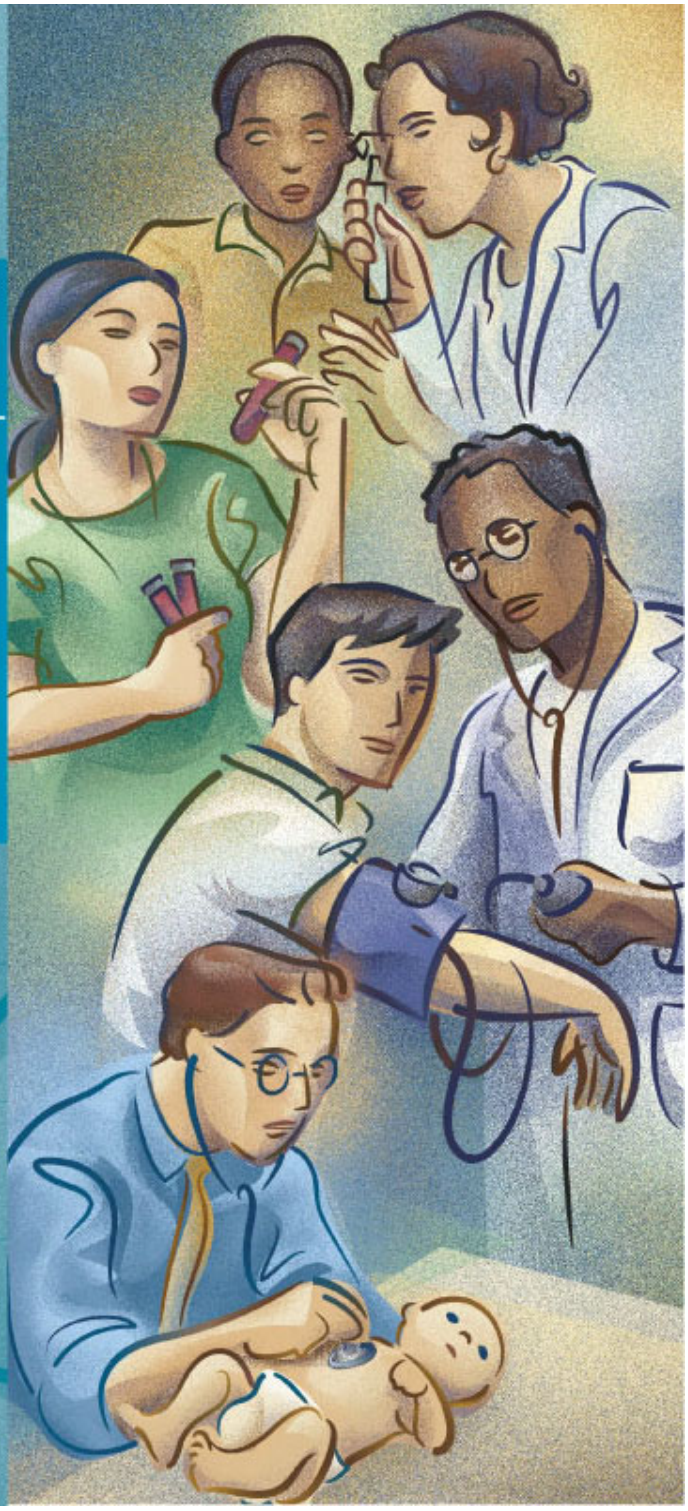
AHRQ

AHRQ QUALITY INDICATORS

Guide to Prevention Quality Indicators



AHRQ Quality Indicators

Guide to Prevention Quality Indicators: Hospital Admission for Ambulatory Care Sensitive Conditions

Department of Health and Human Services
Agency for Healthcare Research and Quality
http://www.qualityindicators.ahrq.gov

October 2001
AHRQ Pub. No. 02-R0203
Revision 4 (November 24, 2004)

Citation

AHRQ Quality Indicators—Guide to Prevention Quality Indicators: Hospital Admission for Ambulatory Care Sensitive Conditions. Rockville, MD: Agency for Healthcare Research and Quality. Revision 4. (November 24, 2004). AHRQ Pub. No. 02-R0203.

Preface

In health care as in other arenas, that which cannot be measured is difficult to improve. Providers, consumers, policy makers, and others seeking to improve the quality of health care need accessible, reliable indicators of quality that they can use to flag potential problems, follow trends over time, and identify disparities across regions, communities, and providers. As noted in a 2001 Institute of Medicine study, *Envisioning the National Health Care Quality Report*, it is important that such measures cover not just acute care but multiple dimensions of care: staying healthy, getting better, living with illness or disability, and coping with the end of life.

The Agency for Healthcare Research and Quality (AHRQ) Quality Indicators (QIs) are one Agency response to this need for a multidimensional, accessible family of quality indicators. They include a family of measures that providers, policy makers, and researchers can use with inpatient data to identify apparent variations in the quality of either inpatient or outpatient care. AHRQ's Evidence-Based Practice Center (EPC) at the University of California San Francisco (UCSF) and Stanford University adapted, expanded, and refined these indicators based on the original Healthcare Cost and Utilization Project (HCUP) Quality Indicators developed in the early 1990s.

The new AHRQ QIs are organized into three modules: Prevention Quality Indicators (PQIs), Inpatient Quality Indicators (IQIs), and Patient Safety Indicators (PSIs). AHRQ has published the three modules as a series. Full technical information on the first two modules can be found in *Refinement of the HCUP Quality Indicators (Summary), May 2001* prepared by the UCSF-Stanford EPC. It can be accessed at http://www.qualityindicators.ahrq.gov/downloads.htm.

This first module focuses on preventive care services—outpatient services geared to staying healthy and living with illness. Researchers and policy makers have agreed for some time that inpatient data offer a useful window on the quality of preventive care in the community. Inpatient data provide information on admissions for "ambulatory care sensitive conditions" that evidence suggests could have been avoided, at least in part, through better outpatient care. Hospitals, community leaders, and policy makers can then use such data to identify community need levels, target resources, and track the impact of programmatic and policy interventions.

One of the most important ways we can improve the quality of health care in America is to reduce the need for some of that care by providing appropriate, high-quality preventive services. For this to happen, however, we need to be able to track not only the level of outpatient services but also the outcome of the services people do or do not receive. The PQIs are intended to facilitate such efforts. The PQIs are already being applied at the national level in the National Healthcare Quality Report (http://qualitytools.ahrq.gov/qualityreport) and National Healthcare Disparities Report (http://qualitytools.ahrq.gov/disparitiesreport.) As always, we would appreciate hearing from those who use our measures and tools so that we can identify how they are used, how they can be refined, and how we can measure and improve the quality of the tools themselves.

Irene Fraser, Ph.D., Director
Center for Organization and Delivery Studies

The programs for the Prevention Quality Indicators (PQIs) can be downloaded from http://www.qualityindicators.ahrq.gov/pqi_download.htm. Instructions on how to use the programs to calculate the PQI rates are contained in the companion text, *Prevention Quality Indicators: Software Documentation.*

Acknowledgments

This product is based on the work of many individuals who contributed to its development and testing.

The following staff from the **Evidence-based Practice Center (EPC) at UCSF-Stanford** performed the evidence review, completed the empirical evaluation, and created the programming code and technical documentation for the new Quality Indicators:

Core Project Team

Mark McClellan, M.D., Ph.D., *principal investigator*
Kathryn M. McDonald, M.M., *EPC coordinator*
Sheryl M. Davies, M.A.
Jeffrey Geppert, J.D.
Patrick Romano, M.D., M.P.H.
Kaveh G. Shojania, M.D.

Other Contributors

Amber Barnato, M.D.
Paul Collins, B.A.
Bradford Duncan M.D.
Michael Gould, M.D., M.S.
Paul Heidenreich, M.D.
Corinna Haberland, M.D.
Paul Matz, M.D.
Courtney Maclean, B.A.
Susana Martins, M.D.
Kristine McCoy, M.P.H.
Suzanne Olson, M.A.
L. LaShawndra Pace, B.A.
Mark Schleinitz, M.D.
Herb Szeto, M.D.
Carol Vorhaus, M.B.A
Peter Weiss, M.D.
Meghan Wheat, B.A.

Consultants
Douglas Staiger, Ph.D.

The following staff from **Social & Scientific Systems**, **Inc.** developed this software product, documentation, and guide:

Programmers
Leif Karell
Kathy McMillan
Fred Rohde

Technical Writer
Patricia Burgess
Graphics Designer
Laura Spofford

Contributors from the **Agency for Healthcare Research and Quality:**

Anne Elixhauser, Ph.D.
Denise Remus, Ph.D., R.N.
H. Joanna Jiang, Ph.D.
Margaret Coopey, R.N., M.G.A, M.P.S.

Support efforts, including refinement and enhancement of the AHRQ Quality Indicators and related products, are provided by the Support for Quality Indicators contract team.

The following individuals from **Battelle Memorial Institute, Stanford University,** and **University of California (UC)** constitute the Support for Quality Indicators-II core team:

Sheryl M. Davies, M.A.
Bruce Ellis, M.S.
Jeffrey Geppert, J.D.
Mark Gritz, Ph.D.
Greg Hubert, B.S.
Elaine Keller, M.Ed.
Kathryn M. McDonald, M.M.
Patrick Romano, M.D., M.P.H
Jeff Schoenborn, BS

The **Agency for Healthcare Research and Quality** Support for Quality Indicators-II team includes:

Denise Remus, Ph.D., R.N., Project Officer
Mary B. Haines, Contract Officer

We also wish to acknowledge the contribution of the peer reviewers of the evidence report and the beta-testers of the software products, whose input was invaluable.

Table of Contents

Preface i

Acknowledgments ii

Introduction to the AHRQ Prevention Quality Indicators 1
What Are the Prevention Quality Indicators? 1
How Can the PQIs Be Used in Quality Assessment? 2
What does this Guide Contain? 3

Origins and Background of the Quality Indicators 4
Development of the AHRQ Quality Indicators 4
AHRQ Quality Indicator Modules 5

Methods of Identifying, Selecting, and Evaluating the Quality Indicators 6
Step 1: Obtain Background Information on QI Use 6
Step 2: Search the Literature to Identify Potential QIs 6
Step 3: Review the Literature to Evaluate the QIs According to Predetermined Criteria 7
Step 4: Perform a Comprehensive Evaluation of Risk Adjustment 8
Step 5: Evaluate the Indicators Using Empirical Analyses 9

Summary Evidence on the Prevention Quality Indicators 11
Strengths and Limitations in Using the PQIs 14
Questions for Future Work 15

Detailed Evidence for Prevention Quality Indicators 17
Diabetes Short-term Complications Admission Rate (PQI 1) 20
Perforated Appendix Admission Rate (PQI 2) 22
Diabetes Long-term Complications Admission Rate (PQI 3) 24
Pediatric Asthma Admission Rate (PQI 4) 26
Chronic Obstructive Pulmonary Disease Admission Rate (PQI 5) 28
Pediatric Gastroenteritis Admission Rate (PQI 6) 30
Hypertension Admission Rate (PQI 7) 32
Congestive Heart Failure Admission Rate (PQI 8) 34
Low Birth Weight Rate (PQI 9) 36
Dehydration Admission Rate (PQI 10) 38
Bacterial Pneumonia Admission Rate (PQI 11) 40
Urinary Tract Infection Admission Rate (PQI 12) 42
Angina without Procedure Admission Rate (PQI 13) 44
Uncontrolled Diabetes Admission Rate (PQI 14) 46
Adult Asthma Admission Rate (PQI 15) 48
Rate of Lower-extremity Amputation among Patients with Diabetes (PQI 16) 50

References 52

Appendix A: Prevention Quality Indicator Definitions A-1

Appendix B: Detailed Methods B-1
Semi-structured Interviews B-1
Quality Indicators Evaluation Framework B-1
Literature Review of Quality Indicators B-8
Risk Adjustment of HCUP Quality Indicators B-13
Empirical Methods B-20
References for Appendix B B-28

Appendix C: Log of Revisions to PQI Documentation and Software Version 2.1, Revision 4 .... C-1

Appendix D: ICD-9-CM and DRG Coding Updates in PQI Release Version 2.1, Revision 4 ..... D-1

List of Tables

Table 1. Prevention Quality Indicators 12
Table 2. Diabetes-related Prevention Quality Indicators 14

Introduction to the AHRQ Prevention Quality Indicators

Prevention is an important role for all health care providers. Providers can help individuals stay healthy by preventing disease, and they can prevent complications of existing disease by helping patients live with their illnesses. To fulfill this role, however, providers need data on the impact of their services and the opportunity to compare these data over time or across communities. Local, State, and Federal policymakers also need these tools and data to identify potential access or quality-of-care problems related to prevention, to plan specific interventions, and to evaluate how well these interventions meet the goals of preventing illness and disability.

The Agency for Healthcare Research and Quality (AHRQ) Prevention Quality Indicators (PQIs) represent one such tool. Local, State, or national data collected using the PQIs can flag potential problems resulting from a breakdown of health care services by tracking hospitalizations for conditions that should be treatable on an outpatient basis, or that could be less severe if treated early and appropriately. The PQIs represent the current state of the art in measuring the outcomes of preventive and outpatient care through analysis of inpatient discharge data.

What Are the Prevention Quality Indicators?

The PQIs are a set of measures that can be used with hospital inpatient discharge data to identify "ambulatory care sensitive conditions" (ACSCs). ACSCs are conditions for which good outpatient care can potentially prevent the need for hospitalization, or for which early intervention can prevent complications or more severe disease.

