EMBRACING ACCOUNTABILITY: PHYSICIAN LEADERSHIP, PUBLIC REPORTING, AND TEAMWORK IN THE WISCONSIN COLLABORATIVE FOR HEALTHCARE QUALITY

Ann Lennarson Greer
University of Wisconsin–Milwaukee

June 2008

ABSTRACT: This case study of the Wisconsin Collaborative for Healthcare Quality (WCHQ) describes the origin and operational tenets of a "bottom-up" physician-led organization that has achieved voluntary public reporting of comparative performance information in both ambulatory and hospital settings. Tenets crucial to the observed success in physician engagement include: 1) an unrelenting focus on quality of care as the goal of reporting; 2) performance data that meet scientific standards of validity and reliability; 3) creation of standard measures to assure applicability to all sites; and 4) mutual sharing of best practices. The WCHQ organization links executives and professional staffs in multi-tier networks that balance national standards with local realities and identify actionable solutions. In 2007, the author conducted 31 personal interviews with executives of the provider organizations that form WCHQ's membership, along with other members of the governing board (business partners and state society officials), and WCHQ staff.

Support for this research was provided by The Commonwealth Fund. The views presented here are those of the authors and not necessarily those of The Commonwealth Fund or its directors, officers, or staff. This and other fund publications are available online at www.commonwealthfund.org. To learn more about new publications when they become available, visit the fund's website and register to receive e-mail alerts. Commonwealth Fund pub. no. 1142.

CONTENTS

About the Author iv

Acknowledgments iv

Executive Summary v

Introduction 1

Findings 2

- The Birth of WCHQ 2
- Engaging Physicians: the Wisconsin Idea 4
- A Call to Leadership 4
- Collaboration to "Lift All Boats" 6
- Physician-Leadership Ready for Prime-Time 8
- The Challenge of Growth 9
- Means of Physician Engagement: Lessons from Wisconsin 12
- The Unit of Reporting: Strong Feelings 14
- Local Provider Sponsorship 16
- A Structure for Collaboration 17
- Diverse Improvement Strategies 18
- Other Users of Public Reporting 19
- Business Case for Quality 21

Conclusions and Implications 22

Notes 25

Appendices 26

- Appendix 1. Methodology 26
- Appendix 2. Wisconsin Collaborative for Healthcare Quality 28
- Appendix 3. WCHQ Performance Data: Standards, Methods of Submission, Validation 31

ABOUT THE AUTHOR

Ann Lennarson Greer, Ph.D., is professor of sociology and urban studies at the University of Wisconsin–Milwaukee and a member of the faculty of the Helen Bader Institute for Nonprofit Management. Over three decades, Greer has conducted personal interviews concerned with the translation of science to medical practice, including adoption of new science and technology, the implementation of evidence-based protocols, support of clinical trials, and attitudes toward the electronic ambulatory medical record. As a Senior International Fellow of the John E. Fogarty International Center for Advanced Study in the Health Sciences, she enjoyed a posting in the United Kingdom with the Health Unit of Brunel Institute for Organization and Social Studies. More recently, she was a Visiting Research Fellow at the Sealy Center for Geriatrics at the University of Texas Medical Branch at Galveston. Greer has served as an advisor to numerous government agencies from the international to the local level, including the World Health Organization–Europe, the British Medical Research Council, the National Institute of Mental Health, the National Institutes of Health Office of Medical Applications Research, and MetaStar, the Wisconsin Quality Improvement Organization. Greer holds a Ph.D. in sociology from Northwestern University. She can be e-mailed at agreer@uwm.edu.

ACKNOWLEDGMENTS

This project was completed in a short timeframe that was further complicated by a concurrent international commitment on the part of the sole investigator. It was achieved only with an extraordinary response from the health care executives who made themselves available for two-hour interviews within the investigator's constrained travel schedule. That they did so is testimony to the value and priority they assigned the Wisconsin Collaborative for Healthcare Quality. This response, however, would not have been enough. Thanks go also to Tammie Cawkins, the wonder woman of this project who, among a number of important tasks that contributed to success, managed scheduling and logistics.

Editorial support was provided by Elizabeth Dossett.

EXECUTIVE SUMMARY

The Wisconsin Collaborative for Healthcare Quality (WCHQ) is a voluntary consortium of health care organizations whose mission is to improve the quality and cost-effectiveness of health care in Wisconsin. WCHQ reports comparative measures of performance in ambulatory and hospital settings using data submitted voluntarily by provider organizations. Its ability to gain the voluntary cooperation of providers in public reporting rests on its self-definition as:

> [A] learning organization whose members continually expand their capacity to create the quality improvement results to which they aspire. Its members do so by working and learning together to foster new and expansive patterns of thinking that drive improvement in our healthcare institutions.[1]

As part of WCHQ's mission, public reporting is used as a tool to jumpstart the learning and quality improvement process. Toward this end, member organizations collaborate in several activities, including: 1) development of scientifically valid ambulatory care measure specifications and an attribution method that enables physician groups and health systems to collect quality data on all patients in their care; 2) open sharing of quality performance data through public reporting; and 3) identifying and sharing of best practices to improve all members' performance. Although WCHQ prides itself on a scalable infrastructure that permits expansion to any number of reporting entities and conditions, early efforts in the area of ambulatory care have focused on conditions that are common, treatable, and costly: diabetes, uncomplicated hypertension, preventive cancer screening, and postpartum care.

The fact that WCHQ is a physician-led effort is critical to its success. Founded in 2003 by chief executives of several large multispecialty practices and their partner hospitals, it has grown rapidly from nine initial members to a membership in 2007 that includes 28 physician groups, hospitals, and health plans. Currently reporting data for more than 50 percent of Wisconsin primary care physicians, WCHQ has set a goal of including 75 percent of the state's physicians by 2010.

Membership consists of medium as well as large physician groups, as neither performance reporting nor participation in improvement efforts require sophisticated data systems, electronic health records, or extensive quality improvement staff. While most members are large organizations, they incorporate diverse practice structures. WCHQ also includes business partners whose participation is a vehicle to maintain direct lines of communication with purchasers of health services in the hope that the measures reported can serve multiple audiences.

This case study relies on 31 lengthy personal interviews conducted from June through September 2007 with persons closely involved with WCHQ. The project sought to interview the universe of persons who lead participating organizations, some of whom also serve on the WCHQ board. Interviews were successfully completed with top executives of 25 medical practices, hospitals, and health systems and 20 of 21 board members. Also interviewed were business partners who serve on the board, chief executives of the state medical and hospital associations (who serve on the board *ex officio),* and WCHQ executive staff.

WCHQ shows great promise as a means to successfully address two problems that have intractably bedeviled efforts to bring community practice into closer alignment with accepted indicators of evidence-based practice. First, it has achieved public performance reporting for ambulatory delivery sites where reporting depends on voluntary submission of clinical data. Second, WCHQ has put into place a dynamic model for translating evidence-based medicine into community practice.

The specific mechanisms by which WCHQ has achieved these goals include: 1) development of performance measure specifications and an attribution method that are accepted by physicians as accurate representations of the quality of care delivered; 2) use of motivational strategies that accord with physicians' desire to practice high-quality medicine, and 3) creation of a multi-level collaborative network that harnesses the knowledge and creativity of frontline professionals.

WCHQ begins with nationally endorsed guidelines and measures but uses grassroots involvement to address crucial barriers to the adoption and development of measurement specifications. One barrier is that current claims-based ambulatory care measures are viewed by physicians as scientifically inaccurate, punitive, and insensitive to actual delivery of care. A second is the common situation where physicians are unable to review, revise, or constructively use the data to achieve improvements. WCHQ's solutions, shaped by physicians familiar with these problems, involve adherence to the standards that physicians demand of data, transparency of measurement methods as well as results, and placement of the entire enterprise into the context of improving practice.

In response to physicians' demand for data that meet scientific standards and for results viewed as accurate benchmarks, WCHQ avoids sampling bias, generates sample sizes sufficient to draw valid conclusions, and reports at level appropriate to a "unit of analysis." Additionally, to be actionable in terms of comparing performance and deficiencies, results

must be timely and as well as revisable in light of new evidence, and must encompass contextual diversity.

Failure to accord with these criteria may be seen as the *sine qua non* of the opposition by many physicians to standards, guidelines, and targets. WCHQ also has a philosophy regarding the use of its performance results: it avoids tactics that blame individuals and encourages a focus on systemic problems, such as failures of communication, teamwork, resources, or system design.

Interviewees agree that these interlinked elements have been successful in achieving a process in which participation in measures development and application leads to 1) acceptance of measures as valid indicators of performance; 2) "apples-to-apples" comparisons with colleagues practicing in similar settings for reliable performance benchmarks; and 3) opportunities to meet with peers to share strategies and practices employed by the high-performing organizations.

