WORKING

PAPER

# Evaluation of the Use of AHRQ and Other Quality Indicators

PETER S. HUSSEY, SOEREN MATTKE, LINDSEY MORSE, AND M. SUSAN RIDGELY

WR-426-HS

October 2006

Prepared for the Agency for Healthcare Research and Quality

This product is part of the RAND Health working paper series. RAND working papers are intended to share researchers' latest findings and to solicit informal peer review. They have been approved for circulation by RAND Health but have not been formally edited or peer reviewed. Unless otherwise indicated, working papers can be quoted and cited without permission of the author, provided the source is clearly referred to as a working paper. RAND's publications do not necessarily reflect the opinions of its research clients and sponsors.

**RAND®** is a registered trademark.

RAND HEALTH

# SUMMARY

Public recognition of health care quality issues has spiked remarkably in the past ten years, driven by a series of high-profile reports from the Institute of Medicine (IOM), the RAND Corporation, and other organizations. [1,2,3,4] These reports showed, among other facts, that preventable medical errors in hospitals result in as many as 98,000 deaths per year; preventable medication errors occur at least 1.5 million times per year; and on average, only 55% percent of recommended care is delivered. In response, a variety of stakeholders from across the spectrum of health care delivery – including providers, professional and hospital associations, accreditation organizations, employers and business groups, insurance companies, and state and federal governments – have focused on monitoring and improving the quality of care. These efforts have focused on avoiding unnecessary deaths and poor health, while also encouraging better quality and value for health care spending. In the current environment, the quality of health care is increasingly recognized as a product of systems, not individuals, and there is widespread agreement that systematic measurement, monitoring, and reporting are needed to make meaningful advances in improving quality.

Health care quality indicators provide an important tool for measuring the quality of care. Indicators are based on evidence of "best practices" in health care that have been proven to lead to improvements in health status and thus can be used to assess, track, and monitor provider performance. While the results of quality measurement were originally not typically shared outside the provider organization conducting the quality improvement project, more recent assessments using the indicators have been included in public reports intended to steer patients toward higher-quality care and drive providers to improve their scores in order to bolster their public reputation. Indicators have also been used to link quality of care to financial incentives, either in the form of pay-for-performance (i.e., paying more for good performance on quality metrics), or in the form of tiered insurance products, which steer patients towards higher-quality providers by charging higher copayments for visits to providers with poorer quality scores.

The Agency for Healthcare Research and Quality (AHRQ) has been a pioneer in the development of health care quality indicators. In 1994 its Healthcare Cost and Utilization Project (HCUP) developed a publicly available set of quality indicators for hospital care based on discharge data. AHRQ updated the HCUP indicators in 2001, which were then

renamed the AHRQ Quality Indicators (AHRQ QIs). Today, AHRQ maintains four sets of QIs: 1) Inpatient Quality Indicators (IQIs), which reflect the quality of care provided in hospitals; 2) Patient Safety Indicators (PSIs), which reflect potentially avoidable complications or other adverse events during hospital care; 3) Prevention Quality Indicators (PQIs), which consist of hospital admission rates for 14 ambulatory care-sensitive conditions; and 4) Pediatric Quality Indicators (PDIs), which combine components of the PSIs, IQIs, and PQIs, as applied to the pediatric population. The AHRQ QIs are publicly distributed and supported by AHRQ, with regular updates. They are widely used by a variety of organizations for many different purposes. Meanwhile, many other organizations, both public and private, have developed and used their own sets of quality indicators.

Given the rapid growth of and robust demand for quality indicators, it is important to assess the position of the AHRQ QIs in the quality indicator "market." Who is using the AHRQ QIs today, and for what purposes? What have users' experiences been, and what unmet needs do they still have? Who else is developing and/or distributing indicators similar to the AHRQ QIs? Most importantly, what has been the impact of the AHRQ QIs on the quality of care delivered to patients? To answer these and related questions, the RAND Corporation was asked to conduct a one-year evaluation to assess user experiences with the AHRQ QIs and to identify users of other quality indicators, vendors of quality measurement products using the AHRQ QIs, and developers of quality indicators comparable to the AHRQ QIs. The results of this study are intended to inform decisions by AHRQ on future priorities for the QI program.

This report has three main objectives:

1. Provide an overview of the market for the AHRQ QIs as well as indicators and quality measurement tools developed by other organizations that are similar to the AHRQ QIs or that incorporate the AHRQ QIs.
2. Provide an overview of the range of ways in which the AHRQ QIs are used by various organizations.
3. Assess the market demand for the AHRQ QIs, identify unmet needs, and discuss implications for future activities by AHRQ.

## THE MARKET AND USES FOR AHRQ QIs

While AHRQ has developed four sets of QIs, all the QIs share certain key characteristics.

- **Based on administrative data.** The AHRQ QIs are based on hospital discharge data and can be used with existing hospital administrative databases.
- **Outcome-focused.** Most of the AHRQ QIs reflect health care outcomes, not the rates at which evidence-based processes of care are followed.
- **Hospital-focused.** Most of the indicators focus on inpatient care, and all of the indicators are based on hospital data.
- **Reviewed for scientific soundness**. The AHRQ QIs were tested by the Evidence-Based Practice Center at the University of California San Francisco and Stanford University, and detailed documentation of technical information is available in the public domain.
- **Available for public use**. The AHRQ QIs and associated code for SAS, SPSS,[a] and Windows are publicly available for download at no cost to the user.

To understand the market for the AHRQ QIs, we conducted a series of interviews with users of AHRQ QIs, users of other products, developers of similar indicator sets, and vendors of quality measurement products that include AHRQ QIs. This review found that the AHRQ QI program fills a unique niche in the market for QIs since there are no other sources of hospital care quality indicators that represent both a national standard and are also publicly available, transparent, and based on administrative data. Many of our interviewees stressed that the AHRQ QIs fill an important void in this respect.

AHRQ's unique place in the market for quality indicators has led to a wide proliferation of uses for the AHRQ QIs. Our environmental scan of users of the AHRQ QIs identified 114 users of the indicators and a range of different uses, including public

---

[a] Support for SPSS will be discontinued by AHRQ in 2007.

reporting, quality improvement/benchmarking, pay-for-performance, and research. The most common uses of the AHRQ QIs include:

- **Research.** We identified 43 uses of the AHRQ QIs for research. For example, Leslie Greenwald and colleagues used the AHRQ QIs to compare the quality of care provided in physician-owned specialty hospitals and competitor hospitals.[5]

- **Quality improvement.** We identified 23 organizations that use the AHRQ QIs as part of a quality improvement activity, including reports benchmarking performance against peers, but do not release the quality information into the public domain.

- **Public reporting.** We identified 20 organizations using the AHRQ QIs for public reporting. We classified an activity as "public reporting" if a publicly available report was published that compares AHRQ QI results between hospitals (IQIs and PSIs) or geographic areas such as counties (PQIs).

- **Pay-for-Performance.** We identified 4 organizations that are using the AHRQ QIs in pay-for-performance programs. Three were health plans and one was a Centers for Medicare and Medicaid Services (CMS) demonstration project.

As part of our environmental scan for users of the AHRQ QIs, we conducted a limited review of international uses. We found that the most visible current endeavor that attempts to make use of the AHRQ QIs is the Organization for Economic Cooperation and Development's (OECD) Health Care Quality Indicators (HCQI) Project. The OECD is an intergovernmental economic research institution headquartered in Paris, France, with a membership of 30 developed countries that share a commitment to democratic government and the market economy. The organization recently convened a meeting to work on the development and implementation of QIs at the international level. Preliminary discussions indicate that there is interest in using the AHRQ QIs internationally as well as sufficient data and technical capability to implement them.

## KEY FINDINGS

The results of our interviews and environmental scan focused on four key factors that can be used as criteria for evaluating quality indicators: importance, usability, scientific soundness, and feasibility.

### Importance

Nearly all of the organizations interviewed stressed the importance of the AHRQ QI program. AHRQ was frequently mentioned as a "national leader" in measurement development and research. Many interviewees stated very strongly that they rely on AHRQ as one of the only sources for publicly available, transparent indicators based on readily available data. They stressed that without the AHRQ QIs, they would have few alternatives and would likely have to drastically change or eliminate their quality reporting and/or measurement activities. Interviewees generally felt that it was important that a federal agency like AHRQ, which is regarded as credible and respected, develop and support a quality indicator set for public use. AHRQ's credibility and the transparency of the AHRQ QI methods were considered to be key in overcoming opposition to quality measurement and reporting by stakeholders, particularly providers. Overcoming this type of opposition is particularly important for public reporting and pay-for-performance initiatives, where providers' reputations and revenues are at stake.

Although only one organization in our interviews had formally measured the impact of AHRQ QIs on the quality of care delivered to patients, many interviewees provided anecdotal evidence of the effect of the use of indicators on quality. When asked whether they had measured the impact of using the AHRQ QIs, many interviewees responded that indicator use began too recently to allow for observation of any impact. However, many interviewees reported anecdotally that their or their clients' use of the AHRQ QIs was having some type of impact on quality of care. The impacts observed usually consisted of an activity such as putting a new quality improvement process in place, rather than an improvement in outcomes.

### Scientific Soundness

On the whole, our interviewees were impressed by the quality and level of detail of the AHRQ documentation on the face validity of the indicators and stated that the indicators captured important aspects of clinical care. Very rarely were indicators challenged on

conceptual grounds. Users largely felt that the AHRQ QIs can be reliably constructed from hospital discharge data, but that there was a certain learning curve during which hospital coding departments had to adjust to the requirements for the QIs. Thus far, coders had mainly been trained to apply coding rules to fulfill reimbursement requirements, but now they had to understand that coding practices also had implications for quality reporting. In selected instances, we heard concerns about ambiguity in the coding rules that would not provide sufficient guidance on whether to code an indicator-relevant diagnosis.

Sample size issues (whether due to the rarity of certain procedures or the infrequency with which some procedures are conducted at certain facilities) were repeatedly mentioned as a threat to the validity of the indicators, particularly the PSIs. Most users stated that the indicators were correctly operationalized within the constraints of the underlying data source. Isolated findings of specification errors were brought to our attention, but interviewees emphasized that AHRQ was always able to address those quickly. The limitations of administrative data were frequently mentioned as a threat to validity.

**Usability**

Most interviewees stated that the AHRQ QIs provide a workable solution for their needs and were very appreciative of the support that the AHRQ QI team provides for implementation and ongoing use. Despite these overall favorable impressions of the usability of the QIs, three needs related to their usability for reporting were raised repeatedly: development of reporting templates, development of composite indicators, and clearer guidance on the use of the AHRQ QIs for public reporting and pay-for-performance programs.

**Standard reporting format.** Nine of 54 interviewees highlighted the need for a standard format for reporting AHRQ QI results as a top priority. At the simplest level, some interviewees requested AHRQ-supported, standard, basic names for the AHRQ QIs in plain language, as some of the current indicator names are difficult for non-clinical audiences to understand. Other interviewees expressed a desire for more guidance and support on other aspects of presentation.

**Composite indicators.** Twelve of 54 interviewees expressed a desire for an AHRQ-supported methodology for constructing a composite indicator. Forming composites would allow the results to be summarized into one statistic, which is easier to grasp and

communicate, in particular for non-expert audiences. Composites would also overcome sample size limitations, as indicators could be pooled.

**Guidance on using AHRQ QIs for public reporting and pay-for-performance.** Interviewees who are currently using the AHRQ QIs for public reporting and pay-for-performance generally felt that they provided a workable basis for their activities. Still, interviewees stated that additional standards and guidance on the reporting of AHRQ QI results were needed. Many interviewees expressed dissatisfaction with the current AHRQ guidance on the appropriateness of the AHRQ QIs for public reporting. They felt that clearer guidance from AHRQ would help to counter opposition from those who argue that the AHRQ QIs should only be used for quality monitoring and improvement and research, but not as a public reporting or pay-for-performance tool. Taking the opposing view were several interviewees (mostly hospitals) who would like to see AHRQ make a clear statement that the AHRQ QIs are *not* appropriate for use in public reporting, pay-for-performance, or other reporting activities because of the limitations of the underlying administrative data.

**Feasibility**

We were told consistently that a major advantage of the AHRQ QIs was the feasibility of their implementation. They require only administrative data in the UB-92 format to which many users have routine access, since those data are already being used for billing and other administrative purposes and have to be collected and reported by hospitals in most states.

Interviewees emphasized that another substantial advantage of the AHRQ QIs is that the indicators have clearly defined and publicly available specifications, which helps with implementation of measurement. These specifications were regarded as of particular importance for hospitals, as the originators of the data, because the specifications enable hospitals to work with their coding departments to ensure that the required data elements were abstracted from medical records consistently and with high reliability. In addition, users who analyze data with the QIs, such as researchers, appreciated the fact that they could dissect the indicator results and relate them back to individual records. That ability helped researchers gain a better understanding of the indicator logic and distinguish data quality issues from actual quality problems.

## LESSONS LEARNED FOR FUTURE ACTIVITIES

Interviewees' perspectives provided lessons in three areas: current, anticipated, and potential development projects involving the QIs; AHRQ's role as a measures developer and the ways in which users speculate this role could evolve; and market demand for quality indicators, in particular, user willingness to pay for QIs.

### Priorities for Future Development of QIs

We asked interviewees to prioritize three categories of AHRQ development projects: (1) improvements in the current QI product line, (2) addition of new product lines, and (3) improved support for the QI products. Improving the current products was most frequently seen as the highest priority, followed by both the addition of new products and improvements in service, outreach, and user support for the measures. The most commonly requested improvement to the current QIs was the addition of data elements to the QI specifications, notably a flag for conditions present on hospital admission (currently in development by AHRQ), a flag for patients under do-not-resuscitate orders, and clinical data elements. The most commonly requested new product line was indicators of ambulatory care. As far as service improvements, the most frequently mentioned activities were a template and guidance for public reporting of the QIs, and guidance on next steps in quality improvement following identification of a potential quality problem using the QIs.

### The Future Role of AHRQ Compared to Other Players

Our interviewees held AHRQ in very high regard. The work of the AHRQ QI team was described as technically sound, sensitive to the limitations of the underlying data, and transparent. AHRQ is regarded as an intellectual leader and "go-to" institution for health services research and the use of administrative data for hospital quality measurement. Several other organizations, especially the Joint Commission on Accreditation of Healthcare Organizations (JCAHO), CMS, the Hospital Quality Alliance (HQA), and the Leapfrog Group, are seen as prominent sources for measures, but their indictors are generally regarded as complements to the AHRQ QIs. Interviewees were quite comfortable with AHRQ having a leading role in national quality indicator development. It was generally viewed as positive that a trustworthy federal institution had defined open-source and well-documented quality measurement standards. These standards were viewed as contributing to the transparency of health care quality measurement and reducing the measurement

burden for health care providers by limiting the number of measurement tools they must use to satisfy various reporting requirements.

We discussed whether it could be a viable option for AHRQ to give up parts of the current QI program in order to free up resources and set different priorities. Almost unanimously, interviewees rejected a model under which AHRQ would develop and distribute the QI software without supporting it. We received mixed reactions to a model under which AHRQ would develop and release indicators and their technical specifications, but no longer provide or support software. Interviewees were generally wary of the idea of delegating user support and/or software development and distributions to vendors, fearing that vendors would be prohibitively expensive or incapable of providing the same quality of support as the original developers.

### Willingness to Pay for the AHRQ QIs

As an alternative to AHRQ realigning current funds, we asked interviewees whether AHRQ might consider financing program growth by generating additional revenues by charging users. Not unexpectedly, this proposal was not met with enthusiasm. However, almost half of interviewees (44%) were willing to pay a "reasonable fee" for access to the full QI resources.

## DISCUSSION

### Limitations

The majority of the interviewees in this evaluation were users of the AHRQ QIs. Non-users may have more negative opinions of the AHRQ QIs. The few non-users we did interview did not have express substantially negative opinions about the AHRQ QI program, but a larger sample of non-users may have produced different results. This study also focused on uses of the AHRQ QIs that were publicly discussed or released, so that the results likely do not fully reflect the use of AHRQ QIs for non-public uses such as confidential quality improvement activities by hospitals.

### What is AHRQ's Current Market Position?

The AHRQ QIs have achieved a strong position in their market segment and no obvious alternative or competitor could be identified, although some organizations (notably

JCAHO, CMS, HQA, and Leapfrog) have complimentary indicator sets. This is unlikely to change: new users have an incentive to adopt the prevailing product, because it makes their results comparable to a large number of other users. Widespread use lends legitimacy to the product, which is critical in the often-politicized debates about selecting quality indicators for public reporting and pay-for-performance.

**Where are the Growth Opportunities for the AHRQ QI Program?**

There are now a substantial number of users of the AHRQ QIs for public reporting and pay-for-performance programs. As the prevalence of those activities increases, we expect the number of users to increase substantially both for the programs themselves and for internal quality improvement programs and projects that will attempt to align their target measures with standards for external accountability. Our interviewees wanted expansion of the AHRQ QI program. They were largely aware and appreciative of AHRQ's current efforts to improve and expand the program, but expressed an interest in scaling up, and speeding up, those activities.

**How Could Growth Be Financed?**

Most interviewees stated that federal funding should be used to support future AHRQ QI activities, even though they realized that this was a difficult proposition given the pressure on public budgets in general, and on AHRQ's budget in particular. Interviewees were reluctant to see AHRQ give up software development and/or user support. As an alternative, we discussed the option of AHRQ continuing to provide specifications, software and user support but starting to charge for those services. While there was little enthusiasm for user fees, only a few stated that they would stop using the AHRQ QI product in that case. Most interviewees seemed to be willing to pay a "reasonable" amount. However, if AHRQ were to implement a charge-based model for the QIs, it would face the challenge of finding a business model that would generate sufficient revenue and still be consistent with AHRQ's mission and values as a public agency.

# LIST OF SYMBOLS

| Symbol | Definition |
|---|---|
| ACOG | American College of Obstetricians and Gynecologists |
| AHRQ | Agency for Healthcare Research and Quality |
| AMI | Acute myocardial infarction |
| APR-DRGs | All patient refined diagnosis related groups |
| BCBSMA | Blue Cross Blue Shield of Massachusetts |
| CMS | Centers for Medicare and Medicaid Services |
| CHSC | Center for Studying Health System Change |
| DFW | Dallas-Fort Worth |
| DFWHC | Dallas-Fort Worth Hospital Council |
| DI | DFWHC Data Initiative |
| DRGs | Diagnosis related groups |
| DVT | Deep vein thrombosis |
| EC | European Commission |
| ESQH | European Society for Quality in Healthcare |
| ETGs | Episode Treatment Groups |
| GIC | Group Insurance Commission (State of Massachusetts) |
| HCQI | Health Care Quality Indicators Project |
| HCUP | Healthcare Cost and Utilization Project |
| HQA | Hospital Quality Alliance |
| ICD-9-CM | International Statistical Classification of Diseases and Related Health Problems – Version 9 – Clinical Modification |
| IHI | Institute for Healthcare Improvement |
| IOM | Institute of Medicine |

| | |
|---|---|
| **ISQua** | **International Society of Quality in Healthcare** |
| **JCAHO** | **Joint Commission on Accreditation of Healthcare Organizations** |
| **IQIs** | **Inpatient Quality Indicators** |
| **Mass-DAC** | **Massachusetts Data Analysis Center** |
| **MHA** | **Maryland Hospital Association** |
| **MQIP** | **Maryland Quality Indicators Project** |
| **NQF** | **National Quality Forum** |
| **OECD** | **Organization for Economic Cooperation and Development** |
| **PDIs** | **Pediatric Quality Indicators** |
| **PE** | **Pulmonary embolism** |
| **PQIs** | **Prevention Quality Indicators** |
| **PSIs** | **Patient Safety Indicators** |
| **QI** | **Quality indicator** |
| **THCIC** | **Texas Health Care Information Collection** |
| **VBAC** | **Vaginal birth after cesarean section** |
| **WHO** | **World Health Organization** |

# 1. INTRODUCTION

## 1.1 BACKGROUND

Public recognition of health care quality issues has spiked remarkably in the past ten years, driven by a series of high-profile reports from the Institute of Medicine (IOM), the RAND Corporation, and others.[6,7,8,9] These reports showed, among other facts, that preventable medical errors in hospitals result in as many as 98,000 deaths per year; preventable medication errors occur at least 1.5 million times per year; and on average, only 55% percent of recommended care is delivered.

In response, a variety of stakeholders from across the spectrum of health care delivery – including providers, professional and hospital associations, accreditation organizations, employers and business groups, insurance companies, and state and federal governments – have focused on monitoring and improving the quality of care. A primary goal has been to avoid unnecessary deaths and poor health – at a minimum, to "first do no harm" by limiting the amount of injury incurred by health care itself. Another strong motivation has been the increasing recognition that value for money is poor in the health care system. Health care costs are high and have been increasing rapidly, without the level or improvement in quality that is expected by those paying the costs.

