing the NQF Hospital Project

On December 7, 1999, the NQF announced in its newsletter¹⁶⁶ that it would launch a project to “standardize hospital performance measures for the nation’s approximately 6,500 general acute care hospitals.” The impetus for the program was the proliferation of measures and the burden of data collection among providers.

Currently, hospitals across the United States use a wide variety of measurement systems and performance indicators to assess their quality of care. The number of such performance measures has increased in recent years. Hospitals are committing substantial, and increasing resources on data collection and measurement as both consumers and purchasers demand greater accountability from health care providers. Unfortunately, since the various measures are neither uniform nor standardized, consumers and purchasers of health care cannot use the data to make comparisons about quality of care (NQF, 2000b).

¹⁶⁶ NQF (2000). *National Quality Forum News*, accessed on May 2, 2001.

To combat this problem, the NQF selected as one its first projects the standardization of measures for the nation's hospitals.¹⁶⁷ According to Ken Kizer, President and CEO of the NQF,

The project will occur in two phases, with the first phase taking about a year and involving an evaluation of measures to be recommended by the Joint Commission, as well as other measures, and identifying gaps in the measure sets. The second phase will be a more prolonged effort that builds on phase one to examine additional measurement opportunities, expand the data set and fill any gaps identified. He also stressed that this would be a widely inclusive effort, involving the Joint Commission, the AHA, and other hospital associations, providers and professional associations, consumer groups, purchasers, and others" (NQF, 2000b, National Quality Forum News).

Backstage: Overture to the NQF Hospital Performance Measures Project

As a prelude to the announcement of the NQF Hospital Project, President Dick Davidson of the American Hospital Association (AHA) convened a task force that was designed to help AHA staff understand how to best help the members in their efforts to improve the quality and safety of care delivered in hospitals. The message from the task force was that the hospitals felt that they were being measured by a number of organizations and that the measures were inconsistent; because of this, the data collected did not yield any useful information. For example "according to one measurement set, our hospital looks as if it's the best in town for heart care, but by another measurement set, we are pretty mediocre." To improve, hospitals need to know what to believe and where to most effectively use limited resources. In essence, hospitals requested that the AHA to try to bring about a consensus on the national level as to what measures ought to be used to assess hospital quality. A second message from the AHA task

¹⁶⁷ This was one of the three projects that NQF undertook after its incorporation in 1999, but unlike the other two, this project was the first that dealt with performance measures.

force was that there was a perception by the general public that the hospitals were unwilling to share information about the quality of care they were providing to the communities they served. The hospitals indicated, however, that they would like to provide information about the quality of care delivered, but there was no reliable mechanism to do that. The membership requested that the AHA work with other national organizations to establish a reliable mechanism to report hospital quality to the public.

Obviously, the AHA had to determine if these concerns were an issue for their membership only, or whether it was applicable to the larger hospital industry. They met with the other two national hospitals associations—AAMC and the FAH and determined that those issues uncovered by the AHA task force were the same for both teaching hospitals and for-profit hospitals. Therefore, it was to their advantage for these organizations to band together and speak with one voice, which provided not only a certain clarity to their message but also additional political clout when negotiating with other large national health care organizations.

The first step on the task force’s recommendations was to convene a meeting of several organizations to discuss the quality measurement agenda for hospitals. The invitation was issued by the President of the AHA, and his organization acted as the convener; however, the letter was very explicit that the request for a meeting was coming from all three hospital associations. Representatives of CMS, AHRQ, the Joint Commission, the NQF, AHA, FAH, AAMC, and the AMA were invited and participated in the discussions.¹⁶⁸ The theme of the meeting was to discuss the state of the current hospital measurement activity, with an eye toward “moving the measurement agenda forward.” The three associations indicated that overall, the hospitals were “struggling with the disparate requirements and measurement activities” that currently exist

¹⁶⁸ This is an example of the ANT element of “punctualisation” or the ability of one entity to speak for a network of entities; in this case, the hospital associations spoke for and actually replaced the 6,500 hospitals in the country at the discussion table.

within the industry. At first, the leaders of these organizations, with the exception of one, thought that the hospitals were “exaggerating the problem.” But according to an interviewee, “it didn’t take long before they began to realize that in fact while they had common goals and common ideas, they didn’t have a common execution of those.” Thus, some of the impetus for the Hospital Project came from the hospitals themselves. These discussions set off a series of events that would ultimately end with the NQF Hospital Steering Committee.

Backstage: Measure Development for Quality Initiatives

On the national scene, the primary initiatives for hospital quality come from the federal government, explicitly CMS, and on the private side, the Joint Commission. Prior to the announcement of the NQF Hospital Project, both organizations were working to develop their own measures for their quality improvement initiatives. In 1998, CMS initiated field testing of a set of hospital measures evaluating performance in three clinical areas of hospital performance within the QIO program. The clinical areas were myocardial infarction, heart failure, and pneumonia. At about the same time, the Joint Commission was interested in field testing its own measures for the same three conditions and found a potential laboratory in the State of Rhode Island.¹⁶⁹ It just so happened that both CMS and the Joint Commission representatives were at the table in Rhode Island discussing their measures and field tests; what became clear during these meetings was that the specifications of the measures for a single condition were not the same. Eventually, Rhode Island chose to pilot-test the Joint Commission measures, but more important, both CMS and the Joint Commission realized that they had to come to some agreement on measure specifications.

¹⁶⁹ Rhode Island was only one of five test sites that participated in the Joint Commission’s field test of hospital performance measures.

The issue of alignment of measures became even more important with the announcement of the NQF Hospital Measures Project. Both organizations would have representatives on the Hospital Steering Committee, both planned to submit their measures for endorsement, and neither wanted to compete with the other, as both organizations would lose. As a result, CMS and the Joint Commission started meeting to align measure specification on heart disease and pneumonia, and according to one interviewee who participated in these talks, “it was *amazingly hard*” [original emphasis]. Both groups, especially the Joint Commission, were entrenched in its position, and the negotiations, characterized as “intensive” by one interviewee, took approximately eight months of regular weekly meetings to achieve agreement. Both groups were “entrenched,” because each had vested interests in making sure that the measures they currently were working on were selected for national reporting; many felt that their own identity and professional judgments were “on the line.”¹⁷⁰ The measures then were presented to the respective organizational leadership, and ultimately jointly submitted to the NQF for endorsement. Thus, alignment was achieved because of a fortuitous meeting between CMS and the Joint Commission with state representatives of RI, and the unintentional pressure exerted by the NQF through its endorsement process.

Also, around the same time, CMS was considering proposed rule-making for hospitals to submit their performance data, much like nursing homes and home health agencies, under the conditions of participation for Medicare and Medicaid reimbursement. The conditions of participation were rewritten to include a stipulation for hospitals to report quality data, and CMS was ready to move forward with the proposed rule. Unfortunately, the timing of CMS was poor

¹⁷⁰ Measure development is very expensive in terms of resources. For example, CMS invested staff, financial resources and time in measure development, which can account for millions of dollars depending upon the measures being developed. Field testing of developed measures is also expensive—CMS invested from \$4-8 million to test its measures through the QIOs.

because in 2000, the George W. Bush Administration came into power and placed a moratorium on all regulation; hence the conditions of participation were never approved or implemented. In essence, CMS had to find another avenue through which to collect hospital performance data.

Backstage: The Department of Health and Human Services Quality Initiatives

To tell this part of the story, one must return to several initiatives launched in November 2001 by the Department of Health and Human Services, specifically its Quality Initiatives, which were undertaken to “...assure quality health care for all Americans through published consumer information...” The initiatives’ intent were to a) empower consumers to make more informed choices about their health care and b) to encourage providers to improve the quality of health care. The Bush Administration took on the initiative in an effort to make the health care market function more like other markets where consumers could select, though easy to understand, and accurate information, those services of high quality and better value, thus eliminating those that were over-priced and of substandard quality. In order to do that, however, there had to be a way to balance the consumer-supplier equation. Traditionally, the knowledge and information regarding health care performance and quality resided with the providers—the suppliers of health care. This initiative was designed to provide accurate information about care provided by hospitals, nursing homes and other entities in an effort to “let the market forces” work in health care. The power of informed markets to improve performance was the centerpiece in these initiatives.

The first effort, known as Nursing Home Compare, focused on nursing homes and provided information about the facility, quality measures, inspection results, and staff information. Consumers could go to the government-sponsored website and select and compare various facilities, based on the information provided. This initiative was followed by Home

Health Compare, which provided information about home health agencies, and then by Hospital Compare, which would focus on hospitals.

The major differences between the first two initiatives and the hospital initiative are threefold. First, CMS had validated standard measure sets for nursing homes and home health agencies with an established data transmission infrastructure, and the hospitals did not. Second, CMS paid nursing homes and home health agencies for submitting data to post on the government website. Although CMS considered incentives for hospital participation in reporting, it could not come up with a feasible plan that included payment. And finally, nursing homes and home health agencies were mandated by CMS, through the conditions of participation, to provide quality data for public reporting. Hospitals did not have such a mandate; nor was there a standardized measure set that could be used to report on hospital quality.

An internal work group at CMS was charged to look into various ways to entice hospitals to provide performance data for comparative public reporting. Since there was no way to mandate that hospitals provide this information, CMS Administrator Thomas Scully announced that the Hospital Reporting Initiative was going to be “voluntary.”¹⁷¹ But he also made it very clear that people “needed the information now” and that if voluntary participation by hospitals was not forthcoming, then “the government will step in.”¹⁷¹ With that announcement, CMS staff began pursuing talks with the AHA, AAMC, and FAH about making this effort happen. Talks ensued, with final agreement reached on the idea of hospitals volunteering to publicly report their performance in August 2002.

¹⁷¹ The Administrator made these remarks at the 2nd Annual National Quality Forum Meeting, held in 2001. However, after discussions with senior staff from the Office of the Secretary, DHHS was very opposed to CMS moving ahead with proposed rule-making.

Backstage: The Network and the Health Care Environment

To better understand the context in which the NQF Hospital Project originated, and in which its Steering Committee had to operate, an environmental scan at about the time the Committee convened is presented in Appendix D. Several areas and factors influencing the health care industry are listed, with some general comments about issues that surfaced in 2000-01; however, the list is not meant to be all-inclusive, but merely to provide a snapshot of issues of concern to the industry at that time.

The environmental assessment of 2000-01 provides a context in which the Hospital Quality Network coalesced and the NQF's Hospital Steering Committee sought to undertake its charge. I selected this particular Committee for examination because it was one of the first projects undertaken by the NQF, and it serves as a microcosm of the larger Hospital Quality Network, which is too unwieldy to examine here. Since the major actors of the Hospital Quality Network are represented within this Committee, one can potentially use the interactions, documents, and artifacts produced by the Hospital Steering Committee to explain phenomena that occur within the larger network.

Performance Measures as a Network Actor

As the curtain falls on this first act, performance measures, or more specifically their proliferation, are viewed as somewhat onerous and a burden to the providers of hospital care. Previously, performance measures were used primarily for internal quality assessment and improvement. More recently, however, with greater demands for accountability from consumer and purchaser groups, hospitals have been expending substantial resources to collect data and to report on their performance. Hospital providers agreed that they would like to measure their performance and to provide information about the quality of care they give to others; however, they wanted to "tell their story in their own way." The sheer number of requests for data to

populate the large number of performance measures and the resources expended in collection were becoming increasingly unwieldy. Measuring performance increasingly was viewed as burdensome, in that it consumed a large amount of hospital resources without providing much in the way of a return. Thus, as this act closes, performance measures, used primarily for quality improvement in the past, seem to be poised to take on another important role, which is providing information on accountability.

Act II: The NQF Hospital Steering Committee and the Consensus Development Process

(March 2001-May 2002)

The Narrator

The second act opens with the selection and formalization of the Hospital Steering Committee. Members of this committee were chosen from the larger pool of actors within the Hospital Quality Network, and this committee can be portrayed as a sub-network of the larger one. The interactions of the actors, including performance measures, take place within this act. And as a central tenet of ANT, the process of social ordering, and patterning within a network is constantly changing and being reconfigured.

An analysis of these processes, called translation, is a key concept of ANT. Specifically it is concerned with how a set of heterogeneous actors with their own inclinations overcome resistance and form a single punctualised actor (Law 1992a). Recall the four major stages of translation: problematisation, interessement, enrolment, and mobilisation. Problematisation is when an actor defines identities and interests of others that are consistent with their own interests, and establishes itself as an obligatory passage point,¹⁷² thus “rendering itself

¹⁷² Obligatory point of passage can be defined as “points that constitute unavoidable conduits through which entities must pass in order to articulate both their identity and their *raison d’être*” (Singleton and Michael 1993: 229-30).

indispensable” (Callon, 1986). The second stage is *interessement*, which is a process by which the focal actor seeks to have its definitions accepted by the other actors within the network. *Enrolment*, the third stage, is a process whereby actors co-opt or enrol each other to pursue individual interests and objectives, or those of the collective network. “Entities mutually enrol each other into a combination of some type, claiming to speak for each other, interpreting, configuring and reconfiguring each other” (Latour 1991: 261). This is where the issue of power becomes evident. Actors can choose to be enrolled by another actor, or they may decline. Obviously, the more actors enrolled in support of a particular cause, the greater the likelihood of success. The final stage of the translation process is *mobilisation*. According to Murdoch (1997), a successful network would only emerge if the fourth stage of the translation process was completed. *Mobilisation* is the process by which the enrolled actors become organized to address the problem identified by the focal actor in the initial phase of translation i.e., *problematization*. And in order for the network to be maintained, the focal actor must sustain its representation of the enrolled entities. As a result of translation, ordering effects are generated, such as devices, agents, institutions, or organizations (Law 1992a), which in turn provide some durability and stability to the network order.

What follows is a recounting of the events of the Hospital Steering Committee and the discussions that ensued as it struggled to define the network order. Careful attention to the role of the NQF and the NQF staff is essential, as they were instrumental in the translation processes of the Committee. This act illustrates the process of translation that occurred during the Committee’s debates to select a draft set of performance measures for hospitals.

On-stage: The Formation of the NQF Hospital Steering Committee

In response to a contractual request from the CMS and the AHRQ, the Hospital Care Performance Measurement Project initiated the NQF's effort to develop consensus around an initial set of quality performance measures for acute care facilities. Many in the industry contended that because there were a number of measures for the acute care setting, this was "low hanging fruit" and the project could be completed in a relatively short timeframe.

To launch the project in March 2001, the NQF appointed the members of the Steering Committee, which consisted of a variety of individuals with knowledge and expertise in hospitals and the hospital industry. This committee was convened in an effort to "establish and demonstrate the ability to achieve voluntary consensus on standards related to public accountability in health care quality, with detailed and balanced participation of hundreds of stakeholders with traditionally oppositional perspectives" (NQF, 2001b). According to the NQF, the "implementation of the CDP is an exercise in democracy with unique characteristics."

As shown in the timeline (Appendix F), the NQF, via an internal selection process, began contacting potential steering committee members prior to the AHRQ contract that was being negotiated but not yet in place. It selected co-chairs for this committee, a physician and member of the NQF, and a nurse who was not a member. According to several interviewees, the Steering Committee co-chairs were not selected due to their clinical expertise, but rather for their ability to work across disciplines with competing interests. The committee itself consisted of 15 participants who represented all four councils of the NQF and five liaison members that included the CMS and AHRQ.¹⁷³ The committee was completely structured by NQF staff, and membership was determined without a call for nominations. In relation to the four councils, the composition of the committee consisted of five representatives from the Provider and Health

¹⁷³ For a list of the Steering Committee members and the organizations they represent, see Appendix G.

Plan Council, three from the Consumer Council, seven from the Research and Quality Improvement Council, three from the Purchaser Council and three representatives who were not members of the NQF.¹⁷⁴

The co-chairs and the committee were provided direction regarding critical tasks as designated by the funders¹⁷⁵ of the project—first selecting a set of quality indicators, and second, establishing a framework for future review and evaluation of measures. The overall scope of the project was to:

1. Assess and endorse an initial set of existing measures that are reasonable indicators of hospital quality, and that are useful to consumers, purchasers, hospitals, and quality improvement organizations alike. Among measures that will be considered, for example, are those being recommended or applied by the Joint Commission, the state Peer Review Organizations, and major private sector purchasers of health care.
2. Develop a comprehensive framework and standardized process for hospital quality measurement, including a process for updating the initial measurement set as new measures become available, and identify gaps where research is needed to develop appropriate measures. The work of the NQF Strategic Framework Board will provide the underpinning for this phase of the project (NQF, 2001b).

¹⁷⁴ According to NQF policy, to obtain the required expertise for Steering Committees, the staff solicits its membership first, and if the needed expertise and willingness to commit to the required time and workload are not found among its membership, then the staff may go outside the membership to fill vacancies on the various committees of active projects.

¹⁷⁵ Funding for this project was provided initially by the U.S. Department of Health and Human Services (AHRQ and CMS), with planning support provided through grants from The Robert Wood Johnson Foundation, California HealthCare Foundation, The Horace W. Goldsmith Foundation, the Department of Veterans Affairs, the United Hospital Fund of New York, and the U.S. Office of Personnel Management.

*On-stage: Beginning Work on an Initial Performance Measurement Set for Hospitals*¹⁷⁶

The first meeting of the Committee took place on March 21, 2001.¹⁷⁷ It was open to the NQF membership (34 members attended) as well as the general public. In that meeting, Dr. Kizer indicated that the task at hand was to “agree on a set of measures that can be used to assess how a hospital is performing; to increase the value of performance indicators currently in use; and to reduce the burden of data collection” (NQF, Hospital Measures Steering Committee Meeting Summary, March 21, 2001). One of the critical tasks of the Hospital Steering Committee was to operationalize what a voluntary consensus development process for health care. Therefore, the objectives, as recorded in the Meeting Summary were to:

1. provide an understanding of the Steering Committee’s roles, tasks, and the current state of hospital performance measurement
2. establish the scope and priorities of the initial measures set
3. provide direction for a process to identify and screen candidate measures
4. identify issues for an evaluation of the process, including criteria to be considered

The timeline presented was short, denoting a sense of urgency associated with this project, and NQF expected an agreed upon hospital measure set by December 2002. As one of the Co-chairs mentioned, due to the urgency of the project, it was important to “not let the perfect be the enemy of the good,”¹⁷⁸ which is a phrase that was used by the various stakeholders throughout the committee deliberations. Also, it should be noted that the Strategic Framework

¹⁷⁶ This section is organized to recount those issues and events that were unique to a particular meeting. A summary of the issues common to all the Committee’s discussions will be provided toward the end of this section.

¹⁷⁷ It should be noted that the first meeting summary of this committee was available to the general public via the NQF website (www.qualityforum.org). However, subsequent meeting materials for this project, while available at the meeting, were only available in the members-only section of the website and eventually, meeting minutes and summaries were only available to committee members and members of the NQF.

¹⁷⁸ There are those in health care, particularly the providers, who feel that if they are to be measured, the instrument should be as precise as possible, and if the tool is considered “less than perfect,” it should not be used. On the other hand, some groups such as consumers and purchasers, want more measures that describe a much broader scope of the quality of care in health care, and believe that it’s not necessary for them to be perfect, but only “good enough.”

Board was in the process of developing a general framework for the NQF and its policies, which would eventually provide some direction to the committee via a draft report expected on May 1st.¹⁷⁹ But at the time the Hospital Steering Committee met, there was no guidance, other than what was stated by Ken Kizer and the NQF staff. Concern was voiced specifically by the providers and researcher members of the Committee about moving forward without the benefit of guidance from the Strategic Framework Board, and the absence of a formal framework. However, as was suggested by the other members of the Committee that “the urgency of addressing the increasing number of measure sets in use outweighed the benefits of waiting, and that it is likely that areas the committee considers to be highest priority are also likely to rank high in whatever framework is proposed” (Hospital Measures’ Steering Committee Meeting Summary, March 21, 2001).

The selection and screening of candidate measures was also discussed, from which several criteria emerged related to consideration of measures for this initial set. Specifically, candidate measures should be:

1. Relevant to high-volume conditions treated in hospitals that are prevalent and causing significant morbidity and mortality
2. Evidence based¹⁸⁰
3. Actionable, i.e., things that can make a difference
4. A balance of condition-specific and non-condition-specific measures
5. Relevant to the population in general, not just one or two subpopulations
6. Related to things the hospital can control or influence

¹⁷⁹ The SFB recommendations were approved in May 2002.

¹⁸⁰ Evidence-based medicine (now known as evidence-based practice) is defined as “practice based upon a foundation of solid science, especially using research that has applied rigorous epidemiological methods and has been published in peer-reviewed journals. It involves increased reliance on formal, systematic analysis and synthesis of the research literature to determine clinical effectiveness” (Eisenberg, 2001: 369).

7. In the public domain, that is not proprietary
8. Focused on the service, not the business of health care¹⁸¹

As with every meeting the NQF conducts, there is an opportunity for the public to comment on the discussions. The comments from the public included a basic agreement that measures should be evidence-based, non-proprietary, and make a difference to those using them. The comments then ranged from consideration of the importance of “never events” measures to inclusion of measures specific to mental disorders, patient functional status and pediatrics. Staffing ratios were also mentioned as constituting an area of importance that should be considered for inclusion in this initial measurement set.

The second meeting of the Committee was held on May 22, 2001, and the agenda included two items for discussion—establishing a framework for hospital measurement, and discussion of candidate measures for the initial set (NQF, Hospital Measures Steering Committee Meeting Materials, May 22, 2001). The framework would guide future efforts to expand, update, and recommend research necessary for the development of a comprehensive set of hospital performance measures. According to one interviewee, it was “sort of like building a road, while we were building the trucks that were driving over the top of the road at the same time.” The starting point for the discussion was the draft framework¹⁸² of the NQF Strategic Framework Board (SFB). The SFB deliberated for approximately 18 months to develop a framework that outlined a national strategy for health care quality measurement and reporting. The framework was offered as general guidance and to begin discussion that would form the basis of the hospital quality measurement framework. The materials were divided into two parts,

¹⁸¹ These and the other sets of criteria that the committee established can be viewed as “obligatory points of passage,” an element of the ANT. Actors, and in this case performance measures, must pass through unavoidable conduits in order to establish their identity and reason for being.

¹⁸² The draft framework given to the Committee had not been reviewed or endorsed by the NQF Member Councils or the Board of Directors, and therefore was subject to change based on the public comments and subsequent voting of the membership.

with the first summarizing the SFB findings and the second including recommendations (as well as the NQF staff assessment of potential implications for the hospital-specific framework).¹⁸³

Table 7 provides an abbreviated summary of the SFB goals and the NQF staff recommendations in several areas.

Table 7: Draft Recommendations of the SFB and NQF Staff Recommendations

	SFB Draft Recommendations	NQF Staff Comments
Goals	<ul style="list-style-type: none"> • Related to products of the healthcare delivery system • Related to conditions that are prevalent or have a high risk of disability, or death • Represent the needs of diverse populations • Based on evidence that effective care strategies exist • Supported by expert groups and compelling to relevant constituencies.¹⁸⁴ 	<ul style="list-style-type: none"> • Hospital specific goals should be consistent with the overall all priorities of the health system—set priorities by what is common and important to the overall health care delivery, not what is common only to hospital care. • The hospital framework should be flexible to accommodate changes over time, with the creation of a mechanism to accommodate local clinical priorities.
Establishing a Common Set of Measures ¹⁸⁵	<ul style="list-style-type: none"> • Should be parsimonious • Linked to national goals • Have clear and compelling use • Be continually improved upon based on feedback • Not impose undue burdens on those who provide data. • Other attributes include: clear, standard definitions; collect data once, as close to the source as possible; and collect data in such a way that is can be combined, analyzed and reported to serve a wide variety of purposes. • Candidate measures should be evaluated using specific criteria (See Table 5). 	<ul style="list-style-type: none"> • Measures might be viewed first from the perspective of usefulness for improvement, and then for selection and accountability. A common measure set could include a variety of measures across not only clinical priority areas but also different elements of the information network. • The roll up of data elements into summary measures appears to have a significant effect on the design and selection of specific measures for a common set.

¹⁸³ Note the disclaimer that was included with the materials: “The items presented here are selected for their immediate salience to a hospital framework; they will not necessarily correspond directly to the SFB’s own summary of recommendations.”

¹⁸⁴ The SFB also indicated that a balance should be sought among national goals, national clinical priorities and local priorities. Specifically, mechanisms could be in place to permit communities to prioritize their efforts according to their needs and resources, while actively participating in the national quality improvement efforts.

	<ul style="list-style-type: none"> • Measure systems should include an audit standard; definitions for each data elements; and be able to transfer seamlessly across different types of institutions. 	
Consumer Roles in Measurement	<ul style="list-style-type: none"> • They are co-evaluators of care including evaluating experiences with care or the outcomes of care. • They are co-producers of care that includes how well a provider is performing in terms of support and utilization of the patient's skill in managing their condition. 	<ul style="list-style-type: none"> • How best to integrate the consumer role in a measurement system based on data important for clinical processes has not been directly addressed by the SFB, and may be a part of the hospital measurement framework.
Reporting	<ul style="list-style-type: none"> • Requires measures that are consistently available to the public • Measures that are understandable, timely, provider specific, and operational • Greater public awareness about quality issues • Widespread availability of quality information 	<ul style="list-style-type: none"> • A comprehensive hospital measurement framework should consider how measures could be presented to the public. The need for understandable measures has implications for selection of measures for the common set.

¹⁸⁵ Measures in the common set should serve both internal quality improvement efforts and for accountability and selection, i.e., to help consumers select health care plans, providers, and treatments.

Keeping the SFB recommendations in mind, the Steering Committee had to make several decisions to develop a workable plan for the project. The issues for discussion included:

- Setting goals and priorities
- Delineating components of a common measure set
- Evaluating candidate measures for a common set
- Auditing the data
- Reporting

Again, the staff made recommendations as to how the Steering Committee should handle the work. Three work groups were established—one for hospital measurement priorities and components, one for evaluation and auditing, and one for reporting.¹⁸⁶ Each work group was to develop the appropriate guidance and principles relevant to its component of the framework and report to the full Steering Committee at the next meeting.

On-stage: Identifying and Screening Candidates for an Initial Measure Set

To begin the Committee's second task, the NQF staff identified a pool of potential measures for inclusion in the core measure set. Although other sources were mentioned as possible sources of candidate measures, the Committee generally was "dubious of the value of an open call for measures, and suggested that any such call should be very specific regarding the criteria for measures that would qualify" (NQF, Hospital Measures' Steering Committee Meeting Summary, March 21, 2001).¹⁸⁷ National-level compilations and measure sets were examined

¹⁸⁶ See Appendix E for a list of each of the work group members.

¹⁸⁷ This decision would have implications for the project at a later date. The reasons cited for not doing an open call for measures was that it would be "very time-consuming and had recently been done in the development of other measure sets (e.g., the Joint Commission's ORYX measures)" (NQF Steering Committee on Hospital Care National Performance Measures, October 2002).

and identification was based on the criteria for prioritizing potential topics that were suggested at the last Steering Committee meeting and are described below.

The scope of the project was limited to acute inpatient and emergency care, and excluded rehabilitative, outpatient and long-term care services. This was done to limit the number of measures to consider, to select measures that would generally be within the control of hospitals, and to provide measures that could be comparable across hospitals. The priority topics for measurement were selected based whether they: a) addressed a common reason for hospitalization or b) addressed a specific high-priority cross-cutting area identified by the Steering Committee. From this, 28 clinical conditions¹⁸⁸ and four cross-cutting topics were identified.¹⁸⁹ Several conditions of importance to public health were not included on this list, as they were not among the top-listed inpatient conditions. The measures then were grouped into categories by the NQF staff. The categories and the measure sets for each are shown in Table 8 below.

Table 8: Measure Sources Identified by NQF Staff

<p>Category A: Sets of measures that are known to have undergone previous extensive evaluation and that are being implemented. They are:</p> <ul style="list-style-type: none"> • HEDIS (developed by NCQA) • ORYX Core Measures (established by The Joint Commission) • PRO 6th Scope of Work (required by HCFA under contract with PROs) • HCUP Quality Indicators (developed by AHRQ) • FACCT measures (developed by the Foundation for Accountability)
<p>Category B: Measure sets being used by large purchasers and that have national reach or recognition. Measures identified in this category were being used by:</p> <ul style="list-style-type: none"> • CARS effort (Coordinated Autos/UAW Reporting System) • National Business Coalition on Health • Pacific Business Group on Health • The Leapfrog Group

¹⁸⁸ These were identified primarily through the HCUP hospital discharge database. See Appendix H for a list of priority areas identified from the HCUP database for the years 1996 and 1997.

¹⁸⁹ These measures were identified by the Steering Committee and/or public comment at the last meeting.

Category C: Other national compilations of measures. These were:

- The National Association of Psychiatric Health Systems Benchmarking Indicators
- Private-sector measures compiled by the Institute of Medicine for its “Envisioning the National Health Care Quality Report”
- Federal performance measure inventory compiled by the Quality Interagency Coordination Task Force (QuIC). This includes the Veterans Health Administration quality indexes
- CONQUEST (AHRQ’s performance measure database)

The staff excluded measures based on certain criteria. Those excluded were measures that:

- appeared to relate primarily to rehabilitation, long-term care, chronic care or non-emergency outpatient services
- reflected a composite of inpatient and outpatient activities
- appeared to relate solely to utilization, except for surgical and cardiac procedure volume measures, where there is strong evidence that volume is linked to quality
- did not appear to be largely within the control of the hospital
- was identified as still in development
- were proprietary

As a result, all diabetes measures and all asthma measures were excluded from consideration, because they generally did not relate to care provided to inpatients. Utilization measures also were excluded as well as some patient safety and end-of-life measures that were still in development. The resultant list of candidate measures was provided to the Steering Committee in a table that included approximately 140 measure descriptions along with the

owner/developer identification. From this list (consisting of two tables¹⁹⁰), the Steering Committee was first to determine if the process the staff used to derive it was reasonable and appropriate, and then to identify which measures were the strongest candidates for potential inclusion in the initial set. The NQF staff made recommendations indicating that candidates selected for further evaluation should be those measures that appeared in category A, i.e., HEDIS, ORYX, PRO 6th, HCUP, or FACCT or in a category B measure set. The staff put together an additional table that presented a narrower list of candidate condition-specific measures using specific criteria. Measures were included in this table based on the value placed on:

- measures that have undergone extensive previous evaluation
- some degree of demonstrated agreement, i.e., measures appearing in more than one measure set
- measures presumably with greater extent of use (measures included in at least two measure sets)

Using this approach, the measure set was limited to 27 measures, which according to the NQF staff was more manageable. Cross-cutting measures such as coordination of care and patient safety consisted of 59 measures that would be considered for inclusion in the core measure set. Appendix I shows the measures that were under consideration at the time as well as the measure developer. From the table, it should be noted that there are several “competing” measures, that is, where two or more measure developers have developed a measure for a specific condition. As the NQF selects only one measure for endorsement in a particular area, competition among measure developers occurs.