Strong leadership was indisputably a crucial element in launching and defining the organizational culture that guides working relationships within WCHQ. The founders rallied member organizations and their staffs with a few simple rules that resolutely defined public reporting as a means to high-quality health care for all patients in all settings—*not* as a vehicle to eliminate practice sites. The code of ethics also defined behavior that was proscribed due to its potential to disrupt the honest sharing of data and knowledge across organizational boundaries (e.g., using results for marketing purposes).

Around these simple rules, a coherent, powerful, quality-focused culture coalesced that members report to be very satisfying. Leaders believe the observed performance improvement is attributable to participation in WCHQ. They also believe the approach is sustainable and self-generating because it is premised on valid, professional goals and standards.

A key principle has been WCHQ's intention to "lift all boats." To lift all boats is to reassert a goal consistent with the medical profession's traditional norms of collegiality and knowledge-sharing against the contrary influences of market competition. Other recurrent phrases are "bottom-up rather than top-down" (to characterize the basis for professional ownership and identification) and "the truth lies in the middle" (to describe collaboration with business partners).

Interviewees have agreed that WCHQ's model of physician-led voluntary reporting has delivered quality improvement to their organizations. They believe that the

collaborative sharing of knowledge within the WCHQ model is conducive not only to improvement by lower-performing teams, but also to much-valued peer recognition of higher-performing teams.

One of the important findings from WCHQ's experience that may dispel fears of public reporting is that published scores did not naturally reveal a hierarchy of organizations. Instead, the reported data revealed many points of excellence scattered among the reporting organizations. This discovery affirms the idea of a bottom-up approach that learns from, as well as seeks to direct, community practice.

As with the publication of scores showing performance shortfalls in comparison to peers, the engagement with business is reported to have unsettled traditional patterns of thinking and acting. This new perspective has led to scrutiny of practices and to innovative solutions. At the same time, business partners and health care providers differ on appropriate levels of reporting. Business partners advocate reporting data for individual physicians as a means to enable consumer selection and economic credentialing by purchasers. But WCHQ's provider members argue that this practice paints a false picture of the quality of care that patients actually receive within the system, and could encourage individual solutions to poor scores (such as patient cherry-picking) that are harmful to teamwork and efficiency. Nonetheless, the debate within WCHQ is viewed as one of "constructive tension."

In sum, WCHQ's ability to enlist and sustain the involvement of medical groups rests on both avoiding strategies that elicit the resistance of physicians and supporting the achievement of professional patient care goals. The primary limit of this study is its dependence on reports of key executives not only for the model's theory and intent but for data on its actual workings. The consistency of reports across state organizations suggests accuracy, but this needs confirmation and elaboration. Also, generalizations about the transferability of WCHQ's model to new sites require additional research, including interviews with affected clinicians and members of the collaborative networks.

EMBRACING ACCOUNTABILITY: PHYSICIAN LEADERSHIP, PUBLIC REPORTING, AND TEAMWORK IN THE WISCONSIN COLLABORATIVE FOR HEALTHCARE QUALITY

INTRODUCTION

For decades, attempts to implement national standards of practice through payment, incentive, and regulation have failed to have widespread or systematic impact. In 2003, the Institute of Medicine (IOM) gave high visibility to deficiencies in the quality of care in America, attributing them to systemic flaws in the organization of care and the failure to systematically incorporate new science and best practices. The quality chasm identified by IOM is in part one between the perspectives of those informed by evidence achieved scientifically and those whose reference points are clinical practice.[2] The latter, who work with patients in diverse practice settings, often resist performance measurement as intrusive and burdensome, not viewing it as accurate or helpful to practice improvement.

Two basic models address gaps between standards and practice. One relies on extrinsic or market-driven approaches (including pay-for-performance and consumer-driven health care). The other, adopted by the Wisconsin Collaborative for Healthcare Quality (WCHQ), seeks reform from within through collaborative strategies that appeal to the intrinsic goals and learning strategies of health care professionals. Prominent examples include the Institute for Healthcare Improvement's (IHI's) Breakthrough Series and the related Veteran Administration's Quality Enhancement Research Initiative.[3] Collaborative models are also the goal of the new National Institutes of Health Centers for Translational Science and the National Institute of Mental Health Centers for Innovation in Services and Intervention Research. Each of these major initiatives highlights the need for professional, patient, and community engagement in achieving quality gains.

WCHQ leaders view their experiment as aligned with IHI models of collaboration but with two important extensions. WCHQ builds the development of performance measures into the collaborative model and also creates a statewide structure for collaboration that offers a self-generating platform for cross-institutional learning.

This case study rests primarily on 31 personal interviews conducted from June through September 2007 with individuals actively involved with WCHQ. Twenty-five interviews were completed with high-level executives of multispecialty practices, hospitals, and integrated health systems that participate in WCHQ. Also interviewed were business partners who serve on the WCHQ board of directors; the chief executive officers (CEOs)

of the Wisconsin Hospital Association (WHA) and the Wisconsin Medical Society (WMS) who serve on the board *ex officio;* and WCHQ staff (director of operations and president/CEO). The project sought to interview executives from all member health care organizations, whether or not they were currently serving on the board. This goal was largely achieved, as only one physician practice was not interviewed. (For complete details on this study's methodology, see the appendix beginning on page 26.)

FINDINGS

The gathering of providers in Wisconsin to engage in voluntary public reporting is two stories, one of innovation and the other of expansion. The first describes how founding collaborators came to undertake a high-stakes, high-risk innovation. The second speaks to the attractiveness of the model they created for other health care organizations that were not similarly interested in policy reform but who shared their concern about quality improvement.

The Birth of WCHQ

Challenge from the Environment. The health care executives who met in 2002 to discuss what was to become WCHQ perceived that health care delivery was at a crossroads. Questions about the quality of care provided in the United States had achieved a crisis level. Payers were demanding a new level of transparency and accountability and were implementing schemes of insurance that urged consumers to shop for health care services. A plethora of non-provider organizations were assembling and publishing performance reports, proposing them as a basis for payment, and using them in public awareness campaigns. A high level of momentum had the potential for provider organizations to be left behind.

WCHQ founders were CEOs of physician-led health care organizations, and all but two were physicians. One explained their perception of the environment in 2002 in this way:

> What we sensed was that there were reports being published by people who had some knowledge, but perhaps not full knowledge of health care and its delivery. . . . The reason for our meeting [in October 2002] was: it looks like people are going to start writing reports, publishing reports on medical performance and that will be followed by dictating type of care and how care should be delivered. Shouldn't we, the people responsible for the care delivery, shouldn't we be involved in that process? Should we *do it*, in other words?. . . We decided that we would form a group, create a report and make it public. We wanted to show that we were interested, we were

> concerned, that we had . . . better understanding of what those measures should be—and that we were willing to be public in our reporting.

Transparency and Public Reporting: Drivers and Barriers. The executives accepted that those using health services and those paying the bills had a legitimate claim to information about the care being delivered. However, they also had their own reasons for concern. As one state-level observer said of this period:

> Payers … and purchasers were trying to make sure they were getting value.… The providers are saying: "They are asking us questions and we don't quite have the answers. I mean: *do* we provide good quality? We know that we are doing so many mammograms but do we really know that we are doing the best we can? Are all of our docs engaged in this process? Are we doing things to impact change?"

In short, they also needed data.

An apparent paradox was that physicians are medical scientists who routinely demand and utilize data but who, at the same time, resisted performance measurement. These health care leaders did not think the problem was transparency of data. The problem was that physicians considered the measures to be invalid and their intention punitive. In addition, existing schemes of measurement, and the incentives that accompanied them, were discordant with the way physicians understood their actual practice and its subsequent improvement. If physician leaders could devise a process that would generate measures that physicians would respect and could use in a collaborative way to improve clinical care, physicians could be allies of those seeking better performance information. The leaders hypothesized that, under the right conditions, transparency could constructively motivate and even invigorate physician staffs.

Conversely, if physicians were not central to making transparency work, it was difficult to see how it could result in a sustainable model of genuine quality improvement. The failure of externally directed managed care in the 1990s provided evidence that attempts to bludgeon physicians into line were neither effective nor efficient.[4] Of what they foresaw for the future, one founder said, "It's coming and it's going to be bad." The opportunity to link transparency to the medical professional's genuine desire to practice quality medicine could be lost. The system remained unsettled but might not be for long.

Engaging Physicians: the Wisconsin Idea

In the view of this group of leaders, the real problem was that while physicians were supporters of evidence-based medicine and desirous of performing well by its standards, they believed that they were performing well. It was difficult to foster a sense of urgency without data that a higher level of performance was possible. They found many reasons for a gap between a published standard achieved in the highly controlled environment of a clinical trial and what can be achieved in local practice where many things are not controlled. In local practice, physicians are dependent on specific colleagues and teams; local facilities and resources; and the cooperation of patients, some of whom do not cooperate well at all. Physicians did not accept data that demonstrated they were doing less well than they could with these patients, with these colleagues and health care teams, in this community, in this regulatory and payment context.[5] From their perspective, as they sought to customize care to patients and adapt it to local circumstances, measurement from external groups was an affront, an intrusion by persons ignorant of medical practice. Who, after all, had committed their lives to employing medical science in the service of patient care?