In this environment, there is widespread agreement that "business as usual" – i.e., exclusively relying on providers to self-police the quality of care they provide without systematic measurement, monitoring, and reporting – is no longer adequate for improving the quality of care.[10] Quality is increasingly recognized as a product of systems, not individuals, and attention has focused on how to analyze and redesign those systems. A consensus has emerged that although solutions will not be easy, "a key component of any solution… is the routine availability of information on performance at all levels" [11] – that is, "quality of care should be measured and reported routinely at both the national and provider-specific (e.g., hospital and physician) levels."[12] More simply, "that which cannot be measured is difficult to improve."[13]

There are a variety of measurement-based approaches to improving quality. The most widespread and longest-used approach, referred to in this report as "quality improvement," is an internal effort by providers or other organizations to measure quality of care, identify

areas of weakness, devise and implement changes, and monitor the effects of the changes. As results are typically not shared outside of the organization that conducts the quality improvement project, this approach can be viewed as formalization of the traditionally prevailing model of improving quality – mutual peer review. More recently, calls for external accountability of health care providers have led to two types of quality monitoring approaches. One is public reporting of quality of health care providers.[b] These reports can steer patients to higher-quality care and drive providers to improve their scores in order to bolster their professional reputation. The other one is linking quality of care to financial incentives, either in the form of pay-for-performance (paying more for good performance on quality metrics) or in the form of tiered insurance products, which steer patients towards higher-quality providers by charging higher copayments for visits to providers with poorer quality scores.

These efforts all depend on reliable, valid, and agreed-upon ways to measure the quality of care. However, health care quality is difficult to measure, in particular because there are many factors contributing to outcomes of health care in addition to the actual quality of the care provided. Many of these factors – for example, patient compliance with treatment – are largely out of providers' control. Despite this and other difficulties, much progress has been made in the science of quality measurement. The underlying knowledge base has greatly expanded. The science of measuring health status has improved, as has the evidence supporting "best practices" that have been proven to lead to improvements in health status. This evidence base has allowed for the development of numerous quality indicators, which then have been tested for reliability, validity, ease of use, and usefulness for improving quality. The quality indicators fall into two main categories: counts of inputs or actions that are known to lead to better health outcomes ("structure" and "process" measures in Donabedian's classic quality measurement framework),[14] and direct measurements of the outcomes of care ("outcomes"). Unfortunately, many of the most rigorous measures are also the most difficult to implement due to their burden of data collection. The data needed for quality measurement must be collected specifically for that purpose, abstracted from (usually) paper medical records, or abstracted from existing electronic data that have been collected for other purposes, notably billing. However, as will

---

[b] Quality comparisons have been made between states, health plans, nations, etc. in addition to between providers. We focus on provider comparisons here since they are the most common.

be shown in this report, there is now a small but growing arsenal of well-tested quality indicators that can be used with existing data sources or data collection processes.

The Agency for Healthcare Research and Quality (AHRQ) was one of the pioneers of quality indicator development. As early as 1994, its Healthcare Cost and Utilization Project (HCUP) developed a publicly available set of quality indicators for hospital care based on discharge data. Congress later tasked AHRQ with developing an annual report on the quality of health care in the United States, the *National Healthcare Quality Report.*[15] In preparation for this report, in 2001 AHRQ invested in an update of the HCUP indicators, which were renamed the AHRQ Quality Indicators (AHRQ QIs).[16] Indicators of patient safety were added in 2002.[17] The AHRQ QIs and associated computer code are now publicly distributed and supported by AHRQ, with regular updates.[18] They are widely used by a variety of organizations for many different purposes. Meanwhile, many other organizations, both public and private, have developed and used their own sets of quality indicators.

Given the rapid growth of and robust demand for quality indicators, it is important to assess the position of the AHRQ QIs in the quality indicator "market." Who is using the AHRQ QIs today, and for what purposes? What have users' experiences been, and what are their unmet needs? Who else is developing and/or distributing indicators similar to the AHRQ QIs? Most importantly, what has been the impact of the AHRQ QIs on the quality of care delivered to patients?

To understand the answer to these and related questions, the RAND Corporation was asked to conduct a one-year evaluation to assess user experiences with the AHRQ QIs, and to identify users of other quality indicators, vendors of quality measurement products using the AHRQ QIs, and developers of quality indicators comparable to the AHRQ QIs. The results of this study are intended to inform decisions by AHRQ on future priorities for the QI program.

## 1.2 DESCRIPTION OF THE AHRQ QUALITY INDICATORS (AHRQ QIs)

AHRQ has developed several types of QIs, all of which share certain core characteristics. There are currently four sets of AHRQ QIs:

1. **Inpatient Quality Indicators (IQIs).** There are 32 IQIs reflecting the quality of care provided in hospitals. The indicators fall in four categories.

   - The first category includes the in-hospital mortality rates for seven specific medical conditions. The conditions were selected because their mortality rates were shown to vary across providers and because evidence indicates that the mortality risk for those conditions may depend on quality of care. The mortality rates, adjusted for patient risk, can be compared to those of other providers or to other benchmarks, with lower rates indicating better quality.

   - The second category includes the in-hospital mortality rates for eight surgical procedures. Like the condition-specific mortality rates, the procedures were chosen because their mortality risk was known to vary across providers and because evidence indicates that mortality for those procedures may be associated with inferior quality of care. The mortality rates can be benchmarked, and lower rates are better.

   - The third category includes utilization rates for eleven procedures for which there is potential for overuse, underuse, or misuse. Four of the utilization rates were designed for comparisons between geographic areas (either counties or Office of Management and Budget Metro Areas), not hospitals. The rates are measured as the number of procedures performed divided by the total number of patients who were potential candidates for that procedure. Utilization rates that are much higher or lower than benchmarks may point to poor quality. The direction of the indicator – i.e., whether higher or lower rates are desirable – depends on the condition. For example, Cesarean delivery has been identified as an overused procedure, so lower rates indicate better quality (IQI 21). In contrast, vaginal delivery following a previous Cesarean delivery (VBAC) is considered an underused procedure, so higher rates indicate better quality (IQI 34).

   - The fourth category includes indicators of the hospital-level volume for six complex procedures for which research suggests a positive impact of case volume on patient outcomes. The procedure volume therefore provides an indirect indication of the provider's expected outcome for the procedure.

The indicators are simple counts of the number of times that the procedures are performed per year. These counts can be compared to thresholds for the minimum annual volume of procedures that providers should perform to ensure an acceptable level of quality. The volume thresholds are based on published research.

2. **Patient Safety Indicators (PSIs).** The PSIs contain 27 indicators of potentially-avoidable complications or other adverse events related to hospital care. For example, an infection resulting from medical treatment (PSI 7) is considered an avoidable complication, and the rate of surgical instruments or other "foreign bodies" left in the patient following surgery (PSI 5) is considered a preventable adverse event. The indicators are measured as the number of complications or adverse events divided by the total number of patients at risk for those events. Twenty of the indicators were designed for comparisons between hospitals, and the remaining 7 were designed for comparisons between geographic areas (either counties or Office of Management and Budget Metro Areas).

3. **Prevention Quality Indicators (PQIs).** The PQIs are based on the hypothesis that appropriate ambulatory care can help to prevent hospital admissions for some conditions. The indicators consist of hospital admission rates for 14 of these ambulatory care-sensitive conditions and are intended for comparisons of geographic areas. Unlike the other AHRQ QIs, they do not reflect quality of hospital care, but the quality of ambulatory care. The indicators are measured as condition-specific admission rates (e.g., the number of admissions for urinary tract infection per 100,000 people living in a Metro Area or county). Each admission to the hospital reflects a failure in the delivery of appropriate care outside the hospital, so lower rates reflect better quality.

4. **Pediatric Quality Indicators (PDIs).** The PDIs contain 18 indicators, each of which is similar to a comparable type of IQI, PQI, or PSI, except that the PDIs in each case refer to the pediatric population. Since these indicators were released during the course of this study, they were not included in the analysis. Several of the PDIs were formerly included in the AHRQ IQI, PSI, and PQI sets. Most of the issues raised in this report would generalize to the PDIs.

There are several key characteristics that apply to all four sets of the AHRQ QIs.

- **Based on administrative data.** The AHRQ QIs are based on hospital discharge data that are collected using the UB-92 standardized claim reporting form and used mainly for billing purposes. The indicators can be used with existing hospital administrative databases, including the HCUP National Inpatient Sample and Medicare MedPAR data.

- **Outcome-focused.** Most of the AHRQ QIs reflect health care outcomes, and not the rates at which evidence-based processes of care are followed. The exceptions are the subset of IQIs which reflect utilization or volume of services.

- **Hospital-focused.** Most of the indicators focus on inpatient care, and all of the indicators are based on hospital data. The PQIs reflect the quality of ambulatory care, but the measurement focuses on admissions to the hospital that could have been prevented given appropriate, effective ambulatory care.

- **Reviewed for scientific soundness**. The AHRQ QIs were tested by the Evidence-Based Practice Center at the University of California San Francisco and Stanford University. Detailed documentation of technical information, including the methodology, evidence base, and measurement properties of the indicators, is available in the public domain.

- **Available for public use**. The AHRQ QIs and associated code for SAS, SPSS,[c] and Windows are available for public download at no cost to the user. User support is also provided free of charge.

The AHRQ QIs were originally designed to be used for research and quality improvement activities.[19] However, as examples included in this report demonstrate, the QIs are now used for a variety of other purposes, including public reporting, pay-for-performance, and segmenting providers for tiered insurance products. AHRQ has evaluated and endorsed the use of the QIs for these new purposes, given certain caveats.[20]

## 1.3 OBJECTIVES

This report has three main objectives:

---

[c] Support for SPSS will be discontinued by AHRQ in 2007.

1) Provide an overview of the market for the AHRQ QIs as well as indicators and quality measurement tools developed by other organizations that are similar to the AHRQ QIs or that incorporate the AHRQ QIs.
2) Provide an overview of the range of ways in which the AHRQ QIs are used by various organizations.
3) Assess the market demand for the AHRQ QIs, identify unmet needs, and discuss implications for future activities by AHRQ.

## 1.4 ORGANIZATION OF THE REPORT

The remainder of this report is divided into six sections. Section 2 outlines the methodology used in our evaluation. Section 3 assesses the market for quality indicators and AHRQ's role in that market. Section 4 evaluates the AHRQ QI program on the dimensions of importance, scientific soundness, usability, and feasibility. Section 5 presents two case studies of how the AHRQ QIs are used in particular market areas. Section 6 summarizes lessons learned, and Section 7 concludes with a discussion of the implications of the evaluation for the future of the AHRQ QI program.

# 2. METHODS

In this chapter we describe the methods used in this study. There are three main components to our approach: an *environmental scan* to identify and catalog users of the AHRQ QIs and developers or vendors of similar projects; a *series of interviews* with individuals and organizations who use AHRQ or other QIs to understand how and why QIs are used, and what improvements, if any, users might like to see; and *case studies* to illustrate how QIs are used in two particular geographic areas.

## 2.1 ENVIRONMENTAL SCAN

We conducted an environmental scan to identify four types of organizations: (1) users of the AHRQ QIs; (2) "non-users," i.e., organizations that are using an alternative to the AHRQ QIs; (3) developers of similar sets of quality indicators; and (4) vendors of quality measurement products that may or may not include AHRQ QIs. The following six types of sources were used:

1) Databases of published literature
2) Conference presentation abstracts
3) World Wide Web search engines
4) Reviews of quality measurement activities conducted by third parties
5) AHRQ materials
6) Query of RAND Health research staff

The specific sources and the search terms are listed in Tables 2.1 and 2.2.

Table 2.1. Sources Queried in Environmental Scan

| |
|---|
| *Databases of Published Literature* |
| Econlit |
| NYAM Grey Literature |
| Psycinfo |
| PubMed |
| WorldCat |
| *Conference presentation abstracts* |
| AcademyHealth Annual Research Meeting abstracts |
| Papers First |
| Social Sciences Abstracts |
| World of Science |
| *World Wide Web search engines* |
| Google |
| Google Scholar |
| *Reviews of quality measurement activities conducted by third parties* |
| Delmarva Foundation's review of 47 Websites publishing hospital quality information |
| Leapfrog compendium of incentive programs |
| NAHDO survey of state public reporting |
| *AHRQ materials* |
| AHRQ user conference presentations |
| AHRQ QI newsletters |
| Posting on AHRQ Web site requesting that users contact RAND for interview |
| *Query of RAND Health research staff* |

Table 2.2. Search Queries Used in Environmental Scan

| |
|---|
| "AHRQ inpatient quality indicators" |
| "AHRQ IQIs" |
| "AHRQ patient safety indicators" |
| "AHRQ PQIs" |
| "AHRQ prevention quality indicators" |
| "AHRQ PSIs" |
| "AHRQ QIs" |
| "AHRQ quality indicators" |
| "health quality indicator development" |
| "health quality indicator use" |
| "inpatient quality indicator" OR "inpatient quality measure" AND "develop" |
| "inpatient quality indicators" |
| "patient safety indicators" |
| "prevention quality indicators" |
| "quality indicator" AND "AHRQ" OR "Agency for Healthcare Quality and Research" |
| "quality indicator" AND "health care" |
| "quality indicator" AND "health" |
| "quality indicator" AND "patient safety" |
| "quality indicator" OR "quality measure" AND "develop" |
| "quality indicator" AND "inpatient" |

Each organization identified in the search was entered into a database together with the following information:

- Type of organization (e.g., hospital);
- List of quality indicators used;
- Primary use of quality indicators (research, public reporting, pay-for-performance, quality improvement, vendor, developer);
- Description of other uses of quality indicators;
- Citation;
- Contact information;
- Method of identification.

The environmental scan was used to understand the range of ways in which the AHRQ QIs are being used and to select interview candidates.

## 2.2 INTERVIEWS

### 2.2.1 Selection of interviewees

In order to select interview candidates, we first consulted with five members of the AHRQ QI team, which includes both AHRQ staff and contractors. We used information gained from these discussions, together with the results of our environmental scan, to compile a list of potential interviewees. We used the following method to select interviewees from this complete list:

1) We cross-tabulated the environmental scan results by type of organization (hospital association, state government, etc.) and primary indicator use (pay-for-performance, public reporting, etc.) to create a table representing the universe of potential interviewees.
2) We selected the number of interviewees to be chosen in each type-use cell in order to distribute interviews across cells and ensure multiple interviews per cell where possible.
3) Two of the researchers (PH and SM) reviewed the list of organizations independently and chose interviewees for each cell based on the criteria: (a) maximize the variety of uses and types of organizations within each cell; (b) maximize the estimated impact of quality indicator activity. The reviewers compared notes and reconciled differences to come to the final list of suggested interviewees to forward to AHRQ.
4) We met with AHRQ staff to discuss the suggested interviewees and incorporated AHRQ staff comments to arrive at the final list.

We then began the process of identifying the most appropriate respondent(s) for each organization. Identification of the most knowledgeable respondent was a multi-stage process. First, we identified an initial contact through available published materials or a telephone call to the organization. We then requested that the individual we contacted forward our interview request to the most appropriate individual(s). Prior to the interview, interviewees were given a fact sheet on the project, information about how the data would be used and assurances about confidentiality, and a list of sample questions similar to those asked during the interviews. All procedures, including verbal consent procedures, were

approved in advance by the RAND Human Subjects Protection Committee, RAND's Institutional Review Board.

### 2.2.2 Interview procedures

We created interview guides based on input of AHRQ QI team staff, our own knowledge of the AHRQ QI program, and the results of the environmental scan. Separate guides were created for users of AHRQ QIs, developers of similar products, and vendors of quality measurement products. The interviews covered three main topics:

1) How AHRQ QIs (and other quality indicators) have been used;
2) Experiences (including impact of use and lessons learned) from quality indicator use;
3) Suggestions for future priorities for the AHRQ QI program.

Semi-structured interviews were conducted by telephone by one researcher accompanied by a note-taker. Some interviews were also recorded using a digital voice recorder. Interviewees were guaranteed that none of their responses would be reported in an identifiable format. During the interviews, we made note of any supporting documents mentioned by the interviewees and asked for a copy of the documents at the close of the interview. Interview notes were completed and edited in a timely fashion and data from the interviews and supporting materials were coded into the dimensions in the interview guide and entered into a database for analysis. Quotes published in this report have been reconstructed from interview notes or recordings and thus in some cases may differ slightly from the exact wording used by the interviewee.

### 2.2.3 International Users

We took advantage of the (non-project related) participation of one member of the team (SM) in an OECD Expert Group meeting on patient safety. The meeting took place on June 29 and 30, 2006, in Dublin, Ireland. At the meeting, we interviewed five researchers who had used the AHQI QIs in countries other than the United States.

## 2.3 CASE STUDIES

Based on the results of the first round of interviews, we identified two geographic areas in which to conduct in-depth case studies on the use of AHRQ QI indicators for public reporting. The case study is a research strategy that is used in many settings (e.g., policy, political science, and public administration research; organizational and management studies; planning research; etc.). It is an empirical inquiry in which multiple sources of evidence are used to provide data from which judgments can be made about the usefulness of an approach to a problem and the generalizability of the findings to other sponsoring organizations and markets. We selected two geographic areas (Boston, MA and Dallas-Fort Worth, TX) and identified multiple organizations within each of these markets to provide a full picture of the various perspectives on a common application of AHRQ QIs.

These case studies were designed to refine our notions of the critical variables present in AHRQ QI initiatives; obtain evidence regarding how successful organizations were in developing and implementing QI initiatives; and to document the barriers and facilitators to using AHRQ QIs in real-world health care settings. We first identified organizations for case study interviews by (1) analyzing the results of the environmental scan to identify users of the AHRQ QIs in the two geographic areas; (2) asking for nominations from representatives of organizations who had already been asked to participate; (3) taking suggestions from AHRQ staff. Detailed notes were taken for each case study. These notes as well as media and policy reports on health policy issues in the geographic areas were used as the basis for the case study analysis.

# 3. THE MARKET FOR QUALITY INDICATORS

Our environmental scan revealed strong demand for hospital care quality indicators. Demand for indicators for research and quality monitoring is strong and has a relatively long history. Demand is higher and increasing rapidly for quality indicators that can be used for other, newer purposes. These purposes include public reporting to inform consumers' choice of providers and otherwise drive provider improvement; pay-for-performance to reward high-quality providers; the development of tiered insurance products; and using quality indicators to select a network of providers.

This demand has led to a proliferation of quality indicators. In addition to AHRQ, the market leaders in developing hospital quality indicators are the Centers for Medicare and Medicaid Services (CMS), the Joint Commission on Accreditation of Healthcare Organizations (JCAHO), the Hospital Quality Alliance (HQA – a collaboration between CMS, JCAHO, and several other organizations), and the Leapfrog Group. In this section, we discuss these and other developers and vendors of quality indicators, and how the quality indicators developed by each of these agencies/organizations compares to the AHRQ QIs. Our environmental scan identified two main categories of players in the market for quality indicators. The first type, "developers," includes organizations that develop, support, and distribute quality indicators. The second type, "vendors," includes organizations that develop and/or sell quality measurement products to providers, insurers, and others. Their products often include the AHRQ QIs (or variants thereof), indicators from other developers, and/or indicators developed by the vendors themselves.

## 3.1 DEVELOPERS

The environmental scan identified 12 organizations that have developed indicators that are similar in some way to the AHRQ QIs. The organizations that have developed indicators that are widely used and focused on hospitals are summarized in Table 3.1 and described below.

Table 3.1. Developers of Quality Indicators and Comparison with AHRQ QIs

| Developer | Indicators | Similarities to AHRQ QIs | Differences from AHRQ QIs |
|---|---|---|---|
| JCAHO<br>CMS<br>HQA | • Core Measures<br>• Hospital Quality Indicators | • National standard | • Process measures<br>• Clinical data<br>• Implemented through licensed vendors |
| The Leapfrog Group | • Leapfrog Leaps | • National standard<br>• Some outcomes indicators | • Collected through survey<br>• Mostly structure and process measures |
| Institute for Healthcare Improvemen t | • Hospital Standardized Mortality Ratios | • Outcomes indicator (risk-adjusted mortality) | • Mortality not condition-specific<br>• Used in conjunction with specific quality improvement program |
| States (e.g., PA and CA) | • PA Health Care Cost Containment Council Hospital Performance Report Indicators<br>• CA Healthcare Quality and Analysis Division Indicators | • Outcomes indicators<br>• Administrative data | • Uses data elements not available in administrative data in most states |
| Vendors | • Various | • Administrative data<br>• Some outcomes indicators | • Methodology often not transparent |

*Source*: RAND analysis of environmental scan results
*Note:* Indicators were judged to be a "national standard" if they were described that way by any of the study's interviewees.