¹⁹⁰One table contained the full list of candidate condition-specific measures identified, and the second table consisted of cross-cutting measures.

On-stage: The Draft Consensus Reports

An initial performance measure set for hospital care.

The Steering Committee Meeting held on February 26, 2002 initiated the final selection of measures. The objectives of the meeting were to select and recommend a final core measure set for hospitals and to provide feedback on the draft Hospital Performance Measures Report to the NQF membership. The NQF staff activities and recommendations were the basis of discussion, and these materials were provided for Committee members.

Priority areas for further discussion included acute coronary syndrome, cerebrovascular disease, community-acquired pneumonia, congestive heart failure, maternal/neonate care, patient safety, and pediatric heart surgery. Of the 37 measures considered, 31 were included in a preliminary selection, based on detailed staff evaluations. At the time of the last Steering Committee meeting, several other issues were raised that needed further investigation. They included:

- Frequency of key conditions/procedures¹⁹¹
- Risk Adjustment¹⁹²
- Volume-Outcome Relationship¹⁹³ and
- Burden¹⁹⁴

¹⁹¹ An analysis of the frequency of key procedures and conditions was conducted to determine the extent to which hospitals provided care for sufficient numbers of patients with these conditions that would yield sufficient sample sizes to support public reporting of the measure.

¹⁹² Staff reviewed the risk adjustment methodologies that related to the measures under consideration. They recommended that a less burdensome method of risk adjustment be selected and that one method should apply to the whole measure set.

¹⁹³ A literature review of the relationship between volume measures, outcomes and quality was conducted by staff. They recommended to “couple” volume and outcome measures, to risk adjust those measures based on administrative data, to consider defining high and low volume, to explain interpretability when reporting, and to review the measures regularly.

¹⁹⁴ The burden placed on hospitals to collect and report on these measures was investigated. The staff concluded that some burden is associated with performance measurement, that risk adjustment increases burden and that lower burden methods may be as precise as those requiring more complicated methods.

Based on the criteria established, of approximately 140 measures originally considered, 27 were to be considered for further evaluation. Since the time of the last meeting, a number of measures that had been previously unavailable for evaluation were made available to NQF staff. These included measures in priority condition areas of safety, pediatrics and pain management. This brought the total measure set to 61 measures, which were categorized by NQF staff. The results of the evaluation were that three measures were “highly recommended;” 32 were “recommended;” five were “recommended as a pair;”¹⁹⁵ 13 were “recommended with explanation;” and eight were “not recommended at this time.”

The Steering Committee’s task was to come to consensus not only on the measures included, but also on the measures excluded and those not recommended. Other areas that required recommendations included issues relevant to future research, and the development and the validation of measures. In April 2002, the Hospital Steering Committee forwarded a set of hospital measures, its first deliverable to NQF members and others for comment as part of the pre-voting review of the Consensus Development Process (CDP). The report included a summary of all the measures considered by the Hospital Steering Committee (61 measures); however, the memo accompanying the draft report indicated that only the italicized measures were the ones included in the core measure set that were ultimately recommended, resulting in a total of 32 measures. Comments on the draft report were due to the NQF on May 23, 2002.

A comprehensive framework for hospital care performance evaluation.

Originally, the process to establish a framework by which the Steering Committee could select measures was divided into three working groups.¹⁹⁶ However, with subsequent meetings

¹⁹⁵ Based on the strength of the evidence, it was determined to reconsider two mortality measures if they were paired with volume measures for the same condition, e.g., pediatric heart surgery mortality paired with a measure of pediatric heart surgery volume.

¹⁹⁶ See Appendix E for a list of the working groups and their memberships.

of the Steering Committee, “the members of the Committee recognized that while the three working groups would inform and support the Steering Committee’s deliberations on developing a framework, that more focused work by the full Committee was needed on developing a framework.” Key decision points that were considered included:

- Should there be separate frameworks for measurement and reporting? It was determined by the Committee that one framework would guide measure selection, development and reporting.
- Does the adaptation of the Strategic Framework Board’s (SFB) recommendations to this area work as presented? It was determined that while the SFB’s overarching framework was useful as a “filter” for developing a hospital-specific framework; it “lacks sufficient direct applicability and usefulness to be the sole driver of the hospital performance measure framework.” Alternative suggestions from the group included a framework based on key patient questions such as “will I get the care I need, when I need it, at the lowest possible cost?” and developing a framework that is designed to include measures in areas of concern, such as emergency services, and staffing-sensitive measures, such as overcrowding and surge capacity. In the end, there was no consensus of the Committee on what approach would produce a practical and complete framework.
- Should a hospital framework incorporate separate sets of measures for different service areas? While including all service areas in the framework would be problematic, as many hospitals don’t offer many of the services identified and would be incapable of collecting data for the measures in that area. The Steering Committee agreed that for the framework selection of a hospital core set should be limited to measures of inpatient and emergency care settings.

- Should the framework include conditions and/or priorities that are critical to hospitals but are not included in the national goals? It was agreed that national goals should have a central place in the priority areas identified, but that they alone should not dictate hospital priorities.

A draft consensus report entitled *A Comprehensive Framework for Hospital Care Performance Evaluation*, that recommended a long-term framework of what an “ideal”¹⁹⁷ set of hospital performance measures should include, was developed. It includes three guiding principles: standardization, measure set improvement, and implementation. The report also provides detailed descriptions for guidance in the endorsement of a set of consensus standards for hospitals. These areas include:

- Establishing the content of the performance measure set for hospital care
- Evaluating candidate measures
- Improving and updating the performance measure set for hospital care
- Implementing the performance measure set for hospital care
- Reporting results to the public
- Reviewing and evaluating the performance measure set for hospital care (NQF, 2003: 5-21)

The idea was to provide guidance to NQF steering committees and technical advisory panels regarding the selection of measures that would be forwarded to the membership for potential endorsement. This framework was to assure that measures selected for endorsement as voluntary consensus standards were “comprehensive but not unduly burdensome to measure and

¹⁹⁷ “Many comments submitted by members and the public regarding the shortcomings of the Group 1 measure set suggested changes and expansions that are represented as theoretically *ideal* [should be included] in the draft comprehensive framework” (NQF, 2003: v). These “ideals” were not considered by the Hospital Steering Committee because they could not be easily incorporated into deliberations to accommodate the short timeframe for endorsement of the Group 1 measures.

publicly report” (NQF, 2003: vi). Comments on the framework were due in early October 2002.¹⁹⁸

Backstage: Meeting and Conference Call Preparation

Preparation for meeting and conference calls was characterized as “intense” and involved the co-chairs of the Steering Committee as well as NQF staff, who were very influential in the process, according to several interviewees. Staff prepared all the materials for the meetings and the conference calls, and the behind-the-scenes prep sessions included going through the agendas, re-organizing as needed, reframing certain aspects of the materials as necessary, dividing up the work load, and assigning the lead role to certain sections of the agenda, playing to the strengths of the co-chairs.

Additional activities of the NQF staff to support the Steering Committee included researching and collating the evidence of each measure considered, collation of the comments received from the public and the membership, developing responses to those questions and/or comments, developing staff recommendations for the measures, and deleting measures when called for by the membership, as well as other administrative duties. Some interviewees indicated that the NQF staff did not have the requisite experience or expertise to look at the evidence behind a measure, and often described them as a “young” or “junior.” They pointed out that the NQF staff was very much in control of the process and that staffers made most, if not all, of the decisions, particularly in relationship to what was forwarded to the membership.

¹⁹⁸ This is an example of the Steering Committee trying to translate the CDP; that is, trying to construct the network in assuming the role of “spokesperson” and trying to assign what roles the other actors will play. In this case, the Steering Committee was assigning a role via the Hospital Framework it developed to influence the CDP on how another actor, performance measures would be viewed.

Backstage: Reemerging Issues of the Steering Committee Discussions

A number of issues emerged during the discussions of the first two Steering Committee meetings and continued to reemerge consistently in subsequent discussions.

The first issue centered on the use of the endorsed measures. Measures can be used for a variety of things, including quality improvement, payment, credentialing, accreditation and public reporting. Should the criteria for measure selection be accountability or quality improvement? This question led to a related issue that centered on the quality of the measure itself. It reemerged throughout the project: whether a measure was “good enough” to be used for accountability versus being used for quality improvement. According to one interviewee, there was agreement at the time that measures endorsed by NQF should be used for public reporting, which in effect establishes the type of measures that can “pass muster” and sets a “higher bar” for acceptance. While consumers and purchasers routinely thought that most measures were reliable and valid for accountability and could be used for that purpose, most providers and researchers believed that very few measures could be used for accountability. As one interviewee indicated, “for each measure or each candidate measure or each decision that had to be made to establish the boundaries for the measure set, there was just disagreement. It wasn’t hostile disagreement. There were just different perspectives.” The Committee eventually agreed that measures should be for accountability first, and that these same measures could be used for quality improvement as well. Overall, measures should be “relevant to consumers” (patients) and that criterion should be the primary focus of an endorsed measure set.

Another issue that emerged was the subject of ownership and the nature of the measures, whether it was proprietary or whether it resided in the public domain. The decision to consider, for endorsement purposes, only measures that reside in the public domain was made by the Committee, which led to a realization that a large number of measures commonly used to

measure hospital quality were proprietarily owned and could not be considered for endorsement. Measure ownership is not necessarily a negative, because if a measure is owned, then it is more likely to be updated and maintained. These recurring discussions have repercussions for the NQF even to this day.

A final issue that emerged later on in the process was the appropriate role of the NQF staff. Although there is no formal role for the staff in the CDP, other than to support the committees or panels to which they are assigned, the opinions of the interviewees about staff involvement were mixed. Selection of committee members and chairpersons, e.g., the development of committee structure, was done internally at the NQF; the final decision as to membership and who chaired a committee ultimately rested with Ken Kizer, the President and CEO. According to NQF staff, the committees and panels were designed to balance interests among the four membership councils. Beyond that, balance as to geographical locations of potential members, their race, sex, profession, and other factors also were considered when selecting members. Further, the charges to the committees were determined by the project sponsor, which not only narrowed the scope of the project but often narrowed the pool of potential committee members, as well. In this particular project, the NQF staff tried to select from their membership, but because they were new and did not have a large number of members at that time, they had to go outside the membership pool to find “the adequate expertise required” for this particular Committee.

The staff also prepared the meeting materials and provided recommendations to the membership on how to vote on measures and/or recommendations. On the one hand, some interviewees who participated in the process were very grateful to staff for providing guidance and recommendations, as well as for summarizing and condensing information for meetings. Meeting materials and documents can be extensive, and staffs help with distilling the arguments

and evidence for the measures was important as stated by one interviewee because “none of us is expert in all of these different measure sets that have been brought before us. The staff is employed to take these ideas and sift through them to look at the arguments being offered and see how good the research is behind a particular measure.”

On the other hand, committee/panel members did not always appreciate the staff’s assistance, indicating that if NQF staff “wanted to push something through...they would decide...it’s the dynamics of the staff that very much decide the intent.” One interviewee indicated that while synthesizing information for committee discussions was helpful, it was worrisome that “staff would make recommendations to the general membership irrespective of the option of the committee.” Another interviewee took issue with the way materials and documents were presented to committee members and went on to say, “I’ve seen NQF staff give minimal information to inform decision making. I’ve seen selective reviews of the literature. I’ve seen self published papers by for-profit organizations given the same weight as a peer reviewed article in the evaluation of evidence. That makes me really question the process.”

Performance Measures as a Network Actor

In this act, performance measures were essentially redefined by the interactions of the Committee. While all the members of the committee had their own experience with particular measures and had an idea of what constitutes a “good” measure, a certain amount of ordering within the network took place with regard to “how good is good.” And, with this in mind, questions arose as to what the performance measure would be used for—quality improvement or accountability and comparative public reporting, thus maintaining their position of power and dominance in the health care field. This was probably due to the sheer number of Committee members that were providers and researchers. Thus, focal actors such as the AHA and the Joint Commission easily enrolled other actors to support their cause, specifically a limited set of core

hospital measures that were considered reliable and valid by the field. To achieve their ends, an obligatory point of passage was constructed, over the objections of the consumer and purchaser representatives of the Committee i.e., that measures had to be precisely defined, highly reliable and valid. The providers and researchers were successful at translating the Committee in that they adopted very rigid criteria for selecting measures for the purpose of comparative public reporting. What emerged were eight criteria¹⁹⁹ that defined what good measures should look like. In turn, these criteria eliminated a number of measures, and thus, the Committee drafted a small set that would be sent to the membership for comment and voting. As this act closes, performance measures through the actions of the Committee have taken on the identity of being scientific and evidence-based and playing a central role in relation to comparative public reporting on hospital performance.

Act III: The Public and the Consensus Development Process (June 2002-October 2002)

The Narrator

This act continues to illustrate the translation processes, but on a larger scale. In the previous act, the first two stages of translation, problematisation and interessement, were relatively involved discussions, with the eventually enrolment of the Committee by the providers and researchers of their point of view (i.e., a parsimonious, evidence-based set of hospital performance measures). Thus, the translation process was completed as the Committee spoke for all of its members and for the various discussions throughout the time period. In actuality, however, the translation process is not fixed, but is “contestable and often contested” (Law

¹⁹⁹ The criteria were relevant to high-volume conditions treated in hospitals that are prevalent and causing significant morbidity and mortality; evidence-based; actionable, e.g., things that can make a difference; a balance of condition-specific and non-condition-specific measures; relevant to the population in general, not just one or two subpopulations; related to things the hospital can control or influence; in the public domain, that is not proprietary; and focused on the service, not the business of health care.

1992a). When the CDP was open to public comment, the process of translation began again as the providers' and researchers' groups as well as the consumers' and purchasers' groups both were trying to enrol the NQF to support their respective points of view. As the reader will see, all the groups were somewhat successful. The providers and researchers enrolled the NQF and the Steering Committee into re-working the draft report, and they also enrolled the Board of Directors (BOD) to defer endorsing several measures approved by the membership. They were also successful in enrolling the BOD to have an independent workgroup evaluate the CDP used by the initial Hospital Steering Committee, as well as the resultant core measures initially selected, and to make recommendations as to the adequacy of both. The consumers and purchasers were successful in enrolling the BOD in that they were invited to present their point of view, while other constituents of the NQF were not permitted to do so. They also were successful in re-making the CDP, and fashioning it for their own ends to some extent—for example, requesting reconsideration of the measures they submitted to the NQF during the public comment period.

Also within the third act, a new actor emerged in the form of the Hospital Quality Alliance (HQA), which can be considered a network effect from the completion of the translation process. The whole network was reconfigured to accommodate this new actor, and as is illustrated, the NQF lost some of its influence as a result. In sum, this act illustrates that the translation process is a reiterative one, and the ordering of the actors within the network constantly changes as new actors are added to the mix.

On-stage: Public Comment on the Proposed Hospital Core Measure Set

A total of 43 entities provided comments to the NQF on the proposed measures; 28 were from members, and 15 were from non-member organizations or individuals. Comments submitted during the pre-voting review were from a variety of organizations but the majority

came from provider organizations, such as the AMA, the ANA, the American Board of Medical Specialties, Yale New Haven Health System, the FAH and the AHA. From the provider perspective, two basic themes emerged from the comments submitted to NQF: issues related to risk adjustment and scientific evidence of the measures. First, there was concern over the lack of specificity of the measures; they were not precisely defined, yet they would be used for accountability and public reporting. Second, most providers commented that the risk adjustment used was “not felt to be adequately elucidated in light of the potential uses for these measures.” Measures should be evidence-based and that evidence should be documented and communicated to the users so that it could be reviewed and evaluated by the provider community. One provider group even went so far as to indicate that a selected measure was counter to the current scientific evidence. Nearly all of the providers objected to the stated purpose of the measures, i.e., their use for accountability and public reporting. One notable exception from the provider community was the ANA. They argued that, overall, the selected set of measures while small, was appropriate for accountability purposes and that additional measures including a measure of nurse staffing would enhance the core set.

The responses from the national hospital associations are worth mentioning. These associations represent the majority of hospitals the United States and they are in frequent contact with hospitals regarding on-going issues and concerns related to hospitals and hospital care. According to the associations’ comments, national quality goals should be established, first and foremost, and until such time, the selection of measures should not be made until these goals are established and vetted through the consensus process. Their other comments had to do with issues related to the burden placed on hospitals for data collection for accreditation, conditions of participation, and internal quality improvement. Public comments received indicated that the NQF measures selected did not relate to the Joint Commission’s accreditation process or to

CMS’s seventh scope of work for QIOs, and that the hospitals would consider these measures yet “another measure set that may lead to requiring multiple variations of similar measures required by multiple competing organizations.” Further, the hospitals “strongly objected” to the stated purpose of the measure set, accountability, stating that “hospitals *are* publicly accountable for the care they provide today [original emphasis]. Quality improvement will not be fostered by a punitive approach that is proposed under the guise of making hospitals publicly accountable.” Other concerns included issues with adequate risk adjustment and sufficient testing of measures, the timing of data collection and the patient populations involved, measure maintenance and implementation. According to NQF staff, these responses were not unexpected, based on the various discussions that occurred throughout this project among providers, staff, and Steering Committee members.

What was unexpected was the response from the purchaser and consumer stakeholders. A letter from a cross-section of Consumer, Purchaser, and Research Council Members²⁰⁰ was sent to the CEO of NQF; it included general comments and specific recommendations about the Hospital Performance Measurement Project and the proposed initial set of hospital performance measures. The joint letter’s intention was to “enrich the expert process of the NQF, not supersede it.” Its general comments dealt with the need to encompass all six health care quality domains specified by the IOM. The measures selected for the initial set did not address all the areas of quality: safety, timeliness, effectiveness, efficiency, equity and patient-centeredness. These domains “are important to consumers and purchasers and are appropriately central to the scope of the work of the Steering Committee.” Furthermore, of the 32 selected measures all but

²⁰⁰ These included 3M, AFL-CIO*, Buyers Health Care Action Group*, California HealthCare Foundation*, The Employer Health Care Alliance Cooperative*, General Motors Corporation*, The Leapfrog Group*, Midwest Business Group on Health*, Motorola, Inc., National Business Coalition on Health*, National Health Care Purchasing Institute*, National Partnership for Women and Families*, Pacific Business Group on Health*, and Washington Business Group on Health*. An asterisk denotes a member of the NQF.

three needed to be NQF endorsed and expanded, with emphasis on cross-cutting measures²⁰¹ and conditions beyond the seven condition-specific priorities²⁰² adopted by the Steering Committee. The current proposed measure set is relevant, but it is an “incomplete picture of hospital quality.” Additionally, clinical importance should be the primary determinant of NQF endorsement. While minimum levels of validity and feasibility are important, clinically important measures that “guide consumer choice to higher quality at least 75% of the time should be deemed to be scientifically adequate” and should be endorsed by the NQF. Moreover, measures should not be excluded solely due to the burden of medical record review, particularly if clinical importance is relatively high. This measure set “should be designed to accelerate adoption of electronic clinical information systems rather than accommodate handwritten medical records.” To support this assertion, the joint letter cited the example of the Securities and Exchange Commission (SEC) that required “substantial performance reporting by public companies that were not attenuated to accommodate the burden associated with handwritten financial records,” thus initiating and promoting computerized financial reports.

Other recommendations included that consideration of proprietary measures should be given when no alternative exists, such as the experience of care surveys designed by Picker,²⁰³ and modifications to measure specifications should be adopted immediately, instead of waiting for prior widespread demonstration of their viability.

“In summary, the quality equilibrium documented by the IOM is one of serious and widespread defects. Preventable patient suffering is large.

Every year without hospital quality measures is another year in which our

²⁰¹ These measures are not specific to a particular disease or condition but report on important aspects of care, such as patient safety, and coordination of care. See Appendix H.

²⁰² The selection of priorities was based on the prevalence of these conditions among hospital patients.

²⁰³ The Steering Committee excluded the patient experience of care indicator due to the lack of a suitable instrument in the public domain. AHRQ and CMS were actively working on a CAHPS survey for hospitals that would reside in the public domain, which could then be endorsed by NQF.

employees, families and friends are forced to fall back on quality-blind or near-blind hospital selections. For consumers and purchasers to build a quality-sensitive demand curve for hospital care and reduce personal risk of death and disability, we need rapid access to a much wider initial set of less perfect hospital quality measures.”

The specific recommendations of the group included modifications to four measures included in the current NQF measure set, as well as the deletion of the measure pertaining to parenteral fluid usage in gastroenteritis, citing that there was “no evidence that IV fluid administration for hospitalized patients improves outcomes.” The joint letter further recommended the inclusion of an additional 27 measures in the core measure set, which would appeal to a wider consumer audience. To arrive at these additional measures, the group convened an initial panel of performance measure developers and measurement specialists that nominated measures that it believed were suitable for publicly reported quality comparisons among hospitals. A second panel,²⁰⁴ consisting of clinical performance measurement methodologists, evaluated the measures nominated by the initial panel using the same thirteen NQF criteria. This evaluation resulted in a very strong recommendation for further dialogue with regard to the additional 27 measures presented and to explore how these measures could be incorporated into the core measure set.²⁰⁵ The joint letter from the consumers and purchasers concluded: “Our perspective is the ability to rapidly adopt a truly meaningful full dashboard of performance measures through a voluntary, consensus-driven approach is preferred to regulatory or other mandated processes to which we will likely turn in the absence of such action.”

²⁰⁴ Unlike the NQF Hospital Steering Committee that empanelled measure developers as members of the committee, the consumer-purchasers selected experts who had no role in developing or promoting the measures they reviewed and evaluated.

²⁰⁵ The letter was accompanied by an 82 page report of the specific measure recommendations of the panels.

As is routine, the NQF staff in consultation with the Steering Committee members responded to every comment received. As recounted by several interviewees, however, these measures were treated differently, as was demonstrated at the June 2002 meeting of the Hospital Steering Committee.

On-stage: The June 2002 Hospital Steering Committee Meeting

The purpose of this meeting was twofold—first, to review comments submitted and propose action steps on the draft hospital report; and, second, to review and provide feedback on developing a comprehensive measurement and reporting framework for hospitals. For this meeting, the NQF had assembled all the comments into a large grid that summarized the comments and identified the organization, group and/ or individual that provided the comment. Among the meeting materials was a slide presentation that discussed the “range of actions” being considered with regard to the comments received on the proposed hospital measure set. The possible actions included:

- Action to add measure
- Action to delete measure
- No action pending feedback from developer
- Incorporate/revise report narrative/appendices
- Referred to comprehensive framework/phase II
- Incorporate in recommendations for research
- Hold for discussion at improvement of set
- No action

As the presentation indicated, the NQF staff recommended inclusion of the 27 measures submitted by the consumer-purchaser groups and deletion of other measures “based on overwhelming recommendation for deleting” from the public. Of the 34 recommended

measures, only two were deleted.²⁰⁶ Ultimately, the Steering Committee indicated that it was uncomfortable acting on measures that came from outside the established process and that it would review the comments on the core measures set previously considered, but not on new measures. With that, the draft report was revised and finalized for membership voting, which began on July 21 and ended on August 20, 2002.

On-stage: NQF's Response to the Consumers and Purchasers

After NQF's receipt of the joint consumer and purchaser comments, and under some pressure from these groups, the NQF Board invited representatives of the Consumer and Purchaser Councils to "orient the Board of the NQF as to how the viewpoint on measures is different through the eyes of consumers and purchasers than providers and researchers." On July 15, 2002, two representatives of the Consumer and Purchaser Councils gave a presentation entitled "Consumer and Purchaser Perspectives on Hospital Quality Measures." The overall message of the presentation was that the "status quo is unacceptable." Hospital care constitutes a large percentage of the GDP in the United States, and it warrants more than a handful of measures. Consumers require more comprehensive measures, rather than fewer, if they are to take responsibility for their health care. Along these lines, cross-cutting measures are extremely valuable as these measures are indices of how a hospital or doctor is going to do, overall. The third and most controversial point was that from a customer perspective, sometimes imperfect measures were far preferable to no measures. Misclassification of hospitals due to imperfect measures is acceptable to consumers and purchasers, while from a provider perspective it is unacceptable. The final point of the presentation addressed the issue of data collection burden associated with performance measurement. Typically, data are abstracted from the medical

²⁰⁶ The two measures recommended for deletion included 1) Vaginal Birth after Cesarean (VBAC), based on the controversy over this measure as an indicator of quality; and 2) Parenteral fluid for gastroenteritis based on concerns that the measures needed further development.

record by nursing personnel, who collect and manually enter them into a database. Thus, data collection is time consuming and costly to hospitals. From the consumer-purchaser perspective, “we don’t consider that to be our problem. We consider that to be the problem of an industry that long ago resisted getting its critical processes onto an electronic platform.” (One interviewee offered this comparison, “crying measurement burden is sort of like a kid who kills his parents and then complains he is an orphan”).

From the initial measure set consisting of 32 measures proposed by NQF, the consumer-purchaser group suggested that three measures should be deleted and 29 measures should be added, for a total of 58 proposed measures for endorsement. One interviewee described the presentation as an “impassioned plea that basically said, reopen the process. The measures you have selected do not meet our [consumer and purchaser] needs.” Based on the presentation and the discussion that ensued, the NQF Board agreed to reopen the consensus development process and consider endorsing the 27 additional measures that became known as the “Group 2 measures.”²⁰⁷ The originally proposed 32 measures would be designated the “Group 1” measures.

On-stage: Voting on the Hospital Core Measures (Group 1)

The ballot for the Hospital Care National Performance Measures (Group 1) was distributed to the membership pursuant to the NQF CDP, version 1.5, for voting on July 17, 2002, and was due back to NQF on August 20, 2002.²⁰⁸ According to a memo from Kizer to the

²⁰⁷ This is an excellent example of the ANT element of “enrolment,” which is the process by which an actor, here the consumer and purchaser, attempts to enlist other actors for a particular purpose or objective (re-opening the CDP) that involves imposition. The consumers and purchasers are enrolers that were trying to reconstruct the network by purposefully putting consumer and purchaser needs above the needs of the providers and researchers. In this case, they were successful in enlisting the NQF BOD to reopen the process and eventually to expand the initial measure set.

²⁰⁸ Based on the CDP version, it is apparent that several factors influenced this process; thus, the identity of the CDP was made and re-made. For example, the requirement that NQF do an “open call for measures” that is incorporated into subsequent CDP versions is a direct result of the consumer/purchaser joint letter.

NQF membership, a number of documents were revised, based on the comments submitted, and included in the voting package. The memo went on to indicate that “a large number of additional measures were recommended during the comment period, some of which are likely to be recommended for inclusion in the initial Hospital Care National Performance Measurement set after they undergo further evaluation.” The ballot included 37 individual measures and one recommendation that were to be voted on by the membership. The voting options included an “en bloc” vote, which gave the option to approve all measures and recommendations as currently specified, to disapprove all measures and recommendations, or to abstain from voting on all measures and recommendations. The membership could also choose to vote on each individual measure and recommendation; however, they could not choose to vote on each individual measure and also vote en bloc. Along with the ballot, a short description of the measures was provided as well as detailed specifications. Approval of a measure indicated that it should be included in the final core set, while disapproval meant that it should not be. The measures and recommendation are listed in Table 9.

Table 9: NQF Hospital Care National Performance Measures Ballot (2002)

	Measure	Source
1	Aspirin at arrival for acute myocardial infarction (AMI)	CMS/ Joint Commission
2	Aspirin prescribed at discharge for AMI	CMS/ Joint Commission
3	Beta blocker at arrival for AMI	CMS/ Joint Commission
4	Beta blocker prescribed at discharge for AMI	CMS/ Joint Commission
5	AMI inpatient mortality	Joint Commission
6	Angiotensin converting enzyme inhibitor (ACEI) for left ventricular systolic dysfunction (LVSD)	CMS/Joint Commission
7	Percutaneous coronary intervention (PCI) within 120 minutes of arrival for AMI	CMS
8	Thrombolytic agent within 30 minutes of arrival for AMI	CMS
9	Percutaneous coronary intervention (PCI) volume	American College of Cardiology
10	Percutaneous coronary intervention (PCI) mortality	American College of Cardiology
11	Coronary artery bypass graft (CABG) volume	New York State
12	Coronary artery bypass graft (CABG) mortality	New York State
13	Left ventricular function (LVF) assessment	CMS/ Joint Commission
14	Detailed discharge instructions	CMS/Joint Commission
15	ACEI for left ventricular systolic dysfunction	CMS/Joint Commission
16	Oxygenation assessment	CMS/Joint Commission
17	Initial antibiotic consistent with current recommendations	CMS
18	Blood culture collected prior to first antibiotic administration	CMS/Joint Commission
19	Influenza screen or vaccination	CMS
20	Pneumonia screen or pneumococcal vaccination	CMS/Joint Commission
21	Antibiotic timing	CMS
22	Vaginal birth after cesarean delivery rate	Joint Commission

23	Third or fourth degree laceration	Joint Commission
24	Neonatal mortality	Joint Commission
25	Cesarean delivery rate	Healthy People 2010
26	Timing of antibiotic administration (surgical patients)	CMS
27	Selection of antibiotic administration (surgical patients)	CMS
28	Duration of prophylaxis (surgical patients)	CMS
29	Use of relievers for inpatient asthma	Child Health Corp of America
30	Use of systematic corticosteroids for inpatient asthma	Child Health Corp of America
31	Neonate immunization administration	Child Health Corp of America
32	Pediatric heart surgery volume	AHRQ
33	Pediatric heart surgery mortality	AHRQ
34	Smoking cessation advice/counseling for acute myocardial infarction (AMI) patients	CMS/Joint Commission
35	Smoking cessation advice/counseling for heart failure (HF) patients	CMS/Joint Commission
36	Smoking cessation advice/counseling for pneumonia patients	CMS/Joint Commission
37	Carotid endarterectomy volume	AHRQ
	<p><u>Recommendation:</u> The primary purpose of these initial hospital care performance measures is to facilitate quality improvement through public accountability, including the public disclosure of results. In order to drive quality improvement, these accountability measures were selected to address areas where processes to improve care are known. The measures are intended to be used by consumers, purchasers, providers, accreditors, Quality Improvement Organizations, and researchers, among others, to enable performance-based decision making about hospital</p>	

	selection, create incentives to catalyze hospital performance improvement through use by the public of health care quality information, enhance value-based purchasing and stimulate the improvement of care.	
--	---	--

The result of the membership vote was that all four councils approved all 37 measures and the recommendation. Once the membership reached consensus on the measures, they were forwarded to the Board for its vote and endorsement.