An answer would be comparative data from similar health care settings. In contrast to a standard achieved in clinical trials or at a remote site, the performance of physicians whose situations were plausibly very similar to one's own would be quite relevant—as was long ago established by research into the diffusion of innovation.[6]

Comparative public data held the potential of revealing performance gaps and thereby destabilizing settled approaches and solutions. Achieving this result, however, required three things that emerged as key tenets of the WCHQ: first, attention must be refocused on quality; second, physicians must be provided with accurate data, reported appropriately; and third, physicians, clinic managers, and other frontline professionals must be empowered to achieve a local solution.

A Call to Leadership

Unfortunately, what was missing was leadership from providers. This shortfall was easy to understand. Developing an alternative model required two things that were hard to come by—resources (both financial and human) and courage. To take responsibility yourself was very high-risk, for if you developed and endorsed new measures, you could not disclaim the results. Each founding member knew the efforts of others to measure care had been rejected by physicians as invalid, unfair, and unhelpful. While they thought they could do better, there were no guarantees that they could succeed in developing metrics that, unlike precursors, would concord with physician's ideas of clinical quality.

Nor were they guaranteed that the development of accurate measures would lead to success. These CEOs, and anybody who joined them, would have to worry about the scores that would be published for their organizations. Bad scores could harm their organizations in the marketplace. Beyond this, any innovation that affected roles, relationships, and established ways of working in an organization would elicit resistance. The common response is to prefer the evil that is known to the one that may emerge. Why did highly regarded health care executives running highly successful organizations take such a risk? In a public presentation at a later date, members of WCHQ explained the decision with a quotation from Jerry Garcia of The Grateful Dead: "Somebody has to do something and it's just incredibly pathetic that is has to be us."

Why Us? If someone were to rise to the challenge, the Wisconsin leaders might be candidates. They headed large organizations in a state that enjoyed a solid economy, and where employer and state health programs afforded coverage to most of the population. They controlled relatively large budgets and sophisticated, quality staffs from which resources could be deployed. If they joined, they could assemble sufficient resources to attempt a new approach.

In addition, as individuals, they had very long tenure within their organizations—only one executive had served less than 20 years. People in their local communities trusted them as "safe hands," leaders concerned about their colleagues and the communities, not angling for their next job somewhere else at the possible expense of those here. The executives could hope that their physicians and communities would allow them some room for error. In addition, they trusted each other and believed that commitments among them would be respected.

Why Wisconsin? Wisconsin had an unusual advantage as a site for introducing a physician-led model. More than other states, it is distinguished by the presence of many large multispecialty physician-led practices that link to community hospitals in longstanding partnerships. These multispecialty clinics have a long history of physicians managing physicians. A number of practices were founded in early decades of the last century as solutions to the management of expanding knowledge. Wisconsin clinics of this era include the Marshfield Clinic (incorporated 1916), the Dean Clinic (approximately 1919), and the Gundersen Clinic (1930). The Skemp and Luther-Midelfort multispecialty practices that are now part of the Mayo Foundation also date to this period. The clinics sought to improve upon the then (and still) existing methods of physician coordination while retaining physician control, a feature that may enable them to offer a model to smaller practices who cherish this value.

These traits meant two things. First, if the practices did something together, the result could be big enough to be taken seriously. A significant percentage of the physicians and patients in the state would be involved, and the patient and community base would be broad. Many of the charges of exceptionality that faced small demonstrations might be averted. Second, the collaboration might be a model that could be adapted to practice settings across the state and country. To discuss these goals, two leaders developed a strategy for approaching others.

> We ended up … identifying organizations that we thought would be helpful to have at the table at that first meeting. In essence, we sent out an invitation to some of the major multispecialty groups in the state. We had some initial meetings and started to say, "Does this make sense?"… We would identify some quality metrics. We would begin a process of getting our respective staffs working together. We would start producing some data. Then we would draw in others. That is how it got started.

He continued:

> An initial charter was produced. There was an initial set of supportive principles that was generated early on and an agreement that we would all put some money on the table so we could get started, organize ourselves, hire some temporary staff.… It began … very informally in that way and then ultimately it created some momentum.

Collaboration to "Lift All Boats"

The mission statement adopted by WCHQ would state that it was "a voluntary statewide consortium of quality-improvement-driven health care organizations learning and working together to improve the quality of health care in the State of Wisconsin." A fundamental precept of WCHQ was that the solution to deficiencies in quality was the sharing of performance data and collaborative improvement against the contrary influences of market competition—also known as "lifting all boats." For members to participate fully and freely and for all patients in the state to benefit, this goal would have to be implemented. The phrase runs pervasively throughout the interviews, as it was one of several important ground rules. Quality and safety of patient care were too important to be proprietary goods.

Ground Rules for Year One. Putting competition aside in the interests of collaboration was not so easy in a context where competitors might use public information against you.

Founders viewed the risk to themselves and their organizations as high. Consequently, two initial decisions were taken. One was to keep the group small for the first year, during which the model would be developed. The second was to initially include only organizations that did not compete with each other. The first choice was a practical one of avoiding too many cooks in the kitchen, while the second was to create an artificially noncompetitive environment. As one interviewee explained it this way:

> We drew a map and we said, here's Dean and St. Mary's in the middle and there, over against the Mississippi River, is Gundersen. There are other people competing in those areas, but the market areas of Gundersen and Dean don't overlap.… It was self-protective but we promised each other that we would report our data, we would not fudge it, we would have it verified, we would make it public and we would not walk away from whatever we found. But at least we wouldn't have competition within a community saying "I'm better than you."

The founding provider organizations were:

ThedaCare, Appleton

Bellin Health, Green Bay

Dean Heath System, Madison

St. Mary's Hospital Medical Center, Madison

Gundersen Lutheran, La Crosse

Marshfield Clinic, Marshfield

St. Joseph's Hospital, Marshfield

Medical College of Wisconsin Physicians, Milwaukee

Froedtert Memorial Lutheran Hospital, Milwaukee

Invitation to Business. The WCQH founders took a surprising additional step, but one characteristic of innovative organizations. They looked outside of familiar networks for ideas. From among those who were pressing for greater accountability, they chose to open discussions with local employers whom they viewed as purchasers, members of their communities, and consultants. The last idea was that health care could learn from industries that had improved quality and efficiency through the standardization of routine processes. Collaboration with business was thus seen as adding expertise to their project while also informing development of measures that would be useful to multiple stakeholders. The

inclusion of outsiders would also affirm that the intention of WCHQ was one of quality reform, not protection from it. The founders agreed that each would bring with them to the table a business purchaser from their local area. Several mentioned choosing a firm that was known to have an active internal quality improvement program.

The first business partners* were as follows:

Appleton Papers, Appleton (paper manufacturing)

Daimler Chrysler Corporation, Kenosha (auto manufacturing)

Sentry Insurance, Stevens Point (life and casualty insurance)

Schneider Trucking, Green Bay (transportation)

The Employer Health Care Alliance Cooperative (The Alliance), Madison

Trane Corporation, La Crosse (air conditioning)

United Auto Workers, Kenosha (auto manufacturing)

Wisconsin Manufacturers and Commerce (statewide business association)

As with much else, they did not know what to expect. One founder said:

> We didn't have a clue really, what was for sure going to happen.... We thought we didn't have data for the purchasers of health care to make decisions to change clinicians but ... we have enough information for them to start giving us feedback and for us to start looking at how we can improve this process.

They did not expect collaboration with purchasers to be easy, but a phrase used frequently by WCHQ members is that when serious people disagree, "the truth lies somewhere in the middle."

Physician-Leadership Ready for Prime Time

A milestone was reached when a report, the only one that would be offered in print, was released at the end of the first year of meetings. Although embryonic in terms of the goals they held for measures, founders hoped the report was sufficient to 1) show the business community and the public that they were serious about transparency; 2) introduce the idea of a bottom-up approach to performance metrics and the achievement of quality; and 3) serve as a basis for new members to join. Said one hospital executive of the image of WCHQ they wanted to convey:

* For current members see the WCHQ website: www.wchq.org

> It was really an attempt to make sure that this was very much led by physicians … was not a hospital thing. As hospitals we were there participating because of our partnership relationship with our physicians … or in some cases, they were integrated into a health system. We wanted to make sure that this was had a very different flavor to it, and that it would have a flavor linked to more on the outpatient side as opposed to the inpatient side.

They released the report with a splash. In October 2003, they held a well-publicized statewide meeting to which the business community, the health care community, the media, and the public were invited. The meeting was well-attended and well-covered by the press. It was a success.

> I think we were given kudos for the fact that we were able to pull it off and do it. I think we only reported on 10 or 15 parameters, and a lot of it was hospital data but [it included] some of physician practice parameters on efficiency and core metrics.… At that point, we opened it to other potential members.…

With the successful presentation of the first report at a public meeting and its positive receipt, test one had been successfully passed.