Although there are similarities between these indicators and those developed by AHRQ, none of the indicators developed by organizations other than AHRQ were comparable to the AHRQ QIs on all of their major characteristics: based on administrative data, outcome-focused, hospital-focused, based on transparent methodology, and available for public use.

**JCAHO/CMS/HQA.** Both JCAHO and CMS have developed quality indicators of hospital care for common conditions. CMS's measures were originally used for quality improvement initiatives conducted by Medicare Quality Improvement Organizations

(QIOs). JCAHO's Core Measures have been used as part of the JCAHO hospital accreditation process since 2002. They cover five clinical areas: (1) acute myocardial infarction, (2) heart failure, (3) pneumonia, (4) surgical infection prevention, and (5) pregnancy and related conditions. JCAHO-accredited hospitals choose 3 of these 5 areas for reporting, depending on the services they provide. JCAHO publishes the results of the measures publicly on the Web.[21]

Since the measures had significant overlap, CMS and JCAHO agreed in 2004 to align specifications for overlapping measures and to maintain them as a shared set of measures. A subset of the joint CMS-JCAHO measures was later selected by the HQA, a public-private partnership for measuring and reporting hospital quality. Their Hospital Quality Measures are now publicly reported on the Web for both accredited and non-accredited hospitals.[22] They are also used in other CMS activities such as the Premier pay-for-performance demonstration project.[23]

Like the AHRQ QIs, the CMS/JCAHO/HQA measures are widely used and viewed as a national standard.[d] A key difference between those measures and the AHRQ QIs is that they are largely based on clinical data collected from medical records rather than administrative data. JCAHO has estimated that collection of the clinical data for the Core Measures takes an average of 22-27 minutes per case for acute myocardial infarction, heart failure, and pneumonia.[24] A second key difference is that the CMS/JCAHO/HQA measures are process indicators while the AHRQ QIs are outcome indicators. Another difference is that, while the AHRQ QIs reflect a broad range of conditions, the CMS/JCAHO/HQA measures currently reflect only five conditions; however, JCAHO and CMS are currently developing indicators in additional clinical areas.

The method used by JCAHO to implement its Core Measures is also different from that used for the AHRQ QIs. Hospitals pay vendors to measure the JCAHO Core Measures on their behalf using standardized specifications. Hospitals have made a wide variety of arrangements with vendors for Core Measure collection and reporting, according to their specific needs and characteristics. All vendors of the JCAHO Core Measures must undergo

---

[d] Indicators were judged to be a "national standard" if they were described that way by any of the study's interviewees.

a certification process through which JCAHO ensures that they have appropriately implemented the measures.

Due to these differences, the CMS/JCAHO/HQA measures and the AHRQ QIs can be considered complementary in some respects. A number of the users of the AHRQ QIs interviewed (11 of 36) also use the JCAHO/CMS/HQA measures.

The only way in which the CMS/JCAHO/HQA measures and the AHRQ QIs could be considered competitors is as a function of limited hospital resources for quality measurement. Hospitals are required to report the JCAHO Core Measures for accreditation and may have limited resources for other quality measurement activities, including the AHRQ QIs. One interviewee told us:

> AHRQ could do a lot of terrific things with the AHRQ QIs, but facilities are trying to meet requirements right now and don't have time and resources to work with other quality indicators to the exclusion of what they might like to do. Hospitals are doing only what they have to do – either by mandate or by the market.[e]

**Leapfrog.** The Leapfrog Group has developed a set of quality indicators that are widely used and considered to be a national standard. The indicators are intended to improve value in health care purchasing. Provider performance on the indicators is presented in a public report on Leapfrog's Web site. In addition to developing and marketing its own quality indicators, Leapfrog operates a pay-for-performance program, the Leapfrog Hospital Rewards Program, which uses JCAHO Core Measures and an efficiency measure in addition to the Leapfrog indicators. The program is implemented through vendors, who pay Leapfrog for every participating hospital, and then charge hospitals accordingly.

Unlike the AHRQ QIs, most of the Leapfrog indicators are not outcome-focused and require primary data collection. The indicators are organized into four content areas called "Leaps": (1) computerized physician order entry, (2) intensive care unit staffing, (3) high-risk treatments, and (4) safe practices. Data are collected through a survey of hospitals. Leaps 1, 2, and 4 are structure and process indicators, such as use of a computerized physician order entry system or staffing hospital intensive care units with intensivists (physicians who specialize in critical care medicine). Leap 3 (high-risk treatments) overlaps considerably with

---

[e] This and all quotes appearing in this report are reconstructions based on interview notes or recordings.

the AHRQ IQIs. It measures procedure volume and risk-adjusted mortality for selected conditions. Leapfrog is currently standardizing its specifications to those used in the AHRQ IQIs in order to minimize the reporting burden for hospitals.

**Institute for Healthcare Improvement (IHI).** The IHI measures overall hospital mortality as part of its activities to improve hospital quality. This measurement activity is conducted in conjunction with the implementation of a specific set of interventions that are intended to improve quality in participating hospitals. The indicator used is similar to the AHRQ IQIs in that it is based on risk-adjusted mortality associated with hospital stays and is based on the analysis of administrative data. Unlike the AHRQ IQIs, however, the IHI measures the mortality rate for all conditions. Hospital- and area-level characteristics are used in regression models to control for patient risk. This measurement approach originated in the United Kingdom and has also been applied to hospitals in many countries other than the United States.[25]

**States.** We also interviewed representatives from California and Pennsylvania, two states that have developed their own methodologies for measuring quality using administrative data. These states developed their own measurement approaches largely because their public reporting efforts predate the development of the AHRQ QIs. Both states also use data elements that are unavailable in the hospital administrative data collected in most other states. These features include a flag to indicate conditions that were present on hospital admission (California) and detailed data on severity of illness (Pennsylvania). Other states, such as New York, have also developed their own measurement approaches which may predate the AHRQ QIs or use data elements not available in other states.

**Vendors.** We interviewed several vendors who, in addition to implementing existing measures from other developers in their measurement tools, have also developed proprietary indicators. Some of these indicators are similar to the AHRQ QIs in that they are based on administrative data and are outcomes indicators. The key difference is the definitions and specifications of most vendors' indicators are proprietary. The vendors' indicators have also not always been subjected to validation of the same rigor as the AHRQ QIs. In the next subsection, we discuss the vendors identified by the environmental scan in more detail.

## 3.2 VENDORS

The environmental scan identified 12 vendors of quality measurement products that were determined to include the AHRQ QIs.[f] These vendors are listed in Table 3.2.

**Table 3.2. Vendors of Quality Measurement Products That Include the AHRQ QIs**

| |
|---|
| CareScience |
| Consumers' Checkbook |
| Health Benchmarks |
| HealthGrades |
| Innovative Health Solutions |
| Mediqual (Cardinal Health) |
| Medisolv |
| Midas+ |
| Solucient |
| Subimo |
| WebMD Quality Services |

*Source*: RAND analysis of environmental scan results

Typically, the AHRQ QIs are included in software tools that are marketed to hospitals for quality improvement or to insurers or business groups for hospital profiling. The vendors' products offer additional functionality to the basic AHRQ QI software. For example, the vendors' measurement tools often include non-quality indicators that inform hospital administration, such as financial performance indicators. The tools are often designed to offer users a variety of reporting options. These measurement tools may be particularly useful for hospitals that do not have the in-house expertise or staff time to construct indicators of quality and other aspects of care from raw data. Similar tools are used by insurance companies and other organizations.

---

[f] We attempted to determine whether vendors' proprietary products included the AHRQ QIs, but since limited information is available from some vendors, some mistaken attribution is possible. There are also other vendors with similar quality measurement products that do not include the AHRQ QIs, but they were not included in our study.

As mentioned above, many of these tools include proprietary quality indicators developed by the vendors themselves. In addition, many of the vendors are licensed to implement the JCAHO Core Measures, and many also produce indicators from other developers, such as Leapfrog.

Some users of the AHRQ QIs whom we interviewed use vendors for their measurement activities and expressed a high degree of satisfaction with the vendors' services. On the other hand, some users expressed a concern that the AHRQ QIs as implemented by some vendors may differ in key respects from the official AHRQ QI specifications, and that the proprietary nature of the tools makes these differences non-transparent. One hospital association captured this sentiment:

> The AHRQ QIs are standardized measures, risk-adjusted, and not in a "black box" so we can get the numerator and denominator and make them accessible to hospitals. The industry is sick and tired of vendors and consulting firms creating black boxes.

Another interviewee sounded similar themes:

> The problem is that if there's any "black box" methodology, [users] won't touch it - it's politically dead, even if there is an underlying valid scientific process. Hospitals want to check their own numbers. [The vendors'] offers sound nice. The problem is, a hospital can't replicate the findings or understand differences in methodology/calculations. [Users] like transparency, a tool that is open, where everyone can see what is happening, hospitals can replicate the results, then everyone can talk about the differences. It democratizes quality reporting.

## 3.3 AHRQ'S POSITION IN THE MARKET FOR QUALITY INDICATORS

While the quality indicators developed by organizations other than AHRQ share certain characteristics with the AHRQ QI program, there are no other sources of indicators that are viewed as a national standard and are also publicly available, transparent, hospital-focused, outcome-focused, and based on administrative data. Many of our interviewees stressed that the AHRQ QIs fill an important void in this respect. A representative of an integrated delivery system described the process of searching for quality indicators that could be readily used for monitoring quality and guiding quality improvement activities:

> When we started looking for indicators, we really struggled to find valid quality measures based on readily available data and with benchmark potential. Without manually auditing patient charts, and coming up with numerator and denominator definitions on our own, there was no way we could do it by ourselves. AHRQ offered the set of measures prescribed for our needs.

A representative of a state doing public reporting told us:

> If we didn't have the AHRQ QIs, it would be difficult as a state to come up with our own indicators and there are not many other choices that are based on administrative data. Until electronic medical records are commonplace (5-10 years at least), we need to deal with using administrative data.

An insurance company representative highlighted the importance of AHRQ's role in the quality indicator market, stating that more marketing of the QIs is needed:

> AHRQ is doing something that no one else is doing. We have to have a national standard, something used across the country for comparison. [Does AHRQ] realize they're one of the only good options out there? They should really pick up the outreach so that others will pick up using the QIs.

### 3.3.1 Overview of users and uses of the AHRQ QIs

AHRQ's unique position in the market for quality indicators has led to a wide proliferation of uses for the AHRQ QIs. Our environmental scan of users of the AHRQ QIs identified 114 users of the indicators and a range of different purposes, including public reporting, quality improvement/benchmarking, pay-for-performance, and research. Table 3.3 summarizes the number of users of the AHRQ QIs by type of organization and purpose of use.

Table 3.3. Users of the AHRQ QIs

| Type of Organization | Type of Use | | | | | |
|---|---|---|---|---|---|---|
| | Public Reporting | Quality Improvement/ Benchmarking | Pay-for-Performance | Research | Other/ Unknown | Total |
| Business Group | 2 | | | | | 2 |
| Consulting Firm | | | | 2 | | 2 |
| Employer | | 1 | | | | 1 |
| Federal Government | | 1 | 1 | 19 | | 21 |
| Health plan | 1 | 1 | 3 | | 4 | 9 |
| Hospital Association | 1 | 8 | | 2 | | 11 |
| Hospital or Hospital Network | 2 | 3 | | 1 | 9 | 15 |
| Integrated Delivery System | | 2 | | | 7 | 9 |
| Other | 2 | 4 | | | 1 | 7 |
| Research Organization | | 1 | | 14 | 1 | 16 |
| State or Local Government | 12 | 2 | | 5 | 2 | 21 |
| **Total** | **20** | **23** | **4** | **43** | **24** | **114** |

*Source*: RAND analysis of environmental scan results

The most common uses of the AHRQ QIs include:

- **Research.** We identified 43 organizations that use AHRQ QIs for research. For example, Leslie Greenwald and colleagues used the AHRQ QIs to compare the quality of care provided in physician-owned specialty hospitals and competitor hospitals.[26]

- **Quality improvement.** We identified 23 organizations that use the AHRQ QIs as part of a quality improvement activity, including reports benchmarking performance against peers; however, these organizations do not release the quality information into the public domain.[g]

---

[g] Due to the methods used to identify users, the scan is likely to have significantly undercounted the number of organizations (especially hospitals and hospital associations) using the AHRQ QIs for

- **Public reporting.** We identified 20 organizations using the AHRQ QIs for public reporting. We classified an activity as "public reporting" if a publicly available report was published that compares AHRQ QI results between hospitals or geographic areas such as counties. The organizations using the AHRQ QIs for public reporting, with Web links to the reports, are listed in Table 3.4.

- **Pay-for-Performance.** We identified 4 organizations that are using the AHRQ QIs in pay-for-performance programs. Three were health plans and one was a CMS demonstration project.

---

internal quality improvement activities, since this type of use rarely results in publicly available information that could be used to identify the user in an environmental scan.

Table 3.4. Organizations Using the AHRQ QIs for Public Reporting

| Organization Name | Type of Report | Description | QIs used | Citation |
|---|---|---|---|---|
| AFSCME Council 31 | One-time report | The union published a report on quality at Resurrection Health Care hospitals after complaints about quality from workers. | IQIs 15-20 | AFSCME Council 31. The High Price of Growth at Resurrection Health Care: Corporatization and the Decline of Quality of Care. November 2005. Available at: http://www.afscme31.org/cmaextras/qualityofcare.pdf , last accessed January 2006. |
| California Office of Statewide Health Planning & Development | Interactive tool and periodic written reports | A Web site includes an interactive tool for hospital comparison on selected IQIs and other risk-adjusted mortality indicators. | IQIs 1, 2, 4-7, 21-23, 33, 34; PDI 7 | California Office of Statewide Health Planning and Development. "Consumer Information on Quality of Care." Available at: http://www.oshpd.ca.gov/oshpdKEY/qualityofcare.htm, last accessed September 2006. |
| Chicago Department of Public Health | Periodic report | Chicago runs a Web site providing a health profile of city community areas, including PQIs. | PQIs (all except 2,9) | City of Chicago. "Community Health Profiles." Available at: http://www.cchsd.org/cahealthprof.html, last accessed September 2006. |
| Colorado Health and Hospital Association | Periodic report | Hospital reports are shared among hospitals and published on a Web site. | IQIs 15-20, 12-14, 30, 31, 4-7 | http://www.hospitalquality.org/index.php , last accessed November 2005. |
| Connecticut Office of Health Care Access | One-time report | Databook on preventable hospitalizations. | PQIs (all) | Office of Health Care Access Databook. Preventable Hospitalizations in Connecticut: Assessing Access to Community Health Services. FY2000-2004. Available at: http://www.ct.gov/ohca/lib/ohca/publications/acsc_databook00-04.pdf , last accessed November 2005. |
| Excellus Blue Cross/Blue Shield | Interactive tool | Online hospital comparison tool for health plan members only. | Unspecified (members only) | https://www.excellusbcbs.com/guests/find_a_doctor_or_hospital/click_and_compare.shtml#, last accessed September 2006. |
| Exempla Healthcare | Periodic report | Exempla publishes quality | IQIs 12-20, | http://www.exempla.org/about/quality/MortalityRep |

| | | | | |
|---|---|---|---|---|
| | | information on its hospitals on its Web site. (The same results are also reported by the Colorado Health and Hospital Association). | 30, 31 | ortExemplaELMCandESJH51105.htm, last accessed September 2006. |
| Florida State Center for Health Statistics | Interactive tool | Online hospital comparison tool. | PSIs 3, 6-8, 12, 13; IQIs 8-20, 32 | http://www.floridacomparecare.gov/ , last accessed September 2006. |
| Georgia Partnership for Health and Accountability | Periodic report | A periodic report on health in Georgia includes a chapter on avoidable hospitalizations. | PQIs 3, 5, 8, 10, 11, 15 | Georgia Partnership for Health & Accountability. The State of the Health of Georgia, 2004: Ambulatory Care Sensitive Conditions. Available at: http://www.gha.org/pha/publications/stateoftheheal th/2004/ACS112704.pdf , last accessed November 2005 |
| Massachusetts Dept. of Health and Human Services | Interactive tool | Online hospital comparison tool. | IQIs 14, 16-20, 32, 21, 33, 34 | www.mass.gov/healthcareqc , last accessed September 2006. |
| Missouri Department of Health and Senior Services | Periodic report | Comparison of hospital surgery volume to help consumers choose a hospital. | IQIs 1, 2, 4-7; PDI 7 | http://www.dhss.mo.gov/HospitalSurgeryVolume/index.html |
| Niagara Health Quality Coalition and Alliance for Quality Health Care | Interactive tool | Online hospital comparison tool. | IQIs 1-25 | http://www.myhealthfinder.com/, last accessed September 2006. |
| Norton Healthcare | Interactive tool | Health system publishes quality data for its hospitals on its Web site. | PSIs 1-6, 8-16, 18-20; IQIs 1, 2, 4-9, 11-20, 22-24, 30, 31, 34; PDIs 2, 3, | http://www.nortonhealthcare.com/about/qualityreport/index.aspx, last accessed September 2006. |

| | | | | |
|---|---|---|---|---|
| | | | 4-9, 11, 13 | |
| Ohio Department of Health | Periodic report | Online comparison of avoidable hospitalizations by county. | PQIs 1, 4, 5, 7, 8, 11, 14, 15 | http://www2.odh.ohio.gov/Data/CntyPfls/primcare1.htm, last accessed September 2006. |
| Oregon | Interactive tool | Online hospital comparison tool and a report on Oregon' safety net by the Safety Net Advisory Council. | IQIs 11, 12, 15-17, 19, 20, 30; PQIs 3, 5, 8, 10, 11, 12, 15 | http://egov.oregon.gov/DAS/OHPPR/HQ/HospReports.shtml (IQIs) http://www.oregon.gov/DAS/OHPPR/SNAC/SNACWelcomePage.shtml#Power_Point_Presentations (PQIs) |
| Rhode Island | One-time report | Report on hospital procedure volumes. Future reports on IQIs and PSIs in preparation. | IQIs 1-7 | Williams KA, Buechner JS. Health by Numbers Vol. 6, No. 2. February 2004. Available at: http://www.health.ri.gov/chic/statistics/hbn_feb2004.pdf , last accessed December 2005. |
| Texas Health Care Information Collection | Interactive tool | Online hospital comparison tool. | IQIs 1-14, 16-20, 22-25, 30-33; PQIs (all) | http://www.dshs.state.tx.us/THCIC, last accessed September 2006. |
| The Alliance (Wisconsin) | Periodic report | QualityCounts report on hospital safety performance. The report is based on AHRQ PSIs but modifies them for reporting. | PSIs 3, 6, 7, 8, 12, 17; IQI 33 | http://www.qualitycounts.org/, last accessed September 2006 |
| Utah Department of Public Health | Periodic report | Web site providing health information for geographic areas. Three PQIs are included with numerous health status and other measures. State-level IQI results are presented on a one-page poster, available online. | PQIs 11, 4, 1+3+14 combined; IQIs (all) | PQIs: http://ibis.health.utah.gov/indicator/index/alphabetical.html IQIs: http://health.utah.gov/hda/AHRQ2005.pdf |
| Vermont Department of Banking, Insurance, Securities | Periodic report | Online hospital comparison report. | IQIs 1, 2, 4-9, 11, 12, 30, | http://www.bishca.state.vt.us/HcaDiv/HRAP_Act53/HRC_BISHCAcomparison_2006/BISHCA_HRC_compar_menu_2006.htm, last accessed September 2006. |

| | |
|---|---|
| & Health Care Administration | 31; PDIs 6, 7 |

*Source*: RAND analysis of environmental scan results

*Note:* "Public reporting" was defined as a publicly available report that compares AHRQ QI results between hospitals or geographic areas such as counties. Not all of the public reports identified in this table are intended to influence consumers' choice of provider.

"One-time reports" are published comparisons that are not labeled as an ongoing activity.
"Periodic reports" are published comparisons, updated periodically, that are in static format (e.g., documents available as .pdf files online).
"Interactive tools" are online comparisons that allow users to create customized reports (e.g., selection of providers or indicators of interest).

### 3.3.2 Uses of Specific AHRQ QIs

We asked users of the AHRQ QIs, and vendors of quality measurement packages including the AHRQ QIs, which specific QIs they were using. Among organizations we interviewed, the PSIs and IQIs were used more frequently than the PQIs. Of the 42 organizations, 33 were using the PSIs, 30 were using the IQIs, and 17 were using the PQIs.