Even though these indicators are based on hospital inpatient data, they provide insight into the quality of the health care system *outside* the hospital setting. Patients with diabetes may be hospitalized for diabetic complications if their conditions are not adequately monitored or if they do not receive the patient education needed for appropriate self-management. Patients may be hospitalized for asthma if primary care providers fail to adhere to practice guidelines or to prescribe appropriate treatments. Patients with appendicitis who do not have ready access to surgical evaluation may experience delays in receiving needed care, which can result in a life-threatening condition—perforated appendix. The PQIs consist of the following 16 ambulatory care sensitive conditions, which are measured as rates of admission to the hospital:

PQI Number	Prevention Quality Indicators
1	Diabetes short-term complication admission rate
2	Perforated appendix admission rate
3	Diabetes long-term complication admission rate
4	Pediatric asthma admission rate
5	Chronic obstructive pulmonary disease admission rate
6	Pediatric gastroenteritis admission rate
7	Hypertension admission rate
8	Congestive heart failure admission rate
9	Low birth weight rate
10	Dehydration admission rate
11	Bacterial pneumonia admission rate
12	Urinary tract infection admission rate
13	Angina admission without procedure
14	Uncontrolled diabetes admission rate
15	Adult asthma admission rate
16	Rate of lower-extremity amputation among patients with diabetes

Although other factors outside the direct control of the health care system, such as poor environmental conditions or lack of patient adherence to treatment recommendations, can result in hospitalization, the PQIs provide a good starting point for assessing quality of health services in the community. Because the PQIs are calculated using readily available hospital administrative data, they are an easy-to-use and inexpensive screening tool. They can be used to provide a window into the community—to identify unmet community heath care needs, to monitor how well complications from a number of common conditions are being avoided in the outpatient setting, and to compare performance of local health care systems across communities.

How Can the PQIs Be Used in Quality Assessment?

While these indicators use hospital inpatient data, their focus is on outpatient health care. Except in the case of patients who are readmitted soon after discharge from a hospital, the quality of inpatient care is unlikely to be a significant determinant of admission rates for ambulatory care sensitive conditions. Rather, the PQIs assess the quality of the health care system as a whole, and especially the quality of ambulatory care, in preventing medical complications. As a result, these measures are likely to be of the greatest value when calculated at the population level and when used by public health groups, State data organizations, and other organizations concerned with the health of populations.[1]

These indicators serve as a screening tool rather than as definitive measures of quality problems. They can provide initial information about potential problems in the community that may require further, more in-depth analysis. Policy makers and health care providers can use the PQIs to answer questions such as:

- How does the low birth weight rate in my State compare with the national average?
- What can the pediatric indicators in the PQIs tell me about the adequacy of pediatric primary care in my community?
- Does the admission rate for diabetes complications in my community suggest a problem in the provision of appropriate outpatient care to this population?
- How does the admission rate for congestive heart failure vary over time and from one region of the country to another?

State policy makers and local community organizations can use the PQIs to assess and improve community health care. For example, an official at a State health department wants to gain a better understanding of the quality of care provided to people with diabetes in her State. She selects the four PQIs related to diabetes and applies the statistical programs downloaded from the AHRQ Web site to hospital discharge abstract data collected by her State.

Based on output from the programs, she examines the age- and sex-adjusted admission rates for these diabetes PQIs for her State as a whole and for communities within her State. The programs provide output that she uses to compare different population subgroups, defined by age, ethnicity, or gender. She finds that admission rates for short-term diabetes complications and uncontrolled diabetes are especially high in a major city in her State and that there are differences by race/ethnicity. She also applies the PQI programs to multiple years of her State's data to track trends in hospital admissions over time. She discovers that the trends for these two PQIs are increasing in this city but are stable in the rest of the State. She then compares the figures from her State to national and regional averages on these PQIs using HCUPnet—an online query system providing access to statistics based on HCUP data.[2] The

[1] Individual hospitals that are sole providers for communities and that are involved in outpatient care may be able to use the PQI programs. Managed care organizations and health care providers with responsibility for a specified enrolled population can use the PQI programs but must provide their own population denominator data.
[2] HCUPnet can be found at http://hcup.ahrq.gov/HCUPnet.asp and provides instant access to national and regional data from the Healthcare Cost and Utilization Project, a Federal-State-industry partnership in health data maintained by the Agency for Healthcare Research and Quality.

State average is slightly higher than the regional and national averages, but the averages for this city are substantially higher.

After she has identified disparities in admission rates in this community and in specific patient groups, she further investigates the underlying reasons for those disparities. She attempts to obtain information on the prevalence of diabetes across the State to determine if prevalence is higher in this city than in other communities. Finding no differences, she consults with the State medical association to begin work with local providers to discern if quality-of-care problems underlie these disparities. She contacts hospitals and physicians in this community to determine if community outreach programs can be implemented to encourage patients with diabetes to seek care and to educate them on lifestyle modifications and diabetes self-management. She then helps to develop specific interventions to improve care for people with diabetes and reduce preventable complications and resulting hospitalizations.

What does this Guide Contain?

This guide provides background information on the PQIs. First, it describes the origin of the entire family of AHRQ Quality Indicators. Second, it provides an overview of the methods used to identify, select, and evaluate the AHRQ Quality Indicators. Third, the guide summarizes the PQIs specifically, describes strengths and limitations of the indicators, documents the evidence that links the PQIs to the quality of outpatient health care services, and then provides in-depth two-page descriptions of each PQI. Finally, two appendices present additional technical background information. The first appendix outlines the specific definitions of each PQI, with complete ICD-9-CM coding specifications. The second appendix provides the details of the empirical methods used to explore the PQIs.

Origins and Background of the Quality Indicators

In the early 1990s, in response to requests for assistance from State-level data organizations and hospital associations with inpatient data collection systems, AHRQ developed a set of quality measures that required only the type of information found in routine hospital administrative data—diagnoses and procedures, along with information on patient's age, gender, source of admission, and discharge status. These States were part of the Healthcare Cost and Utilization Project, an ongoing Federal-State-private sector collaboration to build uniform databases from administrative hospital-based data.

AHRQ developed these measures, called the HCUP Quality Indicators, to take advantage of a readily available data source—administrative data based on hospital claims—and quality measures that had been reported elsewhere.[3] The 33 HCUP QIs included measures for avoidable adverse outcomes, such as in-hospital mortality and complications of procedures; use of specific inpatient procedures thought to be overused, underused, or misused; and ambulatory care sensitive conditions.

Although administrative data cannot provide definitive measures of health care quality, they can be used to provide indicators of health care quality that can serve as the starting point for further investigation. The HCUP QIs have been used to assess potential quality-of-care problems and to delineate approaches for dealing with those problems. Hospitals with high rates of poor outcomes on the HCUP QIs have reviewed medical records to verify the presence of those outcomes and to investigate potential quality-of-care problems.[4] For example, one hospital that detected high rates of admissions for diabetes complications investigated the underlying reasons for the rates and established a center of excellence to strengthen outpatient services for patients with diabetes.

Development of the AHRQ Quality Indicators

Since the original development of the HCUP QIs, the knowledge base on quality indicators has increased significantly. Risk-adjustment methods have become more readily available, new measures have been developed, and analytic capacity at the State level has expanded considerably. Based on input from current users and advances to the scientific base for specific indicators, AHRQ funded a project to refine and further develop the original QIs. The project was conducted by the UCSF-Stanford EPC.

The major constraint placed on the UCSF-Stanford EPC was that the measures could require only the type of information found in hospital discharge abstract data. Further, the data elements required by the measures had to be available from most inpatient administrative data systems. Some State data systems contain innovative data elements, often based on additional information from the medical record. Despite the value of these record-based data elements, the intent of this project was to create measures that were based on a common denominator discharge data set, without the need for additional data collection. This was critical for two reasons. First, this constraint would result in a tool that could be used with any inpatient administrative data, thus making it useful to most data systems. Second, this would enable national and regional benchmark rates to be provided using HCUP data, since these benchmark rates would need to be calculated using the universe of data available from the States.

---

[3] Ball JK, Elixhauser A, Johantgen M, et al. *HCUP Quality Indicators, Methods, Version 1.1: Outcome, Utilization, and Access Measures for Quality Improvement*. (AHCPR Publication No. 98-0035). Healthcare Cost and Utilization project (HCUP-3) Research notes: Rockville, MD: Agency for Health Care Policy and Research, 1998.

[4] *Impact: Case Studies Notebook – Documented Impact and Use of AHRQ's Research*. Compiled by Division of Public Affairs, Office of Health Care Information, Agency for Healthcare Research and Quality.

AHRQ Quality Indicator Modules

The work of the UCSF-Stanford EPC resulted in the AHRQ Quality Indicators, which are available as three separate modules:

- **Prevention Quality Indicators**. These indicators consist of "ambulatory care sensitive conditions," hospital admissions that evidence suggests could have been avoided through high-quality outpatient care or that reflect conditions that could be less severe, if treated early and appropriately.
- **Inpatient Quality Indicators**. These indicators reflect quality of care inside hospitals and include inpatient mortality; utilization of procedures for which there are questions of overuse, underuse, or misuse; and volume of procedures for which there is evidence that a higher volume of procedures is associated with lower mortality.
- **Patient Safety Indicators**. These indicators also reflect quality of care inside hospitals, but focus on surgical complications and other iatrogenic events.

Methods of Identifying, Selecting, and Evaluating the Quality Indicators

In developing the new quality indicators, the UCSF-Stanford EPC applied the Institute of Medicine's widely cited definition of quality care: "the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge."[5] They formulated six specific key questions to guide the development process:

- Which indicators are currently in use or described in the literature that could be *defined using hospital discharge data*?
- What are the *quality relationships* reported in the literature that could be used to define new indicators using hospital discharge data?
- What evidence exists for *indicators not well represented* in the original indicators—pediatric conditions, chronic disease, new technologies, and ambulatory care sensitive conditions?
- Which indicators have *literature-based evidence* to support face validity, precision of measurement, minimum bias, and construct validity of the indicator?
- What *risk-adjustment method* should be suggested for use with the recommended indicators, given the limits of administrative data and other practical concerns?
- Which indicators perform well on *empirical tests* of precision of measurement, minimum bias, and construct validity?

As part of this project, the UCSF-Stanford EPC identified quality indicators reported in the literature and used by health care organizations, evaluated the original quality indicators and potential indicators using literature review and empirical methods, incorporated risk adjustment for comparative analysis, and developed new programs that could be employed by users with their own hospital administrative data. This section outlines the steps used to arrive at a final set of quality measures.

Step 1: Obtain Background Information on QI Use

The project team at the UCSF-Stanford EPC interviewed 33 individuals affiliated with hospital associations, business coalitions, State data groups, Federal agencies, and academia about various topics related to quality measurement, including indicator use, suggested indicators, and other potential contacts. Interviews were tailored to the specific expertise of interviewees. The sample was not intended to be representative of any population; rather, individuals were selected to include QI users and potential users from a broad spectrum of organizations in both the public and private sectors.

Three broad audiences were considered for the quality measures: health care providers and managers, who could use the quality measures to assist in initiatives to improve quality; public health policy makers, who could use the information from indicators to target public health interventions; and health care purchasers, who could use the measures to guide decisions about health policies.

Step 2: Search the Literature to Identify Potential QIs

The project team performed a structured review of the literature to identify potential indicators. They used Medline to identify the search strategy that returned a test set of known applicable articles in the most concise manner. Using the Medical Subject Heading (MeSH) terms "hospital, statistic, and methods" and "quality indicators" resulted in approximately 2,600 articles published in 1994 or later. After

[5] Institute of Medicine Division of Health Care Services. Medicare: a strategy for quality assurance. Washington, DC: National Academy Press; 1990.

screening titles and abstracts for relevancy, the search yielded 181 articles that provided information on potential quality indicators based on administrative data.

Clinicians, health services researchers, and other team members abstracted information from these articles in two stages. In the first stage, preliminary abstraction, they evaluated each of the 181 identified articles for the presence of a defined quality indicator, clinical rationale, and strengths and weaknesses. To qualify for full abstraction, the articles must have explicitly defined a novel quality indicator. Only 27 articles met this criterion. The team collected information on the definition of the quality indicator, validation, and rationale during full abstraction.

In addition, they identified additional potential indicators using the CONQUEST database; the National Library of Healthcare Indicators developed by the Joint Commission on Accreditation of Healthcare Organizations (JCAHO); a list of ORYX-approved indicators provided by JCAHO; and telephone interviews.

Step 3: Review the Literature to Evaluate the QIs According to Predetermined Criteria

The project team evaluated each potential quality indicator against the following six criteria, which were considered essential for determining the reliability and validity of a quality indicator:

- **Face validity.** An adequate quality indicator must have sound clinical or empirical rationale for its use. It should measure an important aspect of quality that is subject to provider or health care system control.
- **Precision.** An adequate quality indicator should have relatively large variation among providers or areas that is not due to random variation or patient characteristics. This criterion measures the impact of chance on apparent provider or community health system performance.
- **Minimum bias.** The indicator should not be affected by systematic differences in patient case-mix, including disease severity and comorbidity. In cases where such systematic differences exist, an adequate risk adjustment system should be possible using available data.
- **Construct validity.** The indicator should be related to other indicators or measures intended to measure the same or related aspects of quality. In general, better outpatient care (including, in some cases, adherence to specific evidence-based treatment guidelines) can reduce patient complication rates.
- **Fosters real quality improvement.** The indicator should be robust to possible provider manipulation of the system. In other words, the indicator should be insulated from perverse incentives for providers to improve their reported performance by avoiding difficult or complex cases, or by other responses that do not improve quality of care.
- **Application.** The indicator should have been used in the past or have high potential for working well with other indicators. Sometimes looking at groups of indicators together is likely to provide a more complete picture of quality.

Based on the initial review, the team identified and evaluated over 200 potential indicators using these criteria. Of this initial set, 45 indicators passed this initial screen and received comprehensive literature and empirical evaluation. In some cases, whether an indicator complemented other promising indicators was a consideration in retaining it, allowing the indicators to provide more depth in specific areas.

For this final set of 45 indicators, the team reviewed an additional 2,000 articles to provide evidence on indicators during the evaluation phase. They searched Medline for articles relating to each of the six areas of evaluation described above. Clinicians and health services researchers reviewed the literature for evidence and prepared a referenced summary description on each indicator.