The Challenge of Growth

Following the public conference, WCHQ repositioned itself as a membership organization open to any entity that accepted its goals and the evolving rules of participation. Immediately, a number of larger, freestanding, multispecialty practices and partner hospitals became members. The initial recruits report that they had known about the formation of WCHQ and had observed it with a combination of interest and concern. One put it this way:

> We were a bit miffed by it frankly, because we didn't think that anything that truly should be collaborative would be based on competitiveness. Frankly, we felt that there was some agenda here for this initial group to sort of put itself in the spotlight, if you will.

However, they wanted to join. As one member of the Phase II group summarized it:

> From our point of view, this seemed like a very different approach, being collaborative, being physician-led organizations.… The idea of developing

> our own measures and talking about those and reporting those together instead of competitively was very appealing. It was sort of, as if, you look for some movement, this one seemed to have a much greater opportunity of having a significant impact. You really do not want to be outside that sort of thing.… There was the sense that having reported the data, there might be a follow-up opportunity to actually improve quality, based on that data.

This comment captures the image of WCHQ that has made it attractive to others. True to its mission, it was perceived as an initiative of health care professionals to collaboratively share knowledge in the interest of mutual learning and the improved performance of all members. Across the interviews, the principles by which WCHQ would operate were understood, endorsed, and embraced. When interviewed in 2007, many were to say that the key task for the larger, more complex, emerging organization was that it must keep the focus on the quality of work that physicians and other health professionals do, in the settings where they do it, for the patients in their care.

Ground Rules of Membership. The expansion of WCHQ in 2004 was not without risk to the attributes that had made its launch possible. The collaborative spirit in the first year had been agreed with a handshake and protected by the exclusion of competitors. As new members faced the prospects of publishing performance data, the fears of old members became the fears of new members: if public reporting showed them in a poor light, it could harm them competitively. Were they possibly not as good as they had assumed? Would their patient demographics put them at a disadvantage? Would their physicians rebel? To help assuage these concerns, WCHQ introduced a code of ethics (see sidebar).

Fears of Public Reporting Dispelled. For the expanded membership of WCHQ, the publication of data using the newly developed ambulatory measures was a moment of learning—and relief. Fears were not realized. In the published data, nobody emerged as better than anybody else across the board, and nobody found their scores consistently at the bottom. The data did not lend themselves to identification of a winner. Rather, they revealed that each practice had many areas of potential improvement and some areas where its teams might be sought as mentors. The results were a vote for the emerging measurement system and for transparency as a device to improve practice.

Code of Ethics

All WCHQ members were asked to affirm their commitment to honest reporting (along with validation) and to the collaborative model.

In support of this, they promised to:

1) Participate in the development of measures that would be clinically accurate and actionable (i.e., useful for evidence-based clinical care and continuous quality improvement). This meant assuring the participation, in governance and workgroups, of individuals/disciplines needed to accomplish various tasks.
2) Agree to report all measures for which submission of data was possible.
3) Not use results for marketing purposes.
4) Share best practices to "lift all boats."

In turn, they would:

1) Benefit from comparative measurement.
2) Draw on the collective experience of WCHQ members to improve quality in their own organizations.

Later Adopters. WCHQ has continued to expand, including the incorporation of organizations and smaller practices that had been initially reluctant to consider membership. Prominent among this group of new members are three very large integrated (or integrating) systems that resulted from rapid hospital consolidation in the state's urban corridor in the 1990s. Among the reasons cited for their initial reluctance was investment in alternative quality reporting activities and the existence of established linkages with other improvement initiatives. Along with problems of modifying existing systems to fit with WCHQ's model, these systems had objections to overlapping data collection initiatives and the financial cost of WCHQ membership (in contrast to reporting systems like WHA's Checkpoint that do not require separate dues). At least two late joiners would probably not have done so were it not for pressure from the region's business coalition, which, like employer alliances in other parts of the state, had embraced WCHQ as a positive initiative that was producing heretofore absent performance data and quality gains.

Inevitably, the membership had cause for concern about incorporation of organizations that might be culturally dissimilar and had joined under pressure. Historically, the new organizations had less collaborative relationships with their physicians than did the multispecialty practices, something that may hinder easy adoption of the WCHQ model. Consistent with the need to improve physician-organization alignment, however, these organizations have recently raised the profile of physician leadership. For example, one

system that was formerly led by a CEO who oriented aggressively to market share recently appointed a successor who is a physician.

Once again, the fears have not been fulfilled. New members are reported to be submitting their data promptly and with a supportive attitude. In interviews, new members qualify their remarks as based on short tenure but indicate support for the principles of the organization and optimism about the results of membership. In a vote of confidence for WCHQ, one of the large systems is exploring ways to expand the numbers of physicians reporting to include affiliated independent practices as well as employed physicians.

WCHQ's expansion is also at the other end of the spectrum: one new member has joined that is the smallest practice to have done so to date. Only recently has this multispecialty practice attained its current size of 60 physicians, up from fewer than 20 a decade ago, while also partnering with a local hospital to create a small integrated system. This practice lacks the staff resources of larger organizations but views itself as a "like-minded" organization in having previously mounted an extensive internal quality improvement effort, including local development of performance measures against national protocols. Its primary reason for joining was anticipation of collaboration with other WCHQ members.

Means of Physician Engagement: Lessons from Wisconsin

The WCHQ model is built on three pillars judged to be requirements of physician support:

1) Quality of patient care must be seen as the preeminent goal.
2) Data must be respected by the physicians as scientifically sound: meticulous in measurement and attentive to dangers of bias, sample sizes, and units of analysis.
3) The process of generating and using measures to improve care must be collaborative to tap the knowledge and energy of the frontlines and to engage their commitment.

Focus on Quality: Transparency as A Tool of Patient Care. The focus on quality lies at the heart of WCHQ's success. Its challenge is to successfully focus diverse professionals in multiple worksites on common metrics that truly measure what they are doing and to develop strategies to improve it. In asserting and organizing itself around the improvement of quality, WCHQ is described as creating a sense among doctors and other professional workforces "that somebody finally has the priorities straight." As a means to engage professional workers in change, this core idea cannot be overemphasized.

The CEOs committed themselves to large personal roles that included traveling from disparate parts of the state for face-to-face monthly meetings, which, for

some, required a whole day. They were surprised to realize how satisfying they found participation, describing it as restoring a sense of purpose. One founding member said:

> I think in many ways WCHQ became an oasis from this highly competitive environment … a safe harbor where we are not talking about market dominance and control. We are talking about quality.… We said: "Frankly, these have been very enjoyable discussions [because] … we are focusing in on quality measurements, we are talking about improvement. We are trying to develop some best practices. And, we are not talking about how do we compete against someone.… The dynamics are very different than anything else that is going on in health care today.

A chief medical officer (CMO) whose organization joined in the second round said:

> [I enjoy] the sense of collaboration, and what is kind of fascinating, is that the discussion—of how are we doing, how are we doing relative to each other, how can we do better—constantly brings you back to your primary purpose and that is the patient you are taking care of. You are not comparing hardly any financial data, core size, or anything. The only thing we are comparing is how our patients are doing and what measures are useful to tell us how they are doing.

The focus on quality is described as key to successful physician engagement within the ranks as well. One organization's quality officer describes how the CMO has championed performance measurement in his organization:

> He always understands where the physicians are, kind of what is important to them, you know—patient care. He goes back and says how do we want to put this into our workflow? He understands many of the dynamics: "What? I don't have enough time to measure more things, I'm already stretched." He also understands the financial piece … how if you see more patients, you get more patients in an hour and make more money, but that doesn't give you enough time with each patient.… So, it is really balancing that act. You want to be an organization that does well financially. *More importantly, you also want to be a hospital or clinic that gives the best patient care. He takes the time to put it as a priority.* [Emphasis added.]

Filling a Knowledge Vacuum. Within this context, public reporting becomes a tool of fulfilling medical goals. It fills a vacuum in the physician's knowledge as described by a hospital CMO:

> I think that one of the constructs that the Collaborative built on is that physicians want to do a good job … And many physicians … practice in a knowledge vacuum. That is, they really don't know how their diabetics are doing compared to even a person who is practicing one door down, not to mention in other parts of the state or in other parts of their own market place.… By providing information about how physicians each perform you can influence physician's behavior. I would go so far as to say that there is *no other way to influence physician's behavior.* They are driven to change things when their performance does not look good.

For this change to happen, however, physicians had to believe the performance scores were accurate. In addition, WCHQ founders recognized the results also had to be clinically useful to drive improvement.

A Bottom-Up Approach. WCHQ's founding leaders directed that measures should both align with nationally recognized standards and existing data sets, and be simple enough to be acceptable to physicians in all practice settings—including those with limited staff and information technology. The technical achievement of the WCHQ measurement system—measures, data collection, and validation of data—is beyond the scope of this study. The WCHQ measures and methodology for data capture and entry are briefly described in Appendices 1 and 2 (from the WCHQ website).