Within the PSI and IQI sets, some indicators were used more frequently than others. Users of the PQIs, on the other hand, were more likely to use every PQI. There were no meaningful differences in the frequency of use of particular PQIs (data not shown).

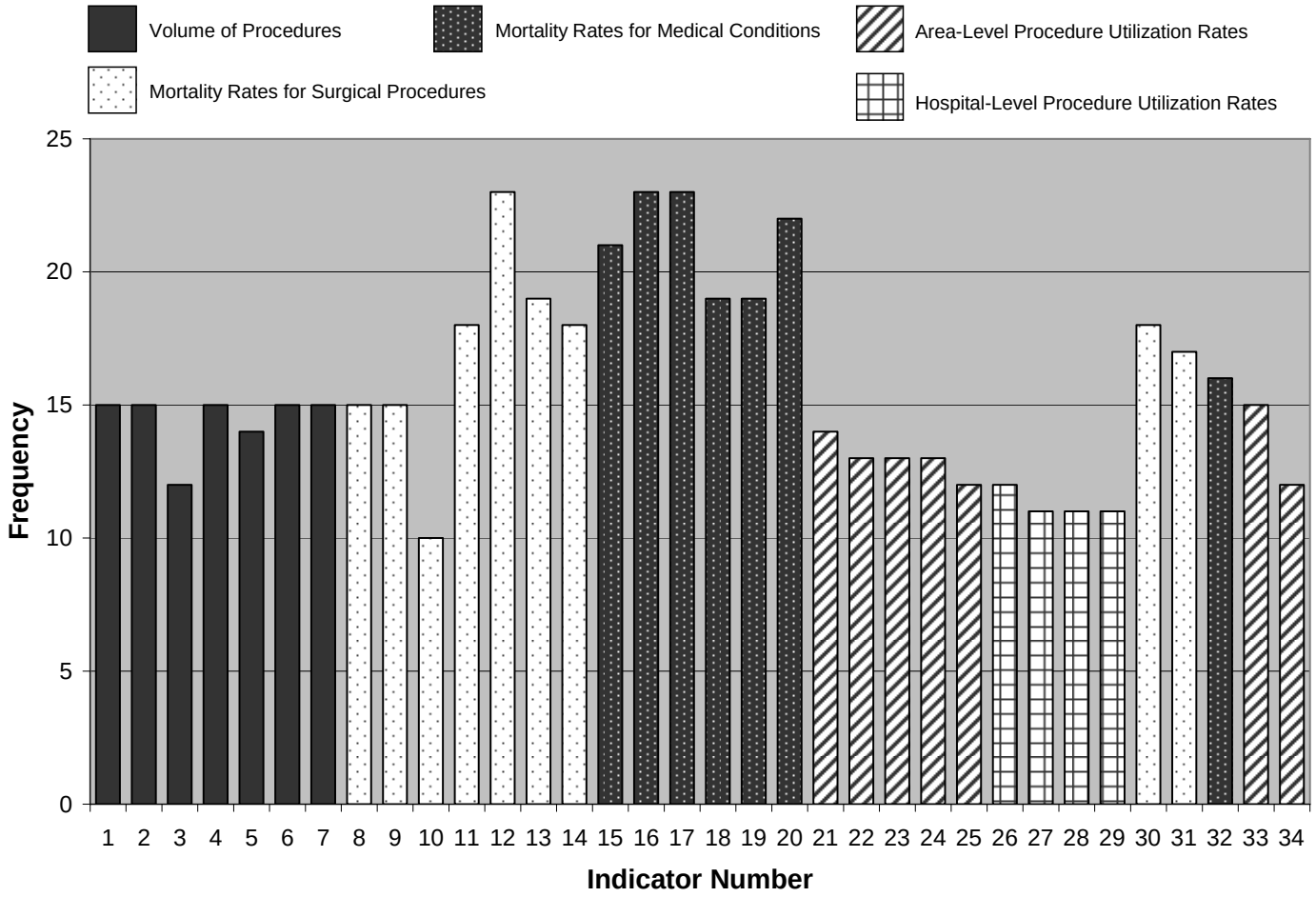
#### 3.3.2.1 Use of IQIs

Figure 3.1 shows the frequency of use of each IQI by the users and vendors we interviewed. The IQIs that reflect mortality rates for medical conditions were used most frequently, particularly IQI 16 – congestive heart failure mortality (23 users), IQI 17 – stroke mortality (23 users), and IQI 20 – pneumonia mortality (22 users). The rates of mortality for certain medical procedures were also commonly used, particularly IQI 12 – coronary artery bypass graft mortality (23 users), IQI 13 – craniotomy mortality (19 users), IQI 11 – abdominal aortic aneurysm repair mortality (18 users), IQI 14 – hip replacement mortality (18 users), and IQI 30 – percutaneous transluminal coronary angioplasty mortality (18 users). The procedure volume indicators were used less frequently, and the procedure utilization rates, both hospital- and area-level, were used least frequently.

#### 3.3.2.2 Use of PSIs

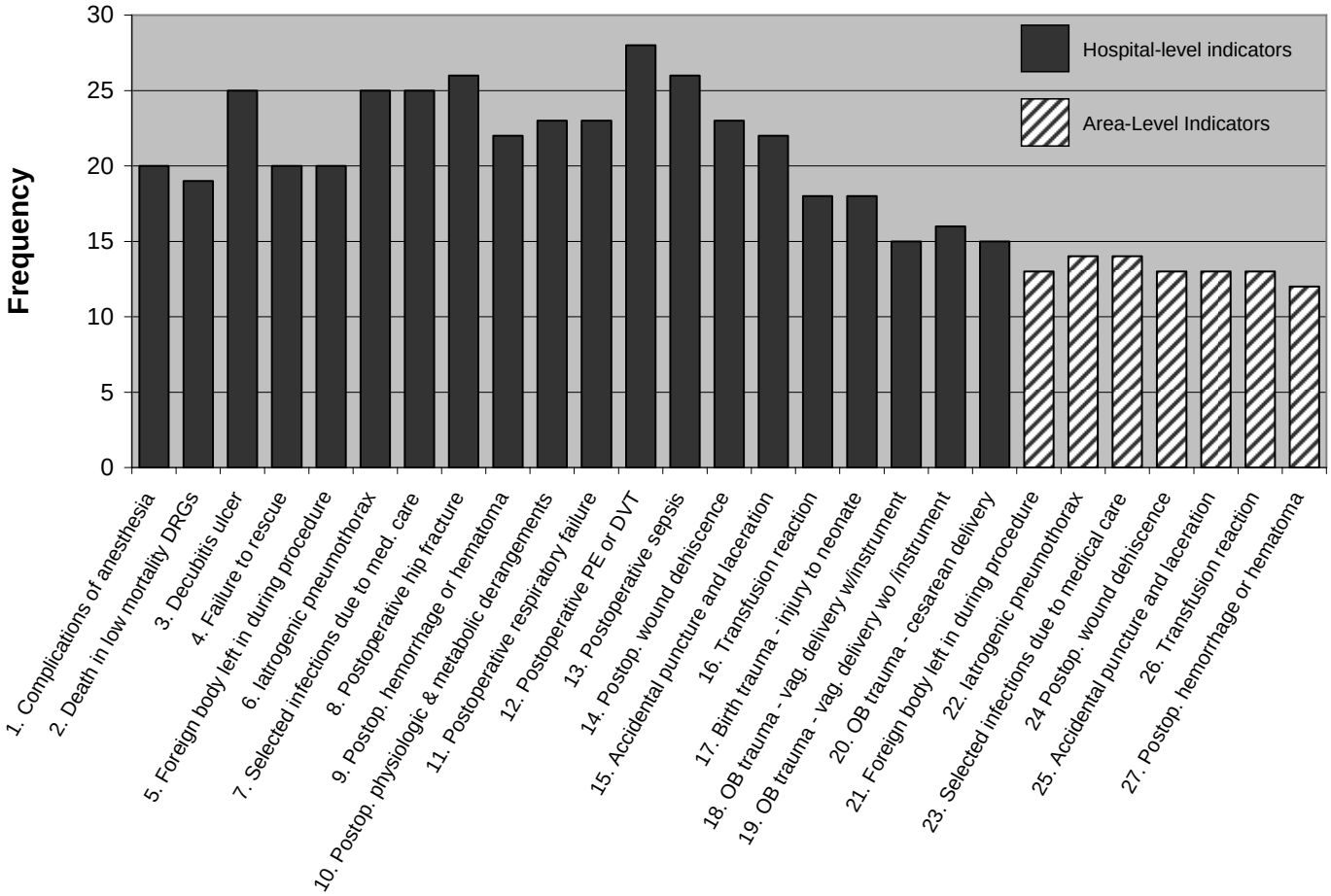
Figure 3.2 shows similar counts of the frequency of use of each PSI. The area-level PSIs were used less frequently than the hospital-level PSIs. Among the hospital-level indicators, there was considerable variation in frequency of use between the indicators. The most frequently used PSIs were PSI 12 – postoperative pulmonary embolism (PE) or deep vein thrombosis (DVT) (28 users), PSI 8 – postoperative hip fracture (26 users), and PSI 13 – postoperative sepsis (26 users). The least frequently used hospital-level PSIs were PSIs 18, 19, and 20 – obstetric trauma with instrument, without instrument, and during cesarean delivery (15, 16, and 15 users, respectively).

Figure 3.1. Frequency of Use of Specific AHRQ IQIs among 42 Users of the AHRQ QIs




*Source*: RAND analysis of environmental scan results

Figure 3.2. Frequency of Use of Specific AHRQ PSIs among 42 Users of the AHRQ QIs




*Source*: RAND analysis of environmental scan results

### 3.3.3 International uses of QIs

Measuring quality of care has become a policy priority in many countries outside of the United States, and numerous countries and international organizations are in the process of instituting requirements for data collection and reporting of quality indicators.[27] The AHRQ QIs are an attractive option for international users, since many countries already require hospitals to report the required administrative data.

Perhaps the most visible international endeavor is the Organization for Economic Cooperation and Development's (OECD) Health Care Quality Indicators (HCQI) Project. The OECD is an intergovernmental economic research institution headquartered in Paris, France, with a membership of 30 countries that share a commitment to democratic government and the market economy. One of its widely used products is OECD Health Data, which provides internationally comparable information on infrastructure, cost and utilization at the health system level,[28] but so far no information on quality of care. In an attempt to bridge this gap, in 2003 the OECD brought 23 of its member countries together with international organizations such as the World Health Organization (WHO) and the European Commission (EC), expert organizations such as the International Society of Quality in Healthcare (ISQua) and the European Society for Quality in Healthcare (ESQH), and several universities.[29] The goal of the meeting was to work on the development and implementation of quality indicators at the international level.

The project initiated its work with two major activities. The first was an effort to introduce a pilot set of quality indicators that can be reported by a substantial portion of the OECD countries. This activity recently led to the 2006 release of an initial list of indicators and corresponding data.[30] The second activity was to identify additional quality indicators for five priority areas: cardiac care, diabetes mellitus, mental health, patient safety, primary care/prevention/health promotion. Through an expert panel process, 86 indicators were selected for the five areas and the OECD is currently investigating the availability and validity of required data.[31] Several AHRQ PSIs were selected for the patient safety area[32] and an indicator similar to the PQIs was selected for the primary care area.[33]

Researchers from several countries have tried to run the PSIs against national discharge data both as part of their participation in the HCQI Project and also for other projects. This has been attempted in Canada, Germany, Italy, and Sweden. In addition, a group in Belgium successfully constructed some of the HCUP indicators, the predecessors of the AHRQ QIs, from national data.[34]

At this point, results from those projects are largely unpublished in English-language journals. But during a recent OECD meeting in Dublin, Ireland, experts from 15 countries discussed issues around the adoption of the AHRQ PSIs in countries other than the United States. Researchers from several countries had cross-walked the AHRQ PSI specifications, which are based on the U.S.-specific ICD-9-CM diagnosis codes, to ICD-10 diagnosis codes, which most countries are now using. This conversion was found to be unproblematic, in particular because only a limited number of diagnosis codes had to be cross-walked to construct the indicators. A greater issue turned out to be the conversion of procedure codes. The AHRQ definitions are based on the ICD-9 procedure classification, whereas other countries use national procedure classification systems. Similarly, other countries use different DRG groupings than those used in the United States. Substantial work on mapping the different coding systems used in the U.S. and in other countries is needed.

In countries that have tested the AHRQ PSIs, the average rates were reported to be similar to those observed in the United States. Countries that do not yet have DRG-based prospective payment systems saw much lower rates, possibly resulting from less thorough coding of secondary diagnoses in the absence of implications for payment.

Our interviews show that there is interest in using the AHRQ QIs internationally and sufficient data and technical capability to implement them. This makes it likely that some AHRQ QIs will be adopted by the OECD HCQI Project for international comparisons of quality of care and patient safety. Furthermore, as several international organizations are

striving to align their measurement and reporting activities,[h] selected AHRQ QIs could become part of an international measurement standard.

### 3.3.4 "Non-users" of QIs

We identified and interviewed representatives of five organizations that are currently using quality indicators other than the AHRQ QIs but that are similar to the AHRQ QIs. Our goal was to understand why these organizations were not using the AHRQ QIs, and specifically whether this decision was based on an evaluation of the merits of the QIs. Three of the organizations were using quality indicators that they had developed themselves and which predated the AHRQ QIs. They did not voice any strong objections to the AHRQ QIs but preferred their own indicators due to various methodological factors and the fact that their indicators were better tailored for their specific needs. The other two organizations had elected not to use the AHRQ QIs because they were not already in use by the hospitals that would be participating in the organizations' quality measurement activities. The JCAHO Core Measures were chosen instead because they were already being collected by hospitals.

### 3.3.5 Examples of uses of QIs

In order to illustrate how the AHRQ QIs are being used, we have chosen examples of specific uses for quality monitoring, public reporting, and pay-for-performance.

#### 3.3.5.1 Example of AHRQ QI use for quality improvement

The following example was drawn from a report provided to hospitals by a hospital association we interviewed. Reports such as the one we reviewed are sent to hospital CEOs quarterly. The reports include all AHRQ IQIs (shown in the example) as well as all AHRQ PSIs. The report also includes indicators from JCAHO and Leapfrog (not shown). Hospitals are presented with targets based on benchmarks calculated by the hospital

---

[h] For example, the European Commission has recently ceded its activities in quality indicator development to the OECD to avoid duplication and is funding part of the HCQI Project.

association. The hospital association works with hospitals to help them explain why they do not meet targets in areas of poor performance.

Figure 3.3. Sample AHRQ QI Report Used by a Hospital Association for Quality Improvement

| AHRQ Quality Indicators | Apr - Jun 2005 (Q2) Relative Performance | Denom | Observed | Target | Median | Jul 2004 - Jun 2005 (recent year) Relative Performance | Denom | Observed | Target | Median |
|---|---|---|---|---|---|---|---|---|---|---|
| | | N | Percent | | | | N | Percent | | |
| **Post Procedure Mortality (%)** | | | | | | | | | | |
| CABG | ⊙ | 74 | 4.3 | 3.2 | 3.5 | ⊙ | 299 | 3.9 | 3.2 | 3.8 |
| PTCA | ⊙ | 181 | 1.3 | 1.3 | 1.1 | ⊙ | 722 | 2.1 | 1.4 | 1.4 |
| Carotid endarterectomy | ⊙! | 16 | 0.8 | 0.6 | 0.0 | ⊙ | 65 | 0.6 | 0.7 | 0.0 |
| AAA repair | ⊙! | 6 | 10.9 | 12.6 | 0.0 | ⊙! | 20 | 13.4 | 11.9 | 6.2 |
| Esophageal cancer resection | ⊙! | 2 | 2.5 | 4.9 | 0.0 | ⊙! | 8 | 7.4 | 6.3 | 0.0 |
| Pancreatic cancer resection | ⊙! | 5 | 4.1 | 4.8 | 0.0 | ⊙! | 18 | 3.7 | 5.1 | 0.0 |
| Pediatric heart surgery | ⊙ | 28 | 6.0 | 5.0 | 2.2 | ⊙ | 90 | 4.6 | 7.1 | 3.1 |
| Craniotomy | ⊙ | 84 | 6.1 | 5.8 | 4.9 | ⊙ | 313 | 6.6 | 5.9 | 5.5 |
| Hip replacement | ⊙ | 36 | 0.1 | 1.2 | 0.0 | ⊙ | 133 | 0.1 | 1.2 | 0.0 |
| **In Hospital Mortality (%)** | | | | | | | | | | |
| AMI - Mortality | ⊙ | 88 | 8.3 | 7.6 | 5.8 | ⊙ | 360 | 8.0 | 7.9 | 7.1 |
| AMI - Mortality w/o transfers | ⊙ | 62 | 7.6 | 7.9 | 5.8 | ⊙ | 248 | 7.7 | 8.4 | 6.9 |
| HF | ⊙ | 137 | 3.0 | 4.3 | 2.9 | ⊙ | 527 | 3.3 | 4.3 | 3.1 |
| Pneumonia | ⊙ | 125 | 6.0 | 6.2 | 5.5 | ⊙ | 492 | 6.0 | 6.0 | 5.7 |
| Acute Stroke | ⊙ | 83 | 12.6 | 13.5 | 12.8 | ⊙ | 316 | 12.8 | 13.2 | 13.5 |
| GI hemorrhage | ⊙ | 55 | 2.7 | 6.1 | 2.6 | ⊙⊙ | 221 | 2.9 | 6.1 | 2.7 |
| Hip fracture | ⊙ | 26 | 3.1 | 3.1 | 1.7 | ⊙ | 98 | 2.8 | 3.0 | 2.6 |
| **Utilization Rates (%)** | | | | | | | | | | |
| C section all | | 527 | 25.8 | | 25.7 | | 2,072 | 26.0 | | 25.6 |
| C section primary | | 452 | 16.6 | | 16.3 | | 1,779 | 17.1 | | 16.6 |
| VBAC all | | 93 | 17.5 | | 16.9 | | 361 | 19.0 | | 17.9 |
| VBAC uncomplicated | | 76 | 18.4 | | 17.6 | | 299 | 19.7 | | 18.9 |
| | | | N | | N | | | N | | |
| **Volumes** | | | | | | | | | | |
| CABG | | | 75 | | 55.5 | ⊙⊙ | | 302 | 100.0 | 230.0 |
| PTCA | | | 182 | | 124.0 | ⊙⊙ | | 725 | 200.0 | 448.0 |
| Carotid endarterectomy | | | 16 | | 12.0 | ⊙ | | 65 | 50.0 | 41.0 |
| AAA repair | | | 6 | | 4.0 | ⊙ | | 20 | 20.0 | 13.0 |
| Esophageal cancer resection | | | 3 | | 2.0 | ⊙⊙ | | 8 | 6.0 | 5.5 |
| Pancreatic cancer resection | | | 6 | | 5.0 | ⊙⊙ | | 18 | 10.0 | 13.5 |
| Pediatric heart surgery | | | 26 | | 16.0 | ⊙ | | 90 | 90.0 | 48.0 |

⊘ No Data Available ● Substantially Worse than Target ◒ Worse than Target ⊙⊙ Better than Target ⊙ Meets Target ! Interpret with Caution (less than 25 cases)

### 3.3.5.2 Example of AHRQ QI use for public reporting

The State of Florida uses the AHRQ QIs as part of a public reporting tool aimed to help consumers choose a hospital. The screen shot below captures a segment of a Web page comparing hospitals in Broward County on one of the AHRQ IQIs, postoperative hip fracture (IQI 19). Users can click on a hospital to get more detailed information on quality as well as the hospital's characteristics (teaching status, non-profit status, etc.) and location.

**Figure 3.4. Sample AHRQ QI Report Used by the State of Florida for Public Reporting**

Postoperative hip fracture
Age 18 years and older
Time Period: July 2004 through June 2005

| Facility / City | Risk Adjusted Rate |
|---|---|
| STATEWIDE | 0.03% |
| Memorial Regional Hospital 100038<br>Hollywood | Higher than Expected<br>0.1% |
| North Broward Medical Center 100086<br>Pompano Beach | As Expected<br>0.07% |
| North Ridge Medical Center 100237<br>Fort Lauderdale | As Expected<br>0.04% |
| Northwest Medical Center 100189<br>Margate | As Expected<br>0.0% |
| Plantation General Hospital 100167<br>Plantation | As Expected<br>0.0% |
| University Hospital and Medical Center 100224<br>Tamarac | As Expected<br>0.0% |
| Westside Regional Medical Center 100228<br>Plantation | As Expected<br>0.09% |

**All copyrights in and to 3M™ APR™DRGs Classification System and 3M™ APR™ Software are owned by 3M. All rights reserved.**

What the Complication/Infection rate means:

The percentage rates reported on this page reflect each hospital's unique population, and should not be compared between hospitals. Instead, it is strongly recommended consumers compare each hospital on the basis of whether their rates are "as expected", "lower than expected", or "higher than expected".