As part of the review process, the team assessed the link between each indicator and health care quality along the following dimensions:

- **Proxy.** Some indicators do not specifically measure a patient outcome or a process measure of quality. Rather, they measure an aspect of care that is correlated with process measures of quality or patient outcomes. These indicators are best used in conjunction with other indicators measuring similar aspects of clinical care, or when followed with more direct and in-depth investigations of quality.
- **Selection bias.** Selection bias results when a substantial percentage of care for a condition is provided in the outpatient setting, so the subset of inpatient cases may be unrepresentative. In these cases, examination of outpatient care or emergency room data may help reduce selection bias.
- **Information bias.** Quality indicators are based on information available in hospital discharge data sets, but some missing information may actually be important to evaluating the outcomes of hospital care. In these cases, examination of missing information may help to improve indicator performance.
- **Confounding bias.** Patient characteristics may substantially affect performance on a measure and may vary systematically across areas. In these cases, adequate risk adjustment may help to improve indicator performance.
- **Unclear construct validity.** Problems with construct validity include uncertain or poor correlations with widely accepted process measures or with risk-adjusted outcome measures. These indicators would benefit from further research to establish their relationship with quality care.
- **Easily manipulated.** Quality indicators may create perverse incentives to improve performance without actually improving quality. Although very few of these perverse responses have been proven, they are theoretically important and should be monitored to ensure true quality improvement.
- **Unclear benchmark.** For some indicators, the "right rate" has not been established, so comparison with national, regional, or peer group means may be the best benchmark available. Very low PQI rates may flag an underuse problem; that is, providers may fail to hospitalize patients who would benefit from inpatient care. On the other hand, overuse of acute care resources may potentially occur when patients who do not clinically require inpatient care are hospitalized.

Step 4: Perform a Comprehensive Evaluation of Risk Adjustment

The project team identified potential risk-adjustment systems by reviewing the applicable literature and asking the interviewees in step 1 to identify their preferences. Generally, users preferred that the system be (1) open, with published logic; (2) cost-effective, with data collection costs minimized and additional data collection being well justified; (3) designed using a multiple-use coding system, such as those used for reimbursement; and (4) officially recognized by government, hospital groups, or other organizations.

In general, diagnosis-related groups (DRGs) seemed to fit more of the user preference-based criteria than other alternatives. A majority of the users interviewed already used the 3M™ All-Patient Refined DRG[6] (APR-DRG) system, which has been reported to perform well in predicting resource use and death when compared to other DRG-based systems.

[6] Information on the 3M™ APR-DRG system is available at http://www.3m.com/us/healthcare/his/products/coding/refined_drg.jhtml.

APR-DRGs were used to conduct indicator evaluations to determine the impact of measured differences in patient severity on the relative performance of providers and to provide the basis for implementing APR-DRGs as an optional risk-adjustment system for hospital-level QI measures. The implementation of APR-DRGs is based on an ordinary least squares regression model. Area indicators (including all PQIs) were risk-adjusted only for age and sex differences. Detailed information on the risk-adjustment methods can be found in Appendix B.

Step 5: Evaluate the Indicators Using Empirical Analyses

The project team conducted extensive empirical testing of all potential indicators using the 1995-97 HCUP State Inpatient Databases (SID) and Nationwide Inpatient Sample (NIS) to determine precision, bias, and construct validity. The 1997 SID contains uniform data on inpatient stays in community hospitals for 22 States covering approximately 60% of all U.S. hospital discharges. The NIS is designed to approximate a 20% of U.S. community hospitals and includes all stays in the sampled hospitals. Each year of the NIS contains between 6 million and 7 million records from about 1,000 hospitals. The NIS combines a subset of the SID data, hospital-level variables, and hospital and discharge weights for producing national estimates. The project team conducted tests to examine three things: precision, bias, and construct validity.

Precision. The first step in the analysis involved precision tests to determine the reliability of the indicator for distinguishing real differences in provider performance. For indicators that may be used for quality improvement, it is important to know with what precision, or surety, a measure can be attributed to an actual construct rather than random variation.

For each indicator, the variance can be broken down into three components: variation within a provider (actual differences in performance due to differing patient characteristics), variation among providers (actual differences in performance among providers), and random variation. An ideal indicator would have a substantial amount of the variance explained by between-provider variance, possibly resulting from differences in quality of care, and a minimum amount of random variation. The project team performed four tests of precision to estimate the magnitude of between-provider variance on each indicator:

- Signal standard deviation was used to measure the extent to which performance of the QI varies systematically across hospitals or areas.
- Provider/area variation share was used to calculate the percentage of signal (or true) variance relative to the total variance of the QI.
- Signal-to-noise ratio was used to measure the percentage of the apparent variation in QIs across providers that is truly related to systematic differences across providers and not random variations (noise) from year to year.
- In-sample R-squared was used to identify the incremental benefit of applying multivariate signal extraction methods for identifying additional signal on top of the signal-to-noise ratio.

In general, random variation is most problematic when there are relatively few observations per provider, when adverse outcome rates are relatively low, and when providers have little control over patient outcomes or variation in important processes of care is minimal. If a large number of patient factors that are difficult to observe influence whether or not a patient has an adverse outcome, it may be difficult to separate the “quality signal” from the surrounding noise. Two signal extraction techniques were applied to improve the precision of an indicator:

- Univariate methods were used to estimate the “true” quality signal of an indicator based on information from the specific indicator and 1 year of data.
- Multivariate signal extraction (MSX) methods were used to estimate the “true” quality signal based on information from a set of indicators and multiple years of data. In most cases, MSX

methods extracted additional signal, which provided much more precise estimates of true hospital or area quality.

Bias. To determine the sensitivity of potential QIs to bias from differences in patient severity, unadjusted performance measures for specific hospitals were compared with performance measures that had been adjusted for age and gender. All of the PQIs and some of the Inpatient Quality Indicators (IQIs) could only be risk-adjusted for age and sex. The 3M™ APR-DRG System Version 12 with Severity of Illness and Risk of Mortality subclasses was used for risk adjustment of the utilization indicators and the in-hospital mortality indicators, respectively. Five empirical tests were performed to investigate the degree of bias in an indicator:

- Rank correlation coefficient of the area or hospital with (and without) risk adjustment—gives the overall impact of risk adjustment on relative provider or area performance.
- Average absolute value of change relative to mean—highlights the amount of absolute change in performance, without reference to other providers' performance.
- Percentage of highly ranked hospitals that remain in high decile—reports the percentage of hospitals or areas that are in the highest deciles without risk adjustment that remain there after risk adjustment is performed.
- Percentage of lowly ranked hospitals that remain in low decile—reports the percentage of hospitals or areas that are in the lowest deciles without risk adjustment that remain there after risk adjustment is performed.
- Percentage that change more than two deciles—identifies the percentage of hospitals whose relative rank changes by a substantial percentage (more than 20%) with and without risk adjustment.

Construct validity. Construct validity analyses provided information regarding the relatedness or independence of the indicators. If quality indicators do indeed measure quality, then two measures of the same construct would be expected to yield similar results. The team used factor analysis to reveal underlying patterns among large numbers of variables—in this case, to measure the degree of relatedness between indicators. In addition, they analyzed correlation matrices for indicators.

Summary Evidence on the Prevention Quality Indicators

The rigorous evaluations performed by the UCSF-Stanford EPC, based on literature review and empirical testing of indicators, resulted in 16 indicators that reflect ambulatory care sensitive conditions (ACSCs). These ACSCs have been reported and tested in a number of published studies involving consensus processes involving panels of expert physicians, using a range of methodologies and decision criteria. Two sets of ambulatory care sensitive conditions are widely used:

- The set developed by John Billings in conjunction with the United Hospital Fund of New York includes 28 ambulatory care sensitive conditions, identified by a panel of six physicians.[7]
- The set developed by Joel Weissman includes 12 avoidable admissions identified through review of the literature and evaluation by a panel of physicians.[8]

Many of the ACSCs have practice guidelines associated with them, including almost all of the chronic conditions and about half of the acute medical or pediatric conditions. Studies have shown that better outpatient care (including, in some cases, adherence to specific evidence-based treatment guidelines) can reduce patient complication rates of existing disease, including complications leading to hospital admissions. Empirically, most of the hospital admission rates for ACSCs are correlated with each other, suggesting that common underlying factors influence many of the rates.

Five of these 16 PQIs were included in the original HCUP QIs—perforated appendix, low birth weight, pediatric asthma, diabetes short-term complications, and diabetes long-term complications—where they were measured at the hospital level. In contrast, the 16 new indicators are constructed at the community level, defined as a Metropolitan Statistical Area (MSA) or a rural county. For each indicator, lower rates indicate potentially better quality.

Table 1 summarizes the results of the literature review and empirical evaluations on the PQIs. It lists each indicator, provides its definition, rates its empirical performance, recommends a risk adjustment strategy, and summarizes important caveats identified from the literature review.

Rating of performance on empirical evaluations, as described in step 5 above, ranged from 0 to 26. (The average score for these 16 PQIs is 14.6.) The scores were intended as a guide for summarizing the performance of each indicator on four empirical tests of precision (signal variance, area-level share, signal ratio, and R-squared) and five tests of minimum bias (rank correlation, top and bottom decile movement, absolute change, and change over two deciles), as described in the previous section and in Appendix B.

The magnitude of the scores, shown in the Empirical Rating column, provides an indication of the relative rankings of the indicators. These scores were based on indicator performance after risk-adjustment and smoothing; that is, they represent the "best estimate" of the indicator's true value after accounting for case-mix and reliability. The score for each individual test is an ordinal ranking (e.g., very high, high, moderate, and low). The final summary score was derived by assigning a weight to each ranking (e.g., 3, 2, 1, 0) and summing across these nine individual tests. Higher scores indicate better performance on the empirical tests.

The Literature Review Findings column summarizes evidence specific to each potential concern on the link between the PQIs and quality of care, as described in step 3 above. A question mark (**?**) indicates that the concern is theoretical or suggested, but no specific evidence was found in the literature.

---

[7]Billings J, Zeitel L, Lukomnik J, et al. Impact of socioeconomic status on hospital use in New York City, Health Aff (Millwood) 1993;12(1):162-73.
[8]Weissman, JS, Gatsonis C, Epstein AM. Rates of avoidable hospitalization by insurance status in Massachusetts and Maryland. JAMA 1992;268(17):2388-94.

A check mark (✔) indicates that the concern has been demonstrated in the literature. For additional details on the results of the literature review, see "Detailed Evidence for the Prevention Quality Indicators."

A complete description of each PQI is included later in the guide under "Detailed Evidence for Prevention Quality Indicators" and in Appendix A. Details on the empirical methods can be found in Appendix B.

Table 1. Prevention Quality Indicators

Indicator Name (Number)	Description	Risk Adjustment Incorporated	Empirical Performance[a]	Literature Review Findings[b]
Diabetes Short-term Complication Admission Rate (PQI 1)	Number of admissions for diabetes short-term complications per 100,000 population.	Age and sex.	Area Rate: 46.7 Area SD: 35.3 Pop. Rate: 50.9 Rating: 14	? Proxy ? Confounding bias
Perforated Appendix Admission Rate (PQI 2)	Number of admissions for perforated appendix as a share of all admissions for appendicitis within an area.	Age and sex.	Area Rate: 32.5 per 100 Area SD: 16.0 Pop. Rate: 30.5 per 100 Rating: 17[c]	? Proxy
Diabetes Long-term Complication Admission Rate (PQI 3)	Number of admissions for long-term diabetes per 100,000 population.	Age and sex.	Area Rate: 112.6 Area SD: 67.6 Pop. Rate: 113.8 Rating: 11	? Proxy ? Confounding bias ? Easily manipulated ✔ Unclear benchmark
Pediatric Asthma Admission Rate (PQI 4)	Number of admissions for pediatric asthma per 100,000 population.	Age and sex.	Area Rate: 164.6 Area SD: 182.5 Pop. Rate: 188.8 Rating: 18	? Proxy ? Easily manipulated ✔ Unclear benchmark
Chronic Obstructive Pulmonary Disease Admission Rate (PQI 5)	Number of admissions for COPD per 100,000 population.	Age and sex.	Area Rate: 344.3 Area SD: 277.7 Pop. Rate: 248.6 Rating: 17	? Proxy ? Confounding bias ? Easily manipulated ✔ Unclear benchmark
Pediatric Gastroenteritis Admission Rate (PQI 6)	Number of admissions for pediatric gastroenteritis per 100,000 population.	Age and sex.	Area Rate: 112.3 Area SD: 137.8 Pop. Rate: 87.7 Rating: 17	? Proxy ? Unclear construct ? Easily manipulated ✔ Unclear benchmark
Hypertension Admission Rate (PQI 7)	Number of admissions for hypertension per 100,000 population.	Age and sex.	Area Rate: 50.2 Area SD: 49.3 Pop. Rate: 44.4 Rating: 14	? Proxy ? Easily manipulated ✔ Unclear benchmark
Congestive Heart Failure Admission Rate (PQI 8)	Number of admissions for CHF per 100,000 population.	Age and sex.	Area Rate: 502.8 Area SD: 250.8 Pop. Rate: 457.7 Rating: 14	? Proxy ? Easily manipulated ✔ Unclear benchmark

Indicator Name (Number)	Description	Risk Adjustment Incorporated	Empirical Performance[a]	Literature Review Findings[b]
Low Birth Weight Rate (PQI 9)	Number of low birth weight births as a share of all births in an area.	Not risk adjusted.	Area Rate: 6.0 per 100 Area SD: 3.8 Pop. Rate: 5.9 per 100 Rating: 11[c] out of 16[d]	? Proxy ? Confounding bias ✔ Unclear construct
Dehydration Admission Rate (PQI 10)	Number of admissions for dehydration per 100,000 population.	Age and sex.	Area Rate: 174.9 Area SD: 131.2 Pop. Rate: 139.9 Rating: 14	? Proxy ? Unclear construct ? Easily manipulated ✔ Unclear benchmark
Bacterial Pneumonia Admission Rate (PQI 11)	Number of admissions for bacterial pneumonia per 100,000 population.	Age and sex.	Area Rate: 503.9 Area SD: 306.8 Pop. Rate: 349.7 Rating: 17	? Proxy ? Unclear construct ? Easily manipulated ✔ Unclear benchmark
Urinary Tract Infection Admission Rate (PQI 12)	Number of admissions for urinary infection per 100,000 population.	Age and sex.	Area Rate: 158.5 Area SD: 97.7 Pop. Rate: 137.9 Rating: 11	? Proxy ? Unclear construct ? Easily manipulated ✔ Unclear benchmark
Angina without Procedure Admission Rate (PQI 13)	Number of admissions for angina without procedure per 100,000 population.	Age and sex.	Area Rate: 82.3 Area SD: 78.0 Pop. Rate: 55.1 Rating: 19	? Proxy ? Unclear construct ? Easily manipulated ✔ Unclear benchmark
Uncontrolled Diabetes Admission Rate[e] (PQI 14)	Number of admissions for uncontrolled diabetes per 100,000 population.	Age and sex.	Area Rate: 27.2 Area SD: 33.9 Pop. Rate: 23.0 Rating: 14	? Proxy ? Confounding bias ? Easily manipulated
Adult Asthma Admission Rate (PQI 15)	Number of admissions for asthma in adults per 100,000 population.	Age and sex.	Area Rate:98.4 Area SD: 79.2 Pop. Rate: 110.9 Rating: 16	? Proxy ? Easily manipulated ✔ Unclear benchmark
Rate of Lower-extremity Amputation Among Patients with Diabetes (PQI 16)	Number of admissions for lower-extremity amputation among patients with diabetes per 100,000 population.	Age and sex.	Area Rate: 37.5 Area SD: 28.2 Pop. Rate: 36.8 Rating: 10[c]	? Proxy ? Unclear construct

[a] Higher scores in the **Empirical Performance** column indicate better performance on the nine empirical tests. Unadjusted means and standard deviations (SD) were calculated using the 2002 SID from 35 states. The area rates are average area rates and area standard deviation based on 2,440 geographic areas (counties) in the 2002 SID. The population rate is based on all discharges in the 2002 SID for 35 states (as opposed to average area rates).