The development of ambulatory care measures held within it the underlying need to align WCHQ standards, drawn from national guidelines and adopted at the level of the state, with local realities where ways of accomplishing goals differed. This need presaged the one down the road to achieve change in these same diverse settings, something that is recognized as requiring local adaptation. [7]

The Unit of Reporting: Strong Feelings

Medical Care as Teamwork. WCHQ members currently limit reporting at the level of the organization, not the individual clinician. The predominant view is that isolating physicians or any element of care contradicts a systems and team approach to health care delivery. In the eyes of providers, this approach leads to an inaccurate portrayal of care and an unsound basis for problem-solving. A cogent summary of this conviction was offered by this CMO:

> A number of us believe that when you are working in a system, like almost everybody … is, you should be reporting it at the system level because it is a collaborative project. It is not an independent product. That is, it is not one disease taken care of by one doctor with one outcome. It is usually a group. It is a disease cared for by a group of people, hopefully more efficiently, with an outcome.

The clinical executives also gave specific examples of the problem of isolating responsibility and hence of physicians' complaint that assignment of responsibility is often wrong. One said:

> There is a philosophy that if you report on the individual doctors data is the greatest motivator for change.… The dilemma is very often that … the data used to attribute it to an individual physician is highly flawed because the patient may see multiple providers of care for their particular condition.… So whose data is it?

Similarly:

> When we talk about the surgical infection prevention measures, we call them physician measures—you know, the timing of the antibiotics and such—but when we do our root cause analysis and our six sigma project on it, it is actually more complicated. The physician can *write* "within one hour of incision time" but it is the OR's staffing and schedule that decides when that surgical incision occurs. They actually do not have control over getting that in within an hour.…

Avoiding Negative Responses. The physician sponsors of WCHQ see in the assignment of responsibility for outcomes to particular individual physicians the potential to elicit undesired behavior. In a context where physicians see themselves penalized individually for team outcomes, they may exercise their considerable power to accept or refuse responsibility for difficult patients, or "cream skimming," to achieve a favorable individual score and attendant benefits.

Desired Result. The policy of reporting at the group level responds to several key points made by the physicians, including: 1) at the level of the individual, numbers become unreliable (a single aberrant case can alter the outcome); 2) assignment of responsibility for particular outcomes is difficult and frequently arbitrary; and 3) health care should

properly be viewed not an individual product but as an outcome of the actions of many different individuals and the systems that support them. When this policy is recognized, interlinked networks of professionals can focus attention on clarifying and improving formal and informal linkages of care coordination. In sum, the preference for group data is methodological, philosophical, and practical.

Local Provider Sponsorship

By virtue of WCHQ's strong physician leadership, its measurement model is bottom-up and collaborative rather than top-down and imposed. Working within a context of a shared desire to improve quality through data sharing and collaboration, interviewees report that WCHQ is respected rather than resisted, embraced as constructive rather than opposed as harmful, and seen to be "ours not theirs."

WCHQ has guided and enabled the development of measures that accurately reflect local clinical care, thereby obviating accusations that measures do not fit clinical practice or local reality. It has achieved voluntary cooperation. Although the author does not have interviews with clinicians, one longtime CMO said persuasively of his experience: "I was fully prepared to be assassinated but I have never gotten an angry response from [any] physician about public reporting of our performance—not yet, not once."

This model contrasts with the perceived motivations of other stakeholders. Insurers are an important example. In the view of many physicians, comparative performance reporting by insurers is inherently an imperfect proxy for actual care delivery, because their data systems are designed for claims payment or competitive advantage rather than clinical purposes. Insurers also may be more willing to accept errors that validation would have caught so long as overall economic goals are met (e.g., curtailment of the number of high-cost physicians billing the company). And finally, because they are private organizations competing in the marketplace, methods of performance calculation remain proprietary—not available for the review scientists expect and physicians seek.

As for the public sector, government has the advantage of not being burdened with distracting profit goals. However, interviewees agree that the Quality Improvement Organization (QIO)—and even The Joint Commission—are unable to cast off the onus of external regulation. To a question regarding WCHQ's rapid success in achieving the support and momentum that had so long eluded QIO, one interviewee said of the failure of the latter:

> … because it [QIO] was regulatory and took a regulatory approach. Same thing with JCAHO [The Joint Commission on Healthcare Organizations]. Because we have JCAHO, our hospitals should be safe, but JHACO is big brother coming in and the response is usually: "Oh there is a JCAHO visit. We will get ready for the JCAHO visit that we will have to pass." Then they leave and they don't come back for 5 years and until the next one comes along nobody thinks about them.

A Structure for Collaboration

The means of achieving goals in WCHQ are multi-tiered networks of collaborating units—workgroups connecting in horizontal networks within which executives communicate at one level, those engaged in measures development at another, physicians contacting other physicians at a third. While pursuing goals identified at the state level that accord with national standards, these networks link them to local realities.

The link to localities allows the development of measures appropriate to all practice locations, enables true benchmarking ("apples to apples"), and structures collaborative learning. With only one exception, the perception of interviewees was that learning opportunities transcended organizational boundaries, with professionals seeking the best solutions to identified gaps in performance.

A constant reiteration is that comparison of scores leads to both competition and collaboration. Unlike in the marketplace, this tenet is expected within the knowledge professions. Evidence of a performance gap is followed by contacting the better performer viewed as, and responding as, a mentor. The following type of statement occurs repeatedly in the interviews, with this example from a CMO:

> WCHQ [provides] the actual benchmarking data for looking at where you are at, and how to improve, and building those connections with other organizations that are similar to you; where you can say "our numbers are not good here, how did you get yours better? What can we learn from how you are doing it?"

A number of interviewees pondered WCHQ's key elements' relative importance to success, but most refused to choose a particular one. Of the success of WCHQ, one CMO posed the question as to whether it was the reporting or the participation that had most contributed to observed quality gains, saying this:

> We were able to bring together groups that in some environments would have been looked at as competitors, to say we are going to put our competition piece behind and aside. Our job is to figure out what it is we can do together to use our information to make the care better.

Diverse Improvement Strategies

The interviews reveal significant diversity from one system or organization to the next as they autonomously developed the interventions they would use for improvement. Common to all strategies was a first step of delivering the data to the physicians and then, typically examining them score by score, and patient by patient. Attachment of patient data to individual physicians at this level was not viewed as a betrayal within the group because perceived errors (e.g., assignment of responsibility, small numerators, patients who are *actually* sicker) can be identified and discussed with a goal of identifying points of system breakdown and solving problems. A principal core is that the data should be used to identify areas of needed improvement and the means to do better.

The CMO of a large independent practice reported that the goal of motivating the doctor to improve—not to evade or cheat—is the reason for both WCHQ's group-level reporting and a supportive change environment within individual practices.

> I think one of the things that is important is the fact that the data that we are reporting publicly is organization data—how is the organization doing as a whole, not how is "Doctor A" doing. Because … when you start getting down to the individual practice level is where there are more legitimate arguments about why something is different in somebody's practice.… We really offer tools and assistance to everybody to raise the entire group rather than to go the QA approach to go step on the poor performers.…

At this point, however, the differences between member organizations are crucial and undocumented. A reservoir of learning is available to be tapped into about what strategies are useful in which environments. Interviewees in well-integrated systems report that for the primary care measures that have been emphasized by WCHQ to date, the physicians play a small role in effecting improvements. Others in the system are better analysts of system problems than are clinical practitioners and, with only a nod from physicians, support staff can implement many corrective steps.

Systems that are this well-integrated remain in a minority, however. Members have different system capacities and are working with physicians with whom they have a variety of arrangements (employed, contracting, referring, and admitting). The idea of data as a tool to a shared goal of improvement is sustained, but change strategies are diverse and evolving. One system's quality leader described the challenge of achieving desired change in the many small practices that are the constituent elements of this organization:

> Having the medical director is helpful but when we embark upon a new project, like diabetes or whatever, we need a champion in order to carry it forward. That is hard sometimes because you have multiple individual practices. You can have a champion for one practice who does not really have a relationship with another. Our job in quality is to provide the individual practices with the resources they need to make the quality changes, but yet allow them enough autonomy so it can be adopted.

The situation also arose of a practice whose CMO has a quality staff of one individual. As at the other site, the quality manager described the usefulness of comparative data for motivating the physicians and generating communication among them, but through mechanisms quite different from those described by large integrated systems:

> We created a "frequently asked question" sheet … so that they are able to share with each other and say "Oh gosh, I had that question, but I never asked."

In sum, WCHQ members and their strategies for behavior change are diverse. What they share is a conviction that the first step toward improvement is data that physicians believe identify genuine problems. The second step is that collaboration ensues: what can you learn from colleagues in your group and the other organizations? What the groups are learning at this level remains for another research project.

Other Users of Public Reporting

To date, providers themselves and large purchasers of health care have been the primary users of data. Interviewees expressed great optimism about achieving alignment of WCHQ and WHA's reporting system(s). Nevertheless, tension persists between providers and business partners, who regard a purpose of performance data to be "economic credentialing" and provider selection, a problem described by a founding hospital executive:

> The business partners are not as active as they used to be. The motivations are a little bit different in that they are looking for data that allows them

> to identify who is the lowest cost provider, which, is different from the motivation of the Collaborative to bring everyone up to a level.… So I think it has been helpful to have some of the business leaders at the table but I think they are frustrated that we haven't clearly been able to identify who is the most efficient provider because, I believe, that isn't our motivation.