On-stage: The Board's Decision and the Consensus Development Process

On October 2, 2002, the NQF Board voted to defer endorsement of six measures that were approved by the membership, because of its inability to come to consensus on these measures. These measures were to be re-considered for endorsement by the Hospital Work Group that would meet on October 17-18, 2002, along with an additional 27 other measures that were submitted to NQF during the comment period of the project. Table 10 lists the measures that were deferred to the work group and their source.

Table 10: Hospital Performance Measures and Source Deferred to Work Group

Percutaneous coronary intervention (PCI) volume	American College of Cardiology
Percutaneous coronary intervention (PCI) mortality	American College of Cardiology
Coronary artery by pass graft (CABG) volume	New York State
Pediatric heart surgery volume	AHRQ
Pediatric heart surgery mortality	AHRQ
Carotid endarterectomy volume	AHRQ

At this meeting a draft Framework for Hospital Performance Measures was discussed, and was sent to the membership for review commencing on October 4th with comments due on November 7th. As if to specifically address the AHA letters, Ken Kizer, who presented an update to the BOD on NQF activities, included in his presentation a slide entitled “Hospital Care

Performance Measures,” which stated “many comments made by organizations when voting were not made during the comment period; the comments could have been acted upon if made during the comment period.”

In a media press release on October 8, 2002, the NQF announced approval of the first set of national voluntary consensus standards for measuring the quality of hospital care, comprising 31 measures in six areas. “Endorsing a first group of voluntary consensus standards for hospital performance is a critical step toward making quality of care information available to consumers via public reporting. Standardized measures also will provide hospitals’ internal quality improvement programs with key data that they can use to improve their care,” Kizer said.

Backstage: The June 2002 Hospital Steering Committee Meeting

Although the formal record reflects a relatively short, concise discussion, in actuality the discussion of the NQF staff’s role and the role of the Steering Committee was protracted and civil, but heated at times. Several members of the Steering Committee indicated that they were “superfluous to the process” because the NQF “staff make all the decisions.” In this view, the staff decides what to include and what not to include for the membership to vote on regardless of the recommendations of the Steering Committee. Many interviewees later stated that the “rules of the game changed as the deliberations went along.” An NQF staff member clarified the staff’s role, stating that “a core set of measures is developed by the Steering Committee, and the staff role is to respond to comments. The BOD wanted the membership to see what the Steering Committee thought of the measures and what their recommendations were.”

Once again, the discussion returned to the measures submitted by the consumers and purchasers. Staff called for the measures submitted to be added to the core set that would be recommended by the Steering Committee, with the next step in the process to be sent to the membership for voting. One concern expressed by the government representatives at the table

indicated that adding measures was by-passing the public comment process; if the intent is to get “buy in” from all groups, was it really prudent to bypass this process for certain measures? And if so, which measures would be considered for special attention? Another concern the government representatives expressed had to do with the NTTAA—specifically, government agencies can participate in consensus development only if there is due process, and in by-passing some elements of the established process, there was concern that the process was not open, and balanced. The representatives went on to say that “if that is the case when we add these measures and by-pass certain aspects of the CDP, then federal agencies such as CMS and AHRQ cannot participate.” What was ultimately decided by a majority of the members of the Committee was that the measures the consumers and purchasers submitted would not be considered for the initial core set of hospital measures.

Backstage: Influence and the NQF Board

On August 26, 2002, two letters arrived at AHRQ addressed to John Eisenberg, the Director of the Agency. One of the letters was on AHA stationery; the other was not, and appeared to have accompanied the NQF Hospital Measures voting ballot, which was due around the time both letters were dated. Staff at AHRQ reviewed the letters, which pertained to the NQF and the consensus process that was currently being followed for the hospital project.²⁰⁹ The AHA letters underscored its belief that these measures were appropriate for quality improvement. It went on to state: “We do, however, have concerns about some of the measures that have been included on this ballot, especially since they are intended to inform the public. We also have concerns about the process by which these measures have been brought forward.” Both letters indicated that there were significant inconsistencies in the consensus development

²⁰⁹ It is unknown if similar letters were sent to the other members of the NQF Board; however since portions of these letters were included in the AHA’s public comments to the NQF, one can assume that they eventually made their way to all NQF members, including those who sat on the BOD prior to the vote for endorsement.

process. Both cite the example of six measures (pediatric heart surgery volume, pediatric heart surgery mortality, carotid endarterectomy volume, CABG volume, PCI volume and PCI mortality) that all members of the Steering Committee rejected based on concerns about validity, usefulness and whether they accurately reflected quality, prior to the draft report issued in May 2002. However, after the public comment period and at the July meeting of the Steering Committee, at which fewer than half of the members of the Committee were present, the six measures again were voted upon, and this time were included in the voting set. The letters went on to state that these six measures had not been reviewed or commented upon by the NQF members or other interested parties because they had not been sent out for review and comment. They stated: “It is inappropriate and a violation of the NQF’s own principles and procedures to ask its membership to vote on measures that have not had the benefit of review and comment by all interested parties” [original emphasis]. The letters urge the Board not to consider these measures for endorsement until they can proceed properly through the consensus development process, citing that changing the CDP process in some cases and circumventing it in others without the involvement of the membership “compromises the very purpose and credibility of the Forum.”

The second letter cited concerns about the role of the Steering Committee of each NQF project, suggesting that in the Hospital Core Measures project, its role “seems to be less central to the development than the Consensus Process document stipulates.” The additional measures received in the public comment period were considered by the Steering Committee and largely rejected because there was insufficient information to assess them to assure scientific validity and reliability. But rather than providing the Steering Committee with additional time and resources to consider these additions, the “Board approved a staff recommendation to create a separate ad hoc workshop to consider the proposed measures outside the framework of the

Steering Committee.” According to the consensus development process, “there is nothing that describes the use of an ad hoc workshop in developing a proposed measure set or that indicates how the recommendations emerging from this workshop will be reconciled with those of the Steering Committee.” Again, the Board was urged to analyze and standardize the current process and to include membership participation in its revision. The letter concluded: “We believe that these constant changes to address unforeseen issues indicate that the process, as currently conceived, is not meeting the needs of the members nor is it achieving the purpose for which the NQF was formed. We believe it is time to step back and rethink exactly how the NQF should take projects from concept to fruition” [original emphasis]. The staff at AHRQ viewed these letters as “informational” and forwarded them to the acting Director of AHRQ, who was serving as a member of the NQF BOD.

Backstage: The Emergence of the Hospital Quality Alliance

In October 2002, the NQF Board of Director endorsed 31 hospital performance measures approved by the membership. Because of the NTTAA, it was likely that these endorsed measures would be encouraged by CMS for use in its hospital reporting initiatives. The major hospital associations began talks with each other in an effort to “slow down some of the performance measurement activity that was underway at CMS and the Joint Commission and gain, to the degree it could, control of it.” Hence the seed that was planted eventually emerged as the Hospital Quality Alliance (HQA)²¹⁰ in December 2002. In the HQA’s statement of intent developed toward the end of October 2002, the goals of the collaborative were to “provide useful and valid information about hospital quality to the public, to give hospitals a sense of predictability about public reporting expectations, and to begin to standardize data and data

²¹⁰ This initiative, originally known as the Hospital Voluntary Reporting Initiative, has had several other names prior to adopting its current one—the Hospital Quality Alliance (HQA).

collection mechanisms.” The HQA proposed a “starter set” of 10 of the Joint Commission /CMS-developed measures that were endorsed by the NQF for three hospital conditions—acute myocardial infarction, heart failure and pneumonia. While some heralded this as a breakthrough in public-private partnerships, and a strong signal that hospitals did want to share performance information with the public, others had a more cynical view. “The AHA’s pressure on CMS to join the HQA under the pretext, ‘we couldn’t possibly implement all these NQF endorsed measures,’ so lets come up with a self-appointed, health industry dominated group that will then pick and chose for ‘immediate implementation’ and thereby enable postponing most of the NQF endorsed measures [for implementation].” As one interviewee put it, “it’s an example of how the public interest gets undermined...so what the American public gets in the near term is 10 measures that apply to only three conditions on which peer review organizations have been working with hospitals for the last 10 years, so the probability of poor scores and variability [among hospitals] is very low.” Along these same lines, another interviewee stated that “in essence this [the HQA] was an end run around the [National] Quality Forum because the Quality Forum had endorsed a number of measures through the development and endorsement process...Not only that, they [the NQF] appeared to have reached a resolution of sorts with the Joint Commission—that they, [the Joint Commission], would not use measures unless they are endorsed by the Quality Forum. The reason the HQA started was that it was a way to do an end run around the NQF dumping too much stuff on hospitals.”

Needless to say, the leadership of the NQF was reluctant to participate in the HQA discussions, due to the similarity of the missions of both organizations. And for a long period of time, the NQF did not participate in the HQA weekly meetings. The NQF leadership, according to one interviewee, “just felt that they should be the convener, and that the hospitals should be implementing all of the measures endorsed by the NQF.” From HQA’s perspective, the hospitals

did not trust the NQF or its processes, for at least two reasons. First, the NQF was a new organization that potentially had a very powerful role within the measurement field, due to the NTTAA. Under the “color of law,” any measure the NQF endorsed was expected to be used in lieu of a government-specific standard, and if the consensus development standard was not used, the federal government could use another only with an explanation as to why. In the words of one interviewee, “this gave CMS cover [for the measures they selected] and also provided the federal government with a way out.” For example, if NQF endorsed a measure and CMS determined that the science behind the measure was flawed and that there was a high potential for unintended consequences, CMS could opt not to use it in its initiatives, but would have to formally explain its reasoning to OMB. Second, the provider communities generally think of the NQF as a consumer-purchaser dominated organization, in which the scientific readiness of measures often was sacrificed in an effort to provide information to the public. The provider communities had concerns about unintended consequences that could potentially cause confusion and harm among consumers, due to flawed data collection methods, poor measure specifications, and faulty interpretation of measure results.

Due to these issues and in particular the distrust of the NQF by the provider community, CMS realized that an effort solely led by the NQF would not have “gone anywhere.” Hence, CMS began meeting with the HQA in an attempt to get hospitals to participate voluntarily in a public reporting initiative of hospital performance planned for 2003.²¹¹ The hospital associations (AHA, AAMC, and FAH) indicated that they would be willing to “sell” the voluntary program to their members; however, they needed to be able to have some control over which measures would be implemented. The hospitals did have some bargaining power within the network, in

²¹¹ CMS Administrator Tom Scully launched an ambitious timeline to publicly report quality performance of nursing homes in 2002, and home care agencies and hospitals in 2003, on a DHHS website.

that they could encourage participation in the CMS effort. Due to the current Bush Administration's disdain of traditional regulation, the hospital associations were in a position to deny recognition of the proposed CMS reporting initiative. Thus, because the hospital associations, as punctualised actors for the entire network of hospitals, were able to use the network effect of power toward their own ends—specifically more control over which measures and how many were implemented. CMS Administrator Tom Scully, however, was not without network power given his organization's stature within the hospital network and in health care. While the current Administration did not favor regulation, Scully had to make it appear that if voluntary participation was not forthcoming, then CMS would move forward with the proposed rule-making to achieve its ends. According to press releases, speeches, and quotes that appeared in the newspapers and health care journals at the time, Scully took every occasion to remind the hospital industry that if hospitals did not voluntarily report their performance, then CMS would consider proposed rule-making to accomplish its goals. At the time, most leaders of health care made it very clear that the time had come to provide this type of information to the public, and that there were efforts underway to figure out how to do this; thus Administrator Scully and CMS were very much part of the push that eventually made it happen.²¹²

Finally, after an extraordinary amount of bargaining and negotiating to re-order the network, the public-private partnership, embodied in a virtual organization, the HQA, was announced in December 2002. Heralded by some in the hospital industry as “the unprecedented voluntary partnership,” others had a less enthusiastic view, particularly some representing consumers and purchasers. These groups observed that “the health care industry is very capable.

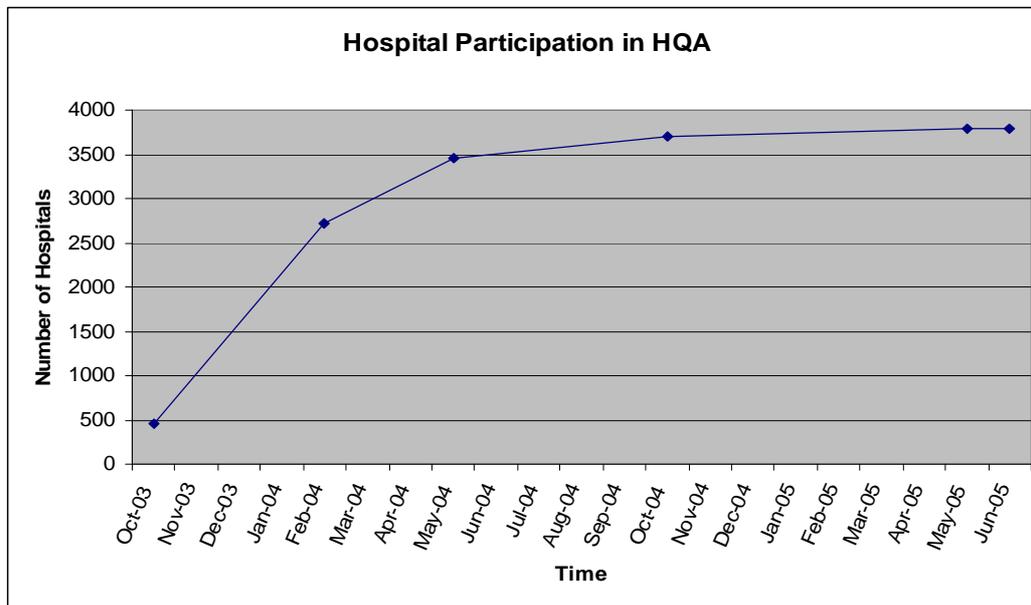
²¹² Scully, prior to his appointment to CMS, was the CEO of the FAH for approximately seven years, which gave him extra credibility with the hospitals. He also could speak with a fair bit of authority about hospitals, so he could see through any argument that hospitals presented very quickly.

If they sense they are losing political ground, they are extremely resourceful at figuring out how to minimize loss,” implying that the formation of the HQA was just such a strategy.

Recruitment of hospitals for the voluntary initiative began shortly after the December announcement; however, the response from the hospitals was “disappointing.” According to one interviewee, “for almost a year, we only had 400 hospitals actually reporting the ten measures, while several thousands had committed to reporting. But you know promises are not action.” At a press conference announcing the HQA reporting initiative, Tom Scully stated that the rates of hospital reporting was disappointing and if it did not improve, mandates would be considered. Not until the Medicare Modernization Act (MMA) was signed into law by President George W. Bush on December 8, 2003 did the participation of hospitals increase. Section 501(b) of the MMA required hospitals to report ten performance measures in order to receive their “full market basket update,” and if they chose not to, they would not receive the additional funds. During discussions with CMS and other federal agencies, Congress at the time the MMA was being drafted was very interested in moving the quality of health care forward. According to one informant, Congress realized that it could get hospitals to publicly report without requiring them to do so, by simply saying, “If you report them [hospital measures], you will get your full market basket. If you don’t, you won’t. So it’s not a regulation, it’s not rule-making, but it was a federal program designed to make public reporting happen.” Figure 9 indicates participation both before and after the MMA was signed into law.²¹³

²¹³ The number of hospitals that were considered “eligible” to report data (that is, general acute care hospitals) was approximately 4,000 out of a total of approximately 6,500 hospitals in the U.S.

Figure 9: Hospital Participation in the HQA Voluntary Initiative



Under the MMA, for inpatient acute care hospitals (not specialty hospitals, such as children’s hospitals, or rehabilitation hospitals) to receive a 0.4 percentage point increase or the market basket update on which Medicare payments are based, they had to begin submission of the data and reporting on the ten performance measure “starter set” endorsed by HQA by July 1, 2004. As one interviewee stated, “It was a surprise quite frankly [that there was language drafted into the MMA specifically attaching financial incentives to these 10 measures], and it’s very difficult—in fact it’s impossible to expand that unless you reopen that can of worms. My supposition is that the hospitals were at the table, but I can’t say for sure. However, if there were broader representation at the table [i.e., consumers and purchasers], I think we would have started with the 10 measures and attached financial incentives but also would have included a stipulation that would allow additional measures to be included beyond the 10.”²¹⁴

²¹⁴ This is a successful example of an actor, the HQA, trying to become an “obligatory point of passage” for NQF endorsed measures and by enrolling other actors such as CMS, achieving its ends.

Performance Measures as a Network Actor

Performance measures in this act appeared to play a contentious role, in that there was disagreement as to what the measures should be used for, which in turn shaped their identity—more rigorous and scientific for accountability and public reporting as opposed to less rigorous for quality improvement. The identities were debated, and thus relations were formed with other actors in the network. With the adoption and endorsement of a core measure set (which was placed in statutory text, thus prolonging the relational patterns and social ties), the durability of the network order was affirmed, thus stabilizing the network. Further, the initial measures selected for the core measure set played a role in decision-making with regard to the HQA. The decision to form this new punctualised actor had to do with the type, number and, to a certain extent, how the measures were selected. Through its emergence, the hospital industry formed an additional obligatory point of passage, the HQA, for which the measures had to pass for national implementation. As the curtain falls in this act, performance measures' primary role is in comparative public reporting for NQF endorsement and HQA endorsement for implementation.

Act IV: The Hospital Care National Performance Measures Workshop

The Narrator

This act demonstrates another round of the translation process, but with a different set of actors. As noted, the compositions of the initial Hospital Steering Committee and the Hospital Workshop members were somewhat different in terms of the addition of more consumer and purchaser members and fewer researchers and providers in the latter. While the group came to some of the same conclusions as the initial Hospital Steering Committee, this work group recommended additional priority areas and additional measures; it also suggested that proprietary measures and risk adjustment could be considered for endorsement if open access for public

scrutiny could be assured. The NQF staff became an obligatory point of passage as will be noted later on because they determined which measures would be given to the membership for a vote as evidenced by the removal of several measures from the voting ballot. The BOD was enroled by the providers and researchers with regard to the proprietary debate, as it decided that only measures that resided in the public domain could be endorsed.

On-stage: Selecting Members and Convening the Workshop

On October 17-18, 2002, the controversial Workshop on Group 2 Hospital Measures convened in a Washington, D.C. suburb, to “address concerns raised during the review period about the adequacy of the Group 1 measure set, and to consider additional measures including several that were voted upon by the NQF membership but deferred by the Board for further consideration” (NQF 2002). Although the workgroup consisted of several members of the Hospital Steering Committee, it did not consist of all of the members. Table 11 shows the differences between the original Hospital Steering Committee Membership and the Group 2 Measure Workshop Members by NQF Council.²¹⁵

*Table 11: Hospital Steering Committee Workshop Member Differences by NQF Council*²¹⁶

	# Providers & HP	# Consumers	# Purchasers	# QI & Research
Hospital SC (19 members)	5(1)	3	2 (1)	6 (1)
Hospital WG (20 members)	4	4 (1)	5 (1)	5

The Group 2 measure set consisted of 34 measures—28 of which were proposed during the public comment period of Group 1 measures and six deferred by the BOD. “The purpose of

²¹⁵ For a complete list of the members of the workshop, see Appendix J.

²¹⁶ These are counted according to NQF council designation of each individual’s organization. Several other individuals served on the Steering Committee and Work Group but were not members of the NQF. Their affiliation was assigned by the researcher and is in parentheses.

the workshop was to 1) consider, discuss, and arrive upon *principles* for resolving fundamental issues raised during the comment period and 2) to apply the *principles* to the 34 candidate measures suggested for inclusion, and recommend Group 2 measures that should enter NQF Consensus Development Process” [original emphasis]. The issues for discussion included:

- Perceived gaps in Group 1 measures²¹⁷
- Burden and parsimony²¹⁸
- Scope and priority areas²¹⁹
- Application of the evaluation criteria²²⁰
- Other²²¹

Over the two-day period, members of the Workshop deliberated on those concerns that surfaced during the initial hospitals measures project. The Workshop participants accepted the recommended priority areas previously identified by the Hospital Steering Committee, but recommended additional measures in areas that were not addressed, to make the measures “more relevant for consumers and purchasers.” Specifically, the measures in Group 1 did not address the priority area of patient safety, whereas the Workshop members recommended eight measures in that area. Again, there was no call for measures and the criteria used to evaluate the measures designated as Group 2 would be the same criteria used for the Group 1 measure set. For Group 2

²¹⁷ To what extent do the IOM aim areas have to be populated? Most of the Group 1 measures address effectiveness, but obvious gaps in the areas of efficiency and patient-centeredness exist.

²¹⁸ Issues discussed here included measures in priority areas with numerous, existing measures; single additional measures in new priority area; measures that rely on additional data collection or chart review; and measures that accelerate adoption of electronic data systems.

²¹⁹ Discussed were the scope/priorities of the Steering Committee as well as the implications for altering it. The scope of the Committee was intentionally narrow, in recognition of burden and parsimony, direct measures of quality, accountability as purpose, wide, national impact, and open, public methodologies.

²²⁰ These criteria were based on the Strategic Framework Board’s draft report and included importance, scientific acceptability, usability, and feasibility. The Steering Committee did not assign a weight to each criterion and after a review of the evidence on all elements for each candidate measures, made informed but subjective recommendations. Also discussed was the inherent tension between what was of clinical importance and what was important to consumers.

²²¹ These included the role of NQF in current/future health systems and the relation of the criteria and processes of this project to other NQF work.

measures, however, the NQF staff made efforts to contact the relevant stakeholder groups, including medical specialty societies to augment the set to be evaluated by the Workshop participants. Further, the Workshop participants indicated that “no principles or criteria were adopted that are inconsistent with the evaluations for the Group 1 measures” (NQF, Voluntary Consensus Standards for Hospital Care: Group 2 Measures, 2002: 3). Table 11 shows the measures that were initially considered for inclusion in the hospital measure set to be recommended by the Workshop participants.²²²

Table 12: Hospital Measures for Discussion by Workshop Participants

	Measure	Measure Source
1	Glycemic control target	American Healthways
2	Use of preoperative beta blockers for Coronary Artery Bypass Graft (CABG)	Society for Thoracic Surgeons (STS)
3	Use of internal mammary artery (IMA) graft for CABG	STS
4	Intensive Care Unit (ICU) mortality index	CALICO-California Intensive Care Outcomes Project
5	Documentation of care preferences	Assessing Care of Vulnerable Elders project-ACOVE
6	Patient participation in life-sustaining care decisions	ACOVE
7	Admission evaluation of cognitive status	ACOVE
8	Advance directives	ACOVE
9	Restraint prevalence	California Nursing Outcomes Coalition Database Project
10	In-hospital hip fracture	HCUP quality indicators
11	Abdominal aortic aneurysm (AAA) repair volume	HCUP quality indicators
12	AAA repair mortality	HealthGrades
13	Co-morbidity adjusted complication rate	Care Science
14	Risk adjusted costs	Care Science
15	Urinary catheter utilization	Centers for Disease Control and Prevention (CDC) National Nosocomial

²²² Any measures recommended for NQF membership consideration are designated as Group 2. Eventually, measures in Group 1 and Group 2 will be merged to comprise the “initial” hospital measure set.

		Infections Surveillance (NNIS) System
16	Central line utilization	CDC NNIS
17	Ventilator utilization	CDC NNIS
18	Central line infection	CDC NNIS
19	Catheter associated urinary track infection (UTI)	CDC NNIS
20	Ventilator associated pneumonia	CDC NNIS
21	Cardiac catheterization in appropriate patients with non-ST elevation MI (NESTEMI)	American College of Cardiology/American Heart Association (ACC/AHA) Task Force on Performance Measurement
22	Heparin in acute coronary syndrome (ACS)	ACC/AHA
23	Lipid profile on admission	ACC/AHA
24	Lipid treatment at discharge	ACC/AHA
25	Falls per 1000 patient days	American Nurses Association (ANA)
26	Prevalence of pressure ulcers	ANA
27	Blood culture within 24 hours of hospital arrival for pneumonia	Centers for Medicare & Medicaid Services QIO
28	Pain assessment	Veterans Health Administration (VHA)
29	CABG volume	HCUP quality indicators/New York State CABG measure
30	Percutaneous Coronary Intervention (PCI) Volume	American College of Cardiology (ACC)
31	PCI Mortality	ACC
32	Pediatric heart surgery volume	HCUP quality indicators
33	Pediatric heart surgery mortality	HCUP quality indicators
34	Carotid endarterectomy volume	HCUP quality indicators

Source: NQF Hospital Measures Workshop, 2002

On-stage: The Hospital Measures Workshop Issues

The Hospital Group 2 Measures Workshop provided an opportunity for consumers and purchasers to push their agenda—NQF endorsement of a more robust hospital measure set. However, the providers had serious concerns about the additional measures, which surfaced during discussions over the two day workshop. To begin with, according to a number of interviewees, the providers and hospital community were “annoyed” with the initial set of hospital measures and how they were selected and then “disgruntled” when the Hospital

Measures Workshop, its purpose and charge, was announced. An additional concern was the potential push back from the hospital community. The recommendations of the Workshop participants would eventually come back to the NQF BOD, and many knew that an inherent bias was built into the BOD, specifically one of “not pushing the health care industry so hard that it leaves the table and does not support the NQF with CMS.” As a result of these concerns, “the Hospital Quality Alliance was ultimately formed which a number of individuals would argue, took some authority away from the NQF.”

At the conclusion of the meeting, the participants of the Workshop reaffirmed the evaluation criteria and measures selected by the Hospital Steering Committee, and used these same criteria to evaluate the 27 measures put forward by the NQF membership. The Workshop participants also agreed with the eight priority areas previously identified i.e., acute coronary syndrome, cerebrovascular disease, heart failure, pneumonia, pregnancy/childbirth/neonatal conditions, patient safety, pediatric conditions, and pain management/control, but also identified an additional priority area—abdominal aortic aneurysm (AAA). Finally, the Workshop participants discussed the issue of whether proprietary measures should be recommended and determined that proprietary measures and risk adjustment methodologies could be considered if the measure developers provided an assurance of open access for public scrutiny. From the initial list of measures that were evaluated, the participants of the Workshop recommended 17 measures for potential endorsement. Table 13 lists the measures recommended.

Table 13: Group 2 Hospital Care Performance Measures and their Source

1	Abdominal Aortic Aneurysm (AAA) repair volume	AHRQ
2	AAA repair mortality ²²³	AHRQ
3	Coronary artery bypass graft (CABG) using internal mammary artery (IMA)	CMS
4	CABG volume	New York State
5	Percutaneous coronary intervention (PCI) volume	American College of Cardiology (ACC)
6	PCI mortality ²²⁴	ACC
7	Carotid endarterectomy volume ²²⁵	AHRQ
8	Restraint prevalence	California Nursing Outcomes Coalition (CalNOC)
9	Urinary catheter utilization	CDC
10	Urinary catheter-associated urinary tract infection (UTI)	CDC
11	Central line catheter utilization	CDC
12	Central line catheter-associated infection	CDC
13	Ventilator utilization ²²⁶	CDC
14	Ventilator-associated pneumonia	CDC
15	Patient falls	ANA
16	Pediatric heart surgery volume	AHRQ
17	Pediatric heart surgery mortality	AHRQ

Source: NQF Hospital Measures Workshop, 2002

The original Hospital Steering Committee also provided input via a conference call, and it recommended endorsement of the 17 measures evaluated by the workshop participants. A draft report was sent to the membership on November 1, 2002 with comments due in early December 2002.

On-Stage: Public Comment and the Hospital Group 2 Measure Set

Approximately 20 organizations provided comments on the 17 measures that were forwarded by the work group. The general themes of the comments had to do with concerns

²²³ Risk adjusted using APR-DRG proprietary model.

²²⁴ Risk adjusted using the American College of Cardiology’s logistic regression model.

²²⁵ A plurality, not a majority of the work group participants recommended this measure be considered by NQF members (8 recommended “yes”, 5 opposing, and 3 abstaining).

²²⁶ A plurality, not a majority recommended this measure (8 recommending “yes”, 5 opposing, and 4 abstaining).

about using these measures for public reporting. As with the comments from the Group 1 measures, the issues of adequate evidence base for a measure and the potential unintended consequences of reporting these measures were the stated concerns of the providers and researchers. As expected, the consumers and purchasers stated “some guidance that, while not perfect, is better than no information whatsoever.”

Some also commented on the process by which these additional measures were brought forward, specifically citing the “use of a workshop to review and recommend a set of measures to the NQF members.” The NQF CDP (version 1.5) states that “proposed standard measures are to be developed through a steering committee or review committee” with no role for a workshop panel or group. The comments went on to say that “since the NQF had a Steering Committee to review and recommend measures for the hospital core set, it would have seemed appropriate and in keeping with the approved Consensus Process to have that Steering Committee review and recommend additional measures to the members rather than establishing a separate two-day workshop with other individuals serving as members of the panel.”

Several comments voiced concerns about endorsing hospital measures without adopting a framework for hospital quality measurement. “It is counter-productive and confusing to try to determine specific measures before we have actually agreed on what the framework for those measures will be...By selecting measures without a framework, we may be approving measures that add to the already substantial measurement burden hospitals experience, without giving potential users the information they really need.”

On-stage: The Proprietary Issue Revisited

As part of the pre-voting public comment period, the NQF received a letter from HealthGrades, a vendor that provides report cards on individual providers and hospitals for a fee. One of the measures submitted by the consumer-purchasers included a HealthGrades’ measure

for AAA repair mortality and the AHRQ measure for AAA repair volume. The Workshop members did not want to have a paired measure set with one volume measure coming from one organization and the mortality measure coming from another, and thus they had to choose between the AHRQ or HealthGrades measures for both volume and mortality. There was considerable discussion by the Workshop members regarding this issue and the selection of proprietary measures; as a result, the Workshop participants set three criteria by which proprietary measures could be endorsed. They included: 1) that the measure resides in the public domain and must be fully transparent to the smallest level of detail; 2) that it must be available at a reasonable cost;²²⁷ and 3) that it must be available to everyone. The letter went on to say that the “HealthGrades AAA repair mortality measure meets all the criteria as it is willing to put the full multivariate regression model into the public domain for use by all interested parties if the measure is selected by NQF.” The Workshop panel chose the AHRQ mortality and volume measures for AAA repair, thus removing the HealthGrades’ measures from consideration, and since “it is very unclear as to why they made this decision, HealthGrades requests that the NQF reconsider this decision and select the HealthGrades AAA repair measure for both volume and mortality.” Three reasons were given to support its request, and included specific criticisms of the AHRQ measures:

- The AHRQ measure did not go through an extensive review process, while the HealthGrades measure did
- AHRQ does not endorse the measure for public reporting
- The AHRQ “AAA repair volume measure is a much simpler measure in that it defines, using ICD-9 codes, what is to be counted as an AAA repair

²²⁷ The issue of “reasonable cost” was to be considered and defined by the NQF BOD at a later date.