The percentage rate is:

- Lower than Expected - Fewer complications/infections than expected given how sick patients were
- As Expected - Expected number of complications/infections given how sick patients were
- Higher than Expected - More complications/infections than expected given how sick patients were

'X' is less than 30 cases, 'Too few cases' is less than 5 cases.

### 3.3.5.3 Example of AHRQ QI use for Pay-for-Performance

The following example is drawn from a report provided to hospitals by an insurer we interviewed. The example extracts one AHRQ PSI (PSI 12), postoperative pulmonary embolism (PE) or deep vein thrombosis (DVT). The report allows hospitals to compare their performance to that of their peers. Good performers earn an incentive payment.

Figure 3.5. Sample AHRQ QI Report Used by an Insurance Company for Pay-for-Performance

| CONFIDENCE INTERVAL DESIGNATIONS | | |
|---|---|---|
| ● | Signif Diff at 80% | ● |
| ●● | Signif Diff at 90% | ●● |
| ●●● | Signif Diff at 95% | ●●● |
| ■ | No Signif Diff vs Cohort Average | ■ |

| 110 - Postoperative PE or DVT | | | | | |
|---|---|---|---|---|---|
| Hospital Nr | Sig | Cases | At Risk | Rate | APO* |
| 2135 | ●●● | 25 | 5294 | 0.47 | 0.68 |
| 2020 | ●●● | 8 | 1835 | 0.44 | 0.70 |
| **10th PERCENTILE** | | | | | **0.79** |
| 2101 | ● | 18 | 2007 | 0.90 | 0.83 |
| 2014 | ● | 31 | 3772 | 0.82 | 0.88 |
| 2007 | ■ | 19 | 2263 | 0.84 | 0.92 |
| 2094 | ■ | 21 | 2226 | 0.94 | 0.92 |
| 2108 | ■ | 8 | 1393 | 0.57 | 0.96 |
| 2010 | ■ | 29 | 3048 | 0.95 | 1.01 |
| 2107 | ■ | 36 | 3541 | 1.02 | 1.06 |
| **COHORT AVERAGE** | | | | | **1.12** |
| 2099 | ■ | 18 | 1494 | 1.20 | 1.19 |
| 2114 | ■ | 29 | 1837 | 1.58 | 1.21 |
| 2040 | ■ | 17 | 1860 | 0.91 | 1.22 |
| 2058 | ■ | 22 | 1797 | 1.22 | 1.32 |
| 2337 | ■ | 61 | 3928 | 1.55 | 1.33 |
| 2118 | ■ | 24 | 1636 | 1.47 | 1.37 |
| 2075 | ●● | 29 | 2870 | 1.01 | 1.55 |
| **90th PERCENTILE** | | | | | **1.55** |
| 2149 | ●● | 60 | 2955 | 2.03 | 1.56 |
| **2225** | **●●●** | **51** | **2126** | **2.40** | **1.97** |

# 4. EVALUATION OF THE AHRQ QI PROGRAM

In this section, we discuss the results of our environmental scan and interviews with regard to the evaluation of the AHRQ QI indicators. We organize the discussion according to four factors that are used as criteria for evaluating quality indicators: importance, scientific soundness, usability, and feasibility. Since this report focuses on the AHRQ QI program as a whole, the comments and insights should be interpreted broadly, and not as critiques of individual indicators. For example, "importance" here refers mainly to interviewees' perceptions of the AHRQ QI program as a whole, not the importance of the constructs underlying specific AHRQ QIs.

## 4.1 IMPORTANCE

### 4.1.1 Users' general views on the importance of the AHRQ QI program

Representatives of nearly all of the organizations stressed the importance of the AHRQ QI program. When asked an open-ended question about the role of AHRQ in quality measurement, 11 of 54 interviewees identified AHRQ as a "national leader" in measurement development and research. The AHRQ QI program was described by a vendor as "a major player, both nationally and internationally…a leader, the top of the pyramid." One interviewee captured this sentiment:

> AHRQ is a very important player and has a rich history of research and evidence basis. The products they provide help everyone develop measures, such as the National Guideline Clearinghouse. The measures they have done to date have an audience, a place and a role – I know states use them.

Interviewees stressed that without the AHRQ QIs, they would have few alternatives and would likely have to drastically change or eliminate their quality reporting and/or measurement activities. As discussed in more detail below, the scientific soundness of the QIs was highly regarded, as was the transparency of the QI evidence review and validation that was conducted as a part of the AHRQ QI development process.

Interviewees generally felt that it was important that a federal agency like AHRQ, which is regarded as credible and respected, develop and support a quality indicator set for public use. AHRQ's credibility and the transparency of the AHRQ QI methods was often

mentioned as a key factor in overcoming opposition to quality measurement and reporting by stakeholders, particularly providers. We were told:

> There is a lot of good documentation regarding how rigorously the indicators have been analyzed by AHRQ, researchers, academics, etc., in a collaborative effort. This is important, especially for hospital administrators, who have to convince medical staff that at least there is a rigorous process behind the indicators.

Overcoming this type of opposition is particularly important for public reporting and pay-for-performance initiatives, where providers' reputations and revenues are at stake. In the scenarios described by many of our interviewees, providers are typically not opposed on conceptual grounds to increasing the transparency of the quality of care they provide. However, providers are sensitive to being evaluated using measures that are unreliable or invalid, and they value the opportunity to be able to review and evaluate the measures they are subjected to and to raise objections to the results, where appropriate.

### 4.1.2 Importance of the Individual AHRQ QIs and Underlying Constructs

Although interviewees were not asked to comment on the importance of the constructs underlying the AHRQ QIs or on individual indicators, a few interviewees raised these issues. When asked why they use the AHRQ QIs, some interviewees mentioned that the AHRQ QIs provide a "good estimate" or that they offer a "reflection of reality."

Several interviewees also remarked that they appreciated having access to the evidence showing that the AHRQ QIs represent important opportunities for quality improvement, which is made available in the AHRQ technical documentation under the headings "face validity" and "fosters real quality improvement."[35] A number of interviewees (10 of 54) mentioned that the availability of this information in the documentation is a key reason why they decided to use the AHRQ QIs, or described the documentation as a factor that facilitated the use of AHRQ QIs in the face of opposition from stakeholders.

### 4.1.3 Impact of AHRQ QI use

Although only one organization in our sample had formally measured the impact of AHRQ QIs on the quality of care delivered to patients, many interviewees provided anecdotal evidence of the effect of the indicators on quality. The one organization that did report conducting a study of the impact of its use of the AHRQ QIs was The Alliance, a

Wisconsin employer-purchasing cooperative that publishes a quality report called QualityCounts. The evaluation of the impact of QualityCounts was conducted by Judith Hibbard and a team from the University of Oregon and was published in *Health Affairs*.[36] The study found that public reporting resulted in increased hospital quality in the clinical areas included in the QualityCounts public report. The improvement appears to be driven by hospital administrators' concerns about their reputation.

When asked whether they had measured the impact of using the AHRQ QIs, a number of interviewees (9 of 29 answering this question) responded that indicator use began too recently to allow for observation of any impact. In addition, several interviewees stated that the results of the AHRQ QIs can be difficult to track longitudinally, since the specifications of the indicators have changed over time.

However, 12 of the 29 interviewees who answered the question on impact reported anecdotal evidence that their or their clients' use of the AHRQ QIs was having some type of impact on quality of care. The impacts observed usually consisted of an activity such as putting a new quality improvement process in place, rather than an improvement in outcomes. Examples of this type of anecdote include:

- A hospital representative reported:

  > We've definitely seen an impact on quality in areas flagged by the AHRQ QIs. Some have been data problems and some have been actual quality improvements. For example, using the infection indicator (PSI 7) we were able to see improvement after implementing the ventilator and central line bundles. Similarly with the sepsis indicator (PSI 13), we implemented the Surgical Care Procedure Practices – a group of interventions to decrease sepsis, and we saw improvements.

- A hospital network representative reported that staff were able to observe the impact of a quality report card on quality improvements in network hospitals. Two interventions introduced in response to the report card were: 1) new guidelines on the angle of the hospital bed for ventilator-assisted pneumonia patients and 2) implementation of a rapid response team.

- From a hospital using a vendor to implement AHRQ QIs:

  > We identified that we had high failure to rescue rates... This was the information we needed to present to our executive team and board to obtain resources to effectively establish and run a rapid response team.

- A hospital association representative reported:

  There have been some changes in [the AHRQ QIs] data [over time], but I don't know if they've been caused [by our use of the AHRQ QIs for quality improvement]. From 2001 to 2004 there is less variation among hospitals, and mortality has decreased for several indicators; on the other hand, fewer hospitals are at or above the volume thresholds. We have looked at trends in other available data and, to the extent there is overlap, there is some correlation and indication that quality is improving.

- A representative of another hospital association provided anecdotal evidence of quality improvements, and also revealed a barrier to conducting more rigorous assessments of impact:

  Hospitals have taken action in terms of identifying individual cases [from the numerator of AHRQ QIs where a problem is flagged], reviewing them [using clinical data], and developing improvement plans (especially moderate cases, such as infection). There are no published impact studies. The climate (in terms of lawsuits, etc.) stands in the way of publishing studies and until the climate is supportive, hospitals will not publish anything.

- A representative of a state that publicly reports AHRQ QIs noted:

  One example of where the report had an immediate impact was one hospital that wasn't hitting the volume threshold for carotid endarterectomy [IQI 7]. They decided to stop performing them. We would like to evaluate effectiveness of reports at some point but don't have specific plans at this point.

- An insurance company representative using the AHRQ QIs for pay-for-performance believes that the program has had an impact by garnering attention for quality improvement from hospital management:

  The indicators for patient safety have raised awareness. Because real money is now on the table, the result has been that the hospitals' financial people now have a more substantive dialogue with the quality people.

- A researcher who participated in a study that used the AHRQ QIs to evaluate a state-wide hospital policy change reported substantial press coverage of the results and an effect on other states considering the same policy.

The primary type of impact observed, however, was improvement to data quality. Representatives of several organizations stated that they viewed improved data quality as a natural progression in the implementation of a quality measurement program. When a potential quality problem is first flagged using the AHRQ QIs, the first response is to investigate whether the observed issue is due to a problem in the data or a problem with the actual quality of care. Once the provider organization has determined that the result in

question is not a data artifact, the provider often examines clinical data and/or performs some other type of quality improvement activity to determine the cause of the quality problem. One state government representative described this process:

> The first step hospitals take, naturally, when they see a potential problem is to ensure that it is not a data artifact. Hospitals found that they were consistently up-coding or down-coding measures. They usually started with initiatives to fix their data. Hospitals in some cases threw up red flags and started quality initiatives but the first step is to answer the question - is it an artifact of data or real issue? One hospital had 3 flags [potential quality problems indicated by the ARHQ QIs]; two turned out to be data problems, but one – stroke mortality – was a quality problem. However, most of the feedback from hospitals has been around trying to make data better. We don't have plans to evaluate the impact of our program because we just don't have the resources.

## 4.2 SCIENTIFIC SOUNDNESS

### 4.2.1 Reliability

Users largely felt that the AHRQ QIs can be reliably constructed from hospital discharge data, but that there was a certain learning curve during which hospital coding departments had to adjust to the requirements for the QIs. Thus far, coders had mainly been trained to apply coding rules to fulfill reimbursement requirements, but now they had to understand that coding practices also had implications for quality reporting. In selected instances, we heard concerns about ambiguity in the coding rules – that the coding rules did not provide sufficient guidance on whether to code an indicator-relevant diagnosis. For example, we heard repeatedly that coders found it difficult to apply coding rules for vaginal trauma during birth (5 of 36 users).

### 4.2.2 Validity

Our interviewees were impressed by the quality and level of detail of the AHRQ documentation on the face validity of the indicators and stated that the indicators captured important aspects of clinical care. Very rarely were indicators challenged on conceptual grounds. One exception were the VBAC measures (IQIs 22 and 34), because a current American College of Obstetricians and Gynecologists (ACOG) guideline[37] recommends VBAC only for facilities with a sufficient infrastructure for emergency C-section, which is commonly not present in smaller hospitals. Two interviewees who do public reporting with AHRQ QIs challenged the validity of the volume-based IQIs, as they did not think the

scientific evidence was unambiguous for a positive impact of high volumes of complex procedures on outcomes.

Sample size issues (whether due to the rarity of certain procedures or events or the infrequency with which some procedures are conducted at certain facilities) were repeatedly mentioned as threat to the validity of the indicators. In particular, the adverse events underlying some of the PSIs (e.g., PSI 5: foreign body left in during procedure) fortunately occur quite rarely, even in larger facilities. Smaller facilities, such as rural hospitals, are commonly only able to report on three QIs, mortality for acute myocardial infarction (AMI) and pneumonia (IQIs 15, 20, and 32), because they do not have the minimum required number of cases (20) for other indicators. While interviewees agreed on the face validity of the indicators, a third of the interviewees (16 of 54) argued that such sample size limitations would render some indicator rates unstable and thus hard to interpret.

On construct validity, most users stated that the indicators were correctly operationalized within the constraints of the underlying data source. Isolated findings of specification errors were brought to our attention, but interviewees emphasized that the AHRQ team was always able to address those quickly. The limitations of administrative data were frequently mentioned as a threat to validity, because the UB-92 format would not provide a sufficient level of clinical detail to account for all factors that should be considered in constructing the measures. Several potential improvements were mentioned, such as the addition of flags for conditions that were present on admission or for do-not-resuscitate orders. The AHRQ QI team is incorporating functionality for a condition present-on-admission flag into the next iteration of QI specifications.

Some users thought that formal validation studies should be conducted to assess the validity of the indicator results in relation to indicators based on medical records data. As discussed above, we learned that hospitals are conducting analyses to find out whether poor performance on a given QI is due to an organization's coding practices or points to a real quality problem. But those efforts are typically driven by unusually poor performance, are not systematically analyzed, and focus on identifying false positive events (i.e., an adverse event was flagged by the indicator that could not be ascertained through chart review). False negative events (i.e., the indicator algorithm failed to identify an actual adverse event) were rarely researched.

### 4.2.3 Risk adjustment

Since the AHRQ IQIs and PSIs generally represent health outcomes, they are sensitive to the morbidity burden of the patient population and must be risk-adjusted to provide a valid comparison of quality. The IQIs and PSIs currently use different risk adjustment methods, although AHRQ will move to a single method for all of the QIs in the future. Currently, the IQIs use the All Patient Refined Diagnosis-Related Groups (APR-DRGs), a proprietary system owned by 3M Health Information Systems. The PSIs use a public-domain risk-adjustment system developed by AHRQ. The current risk adjustment methods for both the PSIs and the IQIs were regarded as adequate. Users particularly emphasized that the AHRQ method for the PSIs had the advantage of being transparent and easy to understand. Even though the APR-DRGs are based on proprietary software, interviewees were generally comfortable with using them for IQI risk adjustment, because they already used the software for other purposes, such as payment, and were familiar with its structure and logic. However, 22% (12 of 54) of interviewees thought that the risk adjustment approach used for the AHRQ QIs should be improved. In particular, interviewees would like the see both PSIs and IQIs using the same risk adjustment method and would like AHRQ's method to be aligned with that of CMS, University Healthsystem Consortium, and other developers.

## 4.3 USABILITY

As discussed in detail above, the AHRQ QIs have been used by many types of organizations and for a variety of purposes. Most interviewees stated that the AHRQ QI products provided a workable solution for their needs and were very appreciative of the support that the AHRQ QI team provides for implementation and ongoing use. Despite these overall favorable impressions of the usability of the QIs, two issues were raised repeatedly: the need for development of reporting templates, and the need for clearer guidance on the use of the AHRQ QIs for public reporting and pay-for-performance programs.

### 4.3.1 Reporting template

A number of interviewees (9 of 54) highlighted as a top priority the need for a standard format for reporting AHRQ QI results. At the simplest level, some interviewees

wanted AHRQ-supported, standard, basic names for the AHRQ QIs in plain language, as some of the current indicator names are difficult for non-clinical audiences to understand. Other interviewees expressed a desire for more guidance and support on other aspects of presentation. Currently, many organizations have developed their own reporting formats. Interviewees were interested in information such as:

- How should indicators be analyzed and reported?
- How should outliers be identified?
- Which indicators are consumers expected to respond to most?
- How should consumers interpret the results of the indicators?
- How do results compare to national, state, or other benchmarks?

### 4.3.2 Composite indicators

Twelve interviewees expressed a desire for an AHRQ-supported methodology for constructing a composite indicator. Forming composites would allow organizations to summarize the results based on multiple indicators into one statistic, which is easier to grasp and communicate, in particular for non-expert audiences. Composites would also help overcome sample size limitations by allowing information to be pooled. Four organizations whose representatives participated in our interviews have developed their own AHRQ QI composite indicators but most would prefer an AHRQ-developed approach. The AHRQ QI team is currently working on the development of composite indicators to meet those needs.

### 4.3.3 Guidance on the use of the AHRQ QIs for public reporting and pay-for-performance

Not surprisingly, our questions on suitability of the AHRQ QIs for public reporting and pay-for-performance programs led to vivid and often emotionally charged discussions and comments. Interviewees who are currently using the AHRQ QIs for public reporting and pay-for-performance generally felt that they provided a workable solution for their needs. The introduction of those programs typically followed a similar sequence: following the initial decision to start a public reporting or pay-for-performance program, a controversial debate would start on the merits of such initiatives in general, and the

suitability of administrative data for quality measurement in this context in particular. Then, hospitals and physicians would slowly start to participate rather than resist. Many interviewees told us that AHRQ's reputation for high quality research, the excellent documentation of the scientific basis of the indicators, the transparency of the method, and the ease of implementation and use were crucial factors in obtaining buy-in. The first release of the data was commonly accompanied by media attention and anxiety on the part of providers. Both would subside in subsequent releases, as all stakeholders became more familiar and comfortable with the program.

Still, half of the interviewees who use AHRQ QIs for public reporting stated that additional standards and guidance on the reporting of AHRQ QI results were needed. Some interviewees (10 of 54) expressed dissatisfaction with the current AHRQ stance on the appropriateness of the AHRQ QIs for public reporting. These interviewees described the current guidance as "difficult to find," "weak," and presenting "mixed messages." The lack of clarity is perceived to be largely due to shifts in AHRQ's stance on appropriate uses of the QIs. Previously published guidance contained much stronger caveats against inappropriate uses than the current guidance. Interviewees felt that clearer guidance from AHRQ would help to counter opposition from those who argue that the AHRQ QIs should only be used for quality monitoring and improvement and research, but not as a public reporting or pay-for-performance tool.

Taking the opposing view were several interviewees (mostly hospitals) who would like to see AHRQ make a clear statement that the AHRQ QIs are *not* appropriate for use in public reporting, pay-for-performance, or other reporting activities. A representative of one hospital told us:

> The AHRQ QIs are fabulous tools, but they are assessment tools, not judgment tools. AHRQ's white paper was very clear in saying that this was not AHRQ's intent. The issue is that AHRQ allowed folks to go too far without a caveat. They tried with that white paper, but now they're endorsing states using it for public reporting – it's not appropriate.

## 4.4 FEASIBILITY

We were told consistently that a major advantage of the AHRQ QIs was the feasibility of their implementation. They require only administrative data in the UB-92 format to which

many users have routine access, since those data are already being used for billing and other administrative purposes and have to be collected and reported by hospitals in most states.[i]

Interviewees emphasized that another substantial advantage of the AHRQ QIs is that the indicators have clearly defined and publicly available specifications, which helps with implementation of measurement. These specifications were regarded as of particular importance for hospitals, as the originators of the data, because the specifications enable hospitals to work with their coding departments to ensure that the required data elements were abstracted from medical records consistently and with high quality. In addition, users who analyze data with the QIs, such as researchers, appreciated the fact that they could dissect the indicator results and relate them back to individual records. That capability helped researchers gain a better understanding of the indicator logic and distinguish data quality issues from actual quality problems.

---

[i] Similarly, many hospitals currently use the APR-DRG grouper, which is the basis for risk adjustment of the IQIs, for billing and rate setting so that they are familiar with its logic.

# 5. FINDINGS FROM THE CASE STUDIES

In this chapter, we report the results of two case studies, one in the area of Boston, Massachusetts, and the other in Texas, with an emphasis on the Dallas-Fort Worth area. The two market areas were selected to represent one market with a long history of using the AHRQ QIs for public reporting (Dallas-Fort Worth) and one market in which public reporting, pay-for-performance, and tiered insurance products using the QIs have been more recently implemented (Boston). For each case study, we provide a discussion of the impact of public reporting, technical lessons learned, political lessons learned, and implications.