[b] Notes under Literature Review Findings:

Proxy – Indicator does not directly measure patient outcomes but an aspect of care that is associated with the outcome; thus, it is best used with other indicators that measure similar aspects of care.

Confounding bias – Patient characteristics may substantially affect the performance of the indicator; risk adjustment is recommended.
Unclear construct – There is uncertainty or poor correlation with widely accepted process measures.
Easily manipulated – Use of the indicator may create perverse incentives to improve performance on the indicator without truly improving quality of care.
Unclear benchmark – The "correct rate" has not been established for the indicator; national, regional, or peer group averages may be the best benchmark available.
? – The concern is theoretical or suggested, but no specific evidence was found in the literature.
✔ – Indicates that the concern has been demonstrated in the literature.

[c] Smoothing recommended (details provided in Appendix B).
[d] Bias was not tested because adequate risk adjustment for low birth weight was not available.
[e] Uncontrolled diabetes is designed to be combined with diabetes short-term complications

Revision 4 of the PQIs provides the option to generate condition-specific rates (e.g., using the number of diabetics in the denominator) by state and age. Table 2 provides the Empirical Performance rates for the four diabetes-related PQIs, expressed per 1,000.

Table 2. Diabetes-related Prevention Quality Indicators

Indicator Name (Number)	Description	Risk Adjustment Incorporated	Empirical Performance[a]	Literature Review Findings
Diabetes Short-term Complication Admission Rate (PQI 1)	Number of admissions for diabetes short-term complications per 1,000 diabetic.	N/A	Area Rate: 7.3 Area SD: 1.3 Pop. Rate: 7.3 Rating: 14	**?** Proxy **?** Confounding bias
Diabetes Long-term Complication Admission Rate (PQI 3)	Number of admissions for long-term diabetes per 1,000 diabetic.	N/A	Area Rate: 15.7 Area SD: 2.7 Pop. Rate: 16.3 Rating: 11	**?** Proxy **?** Confounding bias **?** Easily manipulated ✔ Unclear benchmark
Uncontrolled Diabetes Admission Rate[e] (PQI 14)	Number of admissions for uncontrolled diabetes per 1,000 diabetic.	N/A	Area Rate: 2.8 Area SD: 1.3 Pop. Rate: 3.3 Rating: 14	**?** Proxy **?** Confounding bias **?** Easily manipulated
Rate of Lower-extremity Amputation Among Patients with Diabetes (PQI 16)	Number of admissions for lower-extremity amputation among patients with diabetes per 1,000 diabetic.	N/A	Area Rate: 5.2 Area SD: 0.9 Pop. Rate: 5.3 Rating: 10[c]	**?** Proxy **?** Unclear construct

[a] Based on 35 states in 2002 SID.

Strengths and Limitations in Using the PQIs

The PQIs represent the current state of the art in assessing quality of health services in local communities using inpatient discharge data. These indicators measure the outcomes of preventive care for both acute illness and chronic conditions, reflecting two important components of the quality of preventive care—effectiveness and timeliness. For example, with effective drug therapy in the outpatient setting, hospital admissions for hypertension can be prevented. Likewise, accurate diagnosis and timely access to surgical treatment will help reduce the incidence of perforated appendix. The PQIs are thus valuable tools for identifying potential quality problems in outpatient care that help to set the direction for more in-depth investigation. Because the PQIs are based on readily available data—hospital discharge

abstracts—resource requirements are minimal. With uniform definitions and standardized programs, the PQIs will allow comparisons across States, regions, and local communities over time.

Despite the unique strengths of the PQIs, there are several issues that should be considered when using these indicators. First, for some PQIs, differences in socioeconomic status have been shown to explain a substantial part—perhaps most—of the variation in PQI rates across areas. The complexity of the relationship between socioeconomic status and PQI rates makes it difficult to delineate how much of the observed relationships are due to true access to care difficulties in potentially underserved populations, or due to other patient characteristics, unrelated to quality of care, that vary systematically by socioeconomic status. For some of the indicators, patient preferences and hospital capabilities for inpatient or outpatient care might explain variations in hospitalizations. In addition, environmental conditions that are not under the direct control of the health care system can substantially influence some of the PQIs. For example, the COPD and asthma admission rates are likely to be higher in areas with poorer air quality.

Second, the evidence related to potentially avoidable hospital admissions is limited for each indicator, because many of the indicators have been developed as parts of sets. Only five studies have attempted to validate individual indicators rather than whole measure sets.[9 10 11 12 13] A limitation of this literature is that relatively little is known about which components represent the strongest measures of access and quality. Most of the five papers that did report on individual indicators also used a single variable, such as median area-specific income or rural residence, for construct validation. All but one of these papers[10] included adjustment only for demographic factors (e.g., age, sex, and race).

Third, despite the relationships demonstrated at the patient level between higher quality ambulatory care and lower rates of hospital admission, few studies have directly addressed the question of whether effective treatments in outpatient settings would reduce the overall incidence of hospitalizations. The extent to which the reporting of admission rates for ambulatory care sensitive conditions may lead to changes in ambulatory practices and admission rates also is unknown. Providers may admit patients who do not clinically require inpatient care or they may do the opposite—fail to hospitalize patients who would benefit from inpatient care.

Questions for Future Work

The limitations discussed above suggest some directions for future work on development and use of the PQIs. Additional data and linkages could provide insights into the underlying causes of hospitalization for these conditions and could facilitate the exploration of potential interventions to prevent such events.

- Studies examining health and risk behaviors in a population could illuminate patient factors associated with the incidence of ambulatory care sensitive conditions.
- Examining environmental data, such as air pollution levels, could provide insight into factors outside the direct control of the health care system that are associated with hospitalization for such conditions.
- Exploring differences in disease prevalence in specific areas could help to discern whether

[9]Weissman JS, Gatsonis C, Epstein AM. Rates of avoidable hospitalization by insurance status in Massachusetts and Maryland. JAMA 1992;268(17)2388-94.
[10]Bindman AB, Grumbach K, Osmond D, et al. Preventable hospitalizations and access to health care. JAMA 1995;274(4):305-11.
[11]Billings J, Zeital L, Lukomnik J, et al. Analysis of variation in hospital admission rates associated with area income in New York City. Unpublished report.
[12]Silver MP, Babitz ME, Magill MK. Ambulatory care sensitive hospitalization rates in the aged Medicare population in Utah, 1990 to 1994: a rural-urban comparison. J Rural Health 1997;13(4):285-94.
[13]Millman M, editor. Committee on Monitoring Access to Personal Health Care Services. Washington, DC: National Academy Press; 1993.

variations in hospitalization rates can be attributed to differences in disease burden across communities that would exist even with optimum preventive care.

- Studies could examine the relationship between rural-urban location and distance to health care resources and hospital admission for ambulatory care sensitive conditions. Such studies would require information on patients' residence such as patient ZIP codes.
- Linkages with data on local medical resources could help to illuminate the relationship between hospitalization for ACSCs and the supply of medical services and resources, such as the number of primary care and specialty physicians in a community or the supply of hospital beds. For example, the Dartmouth Atlas provides analyses for the Medicare population that suggest that the supply of hospital beds in a community is linked to ambulatory care sensitive admissions, but reported no relationship with local physician supply.[14]
- Physician office data and outpatient clinic data may provide important information regarding care prior to hospital admission. Outpatient data would enable analyses that examine the processes of care that can prevent hospitalizations due to these conditions.
- Combining inpatient data with emergency department data would support the construction of a more complete picture of quality of care related to ambulatory care sensitive conditions. Some of these conditions are seen in emergency departments without being admitted for inpatient care. This is particularly relevant for the uninsured or underinsured who are more likely to use emergency departments as a routine source of care.

[14] *Dartmouth Atlas of Health Care, 1999*. Center for the Evaluative Clinical Sciences at Dartmouth Medical School, 2000.

Detailed Evidence for Prevention Quality Indicators

This section provides an abbreviated presentation of the details of the literature review and the empirical evaluation for each PQI, including:

- The relationship between the indicator and quality of health care services
- A suggested benchmark or comparison
- The definition of each indicator
- The outcome of interest (or numerator)
- The population at risk (or denominator)
- The results of the empirical testing

Empirical testing rated the statistical performance of each indicator, as described in step 5 in the previous section. Scores ranged from 0 to 26 (mean for these 16 PQIs = 14.6), except for low birth weight for which bias was not tested because adequate risk adjustment was not available. The scores are intended as a guide for summarizing the performance of each indicator on four empirical tests of precision (signal variance, area-level share, signal ratio, and R-squared) and five tests of minimum bias (rank correlation, top and bottom decile movement, absolute change, and change over two deciles), as described in the previous section and in Appendix B. Raw unadjusted rates and SD are calculated using 2002 SID from 35 states.[15] These rates are population rates based on all eligible discharges, as opposed to the average area rates reported in Table 1.

The magnitude of the scores, shown under Empirical Rating, provides an indication of the relative rankings of the indicators. These scores were based on indicator performance after risk-adjustment and smoothing, that is, they represent the "best estimate" of the indicator's true value after accounting for case-mix and reliability. The score for each individual test is an ordinal ranking (e.g., very high, high, moderate, and low). The final summary score was derived by assigning a weight to each ranking (e.g., 3, 2, 1, 0) and summing across these nine individual tests. Higher scores indicate better performance on the empirical tests. The two-page descriptions for each indicator also include a discussion of the summary of evidence, the limitations on using each indicator, and details on:

- Face validity – Does the indicator capture an aspect of quality that is widely regarded as important and subject to provider or public health system control?

[15] The state data organizations that participated in the 2002 HCUP SID: California Office of Statewide Health Planning & Development; Colorado Health & Hospital Association; Connecticut - Chime, Inc.; Florida Agency for Health Care Administration; Georgia: An Association of Hospitals & Health Systems; Hawaii Health Information Corporation; Illinois Health Care Cost Containment Council; Iowa Hospital Association; Kansas Hospital Association; Kentucky Department for Public Health; Maine Health Data Organization; Maryland Health Services Cost Review; Massachusetts Division of Health Care Finance and Policy; Michigan Health & Hospital Association; Minnesota Hospital Association; Missouri Hospital Industry Data Institute; Nebraska Hospital Association; Nevada Department of Human Resources; New Jersey Department of Health & Senior Services; New York State Department of Health; North Carolina Department of Health and Human Services; Ohio Hospital Association; Oregon Association of Hospitals & Health Systems; Pennsylvania Health Care Cost Containment Council; Rhode Island Department of Health; South Carolina State Budget & Control Board; South Dakota Association of Healthcare Organizations; Tennessee Hospital Association; Texas Health Care Information Council; Utah Department of Health; Vermont Association of Hospitals and Health Systems; Virginia Health Information; Washington State Department of Health; West Virginia Health Care Authority; Wisconsin Department of Health & Family Services.

- Precision – Is there a substantial amount of provider or community level variation that is not attributable to random variation?
- Minimum bias – Is there either little effect on the indicator of variations in patient disease severity and comorbidities, or is it possible to apply risk adjustment and statistical methods to remove most or all bias?
- Construct validity – Does the indicator perform well in identifying true (or actual) quality-of-care problems?
- Fosters true quality improvement – Is the indicator insulated from perverse incentives for providers to improve their reported performance by avoiding difficult or complex cases, or by other responses that do not improve quality of care?
- Prior use – Has the measure been used effectively in practice? Does it have potential for working well with other indicators?

Summary of Evidence Reported for All or Multiple PQIs

The literature review of the evidence related to potentially avoidable hospital admissions is limited for each indicator because many of the individual indicators have been developed as parts of sets. This section provides a summary of the general evidence reviewed applicable to all PQIs.

- **Precision.** The precision of avoidable hospitalization rates is likely to depend on the size of the denominator.

- **Minimum bias.** Previous studies have documented several characteristics that are associated with either the risk of an avoidable hospitalization (at the individual level) or the avoidable hospitalization rate (at the area level), including prevalence of the condition, race, socioeconomic status (SES), chronic disease and health of the population.[16,17,18] These characteristics may be confounding factors, but also might be measuring subtle aspects of access to care.