HealthGrades provided its definition of AAA repair volume as part of its AAA repair mortality measure and commented “It is unclear why the Workshop participants would choose the organization with the recommended volume measure and make the leap that their mortality measure should be used as well. The opposite is far more logical and defensible.”

On-Stage: Voting on Group 2 Hospital Measures

In December 2002, a substantially revised report on the Group 2 Measures was sent to the NQF membership, along with a voting ballot due on January 23, 2003. Most notable was the removal of four measures recommended by the Workshop panel: the AHRQ mortality and volume measures for AAA repair (and the removal of this condition’s classification as a priority area), the AHRQ carotid endarterectomy volume (and its removal as a priority area), and the California Nursing Outcomes Coalition (CalNOC) measure of physical restraint prevalence. Explanations for these deletions were based on concerns “raised during the Workshop or actions taken based on new information following the Workshop.” Table 14 lists the deleted measures and the specific reasons for their exclusion.

Table 14: Measures Excluded from Group 2 Measure Ballot

Measure	Exclusion Reason
AAA repair volume and mortality	Due process issues arose subsequent to the Workshop— “because of the ground rules of no new/additional measures for the Workshop; NQF did not become aware of its existence until after the Workshop. Because of this process issue and since proprietary interests are involved, the measures are being excluded because the HealthGrades measures were not accorded equal consideration for evaluation purposes when such was warranted.”
Carotid endarterectomy volume	“Several parties, subsequent to the Workshop recommended that reference to carotid endarterectomy volume at a specific facility being linked to quality be deleted from the draft for the “Safe Practices” document. Staff re-reviewed the literature, and concurred with the recommendation for deletion for the Safe Practices draft and so deleted the measure from the Group 2 set.”
Physical restraint prevalence	Several clarifications were requested of the measure developer. “Following the Workshop, NQF staff contacted the developer, but increased clarification was not achieved. CalNOC did opine that it had used side rails but that side rails should not be considered restraints for a reliable nationally standardized measure; that the potential for non-random selection of the data for calculating the prevalence study was not an issue in their work; and that stratification by bed size was not an inherent component of the measure, but an analytic tool.

Source: NQF Hospital Measures Report-Group 2, 2002

The exclusions prompted a letter from the AHRQ expressing concern about the removal, prior to membership voting, of the AAA repair volume and mortality measures. Specifically, the objection “is that the measures were excluded from the ballot based upon NQF staff decisions... We do not believe the rationale provided by NQF for why the measures were excluded prior to the ballot is sufficient to warrant the exclusion, especially given the support of the measures from the Hospital Group 2 Workshop participants.” The letter goes on to indicate that “if there was a concern about which source (AHRQ or HealthGrades) to use for the specifications for the measure, we believe that the NQF’s own bylaws would clearly indicate that

the choice should be for non-proprietary measures... We believe that there is general concern among the membership about changes made to NQF consensus documents during the comment period that are not specifically based on the consensus of the membership.” Specifically, AHRQ requested that “specific guidance for staff decisions related to changes in consensus documents that are being made prior to membership vote” be developed, which would include criteria for the deletion of voting items.

Of the 13 measures sent to the NQF membership for a vote, eight were approved by the four councils and subsequently by the BOD. These eight measures were merged with the initial 31 measures endorsed earlier for a total of 39 hospital performance measures. The BOD also settled the proprietary measures debate, and rendered a policy that the “NQF will endorse only measures that are in the public domain and can be utilized with respect to specifications and risk adjustment methodologies without charge” (NQF Board of Directors Meeting Minutes, January 2003). The remaining five measures²²⁸ were sent out in February 2003 for a second round of voting. In addition to the ballot, the comments about the five remaining measures were included for member review. In accordance with the CDP, the staff was to “attempt to resolve the matter and submit a revised draft to all Councils for further consideration and a second vote.” However, the Member Council Chairs and NQF staff did not follow the CDP procedures, and elected to forgo the additional review period, “since no apparent changes can be made to the measures themselves that will address member concerns.” They did agree to have an all-member conference call to discuss the fundamental disagreements with the five measures.

After the conference call, it was clear that the perspectives of each of the councils had not changed. Three measures proposed by CDC were felt to “not be useful measures of quality as

²²⁸ These included: Urinary catheter utilization, Central line catheter utilization, Ventilator utilization, Pediatric heart surgery volume, and Pediatric heart surgery mortality.

they were utilization measures.” The two AHRQ pediatric measures, while believed to be very important to consumers for decision making, were “not ready for prime time.” All four Councils disapproved the pediatric measures with a motion to discontinue consideration of these measures for inclusion in the initial hospital measure set.

The three CDC measures were approved by the consumer and purchaser councils, but disapproved by the provider and research councils. On March 26, 2003, the members of the BOD each received a ballot to vote on the five remaining measures along with the voting results of the Councils, a summary of the all-member conference call, and the comments from the membership on these measures that were submitted with the second round of voting ballots. In the end, the NQF BOD voted not to endorse any of the five measures and also concurred with the staff recommendation to discontinue consideration of the AHRQ pediatric measures at that time.

Backstage: Preparing for the Hospital Measures Workshop

During the first round of measure endorsement, the Steering Committee specifically applied the provider-researcher values, which is why a “handful” of measures, often described as “a pretty wimpy set” was endorsed. In round two, the consumers and purchasers were determined to make their wishes known. In particular, one consumer-purchaser representative stated:

“I recognized that one of the reasons consumers and purchaser values had not gotten expressed in the first round was that it’s always the case in any type of multi-stakeholder health care deliberation, that whoever has the most at risk (it’s 100 percent of providers’ paychecks that are involved), invest more in preparation than those for whom health care is just a fraction of their life, like consumers and purchasers. The providers and

researchers make sure that the measure review process is tough and tight and have a very low tolerance for imperfection.”

It was clear from the first round that consumers and purchasers were not prepared to represent their interests. As one interviewee put it, “I realized if we want in round two for the purchasers and consumers to be able to stand their ground and well represent their values and their interests, it was going to require organization, education, training and planning.” First, the leadership of consumers and purchasers was very active in proposing better informed, more assertive candidates to participate in the Workshop. They also planned and implemented a “training camp” that was conducted via phone calls to educate and “prep” the consumer-purchaser representatives appointed to the Workshop. This careful planning resulted in a group of purchaser and consumer representatives who were mobilized, organized, well-informed and as motivated as the providers and researchers. And as one interviewee stated, “you could tell the difference. A lot more measures got approved and there were plenty of times when authoritative people whether from the Joint Commission or a researcher indicated that ‘this measure is just not good enough,’ in which there was a rebuttal from the other side.” In general the votes often were close, but usually favored approval. As one interviewee stated, “this was the only possible way to get a robust flow of measures. It was wonderful to behold.”

While better organization and education on the part of purchasers and consumers helped push their agenda, three other factors influenced the interactions of the members of the Workshop. First, the membership composition differed from the original Hospital Steering Committee, in that some of the more outspoken members of the provider and research community were not selected to participate in the round two Workshop panel. In addition to including more purchaser and consumer representatives (see Table 11 and Appendix J), several

of these representatives also were members of the NQF BOD, who replaced some of the more reticent members of the Hospital Steering Committee.

A second factor that helped the consumer-purchaser agenda move forward was the IOM reports which, according to one interviewee, stated that “society’s prior trustees for health care performance have failed and we currently have a non-quality reliable and very expensive health care delivery system.” Historically, health care professionals shrugged off the occasional purchaser-consumer voice on quality, often stating “we have the best health care system in the world—what are you complaining about?” The IOM reports gave credence to the concerns often voiced by consumer and purchasers, which, prior to 1998, the provider community frequently dismissed.

A third factor that strengthened and advanced the consumer-purchaser agenda was the 2000 elections that resulted in a Republican-controlled White House and Congress with very specific ideas about the role of government. One of their goals for health care included regulatory relief and the belief in informed markets to improve performance. The Administrator of CMS, a political appointee, made it clear that he wanted the Medicare program to lead the way in reporting the performance of providers, which in theory would increase available information, and subsequently would decrease market asymmetry and promote competition.

Backstage: Voting on Group 2 Hospital Measures

Meanwhile, the National Association of Children’s Hospitals and Related Institutions (NACHRI) had some major objections to the proposed pediatric measures. In fact, the staff at NACHRI felt so strongly that these measures were inadequate and could cause potential harm to

hospitals that it circulated a four page letter to key NQF members prior to the close of voting.²²⁹

Based on these and other comments, AHRQ was looking for a way to be “responsive to the needs of the community, while at the same time not backing away from them [the measures] completely.” Ultimately, the measures were not approved, and there was a motion to drop them from further consideration. According to one interviewee:

“This process sucked up enormous amounts of people’s time and energy and created great frustrations and needless animosity. But that’s the tension that I talked about between the desirability of having such a measure, which everybody agreed with, and the availability of a measure to support the desirability which lacked in the belief of those who approached it from a more rigorous scientific perspective. And that may be both the best part of the NQF and the worst part of NQF. It is a truly democratic process but at the same time, it’s a democratic process that only permits those who register to vote to vote. That is, those who pay their money at the table, and there in lies kind of a rub.”

Performance Measures as a Network Actor

In this act, performance measures that were endorsed by the NQF seemed to have crystallized their role, i.e., for public reporting of performance. What also was made clear were the tight criteria for what a measure should look like in terms of scientific evidence, risk adjustment, and an open source available to the public for scrutiny and eventual use.

Furthermore, as this particular view of what performance measures should look like became more prominent, all the other views of performance measures faded into the background. As Law indicates, one good strategy to retain the current network ordering is to “embody a set of

²²⁹ Ballots were received from 62 of 168 (37%) members who were eligible to vote. Ballots were received from five of 10 consumer members (50%), five of 18 purchaser members (28%), 32 of 67 provider/health plan members (48%) and 20 of 73 research and quality improvement members (27%).

relations in durable materials” (1992a). The *National Voluntary Consensus Standards for Hospital Care: An Initial Performance Measure Set* was the vehicle in which the scientific, evidence-based performance measure viewpoint has been inscribed, and is being used in initiatives such as the HQA, thus retaining the current network relationships.

Another role of performance measures was emerging at about this time in the form of pay-for-performance initiatives. While a number of these initiatives were underway, there was scant evidence that supported the effectiveness of this approach to quality improvement. To fill the void, CMS partnered with the Premier Hospital System to test the hypothesis that pay-for-performance improves quality. While this project was starting up, the HQA also was working on recruiting hospitals to “voluntarily report their performance.” Without a financial incentive, only a fraction of the eligible hospitals committed to providing the information. Later in 2003, however, with the help of the MMA legislation, which provided incentives for hospitals to report, almost all of the eligible hospitals were reporting data on at least ten measures, a subset of the initial hospital measure set originally endorsed by the NQF. Although not exactly rewarding providers for high quality performance, the concept of providing incentives through financial means is becoming a popular mechanism in health care, with performance measures taking center stage in these initiatives.

Given that performance measures are central to quality improvement, a currently held belief of the HQN, and as seen in the previous chapters, their identity is re-made by actors in the network in an effort to meet the goal of performance improvement. And to achieve this, all actors within the network interact, which may reshuffle the network order, and bring about different effects. So too, the addition or subtraction of any actors, and the subsequent reshuffling of the network order may also bring about different results. When actors within a network interact, politics as defined by ANT emerge. In the following chapter, the identities of

performance measures are discussed and the role “politics” plays in the turbulent health care environment.

CHAPTER 7:
**FOLLOWING THE “FRAGILE THREAD” OF PERFORMANCE MEASURES—THE
POLITICS OF TRANSLATION**

“The relationship of science and value is a political problem. That is, any particular conception of their relations involves value judgments and cannot be neutrally determined. It is not less ‘political’; to claim these are separate than it is to claim they are indissociably joined; the relation of science and politics is something we invent, more than it is something we discover”

(Proctor 1991: 152, 231; cf. Root 1993:38, 42).

There is a relationship between performance measures and politics as defined by ANT. Performance measures can be viewed as “science” insofar as they are based on research, employ statistical models to make their results more valid and fair, use specific methods for data collection and aggregation, and employ terms like reliability, numerators and denominators to convey impartiality and scientific rigor. In general, politics, has to do with values—who values what, and who can enrol other actors of the network to get what they want. This study of the NQF, its CDP and the Hospital Quality Network illustrates the relationships between values and science, or what Pels refers to as the “natural proximity” of facts and values (2003). It suggests the formation of relationships, be they human or non-human, within the network and their dissolution, resulting in reordering the elements in a new way. Scientific “facts” took on a certain subjectivity based on the different perspectives of the actors and the goals they sought to achieve. Facts, according to Latour and Woolgar “only appear at the end of a long construction chain, in the course of which controversies were progressively closed and subjective claims hardened into self-evident and routinized knowledge...A fact was therefore a statement that was a subjective claim that had been solidified by a successful process of reification and which emerged in the end as reality” (Latour and Woolgar 1986[1979]; Latour 1987, 1988; Woolgar

1988 as cited in Pels, 2003: 101). As ANT relates, the “following the actors” suggests that in different circumstances or before a different audience, the actor may select which “face” to present. For example, in his *Science in Action*, Latour (1987) depicted science as a Janus *bifrons* that alternates between the ready-made official face of science and the “unofficial, pragmatic, and political science in the making” (Pels, 2003: 101).

Turning to the construction of science, where knowledge and interests continually interact, Latour contends that science resembles “a form of politics continued by other means” (as cited in Pels, 2003: 102). Science, represented here by hospital performance measures, becomes politicized in the NQF endorsement process—the CDP took performance measures and attached certain values to them as consensus developed. Ultimately, certain values are silenced while others become more prominent. As one interviewee stated, “it just comes down to interests.” Since the purchasers and consumers have a 51 % majority on the NQF Board, there is a constant tension among the four councils. One interviewee grudgingly admitted that the providers and researchers were “bullied into measures by consumers and purchasers.”

Disagreements occur among actors, and often whose values become prominent and whose values are silenced is a function of strategy and pure power. As one interviewee put it, “when we bring expert panels together, the work of the expert panel is often judged by a) who is the loudest; b) who doesn’t shut up; and perhaps even more importantly, c) who shows up on any given day.” What often happened when the Committee was at a point where it was ready to make a decision, and needed to vote, if “somebody was in the washroom or somebody came late or had to leave early, it often skewed the vote one way or another.”

The actors in the network obviously have different values; the consumers and purchasers want more measures and are more tolerant of misclassification of providers, while the providers and researchers want more precise measures and are highly intolerant of misclassification (which

usually means fewer measures). Regardless, as the various stakeholders spoke for and about performance measures throughout the process, the identity of performance measures changed. Identities attribute meaning (White, 1992), and in this case the identity of performance measures was continually being restructured, based upon its interactions with the other actors in the network. Recall that this is what Law, Callon, Latour and other ANT theorists describe as the second moment of translation—interessement. It is the manner by which the focal actor tries to impose and stabilize other actors' identities. It is the process of convincing the other actors to accept the identities and issues as defined by the focal actor as well as reinforcing the resolve to move through the obligatory point of passage to address the problem. As illustrated above, consumers and purchasers have different ideas from providers and researchers as to what performance measures are and what they should be used for. Thus, even for classifying something as simple as, say, automobiles, it is easy to see that different groups have different models in mind.²³⁰ The same thing can be inferred about identities, as cited by the previous example. Therefore, how we think about an issue often is influenced by our identity, which in turn works to create a certain “reality” which is favorable to our point of view and unfavorable to other conflicting viewpoints.

Using ANT, we can identify the various voices that spoke for performance measures during the CDP, some of which are reflected in the formal record and some of which were silenced through the interactions of the actors within the network.²³¹

²³⁰ This notion is from the field of sociolinguistics. Prototype theory, developed by George Lakoff (1987) and John Taylor (1995), states that “we have a broad picture in our minds of what a chair is; and we extend this picture by metaphor and analogy when trying to decide if any given thing that we are sitting on counts. We call up a best example, and then see if there is a reasonable direct or metaphorical thread that takes us from the example to the object under consideration” (Bowker and Star, 1999: 62-63).

²³¹ This account and the examples cited are not meant to be definitive. For example, while an organization may have a dominant view of performance measures, it should not be implied that that is its only view. Hence an organization can have multiple views of measures and present different faces in different situations.

The Identities of Performance Measures

Performance Measures as a Tool for Quality Improvement

Traditionally in health care, providers thought of performance measurement as a tool for internal quality improvement, and therefore, it took on an almost insignificant role within hospitals because of the role of professionalism. Any occupation could be considered a profession if “it was based on specialized knowledge and a body of formal theory; as providing services oriented to the public interest and welfare; and to think of themselves as the best judges of proper practice” (Brint, 1994: 32). Therefore, professionals within the hospital setting were expected to be responsible for providing high quality of care and to have a “sense of public responsibility” in discharging their duties. Thus, hospitals did not provide a large amount of resources to quality improvement activities. If there were efforts related to measuring performance, the results often were considered confidential and were shared with a very limited number of individuals, such as a hospital board of directors. Because of the assurances of the medical profession, hospitals and other providers, the quality of the care delivered was never in question. In fact, the notion that the quality of care delivered in the U.S. was far superior to other industrialized countries seemed to be supported by the number of scientific and technological breakthroughs credited to researchers in the United States. Performance measures in this scenario are depicted as a tool to measure quality and encourage quality improvement, albeit a weak tool when compared to the judgments rendered by providers, such as hospitals and physicians. Even today in some areas of the country, professional judgments about whether quality exists still dominate, and the idea that performance measures should be used only for internal quality improvement persists.

Throughout the NQF consensus process, performance measures were thought of as a means to encourage quality improvement within hospitals. This was further supported by other

programs, such as the CMS Quality Improvement Initiatives, that used the public reporting of performance measures to “encourage providers to improve.” In fact, the beginning deliberations of the Hospital Steering Committee included a recommendation from NQF staff that measures selected for the core set should focus first on usefulness for improvement. What was revisited in many of the discussions of the Committee was the purpose of the measure set—specifically, whether the measures would be used for quality improvement or for accountability. The discussions ultimately led to the consensus decision that measures in this initial set should be selected with the purpose of provider accountability, which is a higher standard than those used for quality improvement purposes, according to some experts in the field. One interviewee, a provider/researcher stated that “these same measures [used for internal quality improvement] were then used for other purposes such as public reporting, accountability, and payment,” implying that measures developed for quality improvement were not suitable for other purposes. However, others, specifically the consumers and purchasers felt the opposite—that quality improvement measures, precisely specified, could be used for other purposes, particularly since many of these measures had sophisticated risk adjustment and well-defined numerators and denominators. After Committee deliberations, the voice that implied that those measures used for internal quality improvement somehow were less precise than those used for accountability prevailed over those that believed measures used for quality improvement were just as “scientific.”

Performance Measures as an Object of Hope

With the issuance of the IOM’s seminal reports indicating that the quality of health care in this country was not what it should be, the voice of performance measures changed again. Overall, Americans are worried about health care in general, and the IOM reports made it very clear that serious problems existed within the system. Yet while the IOM reports were

particularly troubling, the cost of health care was becoming more and more of an issue. The answer or part of the answer seems to lie in measuring the performance of providers, with the thought being that consumers would seek out better value for their dollar in terms of higher quality services delivered by providers. Thus, performance measures became an object of hope, particularly for consumer groups: measuring performance would provide an incentive to providers to improve the quality of care and the manner in which it was delivered. As stated in their joint letter to the NQF, “for consumers and purchasers to build a quality-sensitive demand curve for hospital care and reduce personal risk of death and disability, we need rapid access to a much wider initial set of less perfect hospital quality measures.” In this case, performance measures were an object of hope in that its use would potentially drive quality improvement among the providers. They also were an object of hope for the purchasers, but for a different reason. While purchasers were interested in quality, they were much more interested in costs and the value derived from health care services. They thought of performance measures as a way to adjust the asymmetry of the health care market, by providing information about quality and performance in report cards that would be made available to the public. By providing information, not only could consumers select higher quality providers, but publicly reporting performance would also in turn, promote competition between providers, thus potentially lowering health care costs. Additionally, purchasers were hopeful that empowering consumers with information for informed decision-making would eventually drive high cost, low quality providers out of the health care system.

While these two groups had different reasons for performance measure use, in this case, both voices merged to form a larger, more vocal group that insisted on using performance measures as a tool to provide information on the quality of provider performance.

Performance Measures as a Tool for Decision-Making

As stated above, performance measures are a tool that can be used to make decisions. Consumers use them for selecting high quality providers who will give them value for their health care dollars. Purchasers, while interested in quality, use performance measures as a way to weed out those providers that cost them too much. Thus, performance measures are a tool for rationing payment. Health plans have long used measures of performance to make decisions as to who is included in their networks and who is not. And government, as one of the largest purchaser of health care, also uses this method, to a certain extent. For instance, with the passage of the MMA, hospitals were required to report to CMS on 10 quality measures in order to receive their market basket update, and if they refused to report, they do not receive the payment increase.

Performance measures also play a powerful role in pay-for-performance initiatives that are currently in vogue within the industry. CMS's Demonstration Project, in which over 270 hospitals are participating, is an example. This project seeks to measure hospital performance, determine the high quality performers, and reward them financially. Initially, the Project just rewarded superior performance; however, its next phase will seek to decrease payment to hospitals for providing inferior care to their patient populations.

Policy makers also use performance measures as a means to make policy decisions. Reports of provider performance, particularly poor performance, often prompt intervention by the government. For example, the Health Care Quality Improvement Act of 1986 sought to control the quality of physicians who provided health care. This Act wanted to "prevent physicians from moving from state to state without a paper trail of malpractice claims and actions taken by hospitals and other health care entities" (Jones and Duncan, 1999: 1052). As stipulated in the Act, a National Practitioner Data Bank was established and is populated by

hospitals and additional health care entities, malpractice insurance providers, and others, in an effort to identify those poorer performing physicians prior to issuing a license or practice privileges within a state.

Performance Measures as a Bone of Contention

With the emergence of the NQF and with the deliberations of the first measurement project, the voice of performance measures changed yet again. Vestiges of its old self emerged in the debate over the purposes for which performance measures would be used—quality improvement or accountability? Ultimately, measures endorsed by NQF meant that they could be used for accountability and public reporting. With that decision, which measures became endorsed was an important issue not only for providers but for consumers and purchasers as well. The quality of measures was discussed frequently during Committee meetings because it continually re-emerged as an issue. Providers or those that would be measured wanted fewer, more precise measures that were based on scientific evidence, specific, and risk-adjusted when necessary. They also wanted a parsimonious set of measures that could be easily collected and reported, often citing the burden already imposed on them by licensing boards, accreditors, and other regulatory bodies. In fact, the impetus of the project was the “proliferation of measures” as cited by the President and CEO of NQF.

Consumers and purchasers, on the other hand, typically wanted more measures rather than fewer, and they believed that imperfect measures were far preferable to no measures at all. In fact, they strongly voiced their objections to the NQF with regard to the initial hospital measure set, stating that “these measures don’t meet our needs.” Prior to the IOM reports, these voices would have been either silenced or considered irrelevant. With the claim that the health care system has serious safety and quality problems, however, and with the emergence of the NQF, the consumer and to a certain extent the purchaser voices became more dominant. This

occurred over the vociferous objections of a provider community that expressed “annoyance” with the consensus process and the “punitive approach” to quality improvement in the form of public reporting of performance.

Performance measures are becoming even more important to providers, because of pay-for-performance initiatives. Better quality practitioners are being rewarded for providing high quality care efficiently, while others may be penalized. Consequently, income and thus market share are affected. Reporting on performance, particularly when comparisons take place, is an issue for practitioners. Many feel that their reputation is “on the line” and that they are being placed at a disadvantage because “their patients are sicker.” That is why gaming and other unintended consequences are issues as performance measures assume its role within the health care system.

Summary

This section has described some of the values of the various actors that participated in the NQF CDP as well as how these values translate into the numerous identities of performance measures and performance measurement. As proposed by ANT, performance measures can exist in multiple venues with different identities all at the same time. Depending upon who the actors are within the network, the role of performance measures is made and re-made. Performance measures themselves have a certain role to play as well. There are many types of measures that measure performance—some well constructed and other not, and in fact, these measures contribute to the conversation about quality. Measures speak to their specificity, the science of their risk adjustment, and the area that they are intended to measure. Providers and researchers would often use these characteristics of measures to support their position, while on the other side, the consumers and purchasers would cite their interpretation of what the measures were

“saying” to them. Measures speak to the state-of-the-art in scientific terms, and are only as good as the science on which they are based. After the merits of the science of measures has been discussed and debated, other actors within the network begin to assign their values and thus identities to the measures. Politics comes into the debate during translation, which is the process of assigning value or identities to actors within the network. It also determines whose vision of the network and the actors within it, prevails over all the others. This transformation of performance measures from irrelevance to highly valued tool is due in part to the network effects of the hospital quality industry that coalesced around the NQF’s hospital project and consensus development process.

PART IV: A DISTINCTIVE VIEW OF HEALTH CARE PERFORMANCE

“Insanity is doing the same thing over and over again and expecting different results”

Albert Einstein

CHAPTER 8:

CONCLUSIONS

In the contemporary health care environment, working within a collaborative network to achieve quality improvement goals is the preferred route of governance, as opposed to government regulation or formal rule-making. The information gathered from this theoretically-informed case study provides insight into ways that can make working within networks more effective and efficient. As previously stated, the selection of measures in part is a political process, not solely based on scientific rigor, as widely believed. My dissertation is different from the prevailing literature because it explored measurement and the measure selection process beyond the science and examined other factors such as values and infrastructure that significantly played into the NQF consensus development process using ANT as a lens.

Viewing Performance Measures and Measurement in a Different Way

Performance and measuring performance are difficult for a number of reasons. It is difficult to determine which behaviors constitute “good” performance as opposed to other behaviors. Even defining what is meant by “good” performance can be a contentious issue. Although there is no general theoretical basis for performance measurement, the principles of informativeness, differentiation, and change, developed by Marshall Meyer, apply broadly to performance measures across disciplines, including health care. And as this study suggests, politics as defined by ANT is one aspect of performance measurement that is ripe for further

research and development, and also has the potential to be developed into an additional principle of performance measurement.

Furthermore, as suggested by the quote at the beginning of this section, we have to begin thinking about performance measures and performance measurement in a different way, particularly in health care. What is currently measured in health care consists largely of process measures that seem to have little relevance to patient outcomes. According to the National Healthcare Quality and the National Healthcare Disparities Reports, submitted annually to Congress by AHRQ, health care quality is improving at a glacial pace with the racial disparity gaps worsening. Considering the numerous quality improvement activities in health care, and several widely reported successes, such as the 100,000 Lives Campaign, it seems that what we are measuring is perhaps not relevant to improving care, or that there are large gaps in how and what we are measuring.

The gap between what we want to measure and what we can measure is what makes measuring performance so difficult. Reasons for this gap include the complexity of organizations, human nature, the science of measurement and as was shown here, values. Organizations, including hospitals, have many departments which at times, have overlapping responsibilities; thus attributing performance can be easier said than done. To further complicate matters, quality performance has multi-dimensional aspects meaning that no one measure can accurately gauge an organization's performance. Thus, the larger and more complex an organization is, the more measures are needed to discriminate between good and poor performance. Which measures are the "right" ones, and how many should be included in determining how an organization is performing are lingering questions. Human nature becomes a factor, not only in the selection of measures, but also when being measured, e.g., gaming the measures or "studying for the test." In addition, the science of measurement has not advanced as

rapidly as the need to measure. Typically, current performance measures quantify past performance, with the idea that it is a predictor of future performance. However, as was demonstrated by the company Enron, is not always the case. And as demonstrated in this research, values play a role in measurement. What gets measured, how it gets measured, and whether there are incentives or disincentives attached to performance depends on what is valued by individuals influential enough to assert their vision of what constitutes good performance. Obviously, there is a need to measure performance differently from what is currently done in health care.

So what should performance measures look like, ideally? According to Meyer, ideal performance measures should have the following characteristics: parsimony, predictive ability, pervasiveness, stability, and applicability to compensation (2002b: 6). These characteristics, to a certain extent, are incorporated into some health care performance measures. The NQF processes and projects have begun to address several aspects of the ideal measure. However, measures that predict outcomes should play a larger role in the health care agenda. While there is exploratory research going on in this area, there should be more resources devoted to it. Additionally there should be research that addresses the multi-dimensional aspects of quality performance and the data sources required that can easily provide a complete picture of the quality of care provided in the United States. Currently in health care, measurement is done as an activity that is separate from patient care and uses measures that require either billing data or abstraction from the medical record, however, with the widespread adoption of the electronic health record, measurement can become more of a by-product of patient care. Yet, in the meantime, the health care industry needs to consider alternative data sources to measure different aspects of quality. The Nurse Sensitive Measure Set endorsed by the NQF takes a step in this

direction by including other data sources, such as information from human resource records, but other measure sets need to be developed.

Additionally, the political aspects of performance measurement as defined by ANT need to be taken into account, which was described in this research study. As resources become more and more limited, organizations will have to be more strategic about what they choose to measure, and how often. Measure developers will also have to focus their efforts to develop measures that are both comprehensive and address gap areas. Collaboration and consensus have become part of the current landscape in the health care measurement environment, and the lessons learned through this study can be useful during consensus development processes and network negotiations.

Core Themes of the Study

From the research study presented, three core themes emerged. These themes provide a simplified view of where we have been, as well as provide the reader with some “take away” points to consider in dealing with performance measures and performance measurement networks.

Theme 1: Politics is a significant and often overlooked factor in the health care performance measurement enterprise.

Measurement and measurement selection based purely on the scientific evidence is a myth. Politics as defined by ANT has to do with the associations of actors in the network and the use of those associations to enrol other actors to support a particular ordering of the network. As a result of this ordering of actors, power, a network effect, is created and used by actors to assure that their assignment of actor roles and their vision of the network prevails over others. As was shown during the deliberations of the Hospital Steering Committee that was dominated by researcher and providers, their point of view of performance measures i.e., evidence based,

precise, and with little or no misclassification of providers, prevailed over the values of the consumers and purchasers. Thus it is easily understood why the Committee recommended relatively few measures. The translation process that occurred during these deliberations at best, made providers and researchers aware of additional points of view, and at worst, maintained the “status quo” in the health care network.