## 5.1 BOSTON

### 5.1.1 Boston Market Area

The Boston health care market is somewhat unique, marked by a large number of well-known academic medical centers, higher-than-average health care costs,[38] and a large number of practicing physicians and other health care workers.[39] Research and training of physicians and nurses are prevalent in the Boston area.

The Boston health insurance market is dominated by three players: Blue Cross Blue Shield of Massachusetts (BCBSMA), Harvard Pilgrim Health Care, and Tufts Health Plan, with BCBSMA experiencing significant membership growth in recent years.[40] Providers are also rather concentrated, with the largest organization being Partners Health Care, which includes two prominent academic medical centers – Massachusetts General Hospital and Brigham and Women's Hospital.

Quality improvement has a long history among Boston-area providers. Indeed, a great deal of the research into quality improvement and quality measurement has been conducted by Boston-area researchers, and prominent quality improvement organizations, such as the Institute for Healthcare Improvement, are located in the area.

A 2005 report by the Center for Studying Health System Change (HSC) found that public reporting and pay-for-performance activities were leading providers to view their performance on quality and efficiency metrics as "a necessary competitive strategy" in

addition to the "mission-driven efforts" that had guided long-standing quality improvement efforts.[41] Many of these recent activities incorporate the AHRQ QIs.

### 5.1.2 Background on Use of Quality Indicators

Although quality measurement and quality improvement activities have a long history in the Boston area, the AHRQ QIs have recently been used for several new purposes, including pay-for-performance, public reporting, and tiered insurance products[j] based on quality indicators. These new activities have been met with stiff resistance from some members of the Boston provider community.

In 2000, the State Department of Health and Human Services (DHHS) began including the HCUP indicators, the predecessor to the AHRQ QIs, in the performance report it distributed to hospitals. This was done in cooperation with the Massachusetts Hospital Association (MHA) and was not intended for public release. However, even at this early stage some participants "saw the writing on the wall" that public reports would eventually be forthcoming and opposed use of the AHRQ indicators in these reports.

In 2005, DHHS started publishing selected AHRQ IQIs[k] on the internet, along with indicators from the Hospital Quality Alliance, Leapfrog, and the Massachusetts Data Analysis Center (Mass-DAC), a Harvard-based state initiative on the measurement of the quality of cardiac care using clinical data. The decision to publish this information on the Web was prompted by impending state legislation that would require public reporting of hospital quality data. The legislation was part of a more widespread movement in Massachusetts towards greater transparency of provider performance data.

Also in 2003, BCBSMA began reporting selected measures from the AHRQ QIs to hospitals in a pay-for-performance program called the Hospital Quality Improvement Program, now known as the Hospital Performance Improvement Program.[l] The first

---

[j] Tiered insurance products charge patients higher cost-sharing levels for providers that perform worse on quality and/or efficiency indicators.

[k] Specifically, IQIs #12, 14, 16-20, 32, 21, 33, 34.

[l] The indicators used include 15 selected QIs (PSIs # 2, 4, 7, 11, 12, 17-19; IQIs # 12, 15-17, 20; PQI # 1; and PDI 14).

payments to hospitals based on their AHRQ QI results occurred in 2006. Ten JCAHO Core Measures and one of the Leapfrog Leaps (computerized physician order entry) are also reported to hospitals as "advisory measures," but are not tied to incentive payments.

A third major development in quality reporting in the Boston area was spurred by the Massachusetts Group Insurance Commission (GIC) – the organization that manages benefits for state employees. The GIC asked participating health insurers to develop insurance products that tiered provider copayments based on efficiency alone, or efficiency in conjunction with quality. These plans were implemented in 2006. Most of the early focus has been on physician efficiency measurement using Ingenix Episode Treatment Groups (ETGs), but one plan (Tufts Navigator) started tiering hospitals in 2004 using selected AHRQ QIs as well as JCAHO, Leapfrog, and other measures. More health plans, such as Harvard Pilgrim Health Care, may also develop methods of tiering hospitals in the future.[m]

### 5.1.3 Impact

Since many of the quality initiatives involving the AHRQ QIs began only in 2005, interviewees told us that it was "too early" to conduct formal assessments of their impact. There was some scattered anecdotal evidence, however, that the initiatives were spurring quality improvements.

In its pay-for-performance program, BCBSMA works with hospitals to help them explore what may be driving their performance on the AHRQ QIs. We were told that in this capacity, they have observed some processes hospitals have put into place that have improved quality, including, for example, standardized order sets for pneumonia, heart attack and heart failure care; routine risk assessments and prophylaxis against blood clots; agreeing to common definitions and documentation standards for obstetrical trauma; and early ambulation for postoperative patients to lessen the chance for pneumonia. The introduction of rapid response teams as part of the IHI 100,000 Lives Campaign is changing the way hospitals evaluate and intervene with patients who are beginning to become unstable. As hospitals implement rapid response teams the expectation is that over time, they will see a reduction in mortality. Additionally, hospitals are increasing the amount and

---

[m] BCBSMA's tiered insurance product will not be used by GIC since BCBSMA does not serve state employees.

sophistication of their own performance measurement and tracking of intermediate processes and outcomes as part of their improvement processes. With these early interventions, BCBSMA representatives related that they are beginning to see significant improvements in several of the AHRQ QI measure results across their provider network.

A hospital system representative had similar anecdotal evidence of quality improvement spurred by the AHRQ QIs in their six acute care hospitals. Although the hospital system opposes the use of AHRQ QIs for pay-for-performance and public reporting, the BCBSMA and DHHS activities have spurred them to begin studying their results and investigating the underlying causes. In areas where a problem is flagged by the AHRQ QIs, a medical chart audit is performed. As a result, we were told that improvements have been noted in areas including iatrogenic complications, infections during medical care, sepsis following surgery, and pressure ulcers.

We were consistently told that improvements in the coding of administrative data have preceded and accompanied quality improvements. In the first stages of the new quality assessment activities, most of the problems flagged were data coding issues. More recently, hospitals have been implementing real improvements in the quality of care. Since the reputation and income of the hospitals is on the line, the activities have succeeded in focusing the attention of hospital administrators on quality improvements in hospitals.

There has also been some evidence that hospitals have been collaborating and sharing their experiences on how to improve quality. As the initiatives mature and hospitals become more familiar with the indicators, more of this type of collaboration may occur.

None of the initiatives we learned about in the case study are being subjected to formal, rigorous evaluations. However, the available anecdotal evidence suggests that AHRQ QIs have had a direct impact on the quality of patient care in the Boston area. It is difficult to judge, however, how much of this impact is due to use of the AHRQ QIs for public reporting, pay-for-performance, and tiered insurance products vis-à-vis quality improvement activities. This is a contentious question since many Boston-area hospital representatives argue that the AHRQ QIs are appropriately used only as part of quality improvement. However, it is uncertain how many hospitals would be using the AHRQ QIs for quality improvement if not for the incentives provided by public reporting, pay-for-performance, and tiered insurance products, especially given hospitals' other quality measurement and improvement activities.

### 5.1.4 Political Lessons Learned

The experience in the Boston area provides a good case study of the politics of quality reporting in a contentious environment. Boston-area hospital representatives, many very knowledgeable in the science of quality measurement, have strongly denounced the use of the AHRQ QIs for any purpose other than confidential research and quality improvement activities. They point out that for several years, AHRQ recommended using the QIs as a "screen", not for purchasing decisions, and only cautiously for public reporting. In contrast, payers and purchasers have promoted the use of the AHRQ QIs for these other purposes as one of the only available, feasible, and scientifically sound options (the other alternatives cited were the HQA, JCAHO, and Leapfrog measures). They note that AHRQ's current published guidance endorses non-quality improvement uses of the QIs.

The debate over appropriate use of the AHRQ QIs was summarized by a Boston-area hospital representative, who said that proponents of public reporting, pay-for-performance, and tiering using the AHRQ QIs "think that anything is better than nothing. We think that nothing can be better than bad." Payers, purchasers, and providers admit that the AHRQ QIs have some limitations, almost all of which are due to the fact that the AHRQ QIs are based on administrative data that were not collected for the primary purpose of quality measurement. For this reason, the data often include important omissions or mistakes that could be very clinically relevant and limit the validity of the quality indicators. On the other hand, use of administrative data-based indicators can lead to improvement in the quality of the underlying data. Furthermore, in the absence of electronic medical records, administrative data are the only affordable choice for quality measurement, given the cost of abstracting clinical data from medical records. The exception is the JCAHO Core Measures (and the closely related HQA Hospital Quality Indicators), the measurement of which is mandatory for accreditation and therefore already part of hospitals' costs. However, the JCAHO/HQA indicators (as well as the Leapfrog quality indicators, the other main alternative), are not considered by Boston-area purchasers and payers to be sufficient on their own for public reporting, pay-for-performance, or tiering, mainly due to the fact that they cover a limited set of conditions and do not measure the outcomes of care.

There is some degree of variability in the opinions of Boston-area organizations about using the AHRQ QIs for any purpose other than quality improvement. At one extreme is the MHA. Although the MHA has been active for some time in regional and national quality measurement and reporting activities, they unilaterally oppose the use of any

administrative data-based indicators for public reporting, pay-for-performance, or tiering on the grounds of poor validity. Their position is that:

> Evaluations of the quality of care used to inform the public, to make purchasing decisions, or to reward/sanction organizations, must rely on a complete clinical picture of the patient and the care delivered. Administrative data bases, because of inherent limitations tied to coding systems and methods – among other issues – are unsuited to this use. Quality of care evaluation tools based on administrative databases were designed to be, and are suitable only as, screening tools for use by health care providers to direct their quality management processes. A complete picture of patient conditions and care delivered is available only in the medical record.

The Massachusetts Medical Society, while generally supportive of increasing the transparency of health care quality, also is wary of the limitations of administrative data-based quality measurement, but stops short of stating that administrative data can only be used as a quality screening tool. This organization is opposed to using inaccurate data and inappropriate use of administrative data, but understands that "administrative data are the best we have right now." Given that reality, the Massachusetts Medical Society has drafted several criteria for the appropriate use of administrative data-based indicators:

- Rigorous, completely transparent methodology
- Meaningful measurements standardized whenever possible across payers and systems
- Opportunity for physicians to review data and make changes when data are inaccurate well before publication
- Collaborative process
- Timely data sharing
- User-friendly format

Other interviewees made distinctions between different sets of the AHRQ QIs. For example, we were told that the mortality-based IQIs were considered more acceptable for public reporting or other activities, but that the remaining IQIs and the PSIs were not.

Representatives of the payers and purchasers who support the use of AHRQ QIs for performance measurement told us that they favor trying to improve the AHRQ QIs rather than try to develop alternatives based on clinical data. They told us that clinical data collection is cost-prohibitive, so that insisting that public reporting, pay-for-performance,

and tiering of providers be based on indicators derived from clinical data effectively limits these efforts to the JCAHO Core Measures. From their point of view, opposition by hospital administrators to reporting administrative data-based indicators is partly a "delay tactic." They pointed out that administrative data are based on medical records, are owned by the hospitals, and can be improved for quality measurement purposes.

Despite these disagreements about the appropriate use of the AHRQ QIs, the new reporting initiatives have gone forward, with some modifications, and the degree of opposition appears to be decreasing. This shift is partly due to growing political support in the state (and nation) for increased transparency in health care. Public reporting, pay-for-performance, and tiering are increasingly viewed as inevitable. Another important factor in overcoming opposition has been the accommodation of some of the concerns about which indicators have been used and how they are used. For example, BCBSMA told us that a key to stakeholder "buy-in" has been to involve the hospitals, build good relationships, and work collaboratively to set clinically important and reasonable goals. Nevertheless, some participants are unlikely to change their conviction that quality indicators based on administrative data are inappropriate for any use other than quality improvement.

### 5.1.5 Technical Lessons Learned

Several common technical issues with the AHRQ QIs were identified by all of our interviewees. The major disagreement focused on whether these technical limitations were sufficient to prohibit use of the AHRQ QIs for non-quality improvement purposes. We were consistently told that the most prominent limitation is that the AHRQ QI specifications and underlying administrative data do not account for conditions that are present at hospital admission. This issue may be remedied, however, since future versions of the AHRQ QIs will accommodate present-on-admission conditions and a new Massachusetts law will require that this data element be included in the state's administrative data. Other commonly mentioned limitations include failure to identify patients under do-not-resuscitate orders and those who stay in the hospital for "comfort care" only.

Other technical problems with the AHRQ QIs are due to variability in coding practices among hospitals. As mentioned earlier in this report, many abnormal results in the AHRQ QIs are found to be due to data-coding issues rather than quality-of-care problems, and the first step in quality improvement using the AHRQ QIs is often improvement in

data-coding practices. These issues underscore the need to allow providers to review their AHRQ QI results prior to use and to have the opportunity to investigate and correct abnormal results that may be due to data-coding issues. Since not all data quality issues can be resolved quickly, it may also be useful to allow hospitals to offer an explanation for abnormal results in a public report.

Another technical issue mentioned by several interviewees is that implementation of the AHRQ QIs by different vendors has led to varying results for the same provider on the same indicator. These vendors may be using non-transparent changes to the specifications of the AHRQ QIs that lead to these divergent results.

In summary, our interviews suggest that many of the technical issues raised with respect to the AHRQ QIs may be amenable to improvement. The most prominent technical issue mentioned, a flag for conditions that are present on hospital admission, provides one example of a problem that is currently being addressed by AHRQ and state legislation. However, while proponents of using the AHRQ QIs stressed the value of working to improve the indicators and data, some interviewees felt that the indicators did not warrant any improvement. We were told by one interviewee that "at this point, any changes are tweaking around the edges…AHRQ could pour a lot more money into the indicators to make them better, but it would not be money well spent because they can't get a whole lot better."

### 5.1.6 Implications

The experience in the Boston area suggests that the AHRQ QIs might be used for public reporting, pay-for-performance, and tiered insurance products without major negative ramifications, despite strong opposition. However, a number of important caveats must be considered. The activities in Boston are in their early stages and could become problematic as they are expanded (e.g., addition of PSIs to the state's public report or additional funds devoted to pay-for-performance programs). We heard strong warnings from some interviewees that these activities were inappropriate and could lead to problems for payers, purchasers, and AHRQ, since AHRQ had endorsed use of the QIs for these purposes. Nevertheless, given the growing political focus on transparency in health care and the lack of viable alternatives to indicators based on administrative data, use of the AHRQ QIs for these purposes appears to be entrenched in the Boston area. The strength of outright opposition

appears to be declining, and opponents appear to be increasingly focused on addressing their concerns by changing the ways the AHRQ QIs are implemented.

The Boston experience underscores the importance of clear, official AHRQ guidance on how AHRQ QIs should be used. None of our interviewees doubted the usefulness of the AHRQ QIs; they disagreed on appropriate uses. Critical to the successful implementation of quality-reporting activities in Boston was AHRQ's written endorsement of use of the AHRQ QIs for non-quality improvement purposes. Payers and purchasers requested that this guidance be more explicit and strongly worded. Other interviewees suggested that AHRQ should not have issued this type of endorsement, and point out that for some time AHRQ endorsed use of the AHRQ QIs only for their originally designed purpose – quality improvement.

Despite this disagreement, respondents in Boston agreed that AHRQ should be the leader of developing quality indicators based on administrative data. In addition, several interviewees suggested that AHRQ should collaborate with other national organizations, including Leapfrog, JCAHO, and CMS, to create a consensus around the use of at least some subset of the ARHQ QIs in a national quality indicator set.

## 5.2 DALLAS-FORT WORTH

### 5.2.1 Dallas-Fort Worth Market Area

The Dallas-Fort Worth (DFW) market can be characterized as an area with well-established hospital reporting. Sophisticated institutional players both drove the introduction of reporting on quality and helped hospitals and purchasers turn results into actions.

The DFW Hospital Council is a regional hospital association of over 70 hospitals. It was founded in 1997 to support hospitals in collaborating and using data to improve patient safety and quality. As part of this association, the Dallas-Fort Worth Hospital Council (DFWHC) Data Initiative (DI)[42] is an education and research foundation. Among its many functions, the DI serves as an expert intermediary between the hospitals and the State for purposes of submitting discharge data from hospitals for generation of the legislatively-mandated report card on the quality of hospital care (see discussion below). It also independently calculates AHRQ QIs for all participating hospitals and feeds back to each hospital its own indicator results so that hospitals can see their performance on the AHRQ

QIs well ahead of the public release. In addition, the DI has developed a sophisticated software tool with which hospitals can analyze and benchmark their own data, as well as identify individual cases that were flagged by an indicator as potential adverse events. Hospitals compare this information against medical records to distinguish coding problems from quality issues.

In addition, the media traditionally gives a lot of attention to the issue of hospital quality in the DFW area, partly because of the DI and employer initiatives and partly because of a local news reporter with a strong personal interest.

### 5.2.2 Background on Use of Quality Indicators

Texas has a long history of using data to encourage informed consumer decisions and competition on quality of care. In 1995, the Texas Legislature created the Texas Health Care Information Collection (THCIC)[43,44] with the primary purpose of providing data to enable Texas consumers and health plan purchasers to make informed health care decisions. THCIC's charge is to collect data and report on the quality performance of hospitals and health maintenance organizations operating in Texas. This same legislative mandate required the THCIC to publicly report on the quality of care in Texas hospitals[n] and explicitly mandated the use of a severity adjustment.

A scientific advisory committee guided THCIC in its decisions on how to implement the mandate, and the process was accompanied by extensive stakeholder consultations. The committee argued that indicators for the report should be based on a national standard and an open source methodology. The AHRQ QIs were the only viable indicator set found, and the option of developing indicators was rejected because of resource constraints and concerns about scientific defensibility.

THCIC decided to use most of the IQIs[o] and all the PQIs, but not the PSIs. THCIC also plans to add the PDIs to the report. The decision not to report the PSIs was motivated by concerns about unstable rates due to small denominators and their controversial nature,

---

[n] All hospitals except critical access hospitals in areas with a population of less than 35,000 inhabitants.

[o] Specifically, IQIs # 1-14, 16-20, 22-25, 30-33.

as those indicators capture severe adverse events in hospitalized patients.[45] The first report was released in 2002, shortly after the AHRQ QIs became available, which made Texas one of the first users of this product. The report has been updated annually since.

### 5.2.3 Impact of Public Reporting

There have been no formal evaluations of the impact of publicly reporting the AHRQ QIs, so that only anecdotal evidence is available. Our interviewees suggested that the indicators have had an impact, at least in the DFW area. One interviewee explained:

> The use of AHRQ QIs probably has affected patient care indirectly, but not directly. Because hospital-level information would be subject to Freedom of Information Act requests, the state does not and cannot work specifically with hospitals in addressing quality of care issues. But, as a result of the QIs, at least one of the regional hospital associations works with member hospitals to look at indicators and quality.

Because of media interest and purchaser pressure, hospitals administrators pay close attention. Hospital CEOs, CFOs, and boards are commonly briefed on their AHRQ QI results. Larger hospitals have quality improvement teams that routinely utilize the tools provided by the DI to analyze their results, and even smaller facilities try to understand their performance along the indicators. The main focus so far has been to work with coding departments on the particular requirements for the AHRQ QIs to make sure coding issues do not distort the results. But some hospitals reported having found, as part of their investigation into coding practices, instances in which a quality problem was detected and addressed.

Overall, hospitals were not enthusiastic about publicly reporting the AHRQ QIs, but accepted that reporting was "here to stay" and considered the AHRQ QIs to provide a reasonable mechanism for meeting this requirement. Hospitals were concerned, however, about the limitations of billing information as the underlying data source. In addition, the AHRQ QIs, because of their visibility, now play a substantial role in setting priorities for quality improvement projects and sometimes drain resources from other initiatives that hospitals see as more urgent. One interviewee complained, "There are so many indicators, but the public availability of the AHRQ QIs forces us to deal with them, even if we don't believe in the results." This is reinforced by the media attention to the AHRQ QIs, because reporters tend to overemphasize negative results and over-interpret the findings.

Our interviewees emphasized that the situation was quite different in other parts of Texas, where there is limited media attention and no expert intermediary to help hospitals with data analysis. We consistently heard that little sustained attention has been paid to the public indicator reports after the excitement around the first release had subsided.

Another important observation is that the QI projects are mainly driven by hospital administrators and not physicians. Physicians are involved on an individual basis, depending on their interest in quality improvement, but the Texas Medical Association has focused its quality agenda on other initiatives, such as the IHI 100,000 Lives Campaign. It does not see a role for organized medicine in the public reporting of the AHRQ QIs, which it considers to be hospital-related. We heard repeatedly that hospital administrators used the AHRQ QIs to convince physicians to work on quality improvement projects (e.g., standing order sets, clinical pathways), which some physicians tended to dismiss as "cookbook medicine." However, if the indicators suggested poor performance in a given clinical area, hospital administrators had additional leverage in convincing physicians to embrace change.