- **Construct validity.** Most previous studies have assessed the validity of an entire set of avoidable hospital conditions, rather than each condition alone, and have used SES as a marker of access to care. These studies have repeatedly shown strong correlations between household income and avoidable hospitalizations, both at the individual level and the area level. At the zip code level, income alone explains 51-84% of the variability in ACS admission rates across 15 metropolitan areas in the US.[19] This association is substantially weaker among persons 65 or more years of age,[20,21] as one would expect if it is driven by access to care rather than underlying social factors. Avoidable hospitalization rates are higher among uninsured or Medicaid-enrolled persons than among privately insured persons, even after adjustment for race and income.[22]

---

[16] Bindman AB, Grumbach K, Osmond D, et al. Preventable hospitalizations and access to health care. JAMA 1995;274(4):305-11.
[17] Culler SD, Parchman ML, Przybylski M. Factors related to potentially preventable hospitalizations among the elderly. Med Care 1998;36(6):804-17.
[18] Blustein J, Hanson K, Shea S. Preventable hospitalizations and socioeconomic status. Health Aff (Millwood) 1998;17(2):177-89.
[19] Billings J, Anderson GM, Newman LS. Recent findings on preventable hospitalizations . Health Aff (Millwood) 1996;15(3):239-49.
[20] Billings J, Zeitel L, Lukomnik J, et al. Impact of socioeconomic status on hospital use in New York City. Health Aff (Millwood) 1993;12(1):162-73.
[21] Pappas G, Hadden WC, Kozak LJ, et al. Potentially avoidable hospitalizations: inequalities in rates between US socioeconomic groups. Am J Public Health 1997;87(5):811-6.
[22] Weissman JS, Gatsonis C, Epstein AM. Rates of avoidable hospitalization by insurance status in Massachusetts and Maryland. Jama 1992;268(17):2388-94.

Fewer studies have tested true measures of access to care. In the best of these studies, Bindman and colleagues[16] showed that self-reported "difficulty in receiving medical care when needed" explained 50% of the variability in hospitalization rates for 5 chronic medical conditions (asthma, CHF, COPD, diabetes, and hypertension). Adjustment for condition prevalence, propensity to seek care, physician admitting style, and ecological measures of income, education, insurance, race, and gender, had little effect on the association. Having a regular source of care, and primary care physician/population ratios, were also independently associated with avoidable hospitalization rates, when substituted for self-reported access.[23] These relationships did **not** hold in two separate studies of rural zip codes, suggesting that avoidable hospitalization rates are invalid indicators of access in rural areas.[24,25]

In other studies, the physician/population ratio for family and general physicians has been more strongly associated with avoidable hospitalization rates than measures that include internists, pediatricians, or all physicians.[26,27] In studies of Medicaid populations, provider continuity in ambulatory care[28] and usual care received from a community health center[29] were associated with lower avoidable hospitalization rates, and not having a primary care physician was associated with higher rates of avoidable hospitalization.[30] However, having a regular source of care (for more than 50% of physician office visits) was not associated with lower avoidable hospitalization rates.[31]

Several studies of Medicare beneficiaries have shown weak and inconsistent associations between access indicators and avoidable hospitalization rates. For example, persons in the Medicare Current Beneficiary Survey who reported problems obtaining health care, or lived in a health professional shortage area, were not at increased risk of preventable hospitalization.[17] Instead, their risk was heavily influenced by clinical factors. However, beneficiaries in fair or poor health reportedly were at increased risk if they lived in a **primary care** shortage area.[32] An area-level analysis based on Medicare claims suggests that the association between admission rates and physician/population ratios is limited to the 10% of health care service areas with the most severe shortage of physicians.[33]

A full report on the literature review and empirical evaluation can be found in *Refinement of the HCUP Quality Indicators* by the UCSF-Stanford EPC, available at http://www.qualityindicators.ahrq.gov/. Detailed coding information for each PQI is provided in Appendix A.

---

[23] Komaromy M, Lurie N, Osmond D, et al. Physician practice style and rates of hospitalization for chronic medical conditions. Med Care 1996;34(6):594-609.
[24] Schreiber S, Zielinski T. The meaning of ambulatory care sensitive admissions: urban and rural perspectives. J Rural Health 1997;13(4):276-84.
[25] Bindman A, Grumbach K, Osmond D, et al. Accuracy of preventable hospitalization rates for measuring access to care in rural communities. JGIM 1996;11[Suppl 1]:64.
[26] Parchman ML, Culler S. Primary care physicians and avoidable hospitalizations . J Fam Pract 1994;39(2):123-8.
[27] Epstein A. The role of the medical market in preventable hospitalizations. Abstract Book/Association of Health Services Research 1998;15(316-7).
[28] Gill JM, Mainous AG, 3rd. The role of provider continuity in preventing hospitalizations. Arch Fam Med 1998;7(4):352-7.
[29] Falik M, Needleman J, McCall N, et al. Ambulatory care sensitive conditions: hospitalization rates by usual source of care. Abstract Book/Association for Health Services Research 1998;15:236-7.
[30] Shi L, Samuels ME, Pease M, et al. Patient characteristics associated with hospitalizations for ambulatory care sensitive conditions in South Carolina. Southern Medical Journal 1999;92(10):989-98.
[31] Gill JM. Can hospitalizations be avoided by having a regular source of care? Fam Med 1997;29(3):166-71.
[32] Parchman ML, Culler SD. Preventable hospitalizations in primary care shortage areas. An analysis of vulnerable Medicare beneficiaries. Arch Fam Med 1999;8(6):487-91.
[33] Krakauer H, Jacoby I, Millman M, et al. Physician impact on hospital admission and on mortality rates in the Medicare population. Health Serv Res 1996;31(2):191-211.

Diabetes Short-term Complications Admission Rate (PQI 1)

Short-term complications of diabetes mellitus include diabetic ketoacidosis, hyperosmolarity, and coma. These life-threatening emergencies arise when a patient experiences an excess of glucose (hyperglycemia) or insulin (hypoglycemia).

Relationship to Quality	Proper outpatient treatment and adherence to care may reduce the incidence of diabetic short-term complications, and lower rates represent better quality care.
Benchmark	State, regional, or peer group average.
Definition	Admissions for diabetic short-term complications per 100,000 population.
Outcome of Interest	Discharges with ICD-9-CM principal diagnosis codes for diabetes short-term complications (ketoacidosis, hyperosmolarity, coma). Age 18 years and older. Exclude patients transferring from another institution, MDC 14 (pregnancy, childbirth, and puerperium), or MDC 15 (newborns and neonates).
Population at Risk	Population in MSA or county, age 18 years and older.
Empirical Results and Rating	Rate (2002): 50.9 per 100,000 population Rating: 14

Summary of Evidence

Hospital admission for diabetes short-term complications is a PQI that would be of most interest to comprehensive health care delivery systems. Short-term diabetic emergencies arise from the imbalance of glucose and insulin, which can result from deviations in proper care, misadministration of insulin, or failure to follow a proper diet.

Although risk adjustment with age and sex does not impact the relative or absolute performance of areas, this indicator should be risk-adjusted. Some areas may have higher rates of diabetes as a result of racial composition and systematic differences in other risk factors.

Areas with high rates of diabetic emergencies may want to examine education practices, access to care, and other potential causes of non-compliance when interpreting this indicator. Also, areas may consider examining the rates of hyperglycemic versus hypoglycemic events when interpreting this indicator.

Limitations on Use

As a PQI, short-term diabetes complication rate is not a measure of hospital quality, but rather one measure of outpatient and other health care. Rates of diabetes may vary systematically by area, creating bias for this indicator. Examination of both inpatient and outpatient data may provide a more complete picture of diabetes care.

Details

Face validity: Does the indicator capture an aspect of quality that is widely regarded as important and subject to provider or public health system control?

High-quality outpatient management of patients with diabetes has been shown to lead to reductions in almost all types of serious avoidable hospitalizations. However, tight control may be associated with more episodes of hypoglycemia, which leads to more admissions.

Precision: Is there a substantial amount of provider or community level variation that is not attributable to random variation?

Based on empirical evidence, this indicator is moderately precise, with a raw area level rate of 36 per 100,000 population and a standard deviation of 24.6.

The signal ratio (i.e., the proportion of the total variation across areas that is truly related to systematic differences in area performance rather than random variation) is moderate, at 51.7%, indicating that some of the observed differences in age-sex adjusted rates do not represent true differences in area performance. Using multivariate signal extraction techniques appears to have little additional impact on estimating true differences across areas.

Minimum bias: Is there either little effect on the indicator of variations in patient disease severity and comorbidities, or is it possible to apply risk adjustment and statistical methods to remove most or all bias?

Minorities have higher rates of diabetes, and higher hospitalization rates may result in areas with higher minority concentrations. Empirical results show that area rankings and absolute performance are not affected by age-sex risk adjustment.

Construct validity: Does the indicator perform well in identifying true (or actual) quality-of-care problems?

Studies of precipitating events of admission for diabetic emergencies often rely on self-report, which may be a biased measurement in and of itself. The results of one study showed that over 60% of patients with known and treated diabetes had made an error in insulin administration or had omitted insulin.[34] In a potentially under-served population of urban African-Americans, two-thirds of admissions were due to cessation of insulin therapy—over half of the time for financial or other difficulties obtaining insulin.[35]

Bindman reported that an area's self-rated access to care report explained 46% of the variance in admissions for diabetes, although the analysis was not restricted to diabetic emergencies.[36] Weissman found that uninsured patients had more than twice the risk of admission for diabetic ketoacidosis and coma than privately insured patients.[37]

Fosters true quality improvement: Is the indicator insulated from perverse incentives for providers to improve their reported performance by avoiding difficult or complex cases, or by other responses that do not improve quality of care?

Because diabetic emergencies are potentially life-threatening, hospitals are unlikely to fail to admit patients requiring hospitalization.

Prior use: Has the measure been used effectively in practice? Does it have potential for working well with other indicators?

Admission for diabetic emergencies was included in both Billings'[38] and Weissman's[39] sets of avoidable hospitalization measures. This indicator, defined as a provider-level indicator, was an original HCUP QI.

[34]Bagg W, Sathu A, Streat S, et al. Diabetic ketoacidosis in adults at Auckland Hospital, 1988-1996. Aust N Z J Med 1998;28(5):604-8.
[35]Musey VC, Lee JK, Crawford R, et al. Diabetes in urban African-Americans. I. Cessation of insulin therapy is the major precipitating cause of diabetic ketoacidosis. Diabetes Care 1995;18(4):483-9.
[36]Bindman AB, Grumbach K, Osmond D, et al. Preventable hospitalizations and access to health care. JAMA 1995;274(4):305-11.
[37]Weissman JS, Gatsonis C, Epstein AM. Rates of avoidable hospitalization by insurance status in Massachusetts and Maryland. JAMA 1992;268(17)2388-94.
[38]Billings J, Zeital L, Lukomnik J, et al. Analysis of variation in hospital admission rates associated with area income in New York City. Unpublished report.
[39]Weissman, et al., 1992.

Perforated Appendix Admission Rate (PQI 2)

Perforated appendix may occur when appropriate treatment for acute appendicitis is delayed for a number of reasons, including problems with access to care, failure by the patient to interpret symptoms as important, and misdiagnosis and other delays in obtaining surgery.

Relationship to Quality	Timely diagnosis and treatment may reduce the incidence of perforated appendix, and lower rates represent better quality care.
Benchmark	State, regional, or peer group average.
Definition	Admissions for perforated appendix per 100 admissions for appendicitis within MSA or county.
Outcome of Interest	Discharges with ICD-9-CM diagnosis code for perforation or abscess of appendix in any field. Exclude patients transferring from another institution, MDC 14 (pregnancy, childbirth, and puerperium), or MDC 15 (newborns and neonates).
Population at Risk	Discharges with diagnosis code for appendicitis in any field within MSA or county.
Empirical Results and Rating	Rate (2002): 30.5 per 100 eligible discharges Rating: 17 (Smoothing recommended)

Summary of Evidence

Hospital admission for perforated appendix is a PQI that would be of most interest to comprehensive health care delivery systems. With prompt and appropriate care, acute appendicitis should not progress to perforation or rupture. Rates for perforated appendix are higher in the uninsured or underinsured in both adult and pediatric populations, which may be caused by patients failing to seek appropriate care, difficulty in accessing care, or misdiagnoses and poor quality care.

Perforated appendix rates vary systematically by race, although the cause is unknown. Areas with high rates of perforated appendix may want to target points of intervention by using chart reviews and other supplemental data to investigate the reasons for delay in receiving surgery. Hospital contributions to the overall area rate may be particularly useful for this indicator, because misdiagnoses and other delays in receiving surgery in an emergency room may contribute substantially to the rate.

Limitations on Use

As a PQI, admission for perforated appendix is not a measure of hospital quality, but rather one measure of outpatient and other health care.

Details

Face validity: Does the indicator capture an aspect of quality that is widely regarded as important and subject to provider or public health system control?

Perforated appendix results from delay in surgery, potentially reflecting problems in access to ambulatory care, misdiagnosis, and other delays in obtaining surgery.

Precision: Is there a substantial amount of provider or community level variation that is not attributable to random variation?

Perforated appendix occurs in one-fourth to one-third of hospitalized acute appendicitis patients.[40] Based on empirical evidence, this indicator is precise, with a raw area level rate of 33.3% and a substantial standard deviation of 14.4%.

Relative to other indicators, a higher percentage of the variation occurs at the area level rather than the discharge level. However, the signal ratio (i.e., the proportion of the total variation across areas that is truly related to systematic differences in area performance rather than

[40] Braveman P, Schaaf VM, Egerter S, et al. Insurance-related differences in the risk of ruptured appendix [see comments]. N Engl J Med 1994;331(7):444-9.

random variation) is low, at 26.5%, indicating that much of the observed differences in age-sex adjusted rates likely do not represent true differences across areas. Applying multivariate signal extraction methods can improve estimation of true differences in area performance.