The selection of measures through a collaborative, consensus-based process is infused with politics, because of the nature of performance measurement. Measures make visible and legitimate the work of providers, as well as promote communication and accountability. Yet at the same time, there are consequences to measure selection. Some of these are ethical or financial, while others are associated with professional identity, autonomy and status which are why arguments about the minutia of measures are so important. While the NQF consensus process does have criteria for measure selection, other factors, some overt and others not so obvious, play a role in this process. The most obvious are the obligatory points of passage i.e., the criteria NQF used to evaluate measures, which include usability, feasibility, importance, and scientific evidence. Other less observable or embedded, although nonetheless actors in the network include existing infrastructure, technology, gaps in measurement sets and inertia (e.g., red tape). According to ANT, all of these actors play a role in the network, and their interaction with other actors in the network often causes a reshuffling of the network order, which changes the network effects produced.

As suggested in this research, performance measures are only based on science and scientific evidence to a point and after that, measure selection becomes one of values. Whose voice dominates and whose values trump others can be an exercise of pure power that was generated by the ordering of the network. For example, when a small set of hospital measures was about to be endorsed the consumer and purchaser groups added another set of actors i.e.,

their own expert panel to the network, thus changing the associations of the actors. Specifically, it disrupted the expert decision-maker model in that it gave the consumer and purchaser groups more knowledge about measures than they previously had and made them savvier in the area of measure identification for a national system, which in turn, put them in a different relationship with the providers and researchers of the network. This example also illustrates a certain invisibility in the consensus process that factors into measure selection, particularly with regard to interactions of actors within a network—the work of the expert panel was not widely known by others in the network. As this research has shown, there appear to be a variety of activities, some of which are readily transparent (front stage) and others that are not (backstage). Some of the backstage activity can be attributed to the long standing relationships of actors within the hospital network. These relationships, such as those between providers and researchers have developed and evolved over years, and thus, they do not have to establish a relationship as they would if new actors were added in the mix. This is one reason why the industry is so resilient—it's because the relationships are established and embedded in the current infrastructure of health care, thus making the established ordering, one that favors these particular actors, more durable. However, when new actors such as consumer and purchasers enter the network, they initiate a reshuffling of all the actors of the network, but as was evident in this research, the reshuffling does not necessary guarantee different network effects. New actors in the health care networks, particularly one that involve consensus building, must establish relationships and must be prepared, active participants in the process, whether consistently or at strategic points throughout the deliberations.²³²

²³² Other areas of potential research include what types of actor activities advance agendas and which ones detract from advancement? Are politics a factor in any type of measure selection process, i.e., one that does not involve the development of consensus? Do politics factor into the selection of measures for other health care sites, such as nursing homes or outpatient surgery departments? Do politics come into play in organizations that chose measures for accreditation programs that allow some discretion in which measures to report?

Theme 2: The environments of collaborative networks are political in nature, but offer a considerable alternative to traditional forms of health care governance.

To a greater extent, networks are becoming more common in the health care landscape. They are significant in that they are increasingly prominent in legislative debates, federal agency discussions, and legislation that are passed by Congress and signed into law. Currently the climate in health care favors collaboration between the various actors in the form of incentives i.e., public reporting performance, and P4P initiatives, for improvement. The idea behind these programs is to encourage competition within the market, which, in turn, will improve quality and drive down health care costs by eliminating those providers that are not efficient or effective about the care they deliver. This is being accomplished without government interference in the form of traditional rule-making and regulation as evident in the MMA, which stipulated that only hospitals that reported their performance publicly would be eligible for the full market basket update. The law gave the hospitals the choice, but nonetheless tied payment to publicly reporting performance. Hence, a new actor entered the hospital network—the MMA, which changed the relationships of all the other actors within the network. What this did was generate a network effect, and in this case a desirable one i.e., that hospitals voluntarily report their performance to the public based upon a standardized measure set.

Collaborative networks and network governance are a reality in health care for a variety of reasons. Networks consist of punctualised actors that represent whole other networks and/or groups of individuals that are linked by what Granovetter describes as the “cohesive power of weak ties” (1973:1360). According to ANT, these social ties are constantly being made and re-made as different actors enter the network. Actors enter and leave networks for a variety of reasons, but mostly because the network provides something of value to the actor. Networks provide information in terms of monitoring the activities of other actors within the network.

Monitoring is one way to assess the network to identify and explore different associations that could potentially re-order the network, thus providing actors with access to network effects such as power or other resources.

Networks also provide a way to pool resources, particularly in the current environment where the resources available to actors is limited. Pooling resources among actors can achieve mutually beneficial goals that would otherwise be impossible to achieve alone. Along these lines, networks provide opportunities for joint learning among actors. For instance, the field of reporting quality information to consumers is relatively new, and because actors within the HQN are concerned about reporting hospital performance the “right” way, actors within the network have shared their results of focus group testing, and additional research with others in the network.

But overall, networks foster flexibility in achieving the goals common to health care stakeholders. Specifically because of the flexibility of this type of governance currently favored in health care, performance measures selected by consensus of the network actors can be refined and easily changed as the evidence base upon which they are based changes and improves. Due to the flexibility and creativity of networks, government to a certain extent relies on these loosely constructed entities, particularly with regard to measuring performance and quality improvement. In these types of collaboratives, the federal government, guides network actors toward common goals. It does so through the addition of actors within the network, i.e., the addition of the MMA to the HQN, P4P initiatives, etc., and through developing additional social ties with other actors, i.e., the talks CMS initiated with the HQA to achieve their goal of reporting hospital performance to the public. In this case, the HQN served to scrutinize, develop consensus and select a core group of hospital measures as part of a national performance reporting and quality improvement initiative for hospitals.

However, networks are political in nature. Recall the meaning of “politics” from the ANT perspective: “the progressive composition of the common world and all the competencies exercised by the collective” (Latour, 2004: 247). Networks consist of relationships that in turn design the network, thus creating a common world for the actors of the network. Altering the composition of the network can change network effects as was illustrated in the example of the MMA. But what is important is *who’s* vision of the common world will prevail over the other visions that are silenced. In the HQN, with the emergence of the HQA, it was clear that the NQF vision of hospitals implementing all the endorsed measures was silenced, and the HQA vision of implementing far fewer became the dominate vision of the common world of the network. This change in the HQN effect could not have been accomplished without new actors admitted to the network, i.e., the HQA, the new Administration, and the additional relationships established around the selection of a core set of measures for hospitals, i.e., a shift in the relationship between the Joint Commission and the three national hospital associations and the emergence of the NQF. Furthermore, the re-ordering of the HQN that occurred with the appearance and relative success of the HQA, made changes in the other quality improvement networks of health care. Specifically with the formation of other Quality Alliances (the Pediatric Quality Alliance, the Ambulatory Quality Alliance, the Pharmacy Quality Alliance, etc), it has become apparent that a collaborative, consensus-based governance system is preferred by a majority of stakeholders within the industry.²³³

²³³ There are other potential areas of research that seem to emerge from this theme, such as: Is collaborative governance the best way to move the health care agenda forward? What are the pros and cons associated with these types of networks? Are there other types of governance that are more efficient or more effective?

Theme 3: Viewing networks from an ANT perspective provides a basis for actors to assess the political nature of a network, and develop strategies to achieve their goals within the network structure.

From the ANT perspective, it is important to take into account all the actors within a network, both human and non-human, as each contributes different perspectives that in turn change the nature of relationships. Some illustrations of this point are obvious and others are not necessarily so. For instance an example of the obvious is the state of the science of risk adjustment methodology of performance measures. The methods of risk adjustment are only as good as the science on which it is based, and the science has only advanced so far. Because measures are associated with certain risk adjustment models, it plays an important role in the network to a certain extent, because it influences the other network actors to respond in a certain way, and in this case, to select particular measures over others. If the science of risk adjustment changed, then perhaps other measures might be selected by the members of the NQF Hospital Steering Committee and the HQA.

The not-so-obvious example has to do with things like materials such as letterhead, buildings, reimbursement forms, hotels, the list of Committee members and the like that may be considered background to most, but in terms of ANT, these materials begin to define the actors and their subjectivities prior to their verbal performance. For instance, the selection of luxurious hotels where NQF meetings were held conveys the importance of the task at hand and the importance of being a member of the Committee. These materials translate the individual actors from their original identities, into members of the Hospital Steering Committee. Identities such as physician, nurse, shopper, bureaucrat, etc., were replaced (and ordered according to Law) with specific NQF Council representation—consumer, purchaser, providers, researchers and because it was an NQF meeting, it further translated the members into representatives of a democratic

process. The Committee list of members along with the other meeting materials provided denotes knowledge, authority as experts, and a certain exclusivity as these materials were not provided to the audience. All these materials contribute to a network ordering, which in this case provides a different kind of logic, specifically one of sharing perspectives and accommodating differences—in short, building consensus. Thus, before one word was uttered at the NQF Hospital Steering Committee Meeting, the members of the Committee were translated into members of the Hospital Steering Committee. And furthermore, the members of the audience were also translated through the materials distributed. Specifically, the audience did not have name tags, or tables on which to write, or have access to the same materials that the Committee members had, thus translating them into mere observers of the process.

The NQF, largely a new actor in the network, was designed intentionally as a largely democratic process, where issues could be debated and consensus developed. As ANT suggests, when a network adds or subtracts different actors either human or non-human to the mix, the relationships and social ties change, and thus the effects of the network change, as this research demonstrated. For example, by changing the actors representing the consumers and purchasers, omitting others representing the providers and researchers, and preparing, developing and clearly articulating their interests and values, the purchasers and consumers were able to expand the core measure set that could be used for hospital performance reporting. As demonstrated by this research, actors associations and disassociations are important as they produce different products or results from the network. And ultimately, this plays out on a national stage among the diverse stakeholders within health care, all of which have different values and ideals for what constitutes

acceptable performance by hospitals, how it should be measured, and which strategies should be employed to improve care.²³⁴

Lessons Learned

From the practitioner's perspective, this research highlights the need for physicians and nurses to take into account other perspectives with regard to performance measurement. As illustrated, the consumers and purchasers have a different point of view, which is important, and which ultimately can advance the quality agenda in health care. Further, practitioners should view the quality improvement enterprise as a collaboration among all the stakeholders, and not as a way to disempower providers. Overall, practitioners would be wise to participate in the discussions of networks such as the HQN as well as promote team functioning in quality improvement initiatives.

From a federal agency perspective, it is important to realize the leverage agencies have over collaborative networks. In its position, the federal government has the opportunity to convene all the stakeholders in health care and to guide the quality agenda to fruition. However, it should be cautious that the guidance is not overpowering, because if that is the case, buy-in from all the stakeholders becomes an issue. Representatives of the federal government should be at the table during discussions, not to lead the discussions but to facilitate discussion among the different parties and support advancing the agenda of improvement in health care delivery.

Scholars of performance measurement should acknowledge that politics plays a significant role in performance measurement. Because measuring performance is based upon values, inclusion in the principals of performance measurement should be considered.

²³⁴ Additional areas of research could include an analysis of which actor characteristics contribute to agenda advancement and which do not. A study of panel participant characteristics could provide additional information about how particular groups go about interacting in a network and during consensus development. For example, do physicians approach the task differently than nurses? Does the age or gender of the participant make a difference in negotiations involved in consensus building?

Additionally, scholars of networks should also pay attention to ANT because it takes into account all actors within a network equally. It acknowledges that all kinds of actors have a role to play and that actor interactions form the structure of the network which, in turn determines the product of the network.

From the NQF perspective, it is important to note that politics play a role in the formal CDP. Selection of committee members is vital to the outcome of the CDP, and involving a number of different actors is imperative—not only in terms of the balance between NQF councils but also who is selected to be at the table. Selection of the same individuals who continually push their agendas does not provide different perspectives, particularly of those who are on the front lines of health care. In addition, the process of selection, as well as all of the other processes associated with measure endorsement, should be as transparent as possible. Overall, health care stakeholders may not support endorsement decisions if they could not follow how decisions were made by NQF committees. Additionally, a reasonable amount of time should be provided for discussion and vetting of decisions prior to their implementation. While it is true that most parties believe that they are not “heard well enough” during endorsement decisions, it is important to provide a venue for differing opinions to be expressed throughout the process.

The research presented in this dissertation uses a novel approach to study the Hospital Quality Network and performance measures. It builds on the current literature of performance measurement by adding an additional chapter about consensus development and the politics involved in measurement and measure selection for national standards. Through this instrumental case study, reasonable generalizations can be drawn. This study can potentially apply to hospital performance measurement systems in other countries. Additional studies could look the differences between countries with a national health care system and those without one, and whether the consensus development is as extensive or conducted in the same way as in the

U.S. Along these same lines, because consensus development is becoming more prevalent at the international level, a study that looks at consensus development on health care issues among countries, e.g., WHO meetings, could be useful and potentially add another dimension to process.²³⁵

As consensus building becomes more prevalent in health care, the areas cited here could provide additional information about the process of consensus development and the actors that participate in a collaborative network.²³⁶

Health Care Policy Implications

In today's health care landscape, it is essential for policy makers to provide flexibility within policies due to the fast paced change currently experienced by the industry. While policy makers feel the need to quickly improve the quality of care delivered in the nation, it is important to get the necessary input from the diverse stakeholders of health care. Building consensus takes time, and is difficult and chaotic at times as evident in this case study of the NQF. However, when consensus is developed within collaborative networks such as the HQN, it allows for a certain amount of flexibility and stakeholder "buy in." While politics have a role to play within networks, the actors and their relationships and associations are vital and can provide certain leavers to promote the health care quality agenda. Careful attention to the actors and the associations and disassociations of a network can bring about different solutions to difficult problems that plague the health care system. Therefore it is important to "keep the open spaces

²³⁵ Additional research could also further evaluate the role of the government at all levels in the collaborative process. Are particular agencies perceived as having more influence on measure selection? Are certain agencies more successful at driving the national agenda forward in health care? If so, what are those organizational characteristics that promote and foster agenda advancement?

²³⁶ A summary of the main research questions of this study can be found in Appendix N.

open,” which means to explore options other than the dominant one, because that is where new ideas emerge.

It is also important for the policy maker to recognize that standards and measures of performance more or less reflect the values, which are both ethical and subjective, of different constituencies. For every value or point of view expressed in measurement sets, another often is silenced, or at the very least altered. Once these measures are incorporated into the health care infrastructure, it can be difficult to retrieve those dissenting voices, which may provide solutions to future problems encountered by the health care field. What’s more is that once these measures become part of the infrastructure (either visible or invisible), in essence, the public loses its ability to participate in policy decisions. Therefore, it is essential to recognize that performance measures and performance measurement systems should be “living” e.g., possess the ability to be flexible to address change, but not so flexible that comparison among providers is impossible. This is the constant challenge of selecting measures and measure sets for national programs designed to improve the quality of care delivered in the U.S.

Postscript: The NQF—An Update

Since I began this research, several changes have taken place at the NQF. Most notable was the departure of Ken Kizer, M.D., the first President and CEO of the organization, who ultimately was responsible for developing its membership base and its policies and structure, and for establishing its legitimacy and reputation within the health care community. In early 2006, Janet Corrigan, Ph.D., assumed the responsibilities as President, at which time she merged her former organization, The National Committee for Quality Health Care, with the NQF, to form a new programmatic area within the NQF.

Upon her arrival, Dr. Corrigan sought extensive input from the health care community regarding the NQF, its structure and processes. The results obtained were very similar to those reported here. A task force was appointed that evaluated the information, and a strategy was developed to address the most pressing issues, most notably the poor communication within the current council structure and the excessively long process for measures to emerge from the current consensus development process. A few changes have occurred since 2006: the NQF has refined and enhanced its mission²³⁷ to include “setting national priorities, endorsing consensus standards, and education and outreach” (NQF, Snapshot of Accomplishment 2007: 1). The governance and committee structure of NQF was changed. The BOD reduced its members from 33 to 27 and now consists of 19 at-large seats and four standing seats (three are dedicated to the federal government and one for the President of NQF). Other changes include the restructuring of the councils, whose number went from four to eight, and the addition of the Consensus Standards Approval Committee (CSAC)²³⁸ to the CDP. The National Priorities Partners Committee was also established to identify a set of “high-leverage” priorities for the nation. Additional changes are currently underway and include the Leadership Network, which is designed to provide guidance to NQF on the various aspects of its projects and developing efforts that focus on education and outreach. It should be noted that the original charge to the NQF and its focus has not changed: that it provide a forum for debate and consensus development, and that it remain consumer-focused.

²³⁷ “The mission of the National Quality Forum is to improve the quality of American healthcare by setting national priorities and goals for performance improvement, endorsing national consensus standards for measuring and publicly reporting on performance, and promoting the attainment of national goals through education and outreach programs” (NQF, Snapshot of Accomplishments 2007: 2).

²³⁸ The CSAC was established in 2007 to improve the quality and timeliness of the CDP, and is the final step in the process prior to BOD approval.

REFERENCES

- AFL-CIO. (2008). *Union facts*. [Online]. Available: <http://www.aflcio.org/aboutus/faq/index.cfm> Accessed January 2008.
- Agency for Healthcare Research and Quality. (1998). *The challenge and potential for assuring quality health care for the 21st century*. [Online]. Available: <http://www.ahrq.gov/qual/21stcena.htm> Accessed June 2008.
- Agency for Healthcare Research and Quality. (2001). *Profile: Quality research for quality health care*. (AHRQ Publication No. 00-P005). Rockville, MD: AHRQ.
- Agency for Healthcare Research and Quality. (2005). *AHRQ annual report on research and financial management, FY 2004*. (AHRQ Publication No. 05-0019). Rockville, MD: AHRQ.
- Agency for Healthcare Research and Quality. (2006). *About the health care report card compendium*. [Online]. Available: <http://www.talkingquality.gov/compendium/index.asp?mode=about> Access November 2006.
- Agency for Healthcare Research and Quality. (2006). *National advisory council for healthcare research and quality*. [Online]. Available: <http://www.ahrq.gov/about/council.htm> Accessed May 2006.
- American College of Surgeons. Hospital Standardization Series: Conference on Hospital Standardization: New York, October 24, 1919, 3-9.
- American Hospital Association (AHA). (2005). *History of the AHA*. [Online]. Available: <http://www.aha.org/aha/about/history.html> Accessed June 2006.
- American Hospital Association. (2006). *About the AHA*. [Online]. Available: <http://www.aha.org/aha/about/index.html> Accessed August 2007.
- American Medical Association (2007). *AMA (About AMA) AMA Mission*. [Online]. Available: <http://www.ama-assn.org/ama/pub/category/1815.html> Accessed August 2007.
- American Medical Association (2007). *AMA (HOD) House of Delegates*. [Online]. Available: <http://www.ama-assn.org/ama/pub/category/print/9873.html> Accessed August 2007.
- American Medical Association (2007). *AMA (About AMA) 2006 annual report*. [Online]. Available: <http://www.ama-assn.org/ama/pub/category/12528.html> Accessed August 2007.
- Atkinson, C.J. (2002). The multidimensional systemic representation of actor networks: Modeling breast cancer treatment decision-making. Presented at the 35th Hawaii International Conference on System Sciences.

- Ayanian, J.Z., Udvarhelyi, I.S., Gatsonis, C.A., Pashos, C.L., & Epstein, A.M. (1993). Racial differences in the use of revascularization procedures after coronary angiography. *JAMA* 269(20), 2642-2646.
- Ayanian, J.Z., Weissman, J.S., Chasan-Taber, S., & Epstein, A.M. (1999). Quality of care by race and gender for congestive heart failure and pneumonia. *Medical Care* 37(12), 1260-1269.
- Azoulay, R. (2004). *The rational choice theory*. [Online]. Available: <http://web.grinnell.edu/courses/pol/S00/POL256-01/wto/theory.html> Accessed August 2004.
- Baicker, K., & Chandra, A. (2004). Medicare spending, the physician workforce, and beneficiaries' quality of care. *Health Affairs*, hlthaff.w4.184.
- Bardini, T. (1997). Bridging the gulfs: From Hypertext to Cyberspace. *Journal of Computer-Mediated Communication* 3(2) [Online]. Available: <http://jcmc.indiana.edu/vol3/issue2/bardini.html> Accessed October 2007.
- Barker-Cummings, C., McClellan, W., Soucie, J.M., & Krisher, J. (1995). Ethnic differences in the use of peritoneal dialysis as initial treatment for end-stage renal disease. *JAMA* 274(23), 1858-1862.
- Barr, J. K., Boni, C. E., Kochurka, K. A., Nolan, P., Petrillo, M., Sofaer, S., et al. (2002). Public reporting of hospital patient satisfaction: The Rhode Island experience. *Health Care Financing Review* 23(4), 51-70.
- Batalden, P. B., & Stoltz, P. K. (1993). A framework for the continual improvement of health care: building and applying professional and improvement knowledge to test changes in daily work. *Joint Commission Journal of Quality Improvement* 19(10), 424-447.
- Bauchner, H., Simpson, L., & Chessare, J. (2001). Changing physician behavior. *Archives of Diseases in Childhood* 84(6), 459-462.
- Berger, S. & Sudman, S. (1991). Making total quality management work. *Healthcare Executive* 6(2), 22-25.
- Berki, S.E. (1972). *Hospital economics*. Lexington, MA: Lexington Books.
- Bernstein, S.J., McGlynn, E.A., Siu, A.L. et al. (1993). The appropriateness of hysterectomy. *Journal of the American Medical Association* 269, 2398-402.
- Berry, S., Spranca, M., & Brown, J. (2001). Consumers and health care quality information: Need, availability, utility. Prepared by RAND for the California HealthCare Foundation.
- Berwick, E.M., Godfrey, A.B. & Roessner, J. (1990). *Curing health care: New Strategies for quality improvement*. San Francisco, CA: Jossey-Bass Publishers.

- Bicheno, J. (1989). Exploring productivity accounting for management and strategy in manufacturing and services.. *International Journal of Operations & Production Management* 5, 56-68.
- Bowker, G. C. & Star, S.L. (1999). *Sorting things out: Classification and its consequences*. Cambridge: The MIT Press.
- Boyd, M., & Haraway, C. (1991). Total quality at Baptist Memorial Hospital. *Journal for Quality and Participation* (January/February), 60-66.
- Brashier, L.D., Sower, V.E., Motioani, J. & Savoie, M. (1996). Implementation of TQM/CQI in the health care industry: A comprehensive model. *Benchmarking for Quality Management & Technology* 3(2), 31-50.
- Brennan, T.A., Leape, L.L., Laird, N.M., Herbert, L., Localio, A.R., Lawthers, A.G., Newhouse, J.P., Weiler, P.C., & Hiatt, H.H. (1991). Incidence of adverse events and negligence in hospitalized patients: Results of the Harvard medical practice study. *New England Journal of Medicine* 324, 370-376.
- Brennan, T. & Berwick, D.M. (1996). *New Rules: Regulation, markets, and the quality of American health care*. San Francisco: Jossey-Bass, Inc.
- Brewer, J. & Hunter, A. (1989). *Multimethod research: A synthesis of styles*. Newbury Park, CA: Sage.
- Brint, S. (1994). *In an age of experts: The changing role of professionals in politics and public life*. Princeton, NJ: Princeton University Press.
- Bodenheimer, T. (1999). The American health care system: The movement for improved quality in health care. *The New England Journal of Medicine* 340 (6), 488-492.
- Booz, Allen, & Hamilton. (2006). *Hospital quality reporting in the United States: A cost analysis for the Hospital Quality Alliance*. [On-line]. Available: www.hospitalqualityalliance.org Accessed August 2007.
- Brooks, D.C. (1995). The Joint Commission on Accreditation of Health Care Organizations. In N. Goldfield & D.B. Nash (Eds.), *Providing quality care: Future challenges, 2nd edition*, (pp. 145-160). Ann Arbor, MI: Health Administration Press.
- Brook, R.H. (1973). *Quality of care assessment: A comparison of five methods of peer review*. Rockville, MD: Department of Health, Education and Welfare, Bureau of Health Research.
- Brook, R.H. (1991). Quality of care: Do we care? *Annals of Internal Medicine* 115, 486-490.
- Brook, R.H., Park, R., Chassin, M, Kosecoff, J., Keesey, J., & Solomon, D. (1990). Carotid endarterectomy for elderly patients” Predicting complications. *Annals of Internal Medicine* 113, 747-753.

- Brown, S.D., & Capdevila, R. (1999). Perpetuum mobile: Substance, force and the sociology of translation. In J. Law & J. Hassard (Eds.), *Actor Network Theory and after*, (pp. 26-50). Blackwell Publishers: Malden, MA.
- Bureau of National Affairs, Inc. (2001). Medicare: hospitals need to quickly develop quality measurement standards, Scully says. *Health Care Policy Report: Federal News* 9(35), 1353-1354.
- Burgess, T.F. (1990). "A review of productivity." *Work Study* (Jan-Feb), 6-9.
- Burney, R. (1994). TQM in a surgery center. *Quality Progress* 27(1), 97-100.
- Callon, M. (1986). Some elements in a sociology of translation: Domestication of the scallops and fishermen of St. Brieuc Bay. In J. Law (Ed.), *Power, action, belief: A new sociology of knowledge?* (pp. 196-233). Routledge and Kegan Paul, London.
- Callon, M. (1991). Techno-economic networks and irreversibility. In J. Law (Ed), *Sociology of monsters: Essays on power, technology and domination*, (pp 132-164). London: Routledge.
- Callon, M. (1993). *Social ordering*. Lecture presented at the University of Illinois, Urbana-Champaign, IL.
- Caporaso, J.A. & Levine, D.P. (1992). *Theories of political economy*. New York: Cambridge University Press.
- Caudle, S.L. (1994). Using qualitative approaches. In J.S. Wholey, H.P. Hatry, and K.E. Newcomer (Eds.), *Handbook of practical program evaluation*, (pp.69-95). San Francisco: Jossey-Bass Publishers.
- Centers for Disease Control (2004). *Health, United States 2004, with chartbook on trends in the Health of Americans*. [Online] <http://www.cdc.gov/nchs/data/hus/hus04.pdf> Accessed July 2007.
- Centers for Medicare & Medicaid Services. (2005). *FY 2005 Budget Brief, Centers for Medicare & Medicaid*. [Online] <http://www.hhs.gov/budget/05budget/centersformed.html> Accessed March 2008.
- Centers for Medicare & Medicaid Services. (2006). *National Health Expenditure Accounts 2006 Highlights*. [Online]. Available: <http://www.cms.hhs.gov/nationalhealthexpenddata/downloads/highlights.pdf> Accessed March 2008.
- Chassin, M.R. (1997). Assessing strategies for quality improvement. *Health Affairs* 16(3), 152-61.

- Chassin, M.R. (1998). Is healthcare ready for six sigma quality? *The Milbank Quarterly* 76(4), 565-591.
- Chassin, M.R. (2002). Achieving and sustaining improved quality: Lessons from New York State and cardiac surgery. *Health Affairs* 21(4), 40-51.
- Chassin, M.R., Galvin, R.W., & the National Roundtable on Health Care Quality. (1998). The urgent need to improve health care quality. *Journal of the American Medical Association* 280, 1000-5.
- Chassin, M.R., Hannan, E.L., & DeBuono, B.A.. (1996). Benefits and hazards of reporting medical outcomes publicly. *New England Journal of Medicine* 334, 394-398.
- Collins, S.R., Doty, M.M., Davis, K., Schoen, C., Holmgren, A.L., & Ho, A. (2004). The affordability crisis in U.S. health care: Findings from the Commonwealth Fund Biennial Health Insurance Survey. New York, NY: The Commonwealth Fund.
- Crabtree, B.F., & Miller, W.L. (1992). *Doing qualitative research: Multiple strategies*. Newbury Park, CA: Sage.
- Craig, C.E. & Harris, C.R. (1973). Total productivity measurement at the firm level. *Sloan Management Review* 14(3), 13-29.
- Creswell, J.W. (1998). *Qualitative inquiry and research design: Choosing among five traditions*. Thousand Oaks, CA: Sage Publications.
- Dean, J.W., & Bowen D.E. (1994). Management theory and total quality: Improving research and practice through theory development. *The Academy of Management Review* 19(3), 392-418.
- Denzin, N.K. & Lincoln, Y.S. (1994). Introduction: Entering the field of qualitative research. In N.K. Denzin & Y.S. Lincoln (Eds.), *Handbook of qualitative research* (1-17). Thousand Oaks, CA: Sage Publications.
- Dixon, J.R., Nanni, A.J. & Vollmann, T.E. (1990). *The new performance challenge: Measuring operations for world class competition*. Homewood, IL: Dow Jones/Irwin.
- Donabedian, A. (1966). Evaluating the quality of medical care. *Milbank Memorial Fund Quarterly* 44, 166-203.
- Donabedian, A. (1980). *Explorations in quality assessment and monitoring, volume 1: The definition of quality and approaches to its assessment*. Ann Arbor, MI: Health Administration Press.
- Eddy, D.M. (1998). Performance measurement: Problems and solutions. *Health Affairs* 17(4), 7-25.

- Eisenberg, J. M., Meyer, G., & Foster, N. (2000). AHRQ update: Medical errors and patient safety. *Health Services Research* 35(3), xi-xv.
- Eisenberg, J.M. (2001). What does evidence mean? Can the law and medicine be reconciled? *Journal of Health Politics, Policy and Law* 26(2), 369-381.
- Ellis, J., Mulligan, I., Rowe, J., & Sackett, D.L. (1995). Inpatient general medicine is evidence based. *Lancet* 346, 407-410.
- Epstein, A.M., Ayanian, J.Z., Keogh, J.H., Noonan, S.J., Armistead, N., Cleary, P.D., Weissman, J.S., David-Kasdan, J.A., Carlson, D., Fuller, J., Marsh, D., & Conti, R.M. (2000). Racial disparities in access to renal transplantation: Clinically appropriate or due to undersue or overuse? *New England Journal of Medicine* 343(21), 1537-1544.
- Everett, W. (1993). An epidemiological approach to quality assurance in the hospital. *Health Care Management Review* 18(1), 91-6.
- FACCT (Foundation for Accountability). (1997). *The FACCT consumer information framework: Comparative information for better health care decisions*. [Online]. Available: <http://www.facct.org/information.html> Accessed January 2004.
- FACCT: Foundation for Accountability. (2003). Quality measures. [Online]. Available: <http://www.facct.org/facct/site/facct/facct/Measures> Accessed January 2004.
- Fahey, T.P., & Peters, T.J. (1996). What constitutes controlled hypertension? Patient based comparison of hypertension guidelines. *British Medical Journal* 313, 93-96.
- Fine, T. & Snyder, L. (1999). What is the difference between performance measurement and benchmarking? *Public Management* 81(1), 24-25.
- Fisher, E.S., Wennberg, D.E., Stukel, T.A., Gottlieb, D.J., Lucas, F.L., Pinder, E.L. (2003a). The implications of regional variations in Medicare spending, Part 1: The content, quality, and accessibility of care. *Annals of Internal Medicine* 138 (4), 273-287.
- Fisher, E.S., Wennberg, D.E., Stukel, T.A., Gottlieb, D.J., Lucas, F.L., Pinder, E.L. (2003b). The implications of regional variations in Medicare spending, Part 2: Health outcomes and satisfaction with care. *Annals of Internal Medicine* 138(4), 288-298.
- Fowles, J. B. (2000). Health plan report cards may influence insurers more than consumers: Their effect on insurer behavior in Minnesota. *Health Care Financing & Organization* 3(3).
- Fowles, J.B., Kind, E.A., Braun, B.L., & Knutson, D.J. (2000). Consumer responses to health plan report cards in two markets. *Med Care* 38(5), 469-481.
- Frist, T. (1992). TQM at HCA. *Health Systems Review* 25(3), 15-18.