### 5.2.4 Political Lessons Learned

In spite of the potentially controversial nature of public reporting for hospitals, its introduction has been unremarkable. Hospital associations and the DI worked together with the Texas legislature to craft and implement the statutory requirements, which mandated transparency and proper risk adjustment. The indicators were selected through a multi-stakeholder consultation process. Stakeholders decided early in the process that only well-established national indicator sets should be considered, as any "home-grown" solution would lack credibility. The AHRQ QIs were the only viable option that met those requirements, and had the additional advantage of not requiring additional data collection. AHRQ's reputation for unbiased and rigorous research carried great weight in this decision process. Still, there was "a great deal of anxiety before the first released report," and hospitals received their own data for review and comment 60 days in advance. A few years into this program, anxiety has subsided, making room for a more reasoned discussion of the strengths and limitations of the indicators and the best ways to use them for quality improvement. As one interviewee stated:

> Hospitals don't love the indicators – for performance improvement, clinical data is the best bet, but for in the short run, especially as they improve, administrative data and the AHRQ QIs are a good solution.

The two main lessons learned from this implementation are the importance of proper involvement of stakeholders, in particular hospitals, in every step of the process, and the importance of using scientifically credible and transparent indicators.

### 5.2.5 Technical Lessons Learned

The implementation of a public reporting system based on the AHRQ QIs did not create significant technical problems for the state of Texas and the DFWHA. Both the THCIC and the DI were accustomed to working with discharge data and with the HCUP indicators (the predecessors of the current AHRQ QIs). AHRQ provided limited feedback and technical assistance in the initial implementation, but the program is largely self-sufficient now.

It proved more challenging for the hospitals to adapt to this new requirement. While they were typically familiar with many of the building blocks of the AHRQ QIs, such as the UB-92 data format and the APR-DRGs, they needed to understand the logic of the indicators and how coding practices for the UB-92 affected their results. They also had to start identifying patients who were flagged by an indicator, retrieve their medical records, and assess whether there was an actual adverse event or a coding problem before responding to the findings. Substantial educational efforts for medical-record coders became necessary to make coders aware of the implications of coding rules for the indicators. Without the software tools and the technical assistance that the DI provides, few hospitals, especially the smaller facilities, would have been in a position to analyze their own data, and to improve both data quality and quality of care. This ability, as we heard over and over, is critical for buy-in by the hospitals and also for a public reporting program to lead to real change.

Those problems have not yet been fully overcome, and new issues continue to surface as hospitals become more familiar with the implications of coding practices for the indicators. For some indicators (e.g., vaginal tears during childbirth), strict interpretation of the coding rules to achieve adequate reimbursement has led to poor performance on the indicators – and vice versa. For others (e.g., post-operative hemorrhage), coding rules are not specific enough, resulting in inconsistencies between physicians and coders. Finally, Texas has not mandated the use of E-codes (ICD codes for external cause of injury) for hospitals. This leads to comparability problems, because hospitals can set their own policy as to whether those codes should be used or not. In general, physicians and hospitals would like

to see more rigorous validation studies to assess the strengths and limitations of the different indicators. While many hospitals in the DFW market have compared their indicator results against medical records, none has done this systematically as part of a research project and the efforts have so far focused on false-positive events.

Sample size remains a particular problem for smaller facilities. While there are fewer small hospitals in this area than in other parts of Texas, some hospitals can report only on mortality for AMI and pneumonia, since they lack the required sample size for any other indicator. In addition, many small facilities routinely transfer most AMI patients to hospitals that are equipped for emergency procedures. This can inflate AMI mortality and lead to poor performance on the indicator, because a greater share of patients who are too unstable for transport, or whose prognosis is too poor to allow for invasive procedures, remain in the smaller hospitals.

### 5.2.6 Implications

There are two main implications for the AHRQ QI program from the experience in Texas and in DFW in particular. First, while the AHRQ QIs were not originally designed for public reporting, their use for this purpose, with the appropriate caveats, seems viable. Hospital administrators have adjusted to the AHRQ QIs as metrics and are beginning to educate their coders about the impact of coding practices on the quality reports. At least anecdotally, the reports are having an impact on quality improvement efforts. As previous RAND research has reported,[46] the driving force to act on the results of performance data are not patients or purchasers but rather hospital administrators and their boards, who are concerned about the reputation of their hospital. The reporting requirements raised the profile of quality of care as a priority issue over finances and provided both data and leverage to introduce quality improvement efforts.

Second, the DI as an expert intermediary was crucial to the implementation and will remain crucial in helping hospitals to turn the reports into action. The DI helped overcome the initial resistance and helped hospitals understand the value of reporting and accountability. For other regions in Texas, which lack such expert support, the indicators remain a "black box" and a source of anxiety rather than a stimulus to improve care. This suggests that it is important for AHRQ to continue supporting intermediaries like the DI.

# 6. "LESSONS LEARNED" FOR FUTURE ACTIVITIES

In this section, we present a number of "lessons learned" during our interviews concerning future directions for AHRQ in the development and modification of QIs. Our discussion is organized in three parts. First, we describe interviewees' perspectives on current, anticipated, and potential development projects involving the QIs. Next, we discuss users' perspectives of AHRQ as a measures developer and the ways in which users speculate this role could evolve or change in the future, especially in relation to other potential providers of this service. Finally, we briefly discuss users' views on the subject of market demand, in particular, user willingness to pay for QIs.

## 6.1 VOICES OF THE CUSTOMER: PRIORITIES FOR FUTURE DEVELOPMENT OF THE QIs

A key function of this study was to provide AHRQ with feedback from interviewees about priorities for future development efforts. In order to explore this topic with users, we first solicited input from members of the AHRQ QI team about current, anticipated and potential development projects. We then used these responses in our interviews, which asked explicitly about interviewees' opinions of the need for these projects as well as their own priorities for future development. We grouped the development projects into three categories and asked interviewees which category they would like to see given priority:

- Improvements in the current product line
- Addition of new product lines
- Improved support for the QI products

Improving the current products was most frequently seen as the highest priority, followed by both the addition of new products and improvements in service, outreach, and user support for the measures (Table 6.1).

**Table 6.1. Interviewees' Priorities for Future Developments in AHRQ's QI Program**

| | Current Products | New Products | Service | No Prioritization Given |
|---|---|---|---|---|
| 1st Priority | 43% (21) | 22% (11) | 12% (6) | 22% (11) |
| 2nd Priority | 8% (4) | 33% (16) | 20% (10) | 39% (19) |
| 3rd Priority | 18% (9) | 12% (6) | 31% (15) | 39% (19) |

*Source*: RAND analysis of interview responses. The number of interviewees giving each priority is in parentheses. 49 out of 54 interviewees were asked to give their priority. The remaining 6 interviewees were not sufficiently familiar with the AHRQ QIs to be asked to give a priority.

Many users told us that it was important to improve the current set of indicators as much as possible and expand their use so that the QIs became more of a national standard. One interviewee summarized this sentiment by saying, "One solid measurement set with everyone's buy-in would be enormously positive." Another user pointed out that "it would be good to focus on shoring up current indicators because there is currently a lot of criticism around using them for public reporting."

Many other users said that their recommendation to focus on improving current AHRQ QIs was driven by a desire to overcome stakeholder opposition to indicators based on administrative data. One interviewee summarized this line of thought:

> There is no way that every hospital in the country is going to do primary quality data collection and even if they did, how could we enforce consistency and timeliness? This is a battle that we have been fighting for years, and we've been struggling because people tend to dismiss out of hand any information based on administrative data. In the short term, until there is progress with the electronic health record, administrative data is all there is, and there is no convincing argument that we have exhausted all possibilities to use this type of data for quality improvement.

Another user was more specific about how the indicators might improve:

> The AHRQ QIs may not be perfect but they are a national standard, based on readily available data that are not going away, and the indicators will get better and better - especially with extended ICD-9 codes (and later the move to ICD-10) and the addition of a flag for condition-present-on-admission and things like that. I think there is an opportunity to improve the QIs gradually over time, as the underlying data sources improve – as new data elements are added with the introduction of the UB-04, and eventually electronic health records.

Despite these sentiments, there is also a strong desire by many interviewees for additional QIs covering new areas. Indeed, it was difficult for many users to choose between adding new products and improving current products as their top priority. Improving service and outreach was most frequently given a low rating. In the following

subsections, we discuss in more detail what changes interviewees would like to see in the AHRQ QI program.

### 6.1.1 Improvements of the current product line

Apart from the expectation that AHRQ maintain and update the current QIs, the most commonly requested improvement was the addition of data elements to increase the specificity of the QIs, such as a flag for conditions present at admission or for "do-not-resuscitate" orders and the addition of clinical data elements (Table 6.2). As mentioned above, the AHRQ QI team is incorporating a flag for conditions present at admission in the next iteration of QI specifications. Other improvements mentioned with some regularity include validation studies on the development of composite measures (a project that AHRQ is currently undertaking) and better risk adjustment, with coordination of risk adjustment methods across the subsets of QIs.

Table 6.2. Suggested Improvements to the Current AHRQ QIs

| Number (percent) of interviewees making the recommendation | | Recommendation |
|---|---|---|
| 27 | (50%) | Develop indicator specifications that rely on incorporating additional data elements with the administrative data (including present-on-admission flag, do not resuscitate order flag, clinical data elements, etc.) |
| 13 | (24%) | Perform validation studies |
| 12 | (22%) | Develop composite indices |
| 12 | (22%) | Improve risk adjustment (alignment with other indicator system, non-proprietary system) |
| 7 | (13%) | Add more analytic tools to the software, such as various levels of significance testing |
| 5 | (9%) | Improve identification of "avoidable" admissions using secondary diagnoses or other methods |
| 5 | (9%) | Improve obstetric PSIs |
| 5 | (9%) | Periodically assess the applicability of some of the IQIs to the inpatient setting, in particular for procedures that are now mostly done on an outpatient basis, like laparoscopic cholecystectomy |
| 4 | (7%) | Provide guidance on how to perform trend analysis over time given changes in indicator definitions |
| 3 | (6%) | Provide guidance on appropriate coding of source of admission |
| 2 | (4%) | Adapt PQIs for hospital-level analyses |
| 2 | (4%) | Calculate and disseminate cost-effectiveness of quality improvement using each indicator |
| 1 | (2%) | Assume Poisson distribution for counts of infrequent events rather than normal distribution |
| 1 | (2%) | Change smoothing procedure for PSIs so results are not over-smoothed |
| 1 | (2%) | Develop a common minimum set of checks of data quality ("common minimum edits") |
| 1 | (2%) | Develop open-source methods for probabilistic data linkages |
| 1 | (2%) | Exclude cancer patients from failure to rescue indicator |
| 1 | (2%) | Exclude patients transferred in from another hospital from all indicators |
| 1 | (2%) | Improve handling of zero numerator events |
| 1 | (2%) | Improve risk adjustment for IQI 33 – it groups women 18 and younger and women 35 and older together despite the fact that the different groups have different risks |
| 1 | (2%) | Provide PSIs for all ages – not split pediatric/adult |
| 1 | (2%) | Study how coding practices vary across hospitals |

*Source*: RAND analysis of interview responses.

### 6.1.2 Adding new product lines

Most interviewees were aware and appreciative of the roll-out of the pediatric QIs as a new module, as this important population had been excluded from many of the initial QIs. Almost half of the interviewees mentioned the additional need for measures for hospital outpatient/ambulatory care, such as day surgery and diagnostic procedures (Table 6.3). About a third of interviewees mentioned the need for efficiency, physician-level, and emergency room care measures. Nearly a quarter of interviewees expressed interest in integrating data and indicators for inpatient and outpatient surgery, since an increasing number of procedures are being shifted to outpatient settings. This, for example, has created a real problem for constructing the laparoscopic cholecystectomy indicator (IQI 23), because nearly all of those procedures are now done on an outpatient basis.

Measures for rural/small hospitals were the next priority group. However, interviewees expressed differing views on the implications of having a dedicated set of indicators for rural/small hospitals. On the one hand, many felt that dedicated indicators were needed, because the low patient volume at rural/small hospitals excludes those institutions from most of the current indicators. Further, interviewees felt that some indicators should not be constructed for those facilities. For example, since current ACOG guidelines do not recommend VBAC for facilities without adequate infrastructure for emergency caesarean section, the VBAC indicators (IQI 22 and 34) should not be used for many of them. On the other hand, some interviewees expressed concern that dedicated indicators would suggest that small and rural hospital were second-class facilities, because common quality standards do not apply.

Table 6.3. Priorities That Were Identified for Development of Additional QIs

| Number (percent) of interviewees making the recommendation | | Recommendation |
|---|---|---|
| 26 | (48%) | Outpatient/ambulatory (including ambulatory surgery) |
| 16 | (30%) | Efficiency |
| 15 | (28%) | Physician |
| 15 | (28%) | ER/ED |
| 12 | (22%) | Linked data sets |
| 10 | (19%) | Small hospital/rural |
| 6 | (11%) | Intensive care |
| 4 | (7%) | Mental health |
| 3 | (6%) | Readmissions |
| 3 | (6%) | Maternal |
| 2 | (4%) | Elective |
| 2 | (4%) | Long-term care |
| 2 | (4%) | Misdiagnoses, delayed diagnoses, missed diagnoses |
| 11 | (20%) | Others (1 mention) |

*Source*: RAND analysis of interview responses.

### 6.1.3 Improved services around QI products

One of the most common priorities for improved service among interviewees (Table 6.4) was more explicit guidance from AHRQ on the use of the QIs for public reporting and pay-for-performance (also discussed in Section 4.3.3). Users were sometimes not aware that AHRQ had recently released documents on those issues; the latest guidance document was released in December 2005, predating these interviews by only a few months.[47]

Table 6.4. Suggested Improvements to Service Supporting the AHRQ QIs

| Number (percent) of interviewees making the recommendation | | Recommendation |
|---|---|---|
| 10 | (19%) | Provide a template and guidance for using the QIs for public reporting |
| 10 | (19%) | Provide guidance on next steps for improving quality once a potential opportunity for improvement is flagged by the QIs |
| 8 | (15%) | Collaborate with other organizations to create a national standard set of measures |
| 8 | (15%) | Make QIs more user-friendly (such as simpler indicator names) |
| 7 | (13%) | Increase responsiveness and speed of user support |
| 6 | (11%) | Expand outreach and marketing to more people and more audiences |
| 6 | (11%) | Provide guidance for how consumers should interpret values of QIs |
| 5 | (9%) | Provide benchmark QI values for various populations |
| 4 | (7%) | Add a greater educational component to outreach and user support |
| 4 | (7%) | Develop a comprehensive list of who is using the QIs, how they are using them, and impact they are having |
| 4 | (7%) | Provide funding for research studies using the QIs |
| 3 | (6%) | Improve the process of incorporating feedback from users on the QIs |
| 3 | (6%) | Provide guidance on how coding of discharge data can be improved |
| 2 | (4%) | Choose 1 statistical software package (SAS or SPSS) and drop the other; also drop/cut back on the PC application |
| 2 | (4%) | Contract with other organizations to provide outreach, education, user support, and to disseminate information |
| 2 | (4%) | Make the Fact book series more accessible |
| 1 | (2%) | AHRQ staff should do communication and relationship-building directly with state agencies rather than using a contractor |
| 1 | (2%) | Develop basic educational materials appropriate for local/hospital level (QI director) rather than a researcher, who is already familiar with material, and disseminate them widely |
| 1 | (2%) | Develop something like the Dartmouth Atlas charts showing geographic variations in QIs – would help choose which QIs to report |
| 1 | (2%) | Make formulas easier to access – some formulas are buried in documents and SAS code and some are missing from documentation |
| 1 | (2%) | Provide code for STATA |
| 1 | (2%) | Provide FAQs, list of questions already answered, for quick reference |
| 1 | (2%) | Provide outreach to promote use of PQIs |
| 1 | (2%) | Provide print outreach materials, not just email |
| 1 | (2%) | Provide software for Macintosh computers |
| 1 | (2%) | Standardize the layout/organization of documentation across measure sets |

*Source*: RAND analysis of interview responses.

Another commonly listed priority for increased AHRQ service was for AHRQ to provide guidance on the process that should be followed to improve quality in areas where the QIs indicate a problem. Users had varying levels of experience with quality improvement and varying levels of access to networks that can be used to share quality improvement knowledge. It would be helpful for users to have further guidance (e.g., a general methodology for analyzing medical records following an abnormally high incidence of PSI 4 - failure-to-rescue, and a summary of available evidence on interventions that could be implemented to lower the rate).

Interviewees suggested that AHRQ collaborate more closely with other organizations in attempt to forge more of a consensus on a "national standard" set of quality indicators. The standardization of some AHRQ and Leapfrog indicators and submission of some of the AHRQ QIs for NQF approval are steps that AHRQ has already taken in this direction. Further efforts along these lines would improve the usability of the AHRQ QIs for users.

Interviewees also suggested making the AHRQ QIs more user-friendly and simpler to understand. A simple suggestion in this regard was to promulgate official, simple names for the QIs in language understandable by people with no clinical knowledge.

We asked users specifically about one aspect of AHRQ service – user support. We received favorable feedback about the current level of AHRQ user support for QI users. Of the 15 users who reported using AHRQ support, all but one explicitly reported a good experience. Interviewees were impressed by the technical competence, accessibility, and responsiveness of the helpdesk staff and argued that this support function had played a major role in advancing the field of quality measurement, because it removed the barriers that non-research institutions face when implementing complex measurement systems. To provide a point of comparison, several of the more experienced users recounted the difficulties they had experienced in working with the HCUP indicator code.

One user reported being able to "feed complicated, technical questions from hospitals to AHRQ," and that AHRQ user support was able to answer those questions "from a greater depth and background" than the user had. This user added that responses to inquiries were "based on evidence, thoroughly considered and thought-through, with a quick turnaround." Another user felt that there was "always someone you could get hold of to voice concerns." A vendor commented that the AHRQ QI technical support had a "pretty quick response time compared to what one would expect from federal agency. They would

open a case the same day, send an email confirmation, assign to a person – all in the same day." On the other hand, other interviewees (7 of 54) did suggest the need for increased responsiveness and speed of user support. These interviewees generally wanted most questions to be answered the same day they were asked.

## 6.2 USER PERSPECTIVES: THE FUTURE ROLE OF AHRQ COMPARED TO OTHER PLAYERS

We explored extensively the issue of how users perceive AHRQ as a measures developer, what they think AHRQ's role should be in this area, and whether some function that AHRQ currently performs could be taken over by other public or private institutions.

Our interviewees held AHRQ in very high regard. They credited AHRQ for its vision in pushing for the use of administrative data for quality measurement well before the research and provider community was ready to exploit this data source. The work of the AHRQ QI team was described as technically sound, sensitive to the limitations of the underlying data, and transparent. AHRQ is regarded as an intellectual leader and "go-to" institution for health services research and the use of administrative data for hospital quality measurement. As shown in our environmental scan for comparable products, no clear comparable alternative to the AHRQ QIs has emerged or is likely to emerge. Several other developers, especially JCAHO, CMS, HQA, and Leapfrog, are seen as prominent sources for measures and may be used as alternatives, but their indicators differ in several important ways and are generally regarded as complements to the AHRQ QIs, not true alternatives. We asked QI users to visualize how the quality measurement landscape would change if the AHRQ QI program disappeared.

One interviewee answered:

> If AHRQ stopped the QI program, pieces would be picked up but there wouldn't be a consistent, cohesive package as big as AHRQ is now. The public domain issue is a big one. Providers have only been in the game because the indicators come from a public source – if that public source goes away, I think providers will stop doing it.

And another one said:

> I can't imagine who else would pick up activities from AHRQ so instead, probably activities would be broken down into orphan activities – any one slice would be a different activity; specialty organizations would take over certain types of measures (pediatric, for example).

Interviewees were quite comfortable with AHRQ having a leading role in national quality indicator development. It was generally viewed as positive that a trustworthy federal institution had defined open-source and well-documented quality measurement standards. These standards were viewed as contributing to the transparency of health care quality measurement and reducing the measurement burden for health care providers by limiting the number of measurement tools they must use to satisfy various reporting requirements. Many emphasized the need for even greater leadership from the federal government in this area, either by developing measures or by orchestrating public-private partnerships, so that standard measure sets for various purposes would become available and accessible to everyone.

AHRQ's leading role was also seen as a challenge for AHRQ, because with it comes the responsibility to maintain the QI program, on which so many programs now depend. Our interviewees looked primarily to AHRQ to fill the obvious gaps in the measurement science. Several commented that current funding levels for AHRQ were not adequate to meet all those needs.