Minimum bias: Is there either little effect on the indicator of variations in patient disease severity and comorbidities, or is it possible to apply risk adjustment and statistical methods to remove most or all bias?

Higher rates of perforated appendix have been noted in males, patients with mental illness or substance abuse disorders, people with diabetes, and blacks,[41] as well as in children under the age of 4 (although appendicitis is rare in this age group).[42]

Some of the observed variation in performance is due to systematic differences in patient characteristics. No evidence exists in the literature that clinical characteristics that would vary systematically increase the likelihood of perforated appendix. Therefore, this indicator is unlikely to be clinically biased. Empirical results show that area rankings and absolute performance are not affected by age-sex risk adjustment.

Construct validity: Does the indicator perform well in identifying true (or actual) quality-of-care problems?

Braveman et al. found that the rate of perforated appendix was 50% higher for patients with no insurance or Medicaid than HMO-covered patients, and 20% higher for patients with private fee-for-service insurance. A follow-up study by Blumberg et al. concluded that the high rate of perforated appendix in the black population at an HMO may be explained by delay in seeking care, rather than differences in the quality of health care.[43] Weissman et al. found that uninsured (but not Medicaid) patients are at increased risk for ruptured appendix after adjusting for age and sex.[44]

Based on empirical results, areas with high rates of perforated appendix admissions tend to have lower rates of admissions for other ACSCs.

Fosters true quality improvement: Is the indicator insulated from perverse incentives for providers to improve their reported performance by avoiding difficult or complex cases, or by other responses that do not improve quality of care?

Use of this quality indicator might lead to more performance of appendectomies in cases of questionable symptoms, in addition to reducing the occurrence of rupture.

Prior use: Has the measure been used effectively in practice? Does it have potential for working well with other indicators?

Perforated appendix was included in the original HCUP QI indicator set, as well as in Weissman's set of avoidable hospitalizations.

[41] Braveman et al., 1994.
[42] Bratton SL, Haberkern CM, Waldhausen JH. Acute appendicitis risks of complications: age and Medicaid insurance. Pediatrics 2000;106(1 Pt 1):75-8.
[43] Blumberg MS, Juhn PI. Insurance and the risk of ruptured appendix [letter; comment]. N Engl J Med 1995;332(6):395-6; discussion 397-8.
[44] Weissman JS, Gatsonis C, Epstein AM. Rates of avoidable hospitalization by insurance status in Massachusetts and Maryland. JAMA 1992;268(17)2388-94.

Diabetes Long-term Complications Admission Rate (PQI 3)

Long-term complications of diabetes mellitus include renal, eye, neurological, and circulatory disorders. Long-term complications occur at some time in the majority of patients with diabetes to some degree.

Relationship to Quality	Proper outpatient treatment and adherence to care may reduce the incidence of diabetic long-term complications, and lower rates represent better quality care.
Benchmark	State, regional, or peer group average.
Definition	Admissions for diabetic long-term complications per 100,000 population.
Outcome of Interest	Discharges with ICD-9-CM principal diagnosis codes for long-term complications of diabetes (renal, eye, neurological, circulatory, or complications not otherwise specified). Age 18 years and older. Exclude patients transferring from another institution, MDC 14 (pregnancy, childbirth, and puerperium), or MDC 15 (newborns and neonates).
Population at Risk	Population in MSA or county, age 18 years and older.
Empirical Results and Rating	Rate (2002): 113.8 per 100,000 population Rating: 11

Summary of Evidence

Hospital admission for diabetes long-term complications is a PQI that would be of most interest to comprehensive health care delivery systems. Long-term diabetes complications are thought to arise from sustained long-term poor control of diabetes. Intensive treatment programs have been shown to decrease the incidence of long-term complications in both Type 1 and Type 2 diabetes.

Sociodemographic characteristics of the population, such as race, may bias the indicator, since Native Americans and Hispanic Americans have higher rates of diabetes and poorer glycemic control. The importance of these factors as they relate to admission rates is unknown. Risk adjustment for observable characteristics, such as racial composition of the population, is recommended.

It is unclear whether poor glycemic control arises from poor quality medical care, non-compliance of patients, lack of education, or access to care problems. Areas with high rates may wish to examine these factors when interpreting this indicator.

Limitations on Use

As a PQI, diabetes long-term complication rate is not a measure of hospital quality, but rather one measure of outpatient and other health care. Rates of diabetes may vary systematically by area, creating bias for this indicator. Examination of both inpatient and outpatient data may provide a more complete picture of diabetes care.

Details

Face validity: Does the indicator capture an aspect of quality that is widely regarded as important and subject to provider or public health system control?

Several observational studies have linked improved glycemic control to substantially lower risks of developing complications in both Type 1 and Type 2 diabetes.[45] Given that appropriate adherence to therapy and consistent monitoring of glycemic control help to prevent complications, high-quality outpatient care should lower long-term complication rates. However, adherence to guidelines aimed at

[45] Gaster B, Hirsch IB. The effects of improved glycemic control on complications in type 2 diabetes. Arch Intern Med 1998;158(2):134-40.

reducing complications (including eye and foot examinations and diabetic education) has been described as modest,[46] with only one-third of patients receiving all essential services.[47]

*Precision: Is there a substantial amount of provider or community level variation tha*t is not attributable to random variation?

Diabetes affects a large number of people, as do diabetic complications. However, few studies have documented hospitalization rates for diabetic complications and the extent to which they vary across areas. Based on empirical evidence, this indicator is moderately precise, with a raw area level rate of 80.8 per 100,000 population and a standard deviation of 58.1.

The signal ratio (i.e., the proportion of the total variation across areas that is truly related to systematic differences in area performance rather than random variation) is high, at 75.6%, indicating that the observed differences in age-sex adjusted rates likely represent true differences across areas.

Minimum bias: Is there either little effect on the indicator of variations in patient disease severity and comorbidities, or is it possible to apply risk adjustment and statistical methods to remove most or all bias?

Rates of diabetes are higher in black, Hispanic, and especially Native American populations than in other ethnic groups. Hyperglycemia appears to be particularly frequent among Hispanic and Native American populations.[48] The duration of diabetes is positively associated with the development of complications. Empirical results show that area rankings and absolute performance are moderately affected by age-sex risk adjustment.

Construct validity: Does the indicator perform well in identifying true (or actual) quality-of-care problems?

Compliance of physicians and patients is essential to achieve good outcomes, and it seems likely that problems with both access to and quality of care, as well as patient compliance, may contribute to the occurrence of complications.

Based on empirical results, areas with high rates of diabetes long-term complications also tend to have high rates of admission for other ACSCs.

Fosters true quality improvement: Is the indicator insulated from perverse incentives for providers to improve their reported performance by avoiding difficult or complex cases, or by other responses that do not improve quality of care?

Providers may decrease admission rates by failing to hospitalize patients who would truly benefit from inpatient care. No published evidence indicates that worse health outcomes are associated with reduced hospitalization rates for long-term complications of diabetes.

Prior use: Has the measure been used effectively in practice? Does it have potential for working well with other indicators?

This indicator, defined as a hospital-level indicator, is an original HCUP QI.

---

[46]Zoorob RJ, Hagen MD. Guidelines on the care of diabetic nephropathy, retinopathy and foot disease. Am Fam Physician 1997;56(8):2021-8, 2033-4.
[47]Hiss RG. Barriers to care in non-insulin-dependent diabetes mellitus. The Michigan Experience. Ann Intern Med 1996;124(1 Pt 2):146-8.
[48]Harris MI. Diabetes in America: epidemiology and scope of the problem. Diabetes Care 1998;21 Suppl 3:C11-4.

Pediatric Asthma Admission Rate (PQI 4)

Asthma is the most common chronic disease in childhood and is one of the most frequent admitting diagnoses in children's hospitals. Most published studies combine admission rates for children and adults; therefore, areas may wish to examine this indicator together with the adult asthma indicator.

Relationship to Quality	Proper outpatient treatment may reduce admissions for asthma in the pediatric population, and lower rates represent better quality care.
Benchmark	State, regional, or peer group average.
Definition	Admissions for pediatric asthma per 100,000 population.
Outcome of Interest	Discharges with ICD-9-CM principal diagnosis codes for asthma. Age less than 18 years old. Exclude patients transferring from another institution, MDC 14 (pregnancy, childbirth, and puerperium), or MDC 15 (newborns and neonates).
Population at Risk	Population in MSA or county, age less than 18 years.
Empirical Results and Rating	Rate (2002): 188.8 per 100,000 population Rating: 18

Summary of Evidence

Hospital admission for pediatric asthma is a PQI that would be of most interest to comprehensive health care delivery systems.

Healthy People 2010 has set a goal to reduce the admission rate for asthma to 2.5 per 10,000 population for children under 5 years, and 7.7 per 10,000 population for people ages 5-65 years.[49] Adherence to the guidelines for asthma management has been associated with lower admission rates.

This indicator is measured with high precision, and the observed variance reflects true differences in area performance. Risk adjustment for age and sex does not appear to affect area rankings. A review of the literature indicates that some children may be at risk for admission due to comorbidities, genetic factors, and environmental triggers. It is unclear which of these factors would vary by area, nor is the impact of parental compliance well understood. Race should be adjusted for in comparing rates across areas.

[49]Healthy People 2010. Office of Disease Prevention and Health Promotion, U.S. Department of Health and Human Services.

Limitations on Use

As a PQI, pediatric asthma is not a measure of hospital quality, but rather one measure of outpatient and other health care.

Providers may reduce admission rates without actually improving quality by shifting care to an outpatient setting. Admission rates that are drastically below or above the average or recommended rates should be examined.

Details

Face validity: Does the indicator capture an aspect of quality that is widely regarded as important and subject to provider or public health system control?

In the United States, asthma affects an estimated 4.8 million children and adolescents, and in 1993, it was the cause of 198,000 admissions and 342 deaths in persons aged 24 and younger.[50] Adherence to the treatment guidelines—which emphasize appropriate diagnosis of asthma, a physician-patient relationship, management of asthma symptoms with medications, appropriate prophylactic and maintenance therapy, and adequate follow-up care—can reduce admission rates.

[50]CDC. Asthma mortality and hospitalization among children and young adults—United States, 1980-1993. MMWR Morb Mortal Wkly Rep 1996;45(17):350-3.

Precision: Is there a substantial amount of provider or community level variation that is not attributable to random variation?

Because asthma is one of the most common reasons for pediatric hospitalization, relatively precise estimates of asthma admission across areas or hospitals can be obtained. Admission rates for asthma tend to be higher during peak times of viral respiratory infections (winter) and allergy seasons (spring and fall), so a consistent time period for measurement must be ensured. Based on empirical evidence, this indicator is precise, with a raw level rate of 154.1 and a standard deviation of 143.9.

The signal ratio (i.e., the proportion of the total variation across areas that is truly related to systematic differences in area performance rather than random variation) is high, at 85.1%, indicating that the observed differences in age-sex adjusted rates likely represent true differences across areas.

Minimum bias: Is there either little effect on the indicator of variations in patient disease severity and comorbidities, or is it possible to apply risk adjustment and statistical methods to remove most or all bias?

Environmental triggers for pediatric asthma include indoor allergens such as tobacco smoke[51] and outdoor air pollution.[52] Race represents one of the most complex potentially biasing factors. Black patients have been shown to have higher asthma admission rates, even when stratifying for income and age.[53] Adjustment for race is recommended. Empirical results show that area rankings are not affected by age-sex risk adjustment.

Construct validity: Does the indicator perform well in identifying true (or actual) quality-of-care problems?

Some admissions with asthma are unavoidable and appropriate. Studies have shown that asthma hospitalization rates are associated with median household income (at the area level) and lack of insurance (at the individual level). Lin et al. showed that admission rates were higher in areas with higher poverty, minority populations, unemployment, and lower education levels.[54]

Fosters true quality improvement: Is the indicator insulated from perverse incentives for providers to improve their reported performance by avoiding difficult or complex cases, or by other responses that do not improve quality of care?

Because some pediatric asthma can be managed on an outpatient basis, an appropriate shift to outpatient care may occur. Providers may decrease their rates by failing to hospitalize patients who would benefit from inpatient care.

Prior use: Has the measure been used effectively in practice? Does it have potential for working well with other indicators?

This indicator was originally developed by Billings et al. in conjunction with the United Hospital Fund of New York.[55] It was adopted by the Institute of Medicine and has been widely used in studies of avoidable hospitalizations.

[51]National Heart, Lung, and Blood Institute/National Asthma Education and Prevention Program. Expert Panel Report 2: Guidelines for the diagnosis and management of asthma. In: NIH pub. no. 97-4051. Bethesda, MD; 1997.
[52]NHLBI/NAEPP, 1997.
[53]Ray NF, Thamer M, Fadillioglu B, et al. Race, income, urbanicity, and asthma hospitalization in California: a small area analysis. Chest 1998;113(5):1277-84.

[54]Lin, S, Fitzgerald E, Hwang SA, et al. Asthma hospitalization rates and socioeconomic status in New York State (1987-1993) J Asthma 1999;36(3):239-51.
[55]Billings, J. Zeital L, Lukomnik J, et al. Analysis of variation in hospital admission rates associated with area income in New York City. Unpublished report.

Chronic Obstructive Pulmonary Disease Admission Rate (PQI 5)

Chronic obstructive pulmonary disease (COPD) comprises three primary diseases that cause respiratory dysfunction—asthma, emphysema, and chronic bronchitis—each with distinct etiologies, treatments, and outcomes. This indicator examines emphysema and bronchitis; asthma is discussed separately for children and adults.