- Gaylin, D.S., Held, P.J., Port F.K., Hunsicker, L.G., Wolfe, R.A., Kahan, B.D., Jones, C.A., & Agoda, L.Y. (1993). The impact of comorbid and sociodemographic factors on access to renal transplantation. *JAMA* 269(5), 603-608.
- Gill, P., Dowell, A.C., Neal, R.D., Smith N., Heywood, P., & Wilson, A.E. (1996). Evidence based general practice: A retrospective study of interventions in one training practice. *British Medical Journal* 312, 819-821.
- Godfrey, A.B., Berwick, D.M. & Roessner, J. (1992). Can quality management really work in health care? *Quality Progress* 25(4), 23-27.
- Goffman, E. (1959). *The presentation of self in everyday life*. New York: Doubleday Press.
- Goodman, J. & Walsh, V. (2001). *The story of Taxol: Nature and politics in the pursuit of an anti-cancer drug*. New York: Cambridge University Press.
- Gormley, Jr., W.T. (1998). Assessing health care report cards. *Journal of Public Administration Research and Theory* 8(3), 325-353.
- Gormley, Jr., W.T., & Weimer, D. L. (1999). *Organizational report cards*. Cambridge, MA: Harvard University Press.
- Granovetter, M.S. (1973). The strength of weak ties. *American Journal of Sociology* 78 (6), 1360-1380.
- Gray, B.H. (1992). The legislative battle over health services research. *Health Affairs*, 38-66.
- Gray, B.H., Gusmano, M.K. & Collins, S.R. (2003). AHCPR and the changing politics of health services research. *Health Affairs*, W3-283-W3-307.
- Hamblin, R. (2007). Publishing 'quality' measures: How it works and when it does not? *International Journal for Quality in Health Care* 19(4), 183-186.
- Hannan, E.L., O'Donnell, J.F., Kilburn, Jr., H., Bernard, H.R., & Yazici, A. (1989). Investigation of the relationship between volume and mortality for surgical procedures performed in New York State hospitals. *Journal of the American Medical Association* 262, 503-10.
- Hannan, E.L., Kilburn, H., Racz, M., Shields, E., & Chassin, M.R. (1994). Improving the outcomes of coronary artery bypass surgery in New York State. *Journal of the American Medical Association* 271, 761-6.
- Hannan, E.L., Siu, A.L., Kumar, D., Kilburn, Jr., H., & Chassin, M.R. (1995). The decline in coronary artery bypass graft surgery mortality in New York State: The role of surgeon volume. *Journal of the American Medical Association* 273, 209-13.
- Hannan, E.L., van Ryn, M., Burke, J., Stone, D., Kumar, D., Arani, D., Pierce, W., Rafii, S., Sanborn, T.A., Sharma, S., Slater, J., & Debuono, B.A. (1999). Access to coronary

- artery bypass surgery by race/ethnicity and gender among patients who are appropriate for surgery. *Medical Care* 37(1), 68-77.
- Health Care Information Center. (2003). Hospital quality: Scully disappointed with hospital reporting. *Medicine & Health*, 57(37), 1-8.
- Henry J. Kaiser Family Foundation. (2004a). *July/August 2004 Health Poll Report Survey*. [Online]. Available: http://www.kff.org/kaiserpolls/upload/KHPR_Aug04_Topline_Release.pdf Accessed August 2008.
- Henry J. Kaiser Family Foundation. (2004b). *Trends and indicators in the changing health care marketplace—section 4: Trends in health insurance benefits*. [Online]. Available: <http://www.kff.org/insurance/7031/print-sec4.cfm> Accessed March 2005.
- Henry J. Kaiser Family Foundation, Agency for Healthcare Research and Quality, Harvard School of Public Health. (2004). *National Survey of consumer's experiences with patient safety and quality information*. [Online]. Available: <http://www.kff.org> Accessed December 2005.
- Henry J. Kaiser Family Foundation/eHealthInsurance. (2004). *Update on individual health insurance, August 2004 [publication #713302]*. [Online]. Available: <http://www.kff.org> Accessed August 2008.
- Herholz, H., Goff, Jr., D.C., Ramsey, D.J., Chan, F.A., Ortiz, C., Labarthe, D.R. & Nichaman, M.Z. (1996). Women and Mexican Americans receive fewer cardiovascular drugs following myocardial infarction than men and non-Hispanic white: The Corpus Christi Heart Project, 1988-1990. *Journal of Clinical Epidemiology* 49(3), 279-287.
- Hibbard, J.H., & Jewett, J.J. (1996). What type of quality information do consumers want in a health care report card? *Med Care Res Rev* 52 (1), 28-47.
- Hibbard, J., Stockard, J., & Tusler, M. (2003). "Does publicizing hospital performance stimulate quality improvement efforts? *Health Affairs* 22(2): 84.
- Hibbard, J, Stockard, J, & Tusler, M. (2005). Hospital performance reports: impact on quality, market share, and reputation. *Health Affairs* 24(4): 1150-1160.
- Hofer, T.P., Hayward, R.A., Greenfield, S. et al. (1999). The unreliability of individual physician "report cards" for assessing the costs and quality of care of a chronic disease. *JAMA* 281, 2098-2105.
- Hoffman, Jr., E.D., Klees, B.S., & Curtis, C.A. (2007). *Brief summaries of Medicare & Medicaid: Title XVII and Title XIX of the Social Security Act*. [Online]. Available: <http://www.cms.hhs.gov/MedicareProgramRatesStats/downloads/MedicareMedicaidSummaries2007.pdf> Accessed August 2008.

- Hospital Quality Alliance (HQA). (2002). *Hospital quality alliance home page*. [Online]. Available: <http://www.hospitalqualityalliance.org/hospitalqualityalliance/index.html> Accessed August 2008.
- Hospital Quality Alliance (HQA). (2007). *Mission and vision*. [Online]. Available: <http://www.hospitalqualityalliance.org/hospitalqualityalliance/aboutus/aboutus.html> Accessed August 2008.
- Hussey, P.S., Anderson, G.F., Osborn, R., Feek, C., McLaughlin, V., Millar, J., & Epstein, A. (2004). How does the quality of care compare in five countries? *Health Affairs* 23(3), 89-99.
- ICD-10. (1992a). *International statistical classification of diseases and related health problems, tenth revision, vol. 1*. Geneva: World Health Organization.
- ICD-10. (1992b). *International statistical classification of diseases and related health problems, tenth revision, vol. 2*. Geneva: World Health Organization.
- Institute of Medicine (IOM). (1999). Kohn, L. T., Corrigan, J. M., & Donaldson, M. S. (Eds.). *To err is human: Building a safer health system*. Washington, D.C.: National Academy Press.
- Institute of Medicine (IOM). (2001). *Crossing the quality chasm: A new health system for the 21st century*. Washington, DC: National Academy Press.
- Institute of Medicine (IOM). (2001). *Envisioning the National Health Care Quality Report*. Hurtado, M. P., Swift, E. K., & Corrigan, J. M. (Eds.). Washington, D.C.: National Academy Press.
- Institute of Medicine (IOM). (2002). *Unequal treatment: Confronting racial and ethnic disparities in health care*. Smedley BS, Stith AY, Nelson BD (Eds.). Washington, DC: National Academy Press.
- Institute of Medicine (IOM). (2003). *Institute of Medicine Strategic Vision 2003-2008*. [Online]. Available: <http://www.iom.edu/CMS/AboutIOM/26047.aspx> Accessed January 2008.
- Institute of Medicine (IOM). (2004). *Insuring Health—Insuring America’s health: Principles and recommendation*. Committee on the Consequences of Uninsurance, (Eds.). Washington, DC: National Academy Press.
- Institute of Medicine (IOM). (2005). *Performance measurement: Accelerating improvement*. Washington, D.C.: National Academy Press.
- Institute of Medicine (IOM). (2007). *About the Institute of Medicine: Advising the nation. Improving health*. [Online]. Available: <http://www.iom.edu/Object.File/Master/36/931/IOM-BROCHURE-FINAL.pdf> Accessed March 2008.

- Irvine, D. M., Leatt, P., Evans, M. G., & Baker, G. R. (2002). The behavioral outcomes of quality improvement teams: The role of team success and team identification. *Health Services Management Research* 13, 78-89.
- Irvine-Doran, D. M., Baker, G. R., Murray, M., Bohnen, J., Zahn, C., Sidani, S., et al. (2002). Achieving clinical improvement: An interdisciplinary intervention. *Health Care Management Review* 27(4), 42-56.
- Johnson, D.E.L. (1991). Thorough planning critical to total quality management. *Health Care Strategic Management* 9(7), 2-3.
- Johnson, P.A., Lee, T.H., Cook, E.F., Rouan, G.W., & Goldman, L. (1993). Effect of race on the presentation and management of patients with acute chest pain. *Annals of Internal Medicine* 118(8), 593-601.
- Joint Commission on the Accreditation of Healthcare Organizations (JCAHO). (1990). *Committed to quality*. Oakbrook Terrace, IL: The Joint Commission on the Accreditation of Healthcare Organizations.
- Joint Commission on the Accreditation of Healthcare Organizations (JCAHO). (1994). *Framework for improving performance: From principles to practice*. Oakbrook Terrace, IL: Joint Commission on the Accreditation of Healthcare Organizations.
- Joint Commission on the Accreditation of Healthcare Organizations (JCAHO). (2004). *Performance measurement*. [Online]. Available: <http://www.jointcommission.org/PerformanceMeasurement/PerformanceMeasurement/> Accessed January 2004.
- Joint Commission on the Accreditation of Healthcare Organizations (JCAHO). (2005). *About us: Facts about the Joint Commission on the Accreditation of Healthcare Organizations*. [Online]. Available: http://www.jointcommission.org/AboutUs/Fact_Sheets/joint_commission_facts.htm Accessed April 2006.
- Joint Commission on the Accreditation of Healthcare Organizations (JCAHO). (2006). *About us: President's message*. [Online]. Available: http://www.jointcommission.org/AboutUs/presidents_message.htm Accessed April 2006.
- Joint Commission on the Accreditation of Healthcare Organizations (JCAHO). (2007). *A comprehensive review of development and testing for national implementation of hospital core measures*. [Online]. Available: <http://www.jointcommission.org/NR/rdonlyres/48DFC95A-9C05-4A44-AB05-1769D5253014/0/AComprehensiveReviewofDevelopmentforCoreMeasures.pdf> Accessed August 2007.

- Jones, F.G. & Duncan, D.S. (1999). Regulation in the era of accountability. In A.O. Kilpatrick and J.A. Johnson (Eds.), *Handbook of health administration and policy*, (pp. 1049-1059). New York: Marcel Dekker, Inc.
- Kaplan, R.S. & Norton, D.P. (1992). The balanced scorecard—measures that drive performance. *Harvard Business Review*, 71-9.
- Kazandjian, V. A., & Lied, T. R. (1998). Cesarean section rates: Effects of participation in a performance measurement project. *Joint Commission Journal on Quality Improvement* 6(4), 201-204.
- Kazandjian, V.A. & Lied, T.R. (1999). *Healthcare performance measurement: Systems design and evaluation*. Milwaukee, WI: ASQ Quality Press.
- Kendrick, J.W. (1984). *Improving company productivity: Handbook with case studies*. Baltimore, MD: John Hopkins University Press.
- Kindleberger, C. (1983). Standards as public, collective and private goods. In *Kyklos* 36 (Fasc.3), pp. 377-396.
- Lasswell, H.D. (1936). *Politics: Who gets what, when and how*. New York: McGraw-Hill.
- Latour, B. (1986). The powers of association. In J. Law (Ed.), *Power, Action, Belief*, (pp. 264-280). London: Routledge and Kegan Paul.
- Latour, B. (1987). *Science in action: How to follow scientists and engineers through society*. Cambridge, MA: Harvard University Press.
- Latour, B. (1988). *The pasteurization of France*. Cambridge, MA: Harvard University Press.
- Latour, B. (1990). Drawing things together. In M. Lynch and S. Woolgar (Eds.), *Representation in Scientific Practice*. Cambridge, MA: MIT Press.
- Latour, B. (1991). Where are the missing masses? Sociology of a few mundane artifacts. In W.E. Bokler and J., Law (Eds.), *Shaping technology/building society: Studies in sociotechnical change*, (pp. 245-278). Mass.: MIT Press.
- Latour, B. (1992). Where are the missing masses: The sociology of mundane artifacts. In W.E. Bijerker and J. Law (Eds.), *Shaping Technology/Building Society*, (pp. 255-258). Cambridge: MIT Press.
- Latour, B. (2004). *Politics of nature: How to bring the sciences into democracy*. Cambridge: Harvard University Press.
- Latour, B. (2005). *Reassembling the social: An introduction to Actor-Network Theory*. New York: Oxford University Press.

- Latour, B. and Woolgar, S. (1986 [1979]). *Laboratory life: The construction of scientific facts*. Princeton, NJ: Princeton University Press.
- Law, J. (1992a). *Notes on the theory of the actor network: Ordering, strategy and heterogeneity*. [Online]. Available: <http://www.lanacs.ac.uk/fass/sociology/papers/law-notes-on-ant.pdf> Accessed October 2007.
- Law, J. (1992b). Notes on the theory of the actor-network: Ordering strategy and heterogeneity. *Systems Practice* 5(4), 379-93.
- Leatherman, S. & McCarthy, D. (1999). Public disclosure of health care performance reports: Experience, evidence and issues for policy. *International Journal for Quality in Health Care* 11(2), 93-105.
- Lee, J.S. (1996). CRS report for Congress: Quality of care issues in Medicare reform. Washington, DC: Congressional Research Service.
- Lesparre, M., Lovinger, G., & Poole, K. (1998). "A century of the AHA." *Hospitals & Health Networks* (January 20), 36-48.
- Litman, T.J. & Robins, L.S. (1997). *Health politics and policy*. New York: Delmar Publishers.
- Luce, J.M., Birdman, A.B. & Lee, P.R. (1994). A brief history of health care quality assessment and improvement in the United States. *West J Med* 160(3), 263-268.
- Luft, H. & Hunt, S. (1986). Evaluating individual hospital quality through outcome statistics. *Journal of the American Medical Association* 255, 2780-84.
- Lyons, A.S., & Petrucelli, R.J. (1978). *Medicine: An illustrated history*. New York: Harry N. Abrams, Inc.
- Marshall, C. & Rossman, G.B. (1999). *Designing qualitative research*. Thousand Oaks, CA: Sage Publication.
- Materna, S. & Rothe, K. (1992). A Canadian hospital implements continuous quality improvement. *Quality Progress* 25(4), 89-91.
- Matherly, L.L. & Lasater, H.A. (1992). Implementing TQM in a hospital. *Quality Progress* 25(4), 81-84.
- Mayer, D. (1992). Florida hospital. *The Healthcare Forum Journal* 35(5), 75-80.
- McGlynn, E.A. (1997). Six challenges in measuring the quality of health care. *Health Affairs* 16(3), 7-21.
- McGlynn, E.A. (2003). Introduction and overview of the conceptual framework for a national quality measurement and reporting system. *Medical Care* 41(1), I-1-I-7.

- McGlynn, E.A., Asch, S.M., Adams, J., Keeseey, J., Hicks, J., DeCristofaro, A., & Kerr, E.A. (2003). The quality of health care delivered to adults in the United States. *New England Journal of Medicine* 348(26), 2635-2645.
- McIntyre, D., Rogers, L., & Heier, E.J. (2001). Overview, history, and objective of performance measurement. *Health Care Financing Review* 22(3), 7-21.
- Meyer, M.W. (1994). Measuring performance in economic organizations. In N.J. Smelser and R. Swedberg, (Eds.), *The handbook of economic sociology*, (pp. 556-578). Princeton, N.J.: Princeton University Press.
- Meyer, M.W. (2002a). Finding performance: The new discipline in management. In A. Neely (Ed.), *Business performance measurement: Theory and practice*, (pp. 51-62). New York: Cambridge University Press.
- Meyer, M. W. (2002b). *Rethinking performance measurement: Beyond the balanced scorecard*. New York: Cambridge University Press.
- Meyer, M.W. & Zucker, L.G. (1989). *Permanently failing organizations*. Newbury Park: Sage.
- Microsoft®Encarta® Online Encyclopedia. (2007). *American Medical Association*. [Online]. Available: http://encarta.msn.com/encyclopedia_761567010/American_Medical_Association.html Accessed August 2007.
- Miller, T. & Leatherman, S. (1999). The national quality forum: A 'me-too' or a breakthrough in quality measurement and reporting? *Health Affairs* 18(6), 233-237.
- Motwani, J., Sower, V.E., & Brashier, L.W. (1996). Implementing TQM in the health care sector. *Health Care Management Review* 21(1), 73-82.
- Mueller, R.A. (1992). Implementing TQM in health care requires adaptation and innovation. *Quality Progress* 25(4), 57-60.
- Mundel, M. (1987). *Measuring total productivity of manufacturing organizations*. White Plains, NY: Unipub-Kraus International.
- Munro, R. (1999). Power and discretion: Membership work in the time of technology. *Organization* 6(3), 429-450.
- Murdock, J. (1997). Inhuman/nonhuman/human: Actor-network theory and the prospects for a nondualistic and symmetrical perspective on nature and society. *Environment and Planning D: Society and Space* 15, 731-756.
- National Business Coalition on Health. (2001). An overview of the work of the National Quality Forum (NQF). Paper presented at the National Health Leadership Council Meeting, convened July 18-20, Beaver Creek, CO.

- National Center for Health Services Research. (1977). *Controlling the cost of health care: NCHSR policy research report 1970-1977*. (DHEW Publication No. [HRA] 77-3182). Hyattsville, MD: National Center for Health Services Research.
- National Committee for Quality Assurance (NCQA). (2007). *About NCQA: Leadership team-Margaret E. O’Kane*. [Online]. Available: <http://web.ncqa.org/tabid/344/Default.aspx> Accessed August 2007.
- National Institute of Health Policy. (2002). *Issue brief—The National Quality Forum*. [Online]. Available: <http://www.nihp.org/Issue%20Briefs/Health%20Care%20Quality%20and%20NQF%20Newsletter.htm> Accessed March 2005.
- National Quality Forum (NQF). (1999). *Quality Forum: The national forum for health care quality measurement and reporting*. Washington, DC: National Quality Forum.
- National Quality Forum (NQF). (2000a). *About the National Quality Forum: History of the Quality Forum*. [Online]. Available: <http://216.122.138.39/about/history/important.asp> Accessed September 2001.
- National Quality Forum (NQF). (2000b). *NQF: Inaugural issue..* Washington, DC: National Quality Forum.
- National Quality Forum (NQF). (2001). *National Quality Forum mission*. [Online]. Available: <http://www.qualityforum.org/about/mission.asp> Accessed January 2004.
- National Quality Forum (NQF). (2001b). *NQF project brief: January 2001*. [Online]. Available: <http://www.qualityforum.org> Accessed May 2001.
- National Quality Forum (NQF). (2003). *A comprehensive framework for hospital care performance evaluation: A consensus report*. Washington, DC: National Quality Forum.
- National Quality Forum (NQF). (2004) *The National Quality Forum’s Consensus Development Process-Version 1.7*. Washington, DC: National Quality Forum.
- National Quality Forum (NQF). (2005). *Pay-for-performance programs: Guiding principles and design strategies*. Washington, DC: National Quality Forum.
- National Quality Forum (NQF). (2007). *Snapshot of Accomplishments*. [Online]. Available: <http://www.qualityforum.org/pdf/about/directors/tab3.pdf> Accessed March 2008.
- Neely, A.D. (1999). The performance measurement revolution: Why now and what next? *International Journal of Operations & Production Management* 19(2), 205-228.
- Neely, A.D. (2002). *Business performance measurement: Theory and practice*. New York: Cambridge University Press.

- Neely, A.D. & Austin, R. (2002). Measuring performance: The operations perspective. In A. Neely (Ed.), *Business performance measurement: Theory and practice*, (pp. 41-50). New York: Cambridge University Press.
- Nicollier-Fahrni, A., Vader, J.P., Froehlich, F., Gonvers, J.J., & Burnand, B. (2003). Development of appropriateness criteria for colonoscopy: Comparison between a standardized expert panel and an evidence-based medicine approach. *International Journal for Quality in Health Care* 15 (1), 15-22.
- Norlund, S. (1991). Implementing total quality management programs in health care organizations. *Hospital Material Management Quarterly* 12(4), 22-26.
- Office of Civil Rights (OCR). (2003). *OCR Privacy Brief: Summary of the HIPAA privacy rule*. [Online]. Available: <http://72.14.205.104/search?q=cache:0--LNcX1fHgJ:www.hhs.gov/ocr/privacysummary.pdf+hipaa&hl=en&ct=clnk&cd=2&gl=us> Accessed June 2008.
- Office of the Vice President. (1998). Vice President releases HHS report documenting problems and ways to improve quality health care. [Online]. Available: <http://clinton6.nara.gov/1998/06/1998-06-17-vp-releases-hhs-report-on-improving-health-care.html> Accessed July 2008.
- Office of Management and Budget (OMB). (1998). *OMB circular A-119*. [Online]. Available: <http://www.whitehouse.gov/omb/circulars/a119/a119.html> Access June 2008.
- Omoigui, N., Miller, D., Brown, K., et al. (1996). Outmigration for coronary bypass surgery in an era of public dissemination of clinical outcomes. *Circulation* 93, 27-33.
- Ostrow, P.C. (1983). The historical precedents for quality assurance in health care. *The American Journal of Occupational Therapy* 37(1), 23-26.
- Pacific Business Group on Health. (2003). *California cooperative on healthcare reporting initiative*. [Online]. Available: <http://www.pbgh.org/programs/cchri/default.asp> Accessed January 2004.
- Patton, M.Q. (1987). *How to use qualitative methods in evaluation*. Newbury Park, CA: Sage Publications.
- Patton, M.Q. (1998). *Enhancing the quality and credibility of qualitative analysis*. Presented at the Qualitative Methods in Health Services Research Conference convened in December, Rockville, MD.
- Pels, D. (2003). *Unhastening science: Autonomy and reflexivity in the social theory of knowledge*. Liverpool: Liverpool University Press.
- Petersen, L.A., Wright, S.M., Peterson, E.D., & Daley, J. (2002). Impact of race on cardiac care and outcomes in veterans with acute myocardial infarction. *Medical Care* 40(Suppl. 1), 86-96.

- Phibbs, C.S., Bronstein, J.M., Buxton, E., & Phibbs, R.H. (1996). The effects of patient volume and level of care at the hospital of birth on neonatal mortality. *Journal of the American Medical Association* 276, 1054-9.
- Phillips, D.P., Christenfeld, N., & Glynn, L.M. (1998). Increase in U.S. medication error deaths between 1983 and 1993. *Lancet* 351(9103), 643-644.
- President's Advisory Commission on Consumer Protection and Quality in the Health Care Industry. (1998). *Quality first: Better healthcare for all Americans*. Washington, D.C.: U.S. Government Printing Office.
- Prout, A. (1996). Actor-network theory, technology and medical sociology: An illustrative analysis of the metered dose inhaler. *Sociology of Health & Illness* 18(2), 198-219.
- Quality Interagency Coordination Task Force (QuIC). (2001). *QuIC fact sheet*. [Online]. Available: <http://www.quic.gov/about/quicfact.htm> Accessed December 2003.
- Reitz, T. (1993). Total quality improvement. *Hospital Material Management Quarterly* 15(2), 32-37.
- Reinhardt, U.E., Hussey, P.S., & Anderson, G.F.. (2002). Cross-national comparisons of health systems using OECD data, 1999. *Health Affairs* 21(3), 169-181.
- Reinhardt, U.E., Hussey, P.S., & Anderson, G.F. (2004). U.S. health care spending in an international context. *Health Affairs* 23(3), 10-25.
- Roberts, J., Coale, J., & Redman, R. (1987). A history of the Joint Commission on Accreditation of Hospitals. *JAMA* 258, 936-940.
- Rhodes, R.P. (1992). *Health Care: Politics, policy, and distributive justice*. Albany, NY: State University of New York Press.
- Roper, W. (1997). Transcription of Meeting. Testimony at the Advisory Commission on Consumer Protection and Quality in the Health Care Industry. Washington, D.C., October 21.
- Rousseau, D.M. (1997). Organizational behavior in the new organizational era. *Annual Review of Psychology* 48, 515-46.
- Sabatier, P.A. (1999). Fostering the development of policy theory. In P.A. Sabatier (Ed.), *Theories of the policy process*, (pp. 261-275). Boulder, CO: Westview Press.
- Scarlett, E.P. (1991). What is a profession? In R. Richard and J. Stone (Eds.). *On doctoring*, (119-133). New York: Simon & Schuster.
- Schuster, M.A., Asch, S.M., McGlynn, E.A., Kerr, E.A., Hardy, A.M., & Gifford, D.S. (1997). Development of a quality of care measurement system for children and adolescents:

- Methodological considerations and comparisons with a system for adult women. *Archives of Pediatrics & Adolescent Medicine* 151(11), 1085-1093.
- Schuster, M.A., Reifel, J.L., & McGuigan, K. (1998). *Assessment of the quality of cancer care: A review for the National Cancer Policy Board of the Institute of Medicine*. Washington, DC: National Academy Press.
- Sink, D.S. (1985). *Productivity measurement-Planning, measurement and evaluation control and improvement*. Chichester: John Wiley and Sons.
- Scott, W. R. (1991). Unpacking institutional arguments. In W.W. Powell and P.J. DiMaggio (Eds.), *The new institutionalism in organizational analysis*, (pp. 164-182). Chicago: The University of Chicago Press.
- Scott, W.R. (2001). *Institutions and organizations, 2nd edition*. Thousand Oaks, CA: Sage Publications, Inc.
- Scott, W.R. & Cole, R.E. (2000). Introduction: The quality movement and organization theory. In R.E. Cole and W.R. Scott (Eds.), *The quality movement and organizational theory*, (xiii-xxix), Thousand Oaks, CA: Sage Publications.
- Scott, W. R., & Meyer, J. W. (1983). The organization of societal sectors. In J. W. Meyer and W. R. Scott (Eds.), *Organizational environments: Ritual and rationality*, (pp. 129-153). Beverly Hills, CA: Sage.
- Shortliffe, E.H., Perreault, L.E., Wiederhold, G., & Fagan, L.M. (2001). *Medical informatics: Computer applications in healthcare and biomedicine*. New York: Springer-Verlag.
- Sidorova, A. & Sarker, S. (2000). Unearthing some causes of BPR failure: An actor-network theory perspective. In H.M. Chung (Ed.), *AMCIS proceedings*. Long Beach, CA.
- Singleton, V. & Michael, M. (1993). Actor-networks and ambivalence: General practitioners in the UK cervical screening programme. *Social Studies of Science* 23(2), 227-264.
- Smith, D.B. & Kaluzny, A.D. (1986). *The white labyrinth: A guide to the health care system*. Ann Arbor, MI: Health Administration Press.
- Smith, R. (1991). Where is the wisdom...? The poverty of medical evidence. *British Medical Journal* 303, 798-799.
- Somerville, I. (1999). Agency versus identity: Actor-network theory meets public relations. *Corporate Communications. Bradford* 4(1), 6-12.
- Soumerai, S.B., McLaughlin, T.J., Spiegelman, D., et al. (1997). Adverse outcomes of underuse of B-blockers in elderly survivors of acute myocardial infarction. *Journal of the American Medical Association* 277, 115-21.

- Sprague, L. (2002). "Contracting for Quality: Medicare's Quality Improvement Organizations." *National Health Policy Forum Brief 774*, 1-15.
- Stake, R.E. (1995). *The art of case study research*. Thousand Oaks, CA: Sage Publications.
- Stake, R.E. (1998). Case Studies. In N.K. Denzin & Y.S. Lincoln (Eds.), *Strategies of qualitative inquiry*, (pp. 86-109). Thousand Oaks, CA: Sage Publications.
- Stone, V., Seage, G., Hertz, T., & Epstein, A. (1992). The relation between hospital experience and mortality for patients with AIDS. *Journal of the American Medical Association* 266, 2655-61.
- Tang, N., Eisenberg, J.M., & Meyer, G.S. (2004). The role of government in improving health care quality and safety. *Joint Commission Journal on Quality and Safety* 30(1), 47-55.
- Testa, M.A., & Simpson, D.C. (1994). Assessment of quality of life outcomes. *New England Journal of Medicine* 334, 835-40.
- The Commonwealth Fund Commission on a High Performance Health System. (2006). *Why not the best? Results from a National Scorecard on U.S. Health System Performance*. [Online]. Available: http://www.commonwealthfund.org/publications/publications_show.htm?doc_id=401577 Accessed February 2008.
- The Consumer-Purchaser Disclosure Project. (2001). *History and mission*. [Online]. Available: <http://healthcaredisclosure.org/about/> Accessed January 2008.
- The Consumer-Purchaser Disclosure Project. (2004). *The state experience in health quality data collection*. [Online]. Available: www.healthcaredisclosure.org Accessed May 2005.
- The Consumer-Purchaser Disclosure Project. (2007). *Consumers, healthcare purchasers applaud agreement on physician performance reporting*. [Online]. Available: <http://healthcaredisclosure.org/docs/files/NYAGAgreementDisclosureRelease102907.pdf> Accessed January 2008.
- Thompson, J.B. (1990). *Ideology and modern culture: Critical social theory in the era of mass communication*. Cambridge, U.K.: Polity Press/Blackwell Publishers, Ltd.
- Thompson, G.F. (2003). *Between hierarchies & markets: The logic and limits of network forms of organization*. New York, NY: Oxford University Press.
- Tolbert & Zucker. (1999). The institutionalization of institutional theory. In S.R. Clegg & C. Hardy (Eds.), *Studying organization: Theory & method*, (pp. 169-184). Thousand Oaks, CA: Sage Publications.
- Udvarhelyi, I.S. (1998). Transcript of Meeting. Testimony at the Advisory Commission on Consumer Protection and Quality in the Health Care Industry. Washington, DC.