We discussed whether it could be a viable option for AHRQ to give up parts of the current QI program in order to free up resources and set different priorities. Specifically, we asked whether AHRQ could or should stop developing software and providing user support in order to focus exclusively on indicator development. Almost unanimously, interviewees rejected a model under which AHRQ would develop and distribute the software without supporting it. There was much concern that lack of user support would create enormous barriers to the implementation of quality measurement initiatives, especially for new users and non-research institutions. Using vendors to provide user support was also not commonly regarded as an alternative, because many feared that vendors would be prohibitively expensive or incapable of providing the same quality of support as the original developers. The latter view was even shared by some of the vendors who would potentially stand to gain from this model: one representative stated that "we do not want to support AHRQ's software since we can't support what we don't write."

We received mixed reactions to a model under which AHRQ would only develop and release indicators and their technical specifications, but no longer provide or support software. Many interviewees were familiar with such an arrangement, as it would mirror the division of responsibilities between JCAHO and the Core Measures vendors. But several drawbacks were brought to our attention, such as vendor and license fees, as well as

potential quality problems and comparability issues (if different vendors implemented AHRQ specifications). Several interviewees stated that such a model would represent a step backwards in the development of a unified quality measurement infrastructure, since a transparent national standard would be transformed into multiple proprietary systems, at the same moment at which many entities, like CMS, JCAHO and NQF, are trying to introduce open-source consensus measures, as recommended by the IOM. At a minimum, a rigorous certification program for vendors would be needed and many interviewees worried about the implications of such a change for the momentum that the hospital quality measurement movement has gathered.

Finally, we asked interviewees which parts of the QI program AHRQ could give up, if (hypothetical) budget cuts were to leave it with no other choice. Most of the 54 interviewees stated that the program represented a unified entity that should not be disassembled, although 12 interviewees said software development and 11 said user support could be discontinued by AHRQ and those functions assumed by others.

## 6.3 USER VIEWS: WILLINGNESS TO PAY FOR THE AHRQ QIs

As an alternative to AHRQ realigning current funds, we asked interviewees whether AHRQ might consider financing program growth by generating additional revenues from charging users. Not unexpectedly, this proposal was not met with enthusiasm. Almost half of our interviewees (20 of 54) did not answer the question. Five out of 36 current users stated that they would stop using the QIs in this case. Three current users replied that they had invested so much into their program based on the AHRQ QIs that they would have to accept charges, but emphasized that they might not have selected the QIs in the first place if they had not been a free resource. However, almost half of the interviewees (44%) expressed willingness to pay a reasonable fee for access to the full QI resources.[P] Two even said that the perceived value of the QIs would *increase* if users had to pay for it: "Marketing 101: If you don't charge anything, people aren't going to proscribe value to it. If there is no cost attached, people can take or leave it because it doesn't represent an investment."

---

[P] We did not elicit specific information on what users would consider to be a reasonable fee. A market study would be required to determine what would be considered "reasonable" among the current and potential users. Such an endeavor was outside of the scope of our study.

A slight majority favored a subscription model (i.e. paying a one-time charge), but some argued for a usage-based payment scheme. Most recommended differential pricing by type of organization and purpose of use (i.e. commercial vendors who resell the QIs or incorporate them into their products should pay a higher rate than state agencies that operate public reporting programs). Interviewees also felt that one-time use for research projects should be less expensive than ongoing use for operative purposes.

# 7. DISCUSSION

In this chapter we discuss the limitations of this evaluation, then briefly review some of the top-level findings from our assessment and discuss their implications for AHRQ's future growth opportunities.

## 7.1 LIMITATIONS

This evaluation had several methodological limitations that should be considered in the interpretation of the results. The most important limitation is that the majority of interviewees were users of the AHRQ QIs. Organizations that have decided not to use the AHRQ QIs for quality measurement may be more likely than users to have negative opinions about the AHRQ QI program. For this reason, we interviewed a small number of non-users. No unique, substantially negative opinions of the AHRQ QIs were expressed by the non-users, but a larger sample may have yielded different results.

A second limitation of this evaluation is that the environmental scan used to identify users of the AHRQ QIs probably failed to identify a large number of organizations that do not publicly release AHRQ QI results or publish descriptions of their AHRQ QIs use. If these organizations differed consistently from those identified in the environmental scan, our results could present a biased view of AHRQ QI users' opinions.

## 7.2 WHAT IS AHRQ'S CURRENT MARKET POSITION?

Our results show that the AHRQ QIs are regarded as the leading product in the area of measuring the quality of hospital care from administrative data. In fact, the QIs are currently the only comprehensive measurement system in this area, as our survey did not identify any comparable offerings. The QIs have gained a leading role not just within the United States, but also increasingly in other countries. While a variety of vendors offer quality measurement systems that use administrative data, their products often embed the AHRQ QIs (either the actual software or the specifications).

The fact that the QIs are a free resource certainly helped them to gain market share, but our interviewees were adamant that the excellent quality of the product combined with AHRQ's reputation were the key prerequisites for the indicators' success. Many interviewees

commended AHRQ for the rigor and unbiased nature of its research and felt that AHRQ would remain the natural home for the development of standards for quality measurement. The AHRQ QIs were widely described as a scientifically sound and well-documented set of measures that were easy to implement because of user-friendly software and good user support. The complete transparency of the indicator specifications, the risk adjustment methodology, and the underlying evidence were all credited as crucial factors for the acceptance of the product by various stakeholders.

Thus, the AHRQ QIs have achieved a strong position in their market segment and no obvious alternative or competitor could be identified, although some organizations (notably JCAHO, CMS, and Leapfrog) have complementary indicator sets. This is unlikely to change: new users have an incentive to adopt the prevailing product, because it makes their results comparable to a large number of other users and because the widespread use lends legitimacy to the product, which is critical in the often politicized debates about selecting quality indicators for such uses as public reporting and pay-for-performance. Indicator development based on rigorous science is also quite costly. As a result, other developers would face substantial barriers to entry if they tried to establish alternative measurement systems.

Our interviewees were quite comfortable with AHRQ having this position, although they pointed to two potential risks. First, the dominance of the AHRQ QIs combined with ease of access to administrative data might stifle innovation for indicators that have more demanding data requirements. Second, the dominance also implied a responsibility for AHRQ to maintain the program and to keep expanding it, which was seen as challenging, given AHRQ's budget limitations.

## 7.3 WHERE ARE THE GROWTH OPPORTUNITIES FOR THE AHRQ QI PROGRAM?

The market for the AHRQ QIs is large, growing rapidly, and changing as indicators are being used in new ways. There are now a substantial number of users of the AHRQ QIs for public reporting and pay-for-performance programs. As the prevalence of those activities increases, we expect the number of users to increase substantially both for the programs themselves and for internal quality improvement programs and projects that will attempt to align their target measures with standards for external accountability.

Our interviewees suggested that there is substantial demand for expansion of the AHRQ QI program. The most common requests were for improvements to the current sets, such as accommodation of non-UB-92 variables (e.g., present on admission flags) and methods for composite formation, followed by additional QI sets to close important measurement gaps, such as hospital outpatient care and emergency room care. Interviewees were largely aware and appreciative of AHRQ's current efforts to improve and expand the program, but expressed an interest in scaling up, and speeding up, those activities.

## 7.4 HOW COULD GROWTH BE FINANCED?

Interviewees recognized that expansions of the AHRQ QI program would require additional resources and largely argued that this would be money well invested. Most believe that federal funding should be used to support those activities, realizing that this was a difficult proposition given the pressure on public budgets in general, and on AHRQ's budget in particular. But the availability of scientifically sound indicators in the public domain was seen as a precondition for quality improvement efforts and policy innovations like public reporting and pay-for-performance so that support by public funds seemed warranted. It was frequently stated that pay-for-performance and public reporting programs in particular should be based on fully transparent methodologies to allow hospitals to understand how they were evaluated and to identify opportunities for improvement. Proprietary indicators, the likely outcome of private funding, were seen an unsuitable for those applications.

We challenged interviewees to brainstorm about alternatives to increased public funding for the QI program. One option was the re-allocation of existing funds by reducing the scope of activities under the program and focusing on core competencies - for example, giving up development and distribution of free software to construct the QIs or stopping user support. Most argued that indicator development was AHRQ's core competency and should never be given up, and interviewees also tended to be reluctant to see AHRQ give up software development and user support. We heard consistently that only the original developer of such specialized software is able to provide adequate support. Even some vendors, who could consider trying to take on the support role themselves, agreed with that assessment. Thus, continuing software development but stopping user support does not appear to be a plausible option for AHRQ as a means to free up funds for development activities.

Nor was there enthusiasm for the even more radical alternative of AHRQ deciding to focus only on specification development while leaving software development and user support to vendors. Concerns centered on high fees, restricted access, and potential problems with quality and comparability if different vendors implemented AHRQ specifications. At a minimum, a rigorous certification program for vendors would be needed. By and large, interviewees felt that this approach would only be slightly superior to stopping the program altogether, and that it would greatly impede the proliferation of quality measurement activities.

As an alternative, we discussed the option of AHRQ continuing to provide specifications, software and user support but starting to charge for those services. While there was little enthusiasm for this prospect, only a few stated that they would stop using the product in that case. Most seemed to be willing to pay a reasonable amount, so charging users would be a viable option to support expansions of the QI program. However, the feasibility of implementing such an option would depend on many yet-to-be-answered questions:

1. How much would users be willing to pay? Our study was not designed to investigate what users considered a reasonable charge. Thus, a market research study would be required to elicit willingness-to-pay.

2. How would a fee for use affect the willingness of new future users to implement the QIs? We talked mainly to current users of the QIs, who have already invested resources into implementing them for their particular purposes. Those current users are unlikely to adopt a different indicator system, unless the cost of the AHRQ QIs was to become prohibitive. But non-users might select another product or abort their quality measurement activities entirely if they had to pay for the AHRQ QIs.

3. What are the rights of existing users who have invested in implementing the QIs under the assumptions that they are a free resource?

4. Should pricing be different for different users (e.g., researchers and re-sellers) and by what degree?

5. What is the best pricing model (e.g., fee-per-use, subscription)?

6. How should fees for international users be handled? These users are probably very price-sensitive because most of the current uses are actually small initiatives of individual researchers. It is more difficult to collect money from researchers, but fairness would require charging them if domestic users are charged.

In summary, if AHRQ were to implement a charge-based model for the QIs, it would face the challenge of developing a comprehensive business plan. The size of the market needs to be determined to make sure that the expected revenue could provide a meaningful contribution to the growth of the program, after the added cost of operating a business is taken into consideration. AHRQ would also need to consider the amount of additional revenue it could expect to obtain in proportion to the potential negative effects on the spread of the program to new users and usages. In a sense, AHRQ is now in a situation comparable to that of other organizations that have started out offering content or services on the Internet for free and are contemplating whether to begin charging users. Thoughtful deliberations would be needed to find a business model that generates sufficient revenue but is still consistent with AHRQ's mission and values as a public agency.

## REFERENCES

[1] Institute of Medicine (IOM). 2001. Crossing the Quality Chasm. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, D.C: National Academy Press.

[2] Institute of Medicine (IOM). 2000. To Err Is Human: Building a Safer Health System. L. T. Kohn, J. M. Corrigan, and M. S. Donaldson, eds. Washington, D.C: National Academy Press.

[3] McGlynn, E.A., S.M. Asch, J. Adams, J. Keesey, J. Hicks, A. DeCristofaro, and E.A. Kerr. 2003. "The quality of health care delivered to adults in the United States." *New England Journal of Medicine* 348(26): 2635-45.

[4] Institute of Medicine (IOM). 2006. Preventing Medication Errors. Philip Aspden, Julie Wolcott, J. Lyle Bootman, Linda R. Cronenwett, eds. Washington, D.C.: National Academies Press.

[5] Greenwald, L., Cromwell, J., Adamache, W., Bernard, S., Drozd, E., Root, E., and K. Devers. 2006. "Specialty versus community hospitals: referrals, quality, and community benefits. *Health Affairs* 25(1): 106-118.

[6] Institute of Medicine (IOM). 2001. Crossing the Quality Chasm. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, D.C: National Academy Press.

[7] Institute of Medicine (IOM). 2000. To Err Is Human: Building a Safer Health System. L. T. Kohn, J. M. Corrigan, and M. S. Donaldson, eds. Washington, D.C: National Academy Press.

[8] McGlynn, E.A., S.M. Asch, J. Adams, J. Keesey, J. Hicks, A. DeCristofaro, and E.A. Kerr. 2003. "The quality of health care delivered to adults in the United States." *New England Journal of Medicine* 348(26): 2635-45.

[9] Institute of Medicine (IOM). 2006. Preventing Medication Errors. Philip Aspden, Julie Wolcott, J. Lyle Bootman, Linda R. Cronenwett, eds. Washington, D.C.: National Academies Press.

[10] Mattke, S. Monitoring and Improving the Technical Quality of Medical Care: A New Challenge for Policymakers in OECD Countries. In: Towards High Performing Health Systems. Policy Studies, OECD, Paris 2004.

[11] McGlynn, E.A., S.M. Asch, J. Adams, J. Keesey, J. Hicks, A. DeCristofaro, and E.A. Kerr. 2003. "The quality of health care delivered to adults in the United States." *New England Journal of Medicine* 348(26): 2635-45.

[12] Steinberg, E.P. 2003. "Improving the quality of care – can we practice what we preach?" *New England Journal of Medicine* 348(26):2681-83.

---

[13] Agency for Healthcare Research and Quality. 2006. Guide to Inpatient Quality Indicators. Available at: http://www.qualityindicators.ahrq.gov/downloads/iqi/iqi_guide_v30.pdf, last accessed September 2006.

[14] Donabedian A. 1988. "The quality of care. How can it be assessed?" *JAMA* 260(12): 1743-8.

[15] Agency for Healthcare Research and Quality. 2005. National Healthcare Quality Report, 2005. Available at: http://www.ahrq.gov/qual/nhqr05/nhqr05.htm, last accessed August 2006.

[16] Agency for Healthcare Research and Quality. 2001. Refinement of the HCUP Quality Indicators. Available at: http://www.qualityindicators.ahrq.gov/documentation.htm, last accessed September 2006.

[17] Agency for Healthcare Research and Quality. 2002. Measures of Patient Safety Based on Hospital Administrative Data - The Patient Safety Indicators. Available at: http://www.qualityindicators.ahrq.gov/documentation.htm, last accessed September 2006.

[18] Agency for Healthcare Research and Quality. AHRQ QI Downloads. Available at: http://www.qualityindicators.ahrq.gov/downloads.htm, last accessed September, 2006.

[19] Agency for Healthcare Research and Quality. 2004. Guidance for Using the AHRQ Quality Indicators for Hospital-level Public Reporting or Payment guidance document. Available at: http://www.qualityindicators.ahrq.gov/downloads/technical/qi_guidance.pdf, last accessed August 2006.

[20] Agency for Healthcare Research and Quality. 2005. AHRQ summary statement on comparative hospital public reporting. Available at: http://www.qualityindicators.ahrq.gov/news/AHRQSummaryStatement.pdf, last accessed September 2006.

[21] JCAHO Quality Check. Available at: www.qualitycheck.org, last accessed August 2006.

[22] United States Department of Health and Human Services. Hospital Compare. Available at: http://www.hospitalcompare.hhs.gov/, last accessed August 2006.

[23] Centers for Medicare and Medicaid Services. Premier Hospital Quality Incentive Demonstration. Available at: http://www.cms.hhs.gov/HospitalQualityInits/35_HospitalPremier.asp, last accessed September 2006.

---

[24] JCAHO. A Comprehensive Review of Development and Testing for National Implementation of Hospital Core Measures. Available at: http://www.jointcommission.org/NR/rdonlyres/48DFC95A-9C05-4A44-AB05-1769D5253014/0/AComprehensiveReviewofDevelopmentforCoreMeasures.pdf, last accessed August 2006.

[25] Jarman B., A. Bottle, P. Aylin, M. Browne. 2006. "Monitoring changes in hospital standardised mortality ratios." *British Medical Journal* 330(7487): 329.

[26] Greenwald, L., Cromwell, J., Adamache, W., Bernard, S., Drozd, E., Root, E., and K. Devers. 2006. "Specialty versus community hospitals: referrals, quality, and community benefits. *Health Affairs* 25(1): 106-118.

[27] Mattke S. Monitoring and Improving the Technical Quality of Medical Care: A New Challenge for Policymakers in OECD Countries. In: Towards High Performing Health Systems. Policy Studies, OECD, Paris 2004.

[28] Organization for Economic Co-operation and Development. 2006. OECD Health Data 2006: Statistics and Indicators for 30 Countries. Available at: http://www.oecd.org/document/30/0,2340,en_2649_37407_12968734_1_1_1_37407,00.html, last accessed September 2006.

[29] Organization for Economic Co-operation and Development. OECD Health Care Quality Indicators Project. Available at: http://www.oecd.org/document/34/0,2340,en_2649_37407_37088930_1_1_1_37407,00.html, last accessed September 2006.

[30] Mattke, S., et al. Health Care Quality Indicators Project Initial Indicators Report. 2006. OECD Health Working Papers. No. 22. Available at: http://www.oecd.org/dataoecd/1/34/36262514.pdf, last accessed September 2006.

[31] Mattke S, Epstein AE, Leatherman S. 2006. "The OECD Health Care Quality Indicators Project: history and background." *International Journal of Quality in Healthcare* 18(S1):1-4.

[32] http://www.oecd.org/dataoecd/53/26/33878001.pdf

[33] http://www.oecd.org/dataoecd/27/52/33865865.pdf

[34] Van den Heede K, Sermeus W, Diya L, Lesaffre E, Vleugels A. 2006. "Adverse outcomes in Belgian acute hospitals: retrospective analysis of the national hospital discharge dataset." *International Journal of Quality in Healthcare* 18(3): 211-219.

[35] Agency for Healthcare Research and Quality. AHRQ QIs technical documentation. Available at: http://www.qualityindicators.ahrq.gov/documentation.htm, last accessed August 2006.

---

[36] Hibbard, J, J. Stockard, M. Tusler. 2005. "Hospital performance reports: impact on quality, market share, and reputation." *Health Affairs* 24(4): 1150-1160.

[37] American College of Obstetricians and Gynecologists (ACOG). Vaginal birth after previous cesarean delivery. Washington (DC): American College of Obstetricians and Gynecologists (ACOG); 2004 Jul. 10 p. (ACOG practice bulletin; no. 54).

[38] The Dartmouth Atlas of Health Care, available at http://www.dartmouthatlas.org/index.shtm, last accessed October 2006.

[39] U.S. Department of Health and Human Services, Health Resources and Services Administration, "The Massachusetts Health Care Workforce: Highlights from the Health Workforce Profile, available at http://bhpr.hrsa.gov/healthworkforce/reports/statesummaries/massachusetts.htm, last accessed October 2006.

[40] Center for Studying Healthsystem Change. 2005. Blue Cross Influence Grows in Boston as State Revisits Reform Debates. Available at: http://www.hschange.com/CONTENT/802/, last accessed September 2006.

[41] Center for Studying Healthsystem Change. 2005. Blue Cross Influence Grows in Boston as State Revisits Reform Debates. Available at: http://www.hschange.com/CONTENT/802/, last accessed September 2006.

[42] Dallas-Fort Worth Hospital Council. Data Services. Available at: http://www.dfwhc.org/data+services/, last accessed September 2006.

[43] Texas Department of State Health Services. Texas Health Care Information Collection Center for Health Statistics. Available at: http://www.dshs.state.tx.us/thcic/, last accessed September 2006.

[44] Tex. Health & Safety Code Ann. §§ 108 available at: http://www.capitol.state.tx.us/cgi-bin/statutes/pdfframe.cmd?filepath=/statutes/docs/HS/content/pdf/hs.002.00.000108.00.pdf&title=HEALTH%20%26%20SAFETY%20CODE%20-%20CHAPTER%20108, last accessed August 2006.

---

[45] Blue Cross Blue Shield of Texas reports the PSIs publicly in a little publicized effort. See under Provider Tools: Hospital Comparison Tools: Professional (Beta) and Consumer at http://www.bcbstx.com/provider/, last accessed August 2006.

[46] Marshall MN, Shekelle PG, Leatherman S, Brook RH. 2000. "The public release of performance data: what do we expect to gain? A review of the evidence." *JAMA* 283:1866-74.

[47] Agency for Healthcare Research and Quality. 2005. AHRQ summary statement on comparative hospital public reporting. Available at: http://www.qualityindicators.ahrq.gov/news/AHRQSummaryStatement.pdf, last accessed September 2006.