Relationship to Quality	Proper outpatient treatment may reduce admissions for COPD, and lower rates represent better quality care.
Benchmark	State, regional, or peer group average.
Definition	Admissions for COPD per 100,000 population.
Outcome of Interest	Discharges with ICD-9-CM principal diagnosis codes for COPD. Age 18 years and older. Exclude patients transferring from another institution, MDC 14 (pregnancy, childbirth, and puerperium), or MDC 15 (newborns and neonates).
Population at Risk	Population in MSA or county, age 18 years and older.
Empirical Results and Rating	Rate (2002): 248.6 per 100,000 population Rating: 17

Summary of Evidence

Hospital admission for COPD is a PQI that would be of most interest to comprehensive health care delivery systems. COPD can often be controlled in an outpatient setting. Areas may wish to use chart reviews to understand more clearly whether admissions are a result of poor quality care or other problems.

This indicator is measured with high precision, and the observed variance likely reflects true differences across areas. Risk adjustment for age and sex appears to most affect the areas with the highest rates. Several factors that are likely to vary by area may influence the progression of the disease, including smoking and socioeconomic status. Risk adjustment for observable characteristics is recommended.

Areas may wish to identify hospitals that contribute the most to the overall area rate for this indicator. The patient populations served by these hospitals may be a starting point for interventions.

Limitations on Use

As a PQI, COPD is not a measure of hospital quality, but rather one measure of outpatient and other health care. This indicator has unclear construct validity, because it has not been validated except as part of a set of indicators.

Providers may reduce admission rates without actually improving quality by shifting care to an outpatient setting. Some COPD care takes place in emergency rooms, so combining inpatient and emergency room data may give a more accurate picture.

Details

Face validity: Does the indicator capture an aspect of quality that is widely regarded as important and subject to provider or public health system control?

Admissions for COPD include exacerbations of COPD, respiratory failure, and (rarely) lung volume reduction surgery or lung transplantation. Practice guidelines for COPD have been developed and published over the last decade.[56] With appropriate outpatient treatment and compliance, hospitalizations for the exacerbations of COPD and decline in lung function should be minimized.

Precision: Is there a substantial amount of provider or community level variation that is not attributable to random variation?

[56] Hackner D, Tu G, Weingarten S, et al. Guidelines in pulmonary medicine: a 25-year profile. Chest 1999;116(4):1046-62.

COPD accounts for a substantial number of hospital admissions, suggesting that the indicator is reasonably precise.[57] Based on empirical evidence, this indicator is very precise, with a raw area level rate of 324.0 per 100,000 population and a standard deviation of 203.8.

The signal ratio (i.e., the proportion of the total variation across areas that is truly related to systematic differences in area performance rather than random variation) is very high, at 93.4%, indicating that the differences in age-sex adjusted rates likely represent true differences across areas.

Minimal bias: Is there either little effect on the indicator of variations in patient disease severity and comorbidities, or is it possible to apply risk adjustment and statistical methods to remove most or all bias?

Factors that have been associated with increased admissions for COPD include disease severity, smoking status, age, and socioeconomic status, which are candidates for risk adjustment. Empirical results show that area rankings and absolute performance are somewhat affected by age-sex risk adjustment.

Construct validity: Does the indicator perform well in identifying true (or actual) quality-of-care problems?

Bindman et al. reported that self-reported access to care explained 27% of the variation in COPD hospitalization rates at the ZIP code cluster level.[58] Millman et al. found that low-income ZIP codes had 5.8 times more COPD hospitalizations per capita than high-income ZIP codes.[59] Physician adherence to practice guidelines and patient compliance also influence the effectiveness of therapy.

Based on empirical results, areas with high rates of COPD admissions also tend to have high rates of admissions for other ACSCs.

Fosters true quality improvement: Is the indicator insulated from perverse incentives for providers to improve their reported performance by avoiding difficult or complex cases, or by other responses that do not improve quality of care?

One study found that higher rates of COPD admission may in part reflect improvements in access to care, which results in more detection of significant respiratory impairment in the community.[60] A decline in COPD admission rates may simply reflect a reverse change in coding practices.

Prior use: Has the measure been used effectively in practice? Does it have potential for working well with other indicators?

This indicator was originally developed by Billings et al. in conjunction with the United Hospital Fund of New York.[61] It was subsequently adopted by the Institute of Medicine and has been widely used in studies of avoidable hospitalizations.

[57]Feinleib M, Rosenberg HM, Collins JG, et al. Trends in COPD morbidity and mortality in the United States. Am Rev Respir Dis 1989;140(3 pt 2):S9-18.
[58]Bindman AB, Grumbach K, Osmond D, et al. Preventable hospitalizations and access to health care. JAMA 1995;274(4)305-11.
[59]Millman M, editor. Committee on Monitoring Access to Personal Health Care Services. Washington, DC: National Academy Press; 1993.
[60]Weinberger M, Oddone EZ, Henderson WG. Does increased access to primary care reduce hospital readmissions? VA Cooperative Study Group on Primary Care and Hospital readmission. N Engl J Med 1996;334(22):1441-7.
[61]Billings J, Zeital L, Lukomnik J, et al. Analysis of variation in hospital admission rates associated with area income in New York City. Unpublished report.

Pediatric Gastroenteritis Admission Rate (PQI 6)

Pediatric gastroenteritis, which is one of the most common reasons for pediatric hospitalizations, can be treated on an outpatient basis.

Relationship to Quality	Proper outpatient treatment may reduce admissions for gastroenteritis in the pediatric population, and lower rates represent better quality care.
Benchmark	State, regional, or peer group average.
Definition	Admissions for pediatric gastroenteritis per 100,000 population.
Outcome of Interest	Discharges with ICD-9-CM principal diagnosis codes for gastroenteritis. Age less than 18 years. Exclude patients transferring from another institution, MDC 14 (pregnancy, childbirth, and puerperium), or MDC 15 (newborns and neonates).
Population at Risk	Population in MSA or county, age less than 18 years.
Empirical Results and Rating	Rate (2002): 87.7 per 100,000 population Rating: 17

Summary of Evidence

Hospital admission for pediatric gastroenteritis is a PQI that would be of most interest to comprehensive health care delivery systems. Gastroenteritis accounts for nearly 10% of all admissions of children under 5 years of age.[62] This indicator is measured with good precision, and most of the observed variation reflects true differences across areas.

Admissions may be precipitated by poor quality care, lack of compliance with care, and poor access to care, or may be due to environmental causes. Clear guidelines have been published by the Centers for Disease Control and Prevention and the American Academy of Pediatrics; however, there is little compelling evidence that adherence to these guidelines reduces admission rates because many admissions appear to be discretionary and inappropriate. Areas with high rates may want to identify disease severity by looking at the degree of dehydration of patients and comorbidities to establish whether or not admissions are discretionary, appropriate, or due to poor quality care.

Limitations on Use

As a PQI, admission for pediatric gastroenteritis is not a measure of hospital quality, but rather one measure of outpatient and other health care. This indicator has unclear construct validity, because it has not been validated except as part of a set of indicators. Providers may reduce admission rates without actually improving quality by shifting care to an outpatient setting.

Details

Face validity: Does the indicator capture an aspect of quality that is widely regarded as important and subject to provider or public health system control?

Gastroenteritis is a common illness in childhood. Treatment guidelines emphasize the importance of appropriate rehydration therapy for mild to moderate dehydration resulting from gastroenteritis to avoid the need for hospitalization. A physician panel agreed that timely and effective ambulatory care would reduce the risk of hospitalization for gastroenteritis.[63]

[62] Burkhart DM. Management of acute gastroenteritis in children. American Family Physician 1999;60(9):2555-63, 2565-6.

[63] Billings J, Zeitel L, Lukomnik J, et al. Impact of socioeconomic status on hospital use in New York City, Health Aff (Millwood) 1993;12(1):162-73.

Precision: Is there a substantial amount of provider or community level variation that is not attributable to random variation?

Relatively precise estimates of gastroenteritis admission across areas or hospitals can be obtained. Gastroenteritis varies seasonally, so care must be taken to ensure a consistent time period for measurement. The wide variation across areas in admission rates may cause random variation in a particular year to be considerable for less populated areas.

The signal ratio (i.e., the proportion of the total variation across areas that is truly related to systematic differences in area performance rather than random variation) is high, at 77.8%, indicating that the observed differences in age-sex-adjusted rates likely represent true differences across areas.

Minimum bias: Is there either little effect on the indicator of variations in patient disease severity and comorbidities, or is it possible to apply risk adjustment and statistical methods to remove most or all bias?

Some admissions for gastroenteritis are unavoidable. However, most children admitted with gastroenteritis appear to have no underlying problems (70%), and most are rehydrated within 12 hours (79%). One study suggests that complicated gastroenteritis admissions may be more common among children of low socioeconomic status.[64] Empirical results show that area rankings are not affected by age-sex risk adjustment.

Construct validity: Does the indicator perform well in identifying true (or actual) quality-of-care problems?

No published studies have specifically addressed the construct validity of this indicator. Millman et al. reported that low-income ZIP codes had 1.9 times more pediatric gastroenteritis hospitalizations per capita than high-income ZIP codes.[65]

Based on empirical results, areas with high rates of pediatric gastroenteritis admissions also tend to have high rates of admissions for other ACSCs.

Fosters true quality improvement: Is the indicator insulated from perverse incentives for providers to improve their reported performance by avoiding difficult or complex cases, or by other responses that do not improve quality of care?

Because the optimal hospitalization rate for this condition has not been defined, providers may decrease their rates by failing to hospitalize patients who would truly benefit from inpatient care or by hospitalizing marginally appropriate patients with other concomitant conditions.

Prior use: Has the measure been used effectively in practice? Does it have potential for working well with other indicators?

This indicator was originally developed by Billings et al. in conjunction with the United Hospital Fund of New York. It was subsequently adopted by the Institute of Medicine and has been widely used in a variety of studies of preventable hospitalizations.

---

[64] McConnochie KM, Russo MJ, McBride JT, et al. Socioeconomic variation in asthma hospitalization: excess utilization or greater need? Pediatrics 1999;103(6):375.

[65] Millman M, editor. Committee on Monitoring Access to Personal Health Care Services. Washington, D.C.: National Academy Press. 1993.

Hypertension Admission Rate (PQI 7)

Hypertension is a chronic condition that is often controllable in an outpatient setting with appropriate use of drug therapy.

Relationship to Quality	Proper outpatient treatment may reduce admissions for hypertension, and lower rates represent better quality care.
Benchmark	State, regional, or peer group average.
Definition	Admissions for hypertension per 100,000 population.
Outcome of Interest	Discharges with ICD-9-CM principal diagnosis codes for hypertension. Age 18 years and older. Exclude discharges with specified cardiac procedure codes in any field, patients transferring from another institution, MDC 14 (pregnancy, childbirth, and puerperium), or MDC 15 (newborns and neonates).
Population at Risk	Population in MSA or county, age 18 years and older.
Empirical Results and Rating	Rate (2002): 44.4 per 100,000 population Rating: 14

Summary of Evidence

Hospital admission for hypertension is a PQI that would be of most interest to comprehensive health care delivery systems. Little evidence exists regarding the validity of this indicator, although one study did relate admission rates to access to care problems. This indicator is measured with adequate precision, but some of the variance in age-sex adjusted rates does not reflect true differences in area performance. Adjustment for age-sex is recommended.

Areas may wish to identify hospitals that contribute the most to the overall area rate for this indicator. The patient populations served by these hospitals may be a starting point for interventions.

Limitations on Use

As a PQI, hypertension is not a measure of hospital quality, but rather one measure of outpatient and other health care. Providers may reduce admission rates without actually improving quality by shifting care to an outpatient setting.

Details

Face validity: Does the indicator capture an aspect of quality that is widely regarded as important and subject to provider or public health system control?

Hypertension is often controllable in an outpatient setting with appropriate use of drug therapy.

Precision: Is there a substantial amount of provider or community level variation that is not attributable to random variation?

Although hypertension is a common condition, hospitalizations for complications of hypertension are relatively uncommon. One study noted that hypertension accounted for only 0.5% of total admissions for ACSCs.[66]

Based on empirical evidence, this indicator is moderately precise, with a raw area level rate of 37.1 per 100,000 population and a substantial standard deviation of 32.2. The signal ratio (i.e., the proportion of the total variation across areas that is truly related to systematic differences in area performance rather than random variation) is moderate, at 69.9%, indicating that some of the observed differences in age-sex adjusted rates likely do not represent true differences in area performance.

Minimum bias: Is there either little effect on the indicator of variations in patient disease severity and comorbidities, or is it possible to apply risk

[66]Blustein J. Hanson K, Shea S. Preventable hospitalizations and socioeconomic status. Health Aff (Millwood) 1998;17(2):177-89.

adjustment and statistical methods to remove most or all bias?

Little evidence exists on potential biases for this indicator. The age structure of the population may possibly affect admission rates for this condition. Weissman et al. reported a reduction of 100% in relative risk for Medicaid patients when adjusting for age and sex.[67] No evidence was found on the effects of comorbidities such as obesity or other risk factors that may vary systematically by area on admission rates for hypertension complications in the area. Empirical results show that age-sex adjustment affects the ranking of those areas in the highest decile.

Construct validity: Does the indicator perform well in identifying true (or actual) quality-of-care problems?

Bindman et al. found that an area's self-rated access to care explained 22% of admissions for hypertension.[68] Weissman et al. found that uninsured patients had a relative risk of admission for hypertension of 2.38 in Massachusetts after adjustment for age and sex, while Maryland had a corresponding relative risk of 1.93.[69] Millman et al. reported that low-income ZIP codes had 7.6 times more hypertension hospitalizations per capita than high-income ZIP codes.[70]

Fosters true quality improvement: Is the indicator insulated from perverse incentives for providers to improve their reported performance by avoiding difficult or complex cases, or by other responses that do not improve quality of care?

Little evidence exists on the impact of this quality improvement measure on the delivery of outpatient care for hypertension. There is no published evidence of worse health outcomes in association with reduced hospitalization rates for hypertension. Such an effect seems implausible, given that only the most serious episodes of accelerated or malignant hypertension are treated on an inpatient basis.

Prior use: Has the indicator been used effectively in practice? Does it have potential for working well with other indicators?