- Underwood, J. (1998). *Not another methodology: What ANT tells us about systems development*. Presented at the 6th International Conference on Information Systems Methodologies, British Computer Society, Salford, UK. [Online]. Available: <http://www-staff.mcs.uts.edu.au/~jim/papers/ismeth.htm> Accessed October 2007.
- United States General Accounting Office. (1994). *Health care reform: "Report cards" are useful but significant issues need to be addressed*. (GAO publication No. GAO/HEHS-94-219). Washington, DC: U.S. GAO.
- Walsham, G. (1997). Actor-network theory and IS research: Current status and future prospects. In A.S. Lee, J. Liebenau, and J.I DeGross (Eds.), *Information systems and qualitative research*. London: Chapman & Hall.
- Waggoner, D.B., Neely, A.D., & Kennerley, M.P. (1999). The forces that shape organizational performance measurement systems: An interdisciplinary review. *International Journal of Production Economics* 60-61, 53-60.
- Wennberg, J.E. (1987). The paradox of appropriate care. *Journal of the American Medical Association* 258(18), 2568-9.
- Wennberg, J.E., Barnes, B.A. & Zubkoff, M. (1982). Professional uncertainty and the problem of supplier-induced demand. *Social Science and Medicine* 16, 811-24.
- Wennberg, J.E. & Cooper, M.M. (1998). *The Dartmouth atlas of health care in the United States*. Chicago, IL: American Hospital Publishing, Inc.
- Wilson, J. (2006). Health insurance portability and accountability act privacy rule causes ongoing concerns among clinicians and researchers. *Ann Intern Med* 145(4), 313-316.
- White, H.C. (1992). *Identity and control: A structural theory of social action*. Princeton, NJ: Princeton University Press.
- Williams, J.F., Zimmerman, J.E., Wagner, D.P., Hawkins, M., & Knaus, W.A. (1995). African-American and white patients admitted to the intensive care unit: Is there a difference in therapy and outcome? *Critical Care Medicine* 23(4), 626-636.
- Woolgar, S. (1988). *Science: The very idea*. London, U.K.: Tavistock.
- World Health Organization. (2000). *The world health report 2000—health systems: Improving performance*. Geneva, Switzerland: World Health Organization, International.
- Yin, R.K. (1989). *Case study research: Design and methods*. Thousand Oaks, CA: Sage Publications.
- Yin, R.K. (1994). *Case study research: Design and methods*. Thousand Oaks, CA: Sage Publications.

Appendices

Appendix B: A Glossary of Actor-Network Theory Key Elements*

Key Element	Definition
Actor-network	“Heterogeneous network of aligned interests”
Enrolment	“A situation when actors accept interests defined for them by the focal actor.”
Immutable Mobile	These are entities like maps, photographs, paintings, textual descriptions, and all kinds of images but can also include devices such as ships, cloths, or scientific instruments (Thompson, 2003). These entities are “fixed” in that they remain stable, but can also be “mobile” in that they can be “rearranged and reconfigured through the network of places and agencies to which they are attached or through which they operate” (Thompson, 2003: 73). Immutable mobiles have the combined properties of mobility, stability, and combinability (Latour, 1987).
Inscription	“A process of creation of technical artifacts that would ensure the protection of certain interests.” (Latour, 1992)
Interessement	“A process of convincing actors to accept definition of the focal actor.” (Callon, 1986) It is the manner by which the focal actor tries to impose and stabilize other actor identities. It is a process of convincing the other actors to accept the identities and issues as defined by the focal actor as well as reinforce the resolve to move through the obligatory passage point to address the problem.
Irreversibility	“Degree to which it is subsequently impossible to return to a point where alternative possibilities exist.” (Walsham, 1997)
Mobilisation	The process by which the enrolled actors become mobilized to address the problem identified by the focal actor in the initial phase of translation i.e., Problematization. (Murdoch, 1997)
Obligatory Point of Passage	“A situation that has to occur for all of the actors to be able to achieve their interests, as defined by the focal actor.” It can be defined as an unavoidable point or a narrative bottleneck through which actors, which the focal actor wants to include in the network, must pass in order to continue to exist and develop and to articulate their identity and reason for being (Singleton and Michael, 1993: 229-230).

Problematisation	“The first moment of translation during which a focal actor defines identities and interests of other actors that are consistent with its own interests, and establishes itself as an obligatory passage point, thus rendering itself indispensable.” (Callon, 1986)
Punctualisation	“Involves ‘the substitution of a network by a point’ (Law, 1992b: 385). This point or node represents a complex network of ordered actors and things that is often not readily visible.”
Translation	Is defined as the struggle to create an ordered network “which generates ordering effects such as devices, agents, institutions, or organizations” (Law, 1992a). Translation is a process—a “multifaceted interaction in which actors 1) construct common definitions and meanings; 2) define representatives; and 3) co-opt each other in the pursuit of individual and collective objectives” (Bardini, 1997: 20). It involves showing how actors with differing interests become aligned.

*This table is adapted from the conference proceedings of Sidorova and Sarker (2000) and from Atkinson (2002).

Appendix C: Comparison of Selected Versions of the NQF CDP

CDP Versions (Date)					
	Ver.1.2 (11/2000)	Ver. 1.3 (2/2001)	Ver 1.5 (1/2002)	Ver 1.6(9/2003)	Ver 1.7 (8/2004)
Principals of the NQF CDP	Project conceptualization & priority setting Purpose of NQF standards project Establishing program priorities	Project conceptualization & priority setting Purpose of NQF standards projects Establishing program priorities	Project conceptualization & priority setting Purpose of NQF standards projects Establishing program priorities	Consensus standard development Program officer Steering Committees/Review Committees Technical Advisory Panels Strategic Advisory Council Evidence Basis Draft Recommendations Availability of Project info	Consensus standard development Program officer Steering Committees/Review Committees Technical Advisory Panels Strategic Advisory Council Evidence Basis Draft Recommendations Availability of Project information
	Project planning Nomination of project topics Project review Notification of projects Project funding	Project planning Nomination of project topics Project review Notification of projects Project Funding	Project planning Nomination of project topics Project Review Notification of projects Project funding	Review Pre-Voting Member Review General public review Consideration of Member and other comments Expedited Consideration	Review Pre-voting member reviews General public review Consideration of member and other comments Expedited consideration

	<p>Product development Program officer Steering Committees Technical Advisory Panels Strategic Framework Board Input Evidence basis Draft Recommendations Availability of project info</p>	<p>Product development Program officer Steering Committees Technical Advisory Panels Strategic Framework Board input Evidence basis Draft Recommendations Availability of project info</p>	<p>Product development Program Officer Steering Committees Technical Advisory Panels Strategic Advisory Council input Evidence Basis Draft recommendations Availability of project info</p>	<p>Member council approval Member Voting Votes with Proposed Modifications and/or conditions Approval by Member Councils Second Round of Member voting</p>	<p>Member Council Approval Member voting Votes with proposed modifications and/or conditions Approval by Member councils Second Round of Member voting</p>
	<p>Product review NQF member review General public review Consideration of member and other comments</p>	<p>Product review NQF member review General public review Consideration of member and other comments</p>	<p>Product review NQF member review prior to voting General public review Consideration of member and other comments</p>	<p>Board of Directors Endorsement Board of Directors Endorsement of Voluntary Consensus Standards Notification of Board Decisions Appeal of Board Endorsement</p>	<p>Board of Directors Endorsement Board of Directors endorsement of voluntary consensus standards Notification of Board decisions Appeal of Board Endorsement</p>

	Product endorsement or approval Approval by member councils Approval by Board of Directors Notification of decisions Appeals	Product endorsement or approval Approval by member councils Approval by Board of Directors Notification of decisions Appeals	Product endorsement or approval Specific process for ballot delivery and counting of votes Votes with proposed modifications and/or conditions Approval by Member Councils Second Round of Member Council Voting Approval by Board of Directors Notification of decisions Appeals	Evaluation Operationalization of NQF-endorsed Voluntary Consensus Standards Evaluation Mechanisms Standing Working Groups NQF Staff Routine Review	Evaluation Operationalization of NQF-endorsed Voluntary Consensus Standards Evaluation Mechanisms Standing working groups NQF staff Routine Review
--	--	--	--	---	---

Appendix D: Environmental Assessment of Health Care in 2001

Societal Factors

In 2001, the average age of the population continued to rise, which, in turn, increased the prevalence of certain diseases and conditions, such as chronic illness and disability. This has increased the demand for services and their related costs. Further, the population has become multiethnic at a faster rate than expected, which places extra demands on the health care delivery system, in terms of the effects of cultural and linguistic factors that may affect care quality. In addition, the American economy slowed, which caused employers, governments, and others to pay more attention to controlling costs, including the cost of health care.

Health Care System Factors

The overall landscape of the health care industry continues to be reconfigured by consolidations, mergers, dissolution of mergers, and closures. Care delivery continues to migrate to the ambulatory care and home care settings, thus reducing inpatient services, as well as the reduction of oversight by either the government or accrediting bodies. The cost of care is increasing, thereby increasing the insurance premiums charged to purchasers, who pass on the additional costs to employees. A major factor in the escalation of health care costs involves growing expenditures on pharmaceuticals, which are principally from the introduction of new drugs and the expanded use of existing drugs. Managed care plan enrollment continues to increase, but there was a shift from health maintenance organizations (HMOs) to preferred provider organizations (PPOs). The number of uninsured remained steady; however, due to the weakened U.S. economy, it is anticipated that the percentage will increase. Rapid introduction of new science and technology into health care raises certain issues including safety, economic and ethical concerns, and continuity of care issues. Further, the practice of applying

management concepts and tools continues to grow, in part, due to the interest of purchasers. The adequacy of nurse staffing is an issue, and to a lesser degree, staffing of other disciplines that work in the health care industry. Staffing issues, the shift to ambulatory/outpatient care, and financial pressures to eliminate unused capacity have decreased hospitals' ability to respond to changes in the health care system. There was concern over the risk of bioterrorism, as well as other potential catastrophes, and at the time, no one organization or system was in place to deal with these situations. The health care organizations were considered to be unprepared, because of their limited capacity to respond to emergencies involving multiple casualties.

In the area of performance measurement in health care, the availability of measures continues unabated, without much understanding about their relevance, reliability, or meaning. The Agency for Healthcare Research and Quality (AHRQ) supports performance measurement through its various initiatives, i.e., Computerized Needs-Oriented Quality Measurement Evaluation System (CONQUEST),²⁴⁰ the Healthcare Cost and Utilization Project (HCUP),²⁴¹ and the National Guidelines Clearinghouse (NGC).²⁴² The Health Care Financing Administration (HCFA, now known as CMS) is also supportive of performance measurement, and engages in several activities, such as the Peer Review Organizations and their 6th Scope of Work; the Minimum Data Set,²⁴³ and the Collection of Outcome and Assessment Information Set (OASIS).²⁴⁴

²⁴⁰ CONQUEST is a quality improvement software tool that can be used to choose and use measures.

²⁴¹ HCUP is a database that consists of administrative longitudinal databases. Derived from the databases are quality measures known as Quality Indicators (QIs) that are included under the HCUP project.

²⁴² The NGC is an Internet website that makes evidence-based clinical practice guidelines available.

²⁴³ This is a federally mandated process for clinical assessment of all residents in Medicare or Medicaid certified nursing homes.

²⁴⁴ This is a data set that is required for all home health agencies providing skilled services to Medicare and Medicaid patients.

The National Quality Forum (NQF) is yet another organization focused on measuring health care quality and reporting, in an effort to improve quality. Its efforts consist of three steering committees to address various performance measurement issues: identify a list of serious, usually preventable, adverse events; identify a compendium of safe practices; and identify hospital quality performance measures.

Other efforts of performance measurement focus on the area of behavioral health and the identification of core measures. Activities in this area were sponsored by the American College of Mental Health Administrators, the National Association of Psychiatric Health Systems, and the federal Substance Abuse and Mental Health Services Administration (SAMHSA). An additional area focused on the development and testing of performance measures for hospices, which was sponsored by the National Hospice and Palliative Care Organization and the National Hospice Workgroup.

A final factor, as emphasized by the Institute of Medicine's report, *Crossing the Quality Chasm*, is the use—or, rather lack of use—of information technology in health care. Health care system evaluation and quality measurement are stymied by the lack of an established information infrastructure. Increasingly, practitioners use information technology, by looking up information on best practices online, in their daily clinical work. Telehealth is also providing access to patient care, especially in rural areas. Barriers to advances in this area include the lack of a consensus standard for electronic communications and transactions in health care, and the initial investment required to introduce new systems or to replace aging ones. Additional concerns about using information technology center on privacy and confidentiality of health care patient-specific information. The Federal Health Insurance Portability and Accountability Act of 1996 (HIPAA) sets forth requirements for security of electronic records and their transmission. This

act applies to all provider organizations, as well as practitioners in health care who collect, maintain, and/or transmit patient-identifiable health information.

Purchaser Factors

Health care purchasing by most employers is based on price, and not on quality of services. However, some purchasers are requiring greater accountability in terms of quality, safety, and efficiency of the health care industry. The Leapfrog Group, a coalition of major employers, encourages the use of certain hospital structural characteristics that could improve patient safety. Thus, leveraging the employers' purchasing power motivates the hospitals to adopt the Leapfrog Group's recommendations. The Big Three automakers are supporting the use of ISO 9000 concepts and/or certification in health care. Others, such as the Pacific Business Groups on Health, the California Public Employees Retirement System (CalPER), and the state of Missouri's Consolidated Plan use competition and financial incentives for organizations to voluntarily submit performance data. General Motors and the Alliance in Colorado have gone a step further, and base financial incentives on aggregated costs and quality scores of organizations. Employers, in competition for workers, have expanded health coverage, but increasingly are requiring more cost sharing by the employee. These trends have created a need to provide comparative information about health care facilities and practitioners, to assist employees in making decisions about their health care.

Consumer Factors

More and more consumerism is affecting the health care system. Consumers are increasingly more engaged in their health care and tend to hold providers accountable for the quality of care they receive. They have more access to information and data from the Internet, and thus are able to be more informed about available treatments and therapies. Patient's rights

legislation has been proposed and/or enacted at the federal and state levels during this time. In addition, media attention, as well as the attention of public advocates, has grown about safety within the health care system. Public concern about the issue crystallized with the publication, in 1999, of the Institute of Medicine's report, *To Err is Human: Building a Safer Health System*. Patient safety and health care errors appear to resonate with the public and cause obvious concern among consumers.

It was estimated that over 60 million people, usually college-educated, affluent women, use the Internet to find information about health care and health care organizations either for themselves or someone else. Information on-line may be available, but is sometimes incomplete and inaccurate, and very often does not meet consumer needs. Information-seeking consumers, in increasing numbers, believe in being actively involved in their care, and search for credible information about clinical care. They agree that clinical performance is central to quality, and is an important selection driver. Clearly presented information about health care issues is embraced by this group of consumers, indicating that this information would influence their selection of health care providers.

Federal Factors

The federal government is becoming increasingly involved in the oversight of health care quality. CMS (formerly the Health Care Financing Administration or HCFA) has morphed from a bill-payer to a "prudent purchaser," in an effort to focus on the value received for dollars paid for health care services, and accordingly used the Peer Review Organizations (PROs) as its quality improvement agents.

The climate on Capitol Hill for advancing policy and legislation was significantly different from the prior year. With a new Republican-controlled White House and Congress,

specific ideas about the role of government, the need to work in a bipartisan manner, and the projected budget surplus were relevant to efforts currently underway in health care. The Senate was evenly divided between Republicans and Democrats, with “control” exerted by the Republican Party through the Vice-President’s ability to break tie votes. The Republican House of Representatives, no longer facing the threat of constant veto from the President, as well as the evenly divided Senate, provided the opportunity for both parties to accomplish their goals. A substantial number of Congressional committees and subcommittees in Congress had a new chairperson, as well as a number of new member appointees, who were generally unfamiliar with health care. And accordingly, health care policy developments depended upon how large the projected budget surplus was, and how much of the surplus would be earmarked for tax cuts.

During this year, the White House and Congressional leaders have expressed concern for a number of health care issues, including the uninsured, the patient’s bill of rights, regulatory relief, medical information privacy, Medicare reform, patient safety, nurse staffing, the National Institutes of Health, and organ donation. Each of these is discussed below.

- The Uninsured—there were incremental approaches being framed to address the problem posed by the almost 44,000 Americans who were uninsured in 2001. Some elements under consideration included a refundable tax credit, expansion of the Medicaid and SCHIP programs, and an increase in the number of community health centers over the next few years.
- Patient Bill of Rights—This effort emerged prior to 2001, but had good prospects for closure this year. The White House was interested in protecting patients, especially those in managed care plans, but was cautious because of the potential to increase health plan malpractice liability.

- **Regulatory Relief**—In general, Republicans are enthusiastic about reducing regulatory burdens. At that time, Congressional committees were in the process of gathering a list of candidate items that could be considered for relief. Other efforts were directed at identifying actions that could potentially be realized by Executive Order.
- **Medical Information Privacy**—The government sought to finalize the HIPAA privacy rule, which surprised many in the health care industry. As a result of 24,000 comments on the rule, Secretary Thompson pledged to offer some revisions in May.
- **Medicare Reform**—A prescription drug benefit was being discussed, but there was no plan under construction at the time because of the inability to construct a benefit that would limit financial exposure. The President indicated that the drug benefit and Medicare reform should be linked, and as a result, an entire review of the Medicare program was up for discussion.
- **Patient Safety**—Of increasing concern to the Executive Branch, it has not yet achieved center stage. The Secretary of HHS proposed a safety network that could accumulate and analyze error information between the public and private sectors. In addition, the Senate was in debate about a proposed bipartisan bill that would establish a quality improvement program for safety, in which participation would be voluntary and confidential. Further, the information would be shared, in a non-identifiable form with AHRQ for further analysis which could then be reported yearly to Congress.

- Nurse Staffing—In light of concerns about quality and safety in health care facilities, many in Congress were attentive to reported nurse shortages. Some ways to address the shortage that were under discussion at the time included scholarship programs, incentives to promote retention of nurses, and nurse staffing ratios.
- National Institutes of Health—On the table was a proposed 13.6 % budget increase over last year’s budget for NIH. This increase was in sharp difference to the proposed 3.6 % budget increase for other research and development efforts of federal agencies.
- Organ Donation—The Republican Administration allocated \$20 million to enhance efforts surrounding organ donation awareness, specifically targeting the workplace but also building awareness in several areas, such as the medical, legal and clerical communities. In addition, Congress was considering legislation that would support living donor efforts through a bill that would provide transportation and reimbursement of related expenses to those who donated their organs.

The Administration is expected to continue financing the Quality Interagency Coordination Task Force (QuIC), whose mission is to bring federal agencies together to coordinate quality improvement and safety activities within the federal sector. The chair, AHRQ Director John Eisenberg, expects a further budget increase (approximately 13.5 %) for patient safety and quality of care activities, in addition to funds allocated for the QuIC. One proposed bill in Congress would establish and fund a center for patient safety within AHRQ. Other activities of the Agency include patient safety demonstration projects, the creation of patient

safety centers of excellence, a national conference on medical errors, and quests for an effective way to translate research into practice. In addition, the AHRQ will begin work on the National Quality Reports, using the IOM reports (*Crossing the Quality Chasm* and *Envisioning the National Health Care Quality Report*) as a template for reporting on the quality and disparities in health care to Congress. The database that will supply the information for these reports will be expanded beyond the Medical Expenditure Survey (MEPS) to include more clinical and condition-specific data, as well as to enhance the information on sub-populations, such as children, the elderly, rural inhabitants, and Hispanics.

The CMS (formerly known as HCFA) is under the new leadership of Thomas Scully, who headed the Federation of American Hospitals, a membership organization of for-profit community hospitals. A large portion of the pressure for change will come from the Secretary of Health and Human Services, who expects more from the Medicaid program, and a Congress that wants to reform the Medicare program. What is expected in the coming year is attention to Medicare contractor concerns, oversight of nursing homes and other related health care facilities, data system improvement, and determination of Medicare/Medicaid coverage.

The Medicare Payment Advisory Commission (MedPac), which advises Congress on Medicare issues, will publish a report in December 2001 that will evaluate the appropriateness of standards for Medicare+Choice and fee-for-service providers. Their June 2001 report will focus on quality of care recommendations for rural areas, as specific concerns have surfaced in relation to the lack of providers available in these areas and the financial stresses of rural living.

Also underway is another report from the IOM centered on federal oversight, improvement, and research programs. Issues raised by the committee responsible for the report include whether there should be more waiver authority under Medicare to permit pay for

performance, the need to re-evaluate data collection requirements due to the fact it takes time away from resident care, and the lack of continuity of care being experienced during transitions to other institutions or home by many of the elderly.

State Factors

Many states believe that they should take a more aggressive role in addressing health care errors, and there is a trend toward expanding state quality and safety oversight activities. State agencies currently evaluate Medicare and Medicaid organizations through an agreement with CMS. In addition, states also assess compliance with state licensing regulations for health care facilities. Currently in 2001, there are 15 states reporting adverse events however, it is widely acknowledged that there is substantial under-reporting of these events to state agencies.

In the prior year, 34 legislative bills related to staffing effectiveness in hospital settings were introduced in 16 states. As of May 2001, approximately 12 new staffing related bills have been introduced in five states. Many of the bills introduced require the development of specific staffing ratios.

Several states have data collection efforts underway that are used to either support public report cards or to stimulate quality improvement efforts. These efforts report physician and hospital-specific quality data, including mortality rates related to coronary artery bypass graft surgery, as well as other conditions and procedures.

Further, State Insurance Commissioners are more active in the regulation of managed care plans that surpass the scope of financial solvency. The National Association of Insurance Commissioners developed a regulation model that focuses efforts on quality improvement, patient rights, and patient confidentiality.

International Factors

Health care internationally parallels that of the United States, and in varying degrees, what has surfaced abroad is the need for objective quality evaluation mechanisms for health care organizations. In other countries, the quality of care and readiness for external evaluation are at modest levels, compared to the U.S. Yet interest in international standards is increasing so much that there is a relatively high interest in developing country-specific accreditation programs, particularly in the Middle East, Latin America, and Western Europe. Many European countries are assessing the various evaluation mechanisms such as the ISO 9000, the European Foundation for Quality Management Award, the Malcolm Baldrige Quality Award, and the Joint Commission International (JCI) accreditation in health care. Further, economic globalization is likely to step up health care performance expectations in organizations around the world, which in turn would intensify the demand for quality evaluation mechanisms. To meet this heightened demand, the ISO 9000, the established accrediting bodies in the U.S. and other countries, are prepared to meet the needs for direct quality evaluation of organizations and provide support in the development of country-specific accreditation programs. However, deeming relationships and other types of reliance of government, purchasers, and others on accreditation as an *imprimatur* of quality is essentially non-existent in other countries, which may hamper the growth of new international accreditation ventures overseas.

Appendix E: Hospital Steering Committee Work Groups

Priorities and Scope	Implementation and Improvement	Reporting
Stuart Baker, M.D. (Chair) VHA, Inc. Irving, TX	Cary Sennett, M.D., Ph.D. (Chair) American College of Cardiology Bethesda, MD	Diane L. Bechel, Dr.P.H. (Chair) Ford Motor Company Dearborn, MI
Ronald R. Blanck, D.O. American Osteopathic Association University of North Texas Health Science Center at Fort Worth, Texas College of Osteopathic Medicine Fort Worth, TX	Dawn M. FitzGerald, M.S. Mid-South Foundation for Medical Care, Inc. Memphis, TN	Jennifer Eames California HealthCare Foundation Oakland, CA
Mary Foley, M.S., R.N. American Nurses Association Washington, DC	Judy Levy, R.N., CPHQ Tufts Health Plan Watertown, MA	Ted von Glahn Pacific Business Group on Health San Francisco, CA
William E. Golden, M.D. American health Quality Association/University of Arkansas, Medical Sciences Little Rock, AR	Samuel R. Nussbaum, M.D. Anthem, Inc. Indianapolis, IN	Gina Rocha, R.N., M.P.H. Rhode Island Department of Health Providence, RI
Jerod M. Loeb, Ph.D. Joint Commission on the Accreditation of Healthcare Organizations Oakbrook Terrace, IL	Louise Probst Gateway Purchasers Coalition for Health St. Louis, MO	Ellen Severoni, R.N. California Health Decisions Orange, CA
Arnold Milstein, M.D. Pacific Business Group on Health San Francisco, CA	Bernard Rosof, M.D., F.A.C.P. North Shore Long Island Jewish Health System Great Neck, NY	Marina L. Weiss, Ph.D. March of Dimes Washington, DC
Donald Nielsen, M.D. American Hospital Association Chicago, IL	Shelly Voelz, R.N., BSN St. Francis Hospital and Health Centers Beech Grove, IN	Robert Weiser Medical Review of North Carolina, Inc. Cary, NC

<p>Irwin Press, Ph.D. Press, Ganey Associates South Bend, IN</p>	<p>Liaison Member Gregg Meyer, M.D. Agency for Healthcare Research and Quality Rockville, MD</p>	
<p>Liaison Member Steven B. Clauser, Ph.D. Centers for Medicare and Medicaid Services Baltimore, MD</p>		

Appendix F: Milestones for the NQF Hospital Performance Measures Project

1998		
	RI/QIO projects began—field test for selected hospital measures (AMI, HF, Pneumonia)	QIO funded by CMS; The Joint Commission measures used by RI
1999		
Dec 7	Announcement of Hospital Project by NQF	
2000		
Pre-contract	Established Hospital Performance Measures Steering Committee	
	RI/QIO projects completed; CMS/The Joint Commission begin meetings to align measures for submission to NQF Hospital Project	CMS & The Joint Commission measures were of the same condition, but specs were not similar
Dec 11	Start date of AHRQ contract (with funds from CMS)	
2001		
Mar 21	Steering Committee meeting	
June 21	Project Director hired	
May 22	Steering Committee meeting	
August	Three Working Groups established	
Oct 24	Steering Committee meeting	
Nov	Launch of DHHS Quality Initiatives	
2002		
Feb 26	Steering Committee meeting	Final selection of measures
Apr 18	Draft report (Group 1 measures) review period begins (32 measures)	
May 16/23	Public/member comments due	
June 14	Steering Committee meeting	Discussion of handling measures that comes in from outside the established process.
July	Group 1 report voting period begins (37 measures)	

July 15	Consumer-Purchaser presentation to the Board of Directors	
August	CMS gets agreement w/ national hospital associations	On voluntarily report hospital performance
Aug 20	Group 1 ballots due	
Aug 27	Steering Committee meeting	
Oct 1	NQF annual meeting; discussion of hospital measures and framework	
Oct 2	Board of Directors approval of 31 of 37 Group 1 measures	
Oct 4	Draft report (Hospital measurement framework) review period begins	
Oct 17-18	Workshop on Group 2 measures	Evaluate the 6 measure forwarded by DOD; and the 28 from the comment period. WG recommended 17 measures for potential endorsement.
Oct 31/Nov 7	Public/member comments due on framework	
Nov 1	Draft report (Group 2 measures) review period begins (17 measures)	
Nov 26/ Dec 3	Public/member comments due on Group 2 measures	
Dec	HQA is formed.	
Dec 18	Hospital framework report voting period begins	
2003		
Jan 27	Group 2 and Framework ballots due	
Jan 29	Board of Directors approved Framework and 8 of 17 Group 2 measures	
Feb 26	Second vote on remaining 5 measures begins	
Feb 28	Final report on Task 1 delivered to AHRQ	
Dec	MMA Signed into law.	Provisions included the full market basket update for hospital that reported on the HQA ten measure starter set.