The success with consumer-responsive measures is mixed. Consumer representatives successfully advanced the case for expanding measures from the clinical concerns of physicians (and even the encounter-focused measures of "patient-centeredness" developed by others) to indicators of the ease of accessing and using the services. Consumer input is responsible for a WCHQ measure reporting "time to third-next available appointment."

At the same time, WCHQ has elected not to report its measures in formats that will allow consumers to choose physicians (e.g., five-star surgeons or a table of scores such as that provided by *Consumer Reports)*. Again, the unit of analysis is an issue with providers objecting to extracting the physician from the interacting elements of the medical care system. Providers hope to persuade consumers that simple scores do not serve them well while, at the same time, trying to discern what would help consumers in their decision-making. (No consumers other than business partners were interviewed in this project.)

Missing Stakeholder. Donald Berwick, president and CEO of the Institute for Healthcare Improvement, has suggested that the public is less interested in consumer information than in evidence that those they are paying to deliver high-quality health care in their communities are accountably doing what they are paid for.[8] If this boat is leaking, are you fixing it, and do you have with you the resources of the state? Community priorities might be quite different from those of individual consumers. Community representatives, for example, might wish to know the quality of *local* diabetes care and/or the progress of local providers on other goals such as reducing health care disparities. Despite the historically close relationship of nonprofit health care organizations to local communities (albeit one that is now strained, as illustrated by challenges to the tax status of hospitals), and frequent mention in the interviews of the responsibility to local communities, no initiatives currently go beyond data relevant to individual consumers.

The local community is very likely a beneficiary of a collaborative effort dedicated to mutual assistance, something expressed ironically by an executive (who characterized himself as a layperson) who found this result somewhat perplexing:

> The sense I have [of WCHQ] is that there is a very strong, predominantly physician, coordinated board … [composed of individuals with] a very strong medical commitment to making this work.… I think the more lay administrators there are in the room, the more we will be uncomfortable sharing data potentially.… It may not present us in the best possible light to another hospital in town and perhaps they might use that against us in some way.

He concluded, however: "The mentality that I see here does, I think, achieve more for the community."

Business Case for Quality

Interviewees believe that savings will be generated if medical care is evidence-based, error-free, efficient, and effective. They expect long-term cost savings from standardization of routine processes, enhanced care coordination, timely treatment, attention to correctable problems such as poor compliance, and avoidance of error and iatrogenic illness. While recognizing the tendency of the fragmented U.S. payment system to focus on short-term rather than long-term financial returns, they believe that such savings will benefit purchasers of health care.

The most easily quantifiable cost associated with participation in WCHQ is the annual membership dues, which are currently undergoing reappraisal in light of debate about options for organizational growth and an increasingly diverse and consolidated membership. At one extreme are multi-hospital systems that have entered their ambulatory practices into the database but balk at the cost of entering multiple hospitals, particularly given the compulsory submission of data to the Joint Commission, the QIO, the Centers for Medicare and Medicaid Services, and insurance companies. At the other extreme are small physician groups for which the financial and human costs of participation are a recognized barrier to joining.

The true cost of participating is pragmatically difficult to isolate, partly because it depends on evolving skills and information technology infrastructure, but more importantly because member organizations pursue multiple goals simultaneously. In these interviews, WCHQ was not measured against one goal alone but for its potential impact on many. Estimates of the in-kind cost of participating in WCHQ ran as high as a half-million dollars per year given the cost of regular data collection, validation for submission, and deploying highly skilled staff to WCHQ workgroups. Despite such estimates, executives often represented WCHQ as essentially being "no cost," because it was a redeployment of staff from less effective to more effective means of achieving the same

goal—measuring and improving quality. Additionally, WCHQ encourages such intangibles as professional standards, collegiality, good morale, and personal growth that are often described as in decline.[9] Many interviewees believe that WCHQ is not only producing results where others have failed but also promises long-term savings because it is creating a positive culture and nurturing internal strengths. Speaking of the cost of membership in WCHQ, one CMO said:

> There is a lot of professional development that you would pay a fair amount should you do it outside of WCHQ. You'd pay a fair amount if you sent three people to the IHI conference.... And I am not talking just money.... It is also a great way of getting more people involved, so, you can start growing your talent for this area—and physicians are included in that. We can introduce physicians into quality in a way that is not beat 'em over the head with it. It isn't intimidating. It is an extension....

In other words, WCHQ provides a vehicle by which quality improvement gained a structure conductive to both sustainability and internal self-generation.

CONCLUSIONS AND IMPLICATIONS

WCHQ's success in linking evidence-based measures to local practice settings rests on physician understanding of quality, clinical processes, local practice, and physician approaches to learning and behavior change. In this physician-led, bottom-up model, transparency emerges as a tool that can motivate and empower collaborative problem-solving at the local level to accord with national guidelines. The multi-tier networking coordinated by WCHQ at the state level comprises a corporate model of professional accountability that has been heretofore elusive.

Comparative performance measurement within WCHQ is legitimated by:

1) Insistence that performance measures are:
    a) scientifically sound;
    b) aligned with nationally recognized standards;
    c) relevant to clinical decisions ("actionable");
    d) believably related to local practice;
    e) produced in a transparent and reviewable process; and
    e) generative of collaborative solutions.

2) An ethos of quality improvement.
3) A bottom-up approach that is more efficient and effective than efforts imposed from outside (managed care organizations or government-sponsored entities) and which harnesses local knowledge and creativity to the process of harmonizing local care to statewide goals.

Crucial to these achievements were:

4) Physician leadership in building and sustaining commitment and support, including the human and financial resources, to devising and continually enhancing methods of performance measurement and data collection.
5) A structure of multi-tier deliberative networks for generating common goals while protecting organizational autonomy in using results.
6) Inclusion of multi-stakeholder collaborators for a creative tension that moves all parties forward in pursuit of a shared goal.

The expansion of WCHQ reporting to include small practices will require the creation of methods for aggregating clinics or the development of strategies that reflect the unique cultural and organizational attributes that are characteristic of small practice settings. These actions might include the reinvigoration of indigenous entities such as county medical societies or the deployment of "quality coaches" who assist frontline practitioners and staff in the collection of data and calculation of comparative performance information.[10]

The balance between community and purchaser priorities is not a current topic at WCHQ, although community involvement is consistent with public obligations as perceived by health care executives. Placing a greater emphasis on this issue might represent a positive response to the community-benefit challenges being leveled against health care organizations nationwide. Techniques for identifying community priorities developed in Canada and Europe could serve as a source of indicators.

Critics have faulted enthusiasts of collaborative methods of quality improvement for failing to undertake sound and sophisticated research into claimed successes.[11] WCHQ has created a valuable laboratory for such research. The statewide collaborative structure pioneered by WCHQ offers a unique quasi-experimental situation in which general goals and data are controlled at the state level, but autonomous practices located in diverse settings select and implement priorities and interventions. Further, a feedback loop exists from sites of implementation to WCHQ and thereby to national generators of evidence-based measures and guidelines.

WCHQ has benefited immeasurably from the widespread perception of it as an organization committed first and foremost to quality of care in member organizations. As new opportunities are pursued, WCHQ must assure that this core focus on quality is protected.

It is desirable to confirm and elaborate findings of this study by interviews with 1) quality staff and others who collaborated in the workgroups to create measures that harmonized national indicators to local realities; and 2) physicians and other health care professionals who are targets of performance measurement and change.

Overall, physician-led development of performance measures and quality collaboration are strongly supported by this study as means to long-term sustainable quality improvement. Gains have been recognized by a number of publicly reported measures and by employers, purchasers, professional societies, and others concerned with health care quality. A recommendation following from this case study is for these groups to encourage and facilitate leadership by physicians along the lines demonstrated in Wisconsin.

NOTES

[1] Wisconsin Collaborative for Healthcare Quality, www.wchq.org.

[2] S. Lewis, "Toward a General Theory of Indifference to Research-Based Evidence," *Journal of Health Services Research and Policy,* July 2007 12(3): 166–72; D. Berwick, "Disseminating Innovations in Health Care," *Journal of the American Medical Association,* Apr. 16, 2003 289(15): 1969–75; A. L. Greer, J. S. Goodwin, J. L. Freeman, and Z. H. Wu, "Bringing the Patient Back In: Guidelines, Practice Variations and the Social Context of Medical Care," *International Journal of Technology Assessment in Health Care,* Fall 2002 18(4):747–61.