This indicator was included originally developed by Billings et al. in conjunction with the United Hospital Fund of New York.[71] It was subsequently adopted by the Institute of Medicine and has been widely used in a variety of studies of avoidable or preventable hospitalizations.[72] This indicator was also included in Weissman's set of avoidable hospitalizations.

[67]Weissman JS, Gatsonis C, Epstein AM. Rates of avoidable hospitalization by insurance status in Massachusetts and Maryland. JAMA 1992;268(1):2388-94.
[68]Bindman AB, Grumback K, Osmond D, et al. Preventable hospitalizations and access to health care. JAMA 1995;274(4):305-11.
[69]Weissman, et al. 1992.
[70]Millman M, editor. Committee on Monitoring Access to Personal Health Care Services. Washington, DC: National Academy Press; 1993.

[71]Billings J, Zeitel L, Lukomnik J, et al. Impact of socioeconomic status on hospital use in New York City. Health Aff (Millwood) 1993;12(1):162-73.
[72]Access to Health Care in America. Washington, DC: National Academy Press; 1993.

Congestive Heart Failure Admission Rate (PQI 8)

Congestive heart failure (CHF) can be controlled in an outpatient setting for the most part; however, the disease is a chronic progressive disorder for which some hospitalizations are appropriate.

Relationship to Quality	Proper outpatient treatment may reduce admissions for CHF, and lower rates represent better quality care.
Benchmark	State, regional, or peer group average.
Definition	Admissions for CHF per 100,000 population.
Outcome of Interest	Discharges with ICD-9-CM principal diagnosis codes for CHF. Age 18 years and older. Exclude patients discharged with specified cardiac procedure codes in any field, patients transferring from another institution, MDC 14 (pregnancy, childbirth, and puerperium), or MDC 15 (newborns and neonates).
Population at Risk	Population in MSA or county, age 18 years and older.
Empirical Results and Rating	Rate (2002): 457.7 per 100,000 population Rating: 14

Summary of Evidence

Congestive heart failure is a PQI that would be of most interest to comprehensive health care delivery systems. This indicator is measured with high precision, and most of the observed variance reflects true differences across areas.

Risk adjustment for age and sex appears to affect the areas with the highest and lowest raw rates. Areas with high rates may wish to examine the clinical characteristics of their patients to check for a more complex case mix. Patient age, clinical measures such as heart function, and other management issues may affect admission rates.

As the causes for admissions may include poor quality care, lack of patient compliance, or problems accessing care, areas may wish to review CHF patient records to identify precipitating causes and potential targets for intervention.

Limitations on Use

As a PQI, CHF is not a measure of hospital quality, but rather one measure of outpatient and other health care. Providers may reduce admission rates without actually improving quality by shifting care to an outpatient setting.

Some CHF care takes place in emergency rooms. As such, combining inpatient and emergency room data may give a more accurate picture of this indicator.

Details

Face validity: Does the indicator capture an aspect of quality that is widely regarded as important and subject to provider or public health system control?

Physician management of patients with congestive heart failure differs significantly by physician specialty.[73 74] Such differences in community practices may be reflected in differences in CHF admission rates.

Precision: Is there a substantial amount of provider or community level variation that is not attributable to random variation?

Relatively precise estimates of admission rates for CHF can be obtained, although random variation may be important for small hospitals and rural areas. Based on empirical evidence,

[73] Edep ME, Shah NB, Tateo IM, et al. Differences between primary care physicians and cardiologists in management of congestive heart failure: relation to practice guidelines. J Am Coll Cardiol 1997;30(2):518-26.

[74] Reis, SE, Holubkov R, Edmundowicz D, et al. Treatment of patients admitted to the hospital with congestive heart failure: specialty-related disparities in practice patterns and outcomes. J Am Coll Cardiol 1997;30(3):733-8.

this indicator is very precise, with a raw area level rate of 521.0 per 100,000 population and a standard deviation of 286.5.

The signal ratio (i.e., the proportion of the total variation across areas that is truly related to systematic differences in area performance rather than random variation) is very high, at 93.0%, indicating that the observed differences in age-sex adjusted rates very likely represent true differences across areas.

Minimum bias: Is there either little effect on the indicator of variations in patient disease severity and comorbidities, or is it possible to apply risk adjustment and statistical methods to remove most or all bias?

Important determinants of outcomes with CHF include certain demographic variables, such as patient age; clinical measures; management issues; and treatment strategies.[75] Limited evidence exists on the extent to which these factors can explain area differences in CHF admission rates. Empirical results show that area rankings and absolute performance are somewhat affected by age-sex risk adjustment.

Construct validity: Does the indicator perform well in identifying true (or actual) quality-of-care problems?

Billings et al. found that low-income ZIP codes in New York City had 4.6 times more CHF hospitalizations per capita than high-income ZIP codes.[76] Millman et al. reported that low-income ZIP codes had 6.1 times more CHF hospitalizations per capita than high-income ZIP codes.[77]

Based on empirical results, areas with high rates of CHF also tend to have high rates of admission for other ACSCs.

Fosters true quality improvement: Is the indicator insulated from perverse incentives for providers to improve their reported performance by avoiding difficult or complex cases, or by other responses that do not improve quality of care?

Outpatient interventions such as the use of protocols for ambulatory management of low-severity patients and improvement of access to outpatient care would most likely decrease inpatient admissions for CHF.[78]

Prior use: Has the measure been used effectively in practice? Does it have potential for working well with other indicators?

This indicator was originally developed by Billings et al. in conjunction with the United Hospital Fund of New York. It was subsequently adopted by the Institute of Medicine and has been widely used in a variety of studies of avoidable hospitalizations.

[75]Philbin EF, Andreaou C, Rocco TA, et al. Patterns of angiotensin-converting enzyme inhibitor use in congestive heart failure in two community hospitals. Am J Cardio. 1996;77(1):832-8.
[76]Billings J, Zeital L, Lukomnik J, et al. Analysis of variation in hospital admission rates associated with area income in New York City. Unpublished report.
[77]Millman M, editor. Committee on Monitoring Access to Personal Health Care Services. Washington DC: National Academy Press.

[78]Rosenthal GE, Harper DL, Shah A, et al. A regional evaluation of variation in low-severity hospital admissions. J Gen Intern Med 1997;12(7):416-22.

Low Birth Weight Rate (PQI 9)

Infants may be low birth weight because of inadequate interuterine growth or premature birth. Risk factors include sociodemographic and behavioral characteristics, such as low income and tobacco use during pregnancy.

Relationship to Quality	Proper preventive care may reduce incidence of low birth weight, and lower rates represent better quality care.
Benchmark	State, regional, or peer group average.
Definition	Number of low birth weight infants per 100 births.
Outcome of Interest	Number of births with ICD-9-CM diagnosis codes for birth weight less than 2500 grams in any field. Exclude patients transferring from another institution.
Population at Risk	All births (discharges in MDC 15, newborns and neonates) in MSA or county.
Empirical Results and Rating	Rate (2002): 5.9 per 100 eligible births Rating: 11 out of 16 (Bias was not tested because adequate risk adjustment for low birth weight was not available.) (Smoothing recommended)

Summary of Evidence

Low birth weight is a PQI that would be of most interest to comprehensive health care delivery systems. Healthy People 2010 has set a goal of reducing the percentage of low birth weight infants to 0.9%.[79]

Mothers who give birth to low birth weight infants generally receive less prenatal care than others, and prenatal care persists as a risk factor for low birth weight when adjusting for potential confounds. However, comprehensive care programs in high-risk women have failed to reduce low birth weights. In some studies, specific counseling aimed at reducing a specific risk factor in a specific population may have some impact on reducing low birth weight.

Adequate risk adjustment may require linkage to birth records, which record many of the sociodemographic and behavioral risk factors noted in the literature review (race, age, drug use, stress). Birth records in some States are a rich source of information that could help to identify causes of low birth weight and help to delineate potential areas of intervention.

Where risk adjustment is not possible, results may provide some guidance to case mix in the area if considered in light of measures of socioeconomic status (as determined by insurance status or ZIP code).

Limitations on Use

As a PQI, low birth weight is not a measure of hospital quality, but rather one measure of outpatient and other health care. This indicator could have substantial bias that would require additional risk adjustment from birth records or clinical data.

Details

Face validity: Does the indicator capture an aspect of quality that is widely regarded as important and subject to provider or public health system control?

Risk factors for low birth weight may be addressed with adequate prenatal care and education. Prenatal education and care programs have been established to help reduce low birth weight and other complications in high-risk populations.

Precision: Is there a substantial amount of provider or community level variation that is not attributable to random variation?

Although low birth weight births account for only a small fraction of total births, the large number of births suggest that this indicator should be

[79] Healthy People 2010. Office of Disease Prevention and Health Promotion, U.S. Department of Health and Human Services.

precisely measurable for most areas. Based on empirical evidence, this indicator is precise, with a raw area level rate of 3.9% and a standard deviation of 2.3%. The signal ratio (i.e., the proportion of the total variation across areas that is truly related to systematic differences in area performance rather than random variation) is moderate, at 67.1%, indicating that some of the observed differences in age-sex adjusted rates do not represent true differences in area performance.

*Minimum bias: Is there either little effect on the indicator of variations in patient disease severity and comorbidities, or is it possible to apply risk adjustment and statistical me*thods to remove most or all bias?

Socioeconomic measures such as parental education and income have been shown to be negatively associated with rates of low birth weight infants.[80 81] Demographic factors such as age and race also appear important, and may be correlated with socioeconomic factors. Mothers under 17 years and over 35 years are at a higher risk of having low birth weight infants.[82 83] One study of all California singleton births in 1992 found that after risk adjustment, having a black mother remained a significant risk factor.[84] Little evidence exists on the extent to which each of these factors contributes to differences in the rate of low birth weight births across geographic areas.

Construct validity: Does the indicator perform well in identifying true (or actual) quality-of-care problems?

While specific studies have demonstrated an impact of particular interventions, especially in high-risk populations, evidence on the impact of better prenatal care on low birth weight rates for area populations is less well developed. In one study, the use of prenatal care accounted for less than 15% of the differences between low birth weight in black and white mothers enrolled in an HMO. However, increasing the level of prenatal care was associated with lower rates of low birth weight, particularly in the black patient population.[85]

Low birth weight is inversely related to the other ACSCs and is positively related to perforated appendix rate. Empirical evidence suggests that this indicator at an area level could be potentially biased.

Fosters true quality improvement: Is the indicator insulated from perverse incentives for providers to improve their reported performance by avoiding difficult or complex cases, or by other responses that do not improve quality of care?

Use of this indicator is unlikely to lead to apparent reductions in the rate of low birth weight births that did not represent true reductions.

Prior use: Has the measure been used effectively in practice? Does it have potential for working well with other indicators?

Low birth weight is an indicator in the Health Plan Employer Data and Information Set (HEDIS) measure set for insurance groups and is used by United Health Care and the University Hospital Consortium. This indicator, along with very low birth weight, was previously an HCUP QI.

[80]Hessol NA, Fuentes-Afflick E, Bacchetti P. Risk of low birth weight infants among black and white parents. Obstet Gynecol 1998;92(5):814-22.
[81]O'Campo P, Xue X, Wang MC, et al. Neighborhood risk factors for low birthweight in Baltimore: a multilevel analysis. Am J Public Health 1997;87(7):1113-8.
[82]Hessol, et al. 1998.
[83]O'Campo, et al. 1997.
[84]Hessol, et al. 1998.
[85]Murray JL, Bernfield M. The differential effect of prenatal care on the incidence of low birth weight among blacks and whites in a prepaid health care plan. N Engl J Med 1988;319(21):1385-91.

Dehydration Admission Rate (PQI 10)

Dehydration is a serious acute condition that occurs in frail patients and patients with other underlying illnesses following insufficient attention and support for fluid intake. Dehydration can for the most part be treated in an outpatient setting, but it is potentially fatal for elderly, very young children, frail patients, or patients with serious comorbid conditions.

Relationship to Quality	Proper outpatient treatment may reduce admissions for dehydration, and lower rates represent better quality care.
Benchmark	State, regional, or peer group average.
Definition	Admissions for dehydration per 100,000 population.
Outcome of Interest	Discharges with ICD-9-CM principal diagnosis code for hypovolemia (276.5). Exclude patients transferring from another institution, MDC 14 (pregnancy, childbirth, and puerperium), or MDC 15 (newborns and other neonates).
Population at Risk	Population in MSA or county.
Empirical Results and Rating	Rate (2002): 139.9 per 100,000 population Rating: 14

Summary of Evidence

Hospital admission for dehydration is a PQI that would be of most interest to comprehensive health care delivery systems. Admission for dehydration is somewhat common, suggesting that the indicator will be measured with adequate precision, and most of the observed variation is likely to reflect true differences in admission rates.

This indicator is subject to minimal bias. Risk adjustment appears to affect modestly the areas with the highest and lowest rates. Age may be a particularly important factor, and the indicator should be risk-adjusted for age. Areas with high rates of dehydration admissions also tend to have high rates of admission for other ACSCs.

The considerable variations across areas suggest opportunities for quality improvement in care for patients at risk for dehydration. When high rates of dehydration are identified for a particular hospital, additional study may uncover problems in primary or emergency care in the surrounding area. Appropriate interventions can be developed to address those problems.

Limitations on Use

As a PQI, dehydration is not a measure of hospital quality, but rather one of the measures of outpatient and other health care.

This indicator has unclear construct validity, because it has not been validated except as part of a set of indicators. Providers may reduce admission rates without actually improving quality by shifting care to an outpatient setting. Some dehydration care takes place in emergency rooms. As such, combining inpatient and emergency room data may give a more accurate picture of this indicator.

Details

Face validity: Does the indicator capture an aspect of quality that is widely regarded as important and subject to provider or public health system control?

Dehydration is a potentially fatal condition, and appropriate attention to fluid status can prevent the condition. If left untreated in older adults, serious complications, including death (over 50%), can result.[86]

Precision: Is there a substantial amount of provider or community level variation that is not attributable to random variation?

Little evidence exists in the literature on the precision of this indicator. Based on empirical

[86]Weinberg AD, Minaker KL. Dehydration