Appendix G: Participants of the Hospital Measures Steering Committee

Participant*	Organization	Council Represented
Cary Sennett, M.D., Ph.D. (co-chair)	American College of Cardiology Bethesda, MD	Provider and Health Plan
Mary Wakefield, R.N., PhD (co-chair)	University of North Dakota Grand Forks, ND	None
Stuart Baker, M.D.	VHA Inc. Irving, TX	Provider and Health Plan
Diane L. Bechel, Dr.P.H.	Ford Motor Company Dearborn, MI	Purchaser
Robert M. Dickler ²⁴⁵	Association of American Medical Colleges Washington, DC	Research and Quality Improvement
Joyce Dubow	AARP Washington, DC	Consumer
Mary Foley, M.S., R.N.	American Nurses Association Washington, DC	Provider and Health Plan
Vanessa N. Gamble, M.D., PhD ²⁴⁶	Association of American Medical Colleges Washington, DC	Research and Quality Improvement
William E. Golden, M.D.	American Health Quality Association/University of Arkansas, Medical Sciences Little Rock, AR	Research and Quality Improvement
Brian W. Lindberg, M.M.H.S.	Consumer Coalition for Quality Health Care Washington, DC	Consumer
Jerod M. Loeb, Ph.D.	Joint Commission on Accreditation of Healthcare Organizations Oakbrook Terrace, IL	Research and Quality Improvement
John R. Lumpkin, M.D., M.P.H.	Illinois Department of Public Health Springfield, IL	Research and Quality Improvement
Donald Nielsen, M.D.	American Hospital Association Chicago, IL	Provider and Health Plan
Jonathan Perlin, M.D., Ph.D., MSHA	US Veterans Health Administration Washington, DC	Provider and Health Plan
Louise Probst	Gateway Purchasers Coalition for Health St. Louis, MO	None
Blair L. Sadler, J.D.	Children's Hospital and Health Center	None

²⁴⁵ Since February 2002

²⁴⁶ Through January 2002

	San Diego, CA	
Marina L. Weiss, Ph.D.	March of Dimes Washington, DC	Consumer
Liaison Members		
Steven B. Clauser, Ph.D. ²⁴⁶	Centers for Medicare & Medicaid Services Baltimore, MD	Purchaser
Irene Fraser, Ph.D. ²⁴⁷	Agency for Healthcare Research and Quality Rockville, MD	Research and Quality Improvement
Beth Kosiak, Ph.D. ^{245,248}	Centers for Medicare & Medicaid Services Baltimore, MD	Purchaser
Gregg Meyer, M.D., M.Sc. ²⁴⁸	Agency for Healthcare Research and Quality Rockville, MD	Research and Quality Improvement
Barbara Paul, M.D. ²⁴⁷	Centers for Medicare and Medicaid Services Baltimore, MD	Purchaser

*In December 2002 and March 2003, when voting under the NQF Consensus Development Process occurred for the NQF Report entitled “*National Voluntary Consensus Standards for Hospital Care: An Initial Performance Measure Set.*”

²⁴⁷ Since August 2002

²⁴⁸ Through July 2002

Appendix H: Priority Topics Identified for Measures using HCUP

High-Priority Clinical Inpatient Conditions

Pregnancy/childbirth/maternal and neonatal care
Coronary atherosclerosis
Pneumonia
Congestive heart failure
Acute myocardial infarction
Acute cerebrovascular disease
Mental and substance abuse-related disorders
Cardiac dysrhythmias
Chronic obstructive pulmonary disease and bronchiectasis
Spondylosis, intervertebral disc disorders, other back problems
Nonspecific chest pain
Complication of device, implant or graft
Fluid and electrolyte disorders
Biliary tract disease
Osteoarthritis
Asthma
Septicemia
Diabetes mellitus
Urinary tract infections
Fitting of prostheses and adjustment of devices
Fracture of neck of femur (hip)
Hypertension
Anemia
Heart valve disorders
Complications of surgical procedures or medical care
Bacterial infection
Hyperlipidemia
Other gastrointestinal disorders

Additional Crosscutting High-Priority Area

Pain management
Patient safety
Patient-centered experience/outcomes measures
Surgical outcomes (including volume proxies)

Appendix I: Measures Recommended by NQF Staff for Further Evaluation

Measure	Area	ORYX ²⁴⁹	Pro ²⁵⁰	HEDIS ²⁵¹	HCU P ²⁵²	FAC CT 253	CARS ²⁵⁴	PBGH ²⁵⁵	NBCH ²⁵⁶	Leapfrog ²⁵⁷
Smoking cessation advice/counseling (inpatient)	AMI ²⁵⁸	X								
Aspirin at arrival	AMI	X								
Timing of reperfusion therapy	AMI	X								
Aspirin at discharge	AMI	X					X			
Beta Blocker at arrival	AMI	X								

²⁴⁹ ORYX Initial Core Measures (The Joint Commission)

²⁵⁰ Peer Review Organizations' 6th Scope of Work (CMS)

²⁵¹ Health Plan Employer Data and Information Set (NCQA)

²⁵² Healthcare Cost and Utilization Project, newly evaluated measures (AHRQ)

²⁵³ Foundation for Accountability

²⁵⁴ Coordinated Autos/UAW Reporting System

²⁵⁵ Pacific Business Group on Health

²⁵⁶ National Business Coalition on Health

²⁵⁷ Leapfrog Group

²⁵⁸ Acute myocardial infarction

Left ventricular ejection fraction (LVEF) < 40% prescribed angiotensin-converting enzyme inhibitor (ACEI) at discharge	AMI	X								
Beta Blocker at discharge	AMI	X		X						
Acute myocardial infarction (AMI) mortality	AMI	X			X		X	X		
Appropriate use/non-use of ACEI at discharge	CHF ²⁵⁹			X			X			
Congestive Heart Failure (CHF) mortality	CHF			X	X		X			
CABG volume	Corath ²⁶⁰			X	X		X	X	X	X

²⁵⁹ Congestive heart failure

²⁶⁰ Coronary atherosclerosis

CABG mortality	Corath			X	X		X			
PTCA volume	Corath			X	X		X	X	X	X
Abdominal aortic aneurism (AAA) repair volume	Corath			X	X			X	X	X
Carotid endarterectomy (CEA) volume	CVD ²⁶¹			X	X			X	X	X
Stroke mortality	CVD				X		X			
Vaginal birth after Cesarean section (VBAC) rate	Mom ²⁶²	X		X	X		X			
Third or fourth degree laceration	Mom	X					X			
Cesarean section rate	Mom				X		X			
Pneumonia screen or pneumococcal vaccination	PNEU ²⁶³	X	X							
Initial antibiotic consistent with current recommendations (ICU/non-ICU)	PNEU	X	X							

²⁶¹ Acute cerebrovascular disease

²⁶² Pregnancy/childbirth/maternal and neonatal care

²⁶³ Pneumonia

Blood culture collected prior to first antibiotic administration	PNEU	X	X							
Antibiotic timing	PNEU	X	X				X			
Influenza screen or vaccination	PNEU	X	X							
Pneumonia mortality	PNEU				X		X			
Follow-up/loss to follow-up after diagnosis of depression	PSYC 264			X		X				
Esophageal resection volume	SURG 265				X			X	X	X

²⁶⁴ Mental and substance abuse-related disorders

²⁶⁵ Surgical procedures

Appendix J: Hospital Care National Performance Measures Workshop Participants

Participant*	Organization	Council Represented
William A. Gillespie, M.D. (co-chair)	Provider and Health Plan Council Chair Durham, CT	Provider and Health Plan
Debra L. Ness (co-chair)	National Partnership for Women and Families and Consumer Council Vice Chair Washington, DC	Consumer
William M. Barron, M.D.	Loyola University Medical Center, Center for Clinical Effectiveness Chicago, IL	Research and Quality Improvement
Marissa Clifford, M.P.H.	GlaxoSmithKline Research Triangle Park, NC	Research and Quality Improvement
Joyce Dubow	AARP Washington, DC	Consumer
Karen M. Fernandes, R.N.	Tenet Healthcare Dallas, TX	Provider and Health Plan
William E. Golden, M.D.	American Health Quality Association/University of Arkansas, Medical Sciences Little Rock, AR	Research and Quality Improvement
Stephen Grossbart, Ph.D.	Premier, Inc. Charlotte, NC	Provider and Health Plan
Stephen J. Horner	HCA, Inc. Nashville, TN	Provider and Health Plan
Robert Krughoff, J.D.	Consumers' CHECKBOOK Washington, DC	None
Jack Mahoney, M.D., M.P.H.	Washington Business Groups on Health/Pitney Bowes Stamford, CT	Purchaser
Arnold Milstein, M.D., M.P.H.	Pacific Business Group on Health San Francisco, CA	Purchaser
Christopher J. Queram	Employer Health Care Alliance Cooperative Madison, WI	Purchaser
Louise Probst	Gateway Purchasers Coalition for Health St. Louis, MO	None

Charles P. Sabatino, J.D.	Consumer Coalition for Health Care Quality/American Bar Association Commission on Law and Aging Washington, DC	Consumer
Paul M. Schyve, M.D.	Joint Commission on accreditation of Healthcare Organizations Oakbrook Terrace, IL	Research and Quality Improvement
Cary Sennett, M.D., Ph.D.	American College of Cardiology Bethesda, MD	Research and Quality Improvement
Gerald M. Shea	AFL-CIO Washington, DC	Consumer
Liaison Members		
Kevin B. Piper	National Health Care Purchasing Institute Washington, DC	Purchaser
Barbara Paul, M.D.	Centers for Medicare & Medicaid Services Baltimore, MD	Purchaser

*Those participants that appear in bold type were also members of the original Hospital Steering Committee.

Appendix K: Informed Consent Form

Identification of Project: Performance Measurement System Design: The Case of the National Quality Forum's Hospital Steering Committee.

Introduction

The Center for Public Administration and Policy at the Virginia Polytechnic Institute and State University (Virginia Tech) supports the practice of protection for human subjects participating in research. The following information is provided for you to decide whether you wish to participate in this study. You may refuse to sign this form and not participate in this study. You should be aware that even if you agree to participate, you are free to withdraw at any time.

Purpose of the Research:

The purpose of this research is to examine the organizational network that has emerged around the development of a national performance measurement system for hospitals and the factors that influence measurement selection.

Procedures:

Participation in this study will take 60 minutes of your time and consists of a one-on-one interview comprised of unstructured, open ended questions that will be audio taped with your permission.

Risks and/or Discomforts:

There is no more than a minimal risk or discomfort associated with this research.

Benefits:

The information gained from this study will contribute to the understanding of the factors that influence the development of a national performance measurement system and the work of the organizational network involved in measurement selection.

Confidentiality:

Any information obtained during this study which could identify you will be kept strictly confidential. The data will be stored in a locked cabinet in the investigator's home office and will only be seen by the investigator and her faculty advisor during the course of the research. The information obtained in this study may be published in peer reviewed journals or presented at academic meetings but the data will be reported as aggregated data. The audiotapes will be erased after transcription.

Opportunity to Ask Questions:

You may ask any questions concerning this research and have those questions answered before agreeing to participate in or during the study. Or you may call the investigator at any time, cell phone, (703) 727-7503, or after hours, (703) 421-1260. If you have questions concerning your rights as a research subject that have not been answered by the investigator or to report any concerns about the study, you may contact Virginia Tech Institutional Review Board, telephone (540) 231-4991.

Freedom to Withdraw:

You are free to decide not to participate in this study or to withdraw at any time without adversely affecting your relationship with the investigators or Virginia Tech.

Consent, Right to Receive a Copy:

You are voluntarily making a decision whether or not to participate in this research study. Your signature certifies that you have decided to participate having read and understood the information presented. At your request, you will be given a copy of this consent form to keep.

Signature of Participant:

Signature of Research Participant

Date

Name and Phone Number of Investigator:

Marybeth Farquhar, R.N., M.S.N., Principal Investigator Office: (301) 427-1317

Joseph V. Rees, Ph.D., Faculty Advisor Office: (540)231-6034

Appendix L: Interview Questionnaire

General Information

Name

Title

Organization Affiliation

Organization Classification

Number of years at current position

Education

General info. (Internet search)

- Tell me about yourself and any roles you played in the development of hospitals performance measures or the development of a hospital performance measurement system.
- Tell me about your involvement in the development of the NQF hospital performance measurement system?

1. Role of the Federal Agencies:

- What is the role of federal agencies, specifically AHRQ and CMS, in the development of a national performance measurement system for hospitals? How has it changed over the course of the development process?
- What role do you think the federal government currently plays in the development of performance measures and PMS for hospitals? What role should they play?

- What is CMS' role in the development of PM and PMS for hospitals? Has it changed over the past few years? What should they be doing that they are not doing now?
- What is AHRQ's role in the development of PM and PMS for hospitals? Has it changed over the past few years? What should they be doing that they are not doing now?
- How influential is CMS in the selection of PM? (scale of 1-10)
- How influential is AHRQ in the selection of PM? (scale of 1-10)

2. NQF Steering Committee Critical Tasks

- What have been the NQF Hospital Steering Committee's critical tasks in the development of a national performance measurement system for hospitals between March 2001 and December 2004?
- What have been the NQF Hospital Steering Committee's critical tasks in the development of a national performance measurement system for hospitals?
- What were some of the issues and debates of the committee regarding hospital performance and measure selection?

3. Physician's role

- What role do physicians, particularly those employed by AHRQ and CMS, play in the development of performance measurement systems in this process?
- What role do physicians play in the selection of performance measures? In performance measurement system development?
- How important are they to the process?

- Is the role of government employed physicians different from other physicians in the selection and development process?
- On a scale of 0-10 with zero being no influence and ten being the most influential, how influential are physicians in the decision making process for performance measurement selection for a national system?

4. NQF's interaction with the hospital industry

- How has the NQF Hospital Steering Committee interacted with other industry groups in the development of the performance measurement system between March 2001 and December 2003?
- How has the NQF Hospital Steering Committee interacted with other industry groups in the development of the performance measurement system between March 2001 and December 2003?
- What do you think NQF's role should be in the selection and development of a hospital performance measurement system?

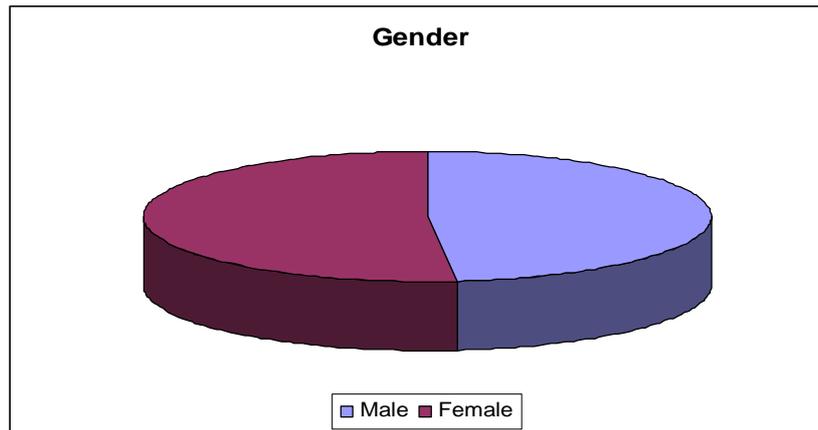
5. Influences of PM selection

- How has the field of performance measurement and the performance measures themselves, influenced the selection of a core set of performance measures for hospitals between March 2001 and December 2003?
- How has the field of performance measurement and the performance measures themselves influenced the selection of a core set of performance measures for hospitals?
- In your opinion, what factors influence the selection of performance measurement?

Appendix M: Selected Characteristics of the Sample

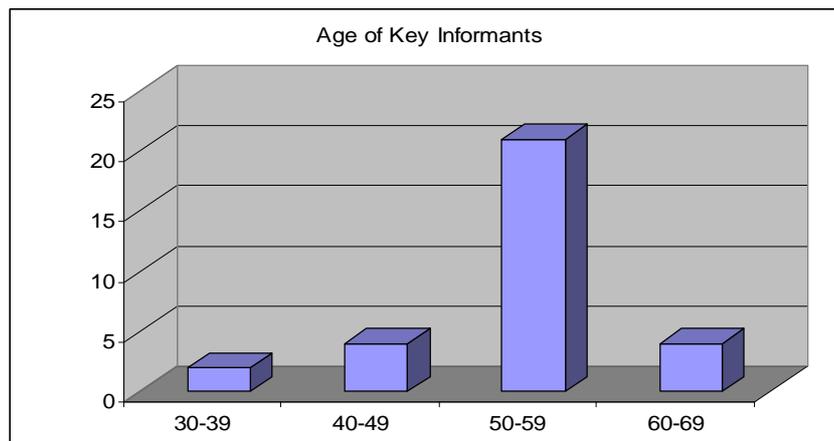
Of the total number of individuals interviewed, 15 were male (48 percent) and 16 were female (52 percent).

Figure 10: Gender Profile



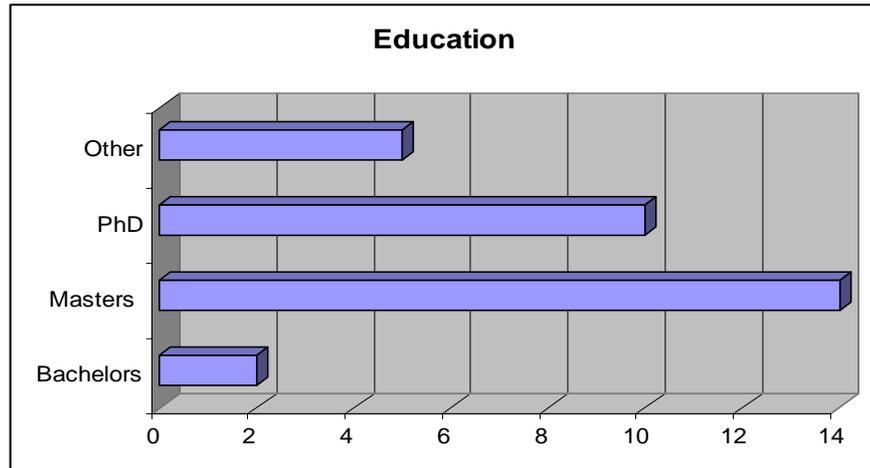
The age of the interviewees varied but the majority interviewed for this research were between 50-59 years old and had been working in the health care industry or a related industry for the majority of their careers.

Figure 11: Age of Key Informants



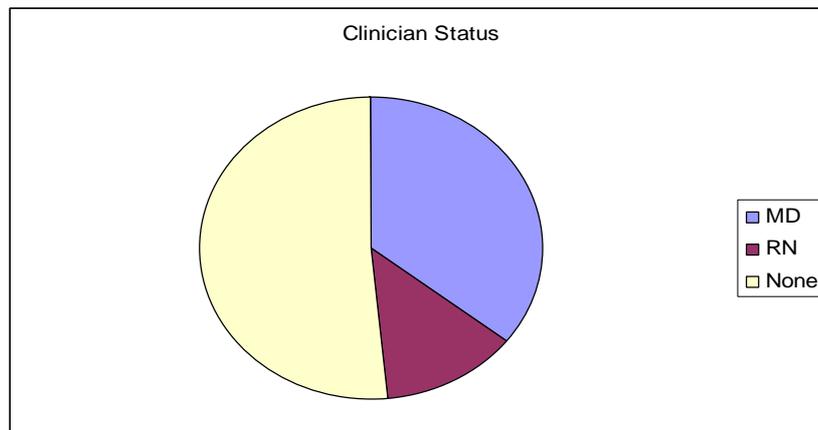
Of those interviewed, 14 had earned a Masters degree; 10 had a Ph.D.; two had earned Bachelors' degrees, and five interviewees were placed in the category of "other," denoting those individuals who had an M.D. degree only.

Figure 12: Education of Key Informants



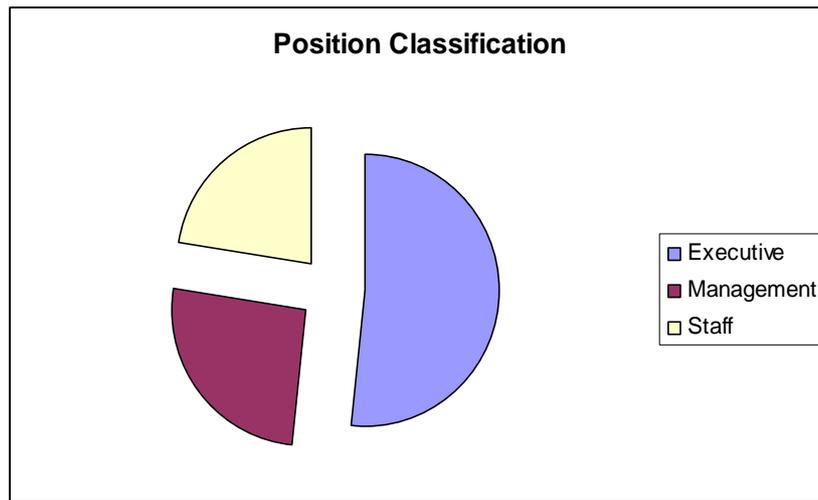
The majority (52 percent) of those interviewed were not clinicians. Of the remainder, 11 were physicians (35 percent) and 4 were registered nurses (13 percent).

Figure 13: Clinician Status



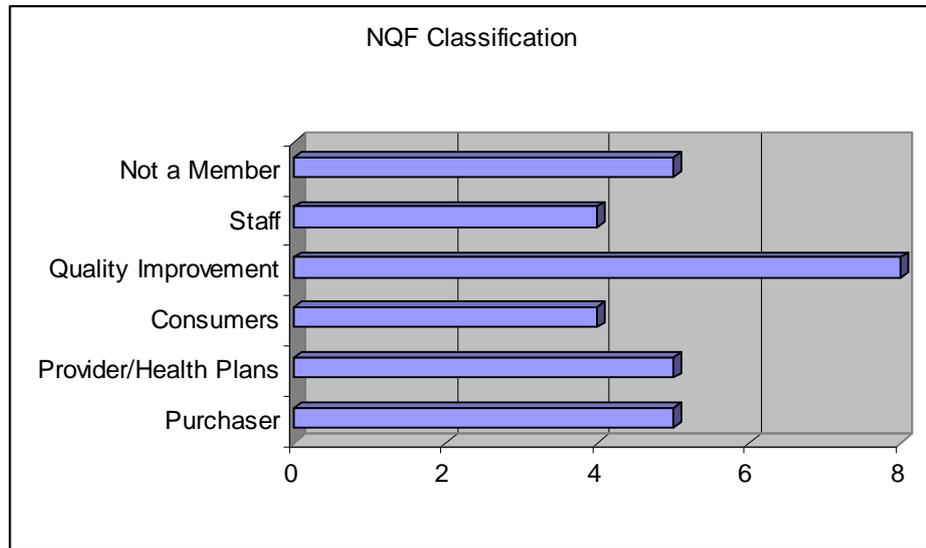
The majority of the key informants (51 percent) held executive positions within their organizations, while 26 percent held management positions and 23 percent held staff level positions.

Figure 14: Position Classification



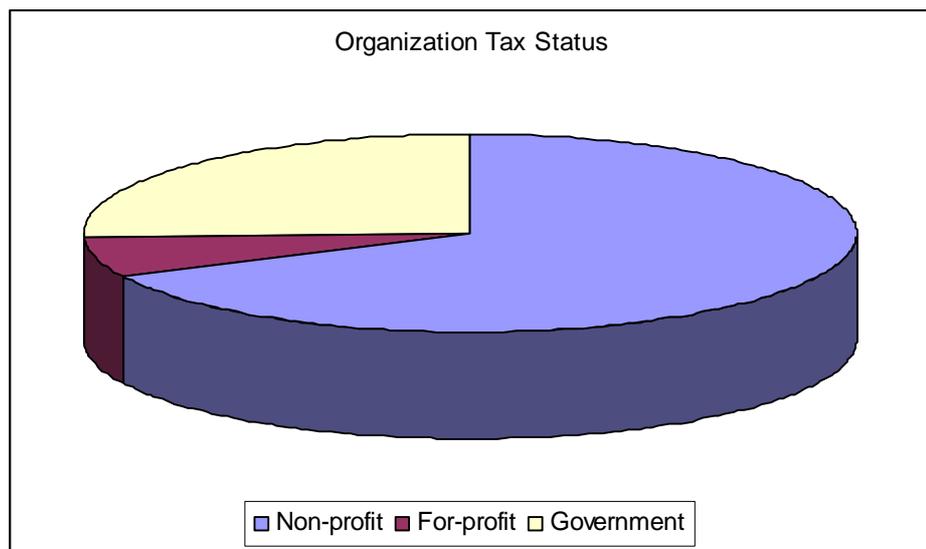
The interviewees were classified by membership in NQF Councils. The majority of the interviewees (eight) were from the Quality Improvement Council. There were five interviewees each from the Purchaser Council and the Provider/Health Plan Council, four from the Consumer Council, and five key informants who were not members of the NQF. The remaining four interviewees were staff members of the NQF.

Figure 15: NQF Classification



The largest number of interviewees was employed by a non-profit organization (68 percent). The federal government employed 26 percent of those interviewed, and the for-profit sector employed six percent of those interviewed.

Figure 16: Tax Status



Appendix N: Summary of the Main Questions of the Study

The Relationship of Science, Values, Politics and Networks

As evident in the case study, the process for selecting performance measures by consensus involved science and scientific evidence, to a point; after that, existing structures and values came into play. Overall, the selection process was infused with practical politics and political maneuverings. It involved negotiation among all the stakeholders within health care, the science of performance measurement and performance measures, a great deal of patience, as well as large amounts of time. Thus the role of politics can be constant, diffused throughout the process and often invisible; it is embedded in the existing structures of a society, and as we learned, dependent upon the actors and their position within the network. The introduction of the NQF to the network as a potentially important player in health care whose focus is consumer/purchaser oriented also helped to rearrange the order of the network actors, which, as has been described, played out during the deliberations of the Hospital Project. In this case study, as different actors were added or subtracted from the mix, the outcomes, i.e., the measures that were selected for endorsement, were significantly altered. For example, when the consumer and purchaser advocates voiced their concerns about the scant number of hospital measures originally selected for endorsement, and became active participants in the process, the process and the outcome changed in the form of the selection of additional measures. As one participant in the process stated, “they applied only provider-researcher values in their first cut, and that’s why they approved so few [measures]. They didn’t weigh the customer’s view point, which has a lower validity standard.”

What also factored into measure selection and seemed to be relatively important was the specific people who were at the table and actively engaged during the Committee deliberations. A clear vision of interests and values was required to articulate arguments for and against certain measures that made preparation essential, e.g., anticipating arguments and devising counter arguments to address key issues that arose. As one key informant stated, “it depends on who is at the table, or who had to leave early; and who wouldn’t shut up...as to which measures are selected.”

The Role of Clinicians in Performance Measure Selection

The actors at the table are of interest. Many were clinicians, physicians or nurses but not all, sent to represent their organizations’ mission, point of view and values, and in some cases to protect their turf. The clinician actors represented the organization that sent them to the table first and foremost, and seemed to have no loyalty conflict with respect to their profession. Overall, the Committee consisted of more physicians than nurses, whose role consisted of informing the discussions with regard to the latest medical research or with anecdotes from the frontlines of health care. In fact, being a clinician only enhanced the credibility of their arguments, particularly when the clinician maintained a current clinical practice. The members of the Committee viewed current practice as highly relevant to their critical tasks, which turned out to be a problem for them in the long run. As they respected the clinician as “knowing their business,” it was relatively easy for providers to articulate arguments for or against the selection of particular measures that were of interest to consumers and purchasers. Thus, consumers and purchasers, who did not have the same expertise in health care, could do little at the beginning of the Committee’s deliberations to counter

provider arguments. As one key informant said, “the reason consumer and purchaser values had not gotten expressed in the first round, which as is always the case in any type of multi-stakeholder health care deliberation, was that whoever has the most at risk, 100 percent of providers paychecks that are involved here, invest more in preparation than those for whom health care is just a fraction of their life...” As the second round of measure selection showed, the consumers and purchasers were much more prepared—substituting more vocal, assertive members at the table and preparing counter arguments that supported their position and interests. One participant realized:

“If we want in round 2 for the purchasers and consumers to be able to stand their ground and well represent their values and their interests, it was going to require organization, education, training and planning. And then we got into the room and I think it was the first time that consumers-purchasers were represented by people who were almost as well informed and certainly as strongly motivated as the providers and the researchers. You could tell the difference. A lot more measures got approved and there were plenty of times when authoritative people said that ‘this measure is just not good enough,’ in which there was a rebuttal from the other side. The votes, often close, generally favored approval.”

Since that time, the consumer and purchaser groups have been very vigilant about nominating individuals for committees and panels, as well as monitoring the NQF projects, to assure that their interests and values are well represented.

The Role of the Federal Government in Performance Measure Selection

The government has several roles in health care. They include providing for those that cannot provide for themselves, providing public and social goods, regulating the market place, and instilling trust and accountability (Tang et al., 2004:48). In his testimony to the President's Advisory Commission, William Roper, M.D., former Administrator of the Health Care Financing Administration stated, "The strength of the Federal Government is its ability to convene and to mobilize a variety of other parties towards a common goal. The visibility, the stature, the power of the Federal Government is unparalleled" (1997: 15). Thus, the federal government plays an intriguing role in the selection of performance measures for endorsement. First and foremost, it depends upon the NQF for producing a supply of endorsed measures that can be used to evaluate health care projects and programs, which are, in turn, vehicles of transformation and change in the larger health care system. CMS in particular, categorized as a purchaser, is a major consumer of NQF-endorsed measures and uses them in its various programs to judge performance, which forms the basis for payment or non-payment to providers. Obviously, measures that decide who will be paid and what they will be paid are highly contentious and, as with most federal programs, bring public scrutiny and criticism. NQF and the consensus process is one way for the federal government to provide a public forum for the vetting of measures by the various stakeholders of health care. As one key informant stated, somewhat cynically, "to get the NQF *imprimatur* for the measures [was important] partly because it gives CMS and JCAHO cover." Measures such as those selected for hospitals have been publicly vetted, studied, and pilot-tested and were considered suitable by the industry.

The federal government also is a supplier of performance measures to the NQF consensus development process. Since many of the measures developed by federal agencies reside in the public domain, they are prime candidates for inclusion in NQF consensus projects. The federal government often sponsors projects so that its measures can be endorsed by the NQF and included in federal programs. Because tax dollars are used to fund NQF projects, the federal sponsor of a NQF project also participates as a liaison in the CDP and other projects, as accountability for dispersion of public tax funds is an issue. However, government participation is an issue for NQF and requires a balance between the government's need to monitor how tax dollars are spent and the integrity of the consensus process. For the NQF process to have legitimacy, there must be the perception that the CDP is fair, unbiased, and free from excessive governmental influence. At the present, NQF has determined that allowing for liaison, non-voting roles in government-sponsored projects is prudent and takes advantage of the specific expertise of individuals employed in federal service.

Other roles of the federal government, played specifically by AHRQ and CMS, that were identified during this research included priority and agenda setting for health care. Through the projects AHRQ, CMS and other federal agencies initiated, and their subsequent requests to NQF for measures, white papers, and the like, the federal government works as a driver of improvement in the field of health care. In this leadership role, it can bring all actors together and incorporate the scientific, practical, and political aspects of health care into a set of national priorities. For example, the selection of hospital measures as one of the first NQF CDP projects assisted in establishing a public reporting website for consumer choice, Hospital Compare (www.hospitalcompare.hhs.gov), which also helped to drive improvement in the health care field.

The federal government also contributes to the scientific knowledge and evidence base in health care through various funding mechanisms. These projects involve not only measure development and refinement, but also measure testing, as well as implementation strategies and subsequent approaches for improvement. In contrast to other efforts, these initiatives focus on meeting national priorities, which set the direction of research funding to develop the evidence base to meet these goals.

The Science of Performance Measurement and Hospital Measure Selection

The state of the science with regard to performance measures and performance measurement is an issue for the selection of hospital measures. Basically, providers and researchers within health care tend to favor highly precise measures, with adequate risk adjustment where appropriate, and are highly sensitive to unintended consequences or perverse incentives. Conversely, purchasers and consumers know what they would like information about and what they want measured, and have a higher tolerance for imprecision. For example, consumers and purchasers believe that a measure that correctly classifies a provider 51 percent of the time is a “good” measure, and that it should be publicly reported or used for payment programs, while providers and researchers feel the same measure is inappropriate for these purposes. The NQF process seeks to strike a balance between “what is possible” and “what we want,” which is a balance between the science of measurement and the need for additional measures to report performance. This is one point that contributes to the arguments for not endorsing particular measures, which can be viewed as a legitimate position. As one key informant indicated, however, “the people who say the science isn’t good enough, some of them, of course, hide behind that because they don’t want to be measured. So you have to

figure that one out. Well is the science really not good enough or is it that they're just looking for every excuse they can?"

Furthermore, when measures that were considered adequate and scored equally on the NQF criteria, according to one key informant, the "nod would go to whichever [measure] had more demonstrated use, and/or, specialty society buy in. You default to the measure with the least resistance." The informant went on to say that "the health care industry is very capable. If they sense they are losing political ground, they are extremely resourceful at figuring out how to minimize loss." For example, the hospital associations came up with the idea of the Hospital Quality Alliance, a "self-appointed, health industry dominated group" that could select NQF endorsed measures for "immediate implementation." One starts "with a ton of measures, and then through a consumer-purchaser majority NQF process, you whittle all hospital activity down to 39 measures. What the American public gets are 10 measures that apply to only three conditions on which peer review organizations have been working with hospitals for the last 10 years. That's what I mean by a health industry that is incredibly resilient and protects the status quo."