[3] C. M. Kilo, "A Framework for Collaborative Improvement: Lessons from the Institute for Healthcare Improvement's Breakthrough Series," *Quality Management in Health Care,* Sept. 1998 6(4):1–13; L.V. Rubenstein, B. S. Mittman, E. M. Yano, and C. Mulrow, "From Understanding Health Care Provider Behavior to Improving Health Care: The QUERI Framework for Quality Improvement," *Medical Care,* June 2000 38(6 Suppl. 1): I-129–I-141.

[4] R. Cunningham, "Professionalism Reconsidered: Physician Payment in a Small-Practice Environment," *Health Affairs,* Nov./Dec. 2004 23(6): 36–47.

[5] A. L. Greer, "The State of the Art vs. the State of the Science," *International Journal of Technology Assessment in Health Care,* 1988 4(1):5–26; A. L. Greer, "The Two Cultures of Biomedicine, *Journal of the American Medical Association*, Nov. 20, 1987 258(19):2739–40.

[6] E. Rogers, *Diffusion of Innovations*, 4th ed. (New York: The Free Press, 1995).

[7] J. E. Fielding and P. A. Briss, "Promoting Evidence-Based Public Health Policy: Can We Have Better Evidence and More Action?" *Health Affairs,* July/Aug. 2006 25(4):969–78; D. Berwick, "Disseminating Innovations in Health Care," *Journal of the American Medical Association,* Apr. 16, 2003 289(15): 1969–75.

[8] R. Galvin, "'A Deficiency of Will and Ambition': A Conversation with Donald Berwick," *Health Affairs,* Jan. 12, 2005 (Web Exclusive): W5-1–W5-9, 2005; R. A. Berenson, "Which Way for Competition" None of the Above," *Health Affairs,* Nov./Dec. 2005 24(6): 1536–42.

[9] D. Mechanic, "Physician Discontent: Challenges and Opportunities," *Journal of the America Medical Association,* Aug. 20, 2003 290(7): 941–46; L. H. Aiken, "Improving Quality through Nursing," in *Policy Challenges in Modern Health Care*, ed. D.Mechanic, L. B. Rogut, D. C. Colby, and J. R. Knickman, (New Brunswick, NJ: Rutgers University Press, 2005): 177–88.

[10] E. S. Fisher, D. O. Staiger, J. P. W. Bynum, and D. J. Gottlieb, "Creating Accountable Care Organizations: The Extended Hospital Medical Staff," *Health Affairs,* Jan./Feb. 2007 26(1):w44–w57; K. Davis and S. Guterman, "Rewarding Excellence and Efficiency in Medicare Payments," *Milbank Quarterly,* Sept. 2007 85(3):449–68.

[11] B. S. Mittman, "Creating the Evidence Base for Quality Improvement Collaboratives," *Annals of Internal Medicine,* June 1, 2004 (140(11):897–901; T. Wilson, D. Berwick, and P. Cleary, "What Do Collaborative Improvement Projects Do? Experience from Seven Countries," *Joint Commission Journal on Quality and Safety,* Feb. 2003 29(2):85–93.

APPENDIX 1: METHODOLOGY

This case study rests primarily on 31 personal interviews conducted from June through September 2007 with individuals actively involved with WCHQ. Twenty-five interviews were completed with high-level executives of multispecialty practices, hospitals, and integrated health systems that participate in WCHQ. Also interviewed were business partners (N=2) who serve on the WCHQ board of directors; the CEOs of the Wisconsin Hospital Association (WHA) and the Wisconsin Medical Society (WMS) who serve on the board *ex officio;* and WCHQ staff (director of operations and president/CEO). The project sought to interview executives from all member health care organizations whether or not serving currently on the board. This goal was largely achieved, as only one physician practice was not interviewed.

Interviewees included eight CEOs, ten chief medical officers, and five executives who are responsible, under other titles, for quality of care within their organizations. Fifteen of the executives interviewed were physicians who had spent many years in clinical practice before assuming management roles. Their medical specialties were internal medicine (N=6); critical care medicine via pediatrics, surgery, and pulmonary medicine (N=3); pediatrics (N=3); cardiology (N=1); family practice (N=1); and psychiatry (N=1).

Interviews were conducted by the author at the offices of the interviewees. A conversational "focused interview" invited subjects to address common themes but to formulate answers from their own perspective and using their own language. A typical interview lasted one-and-a-half to two hours. Each was audiotaped, transcribed verbatim, and entered into a software program that permits retrieval of text across a database. Coding aims to capture all comments on a given theme irrespective of interview position or prompt so that formulations of the same material may be retrieved from all interviews and compared. Analytic techniques are those of grounded theory wherein the analyst works iteratively between conceptualizations and data to identify central tendencies, interrelationships, and exceptions to the rule.

The methodological rigor of inductive research, in contrast to deductive variable-testing research, lies in close examination of all data relevant to an evolving characterization of attitudes or events. This definition means refusal to discard negative evidence, insisting instead that it be accommodated in revised statements. Instances selected for presentation

in this report should be viewed as exemplars of consistent patterns unless otherwise stated. A statement that "interviewees report . . ." indicates a consensus opinion, for example, that physicians will not accept or use performance data they judge to be methodologically deficient. A statement that "many" or "some" interviewees offered particular evidence indicates that while this evidence was presented in only a portion of the interviews, it was not contradicted elsewhere. An example is that "Physicians presented with shortfalls in clinical performance turn to peer mentors for solutions." Actual contradictions necessitate that the analyst abandon this trial statement or revise it to accommodate the deviant evidence. For example, two types of change strategies rather than one may be described along with a suggested explanation for the variation.

The existence of variation within a single case lies at the root of the "theoretical generalizability" that is the goal of inductive research. If, for example, one notes that "Selection of practice change strategies was different in more and less integrated organizations," it is possible to hypothesize what may be expected in circumstances not encompassed in the particular case. This form might include, for example, "what might occur in organizations that are more integrated than the ones studied in Wisconsin" (e.g., Kaiser). So too, the strategies used in the more loosely coupled organizations in WCHQ may offer clues to efficacious models in settings where care coordination depends on physician referral networks and common hospital affiliation.

Limitations

Limitations include the possibility that WCHQ, as a pioneering effort, benefited from a "Hawthorne effect." This possibility cannot be removed without comparative case studies and/or longitudinal data. A second limitation is the inability within the scope of this project to verify that things actually worked as perceived by top executives. The enormous consistency in reports of experiences across many provider organizations lends confidence to major assertions, but three groups of participants are missing. To identify problems and solutions that may not have reached the attention of top executives, interviews are needed with: 1) the professional quality staff who collaborated with physicians and each other in the development of measures satisfactory across sites; 2) the members of clinical and managerial committees within member organizations who, after receiving WCHQ data, identified organizational priorities and implemented change strategies; and 3) frontline clinicians whose work was the object of public reporting.

APPENDIX 2: WISCONSIN COLLABORATIVE FOR HEALTHCARE QUALITY

Source: www.wchq.org

Physician Group Measures Developed by WCHQ

- Breast Cancer Screening: women who should have had at least one mammogram within the previous 24 months.
- Cervical Cancer Screening: women who should have had one or more cervical cancer screening test during the previous 36 months.
- Colorectal Cancer Screening: women and men, at age 50 and older, who should have had a colorectal cancer screening; the optimal interval for screening depends on the test, i.e., fecal occult blood testing, colonoscopy, flexible sigmoidoscopies, and double contrast barium enemas.
- Diabetes–Blood Pressure Control: the percentage of patients 18 to 85 years of age whose most recent blood pressure reading is controlled to a rate of less than 130/80 mmHg in the measurement period.
- Diabetes–Blood Sugar (A1c) Control: the percentage of patients 18 to 85 years of age who have had a diagnosis of diabetes, and whose most recent A1c blood sugar levels were at good control (less than 7.0%); fair to poor control (greater than or equal to 7.0% and less than or equal to 9.0%); uncontrolled (greater than 9.0%); or not tested in the measurement period.
- Diabetes–Blood Sugar (A1c) Testing: the percentage of patients 18 to 85 years of age who have had a diagnosis of diabetes and had two or more A1c tests greater than or equal to 60 days apart, one A1c test, or no A1c tests in the measurement year.
- Diabetes–Kidney Function Monitored: the percentage of patients 18 to 85 years of age who have had a diagnosis of diabetes and were screened and/or monitored for kidney disease in the measurement year.
- Diabetes–LDL Cholesterol Control: the percentage of patients 18 to 85 years of age who have had a diagnosis of diabetes and whose most recent LDL cholesterol tests show good control (less than 100 mg/dl); fair to poor control (greater than or equal to 100 mg/dl and less than 130 mg/dl); uncontrolled (greater than or equal to 130 mg/dl); or not tested in the measurement period.
- Diabetes–LDL Cholesterol Testing: the percentage of patients 18 to 85 years of age who have had a diagnosis of diabetes and received a LDL-cholesterol test in the measurement year.
- Postpartum Care: women who have had a live birth delivery and a postpartum visit 21 to 56 days after delivery.

- Time to Third-Next Available Appointment: the patients' ability to seek and receive care with the provider of their choice, at the time they choose, regardless of the reason for their visit; counting the third-next available appointment is the health care industry standard measure for access to care and indicates how long a patient waits to be seen.
- Uncomplicated Hypertension–Blood Pressure Control: the percentage of patients greater than or equal to 18 years of age who have had a diagnosis of uncomplicated essential hypertension, and whose blood pressure was adequately controlled (less than 140/90 mm Hg).

Hospital Measures Developed by WCHQ

- Heart Attack Care—Hospital Charges and Quality Comparison: a quadrant analysis representing a comparison of acute myocardial infarction charges and quality of care. The charges are risk-adjusted to account for differences in patients, such as their severity of illness and risk of death. The quality score is a composite number that takes into account how well a hospital performed in giving the recommended care proven to give the best results to most adults with a heart attack. The analysis is an attempt to quantify the value each hospital provides when caring for patients with heart attacks.
- Heart Attack Care—Hospital Length of Stay and Quality Comparison: a quadrant analysis representing a comparison of acute myocardial infarction length of stay and quality of care. The length of stay is risk-adjusted to account for differences in patients, such as their severity of illness and risk of death. The quality score is a composite number that takes into account how well a hospital performed in giving the recommended care proven to give the best results to most adults with a heart attack. The analysis is to attempt to quantify the value each hospital provides when caring for patients with heart attacks.
- Heart Failure Care—Hospital Charges and Quality Comparison: a quadrant analysis representing a comparison of congestive heart failure charges and quality of care. The charges are risk-adjusted to account for differences in patients, such as their severity of illness and risk of death. The quality score is a composite number that takes into account how well a hospital performed in giving the recommended care proven to give the best results to most adults with heart failure. The analysis is an attempt to quantify the value each hospital provides when caring for patients with heart failure.
- Heart Failure Care—Hospital Length of Stay and Quality Comparison: a quadrant analysis representing a comparison of congestive heart failure length of stay and quality of care. The length of stay is risk-adjusted to account for differences in

patients, such as their severity of illness and risk of death. The quality score is a composite number that takes into account how well a hospital performed in giving the recommended care proven to give the best results to most adults with heart failure. The analysis is an attempt to quantify the value each hospital provides when caring for patients with heart failure.

- Knee Replacement—Hospital Charge: the average severity adjusted charge for patients who have had a total knee replacement surgery (procedure code 81.54). Only one knee was operated on. The charges are risk-adjusted to account for differences in patients, such as severity of illness and risk of death, using 3M™ APR™ DRG software.
- Knee Replacement—Hospital Length of Stay: the average severity-adjusted length of stay for patients who have had a total knee replacement surgery (procedure code 81.54). Only one knee was operated on. The length of stay is risk-adjusted to account for differences in patients, such as severity of illness and risk of death, using 3M™ APR™ DRG software.
- Normal Vaginal Delivery—Hospital Charge: the average severity charges for patients who have had a normal vaginal delivery (DRG 372, 373). The charges are risk-adjusted to account for differences in patients, such as severity of illness and risk of death, using 3M™ APR™ DRG software.
- Normal Vaginal Delivery—Hospital Length of Stay: the average severity-adjusted length of stay for patients who have had a normal vaginal delivery (DRG 372, 373). The length of stay is risk-adjusted to account for differences in patients, such as severity of illness and risk of death, using 3M™ APR™ DRG software.
- Pneumonia Care—Hospital Charges and Quality Comparison: a quadrant analysis representing a comparison of pneumonia charges and quality of care. The charges are risk-adjusted to account for differences in patients, such as severity of illness and risk of death. The quality score is a composite number that takes into account how well a hospital performed in giving the recommended care proven to give the best results to most adults with pneumonia. The analysis is an attempt to quantify the value each hospital provides when caring for patients with pneumonia.
- Pneumonia Care—Hospital Length of Stay and Quality Comparison: a quadrant analysis representing a comparison of pneumonia length of stay and quality of care. The length of stay is risk-adjusted to account for differences in patients, such as severity of illness and risk of death. The quality score is a composite number that takes into account how well a hospital performed in giving the recommended care proven to give the best results to most adults with pneumonia. The analysis is an attempt to quantify the value each hospital provides when caring for patients with pneumonia.

APPENDIX 3: WCHQ PERFORMANCE DATA: STANDARDS, METHODS OF SUBMISSION, VALIDATION

Source: www.wchq.org

- WCHQ's current and planned measures all align with nationally endorsed measure sets (AQA, NQF).
- Each measure is specified using a series of narrative documents and flowcharts for standard interpretation and use.
- Data are submitted by member organizations through a web-based interface designed in consultation with data submitters.
- WCHQ's data warehouse architecture is dynamic and scalable to support any number of measures and reporting entities without the need to modify the reporting platform. In other words, as currently configured, WCHQ's system has the capacity to collect performance results for one measure or thousands of measures from one entity or thousands.

The key components of the data submission process include a secure, online data submission application; a dedicated Clinical Information Coordinator, who ensures that all results are accurate and valid; and a board-appointed Performance Measures Audit Committee which, modeled on similar committees serving public companies, is charged with overseeing the audit and validation process and certifying results from reporting entities.

Member organizations harvest their data according to a detailed measure flowchart, which specifies all data criteria, including the three questions used to construct the denominator. Physician groups often construct the needed data from administrative databases, electronic data capture, and manual chart review, if necessary. Results are submitted to WCHQ using an innovative, secure, web-based application.

In WCHQ's model, the measure's numerator is identified first. Examples include:

- Mammography every two years (process)
- Blood sugar (A1c) screening every year (process)
- Blood sugar (A1c) result (outcome for diabetes)
- Blood pressure (outcome for hypertension)

For each numerator, parameters of the denominator are carefully constructed to define relevant patients cared for by a physician group. For this denominator to be meaningful,

it must identify every such patient for every payer, not just those who were served by a particular payer. For each measure, WCHQ has identified a standard set of three questions to guide construction of the denominator. When answered, a finite, clearly defined group of patients emerges—the denominator for this measure.

Why Use This Model?

- Data collection efforts are more focused. Because the denominator is defined upfront (e.g., was mammography performed?) data need not be collected about patients who will later be excluded. All exclusions are built into the three-question denominator-building methodology.
- These denominators refute a frequent physician objection of "These are not my patients." Because all qualifying patients are "counted," regardless of payer, physicians can trust that the resulting measures reflect actual performance.
- These carefully constructed, population-based denominators overcome many of the sample size issues insurance companies experience.
- Denominators constructed in this way can also be used as registries of patients with chronic, episodic, or preventive care needs; point out patients who have immediate care needs ("check patient's LDL at tomorrow's visit"); and indicate patients who may need additional attention ("contact patient to schedule follow-up; haven't seen her in 18 months; must check A1c").

RELATED PUBLICATIONS

Publications listed below can be found on The Commonwealth Fund's Web site at www.commonwealthfund.org.

States in Action: A Bimonthly Look at Innovations in Health Policy. Newsletter.

On the Road to Universal Coverage: Impacts of Reform in Massachusetts at One Year (June 3, 2008). Sharon K. Long. *Health Affairs* Web Exclusive.

The North Dakota Experience: Achieving High-Performance Health Care Through Rural Innovation and Cooperation (May 2008). Douglas McCarthy, Rachel Nuzum, Stephanie Mika, Jennifer Wrenn, and Mary Wakefield.

Identifying and Evaluating Equity Provisions in State Health Care Reform (April 2008). Brian Smedley, Beatrice Alvarez, Rea Pañares, Cheryl Fish-Parcham, and Sara Adland.

States' Roles in Shaping High Performance Health Systems (April 2008). Catherine Hess, Sonya Schwartz, Jill Rosenthal, Andrew Snyder, and Alan Weil.

State E-Health Activities in 2007: Findings from a State Survey (February 2008). Vernon K. Smith, Kathleen Gifford, Sandy Kramer, Jennifer Dalton, Patricia MacTaggart, and Michelle Lim Warner.

What Public Employee Health Plans Can Do to Improve Health Care Quality: Examples from the States (January 2008). Aaron McKethan, Terry Savela, and Wesley Joines.

Leading the Way? Maine's Initial Experience in Expanding Coverage Through Dirigo Health Reforms (December 2007). Debra Lipson, Jim Verdier, and Lynn Quincy.

State Health System Performance and State Health Reform (September 18, 2007). Karen Davis and Cathy Schoen (commentary). *Health Affairs* Web Exclusive.

Value-Driven Health Care Purchasing: Four States that Are Ahead of the Curve (August 2007). Sharon Silow-Carroll and Tanya Alteras.

Lessons from Local Access Initiatives: Contributions and Challenges (August 2007). Karen Minyard, Deborah Chollet, Laurie Felland et al.

Denver Health: A High-Performance Public Health Care System (July 2007). Rachel Nuzum, Douglas McCarthy, Anne Gauthier, and Christina Beck.

Aiming Higher: Results from a State Scorecard on Health System Performance (June 2007). Joel C. Cantor, Cathy Schoen, Dina Belloff, Sabrina K. H. How, and Douglas McCarthy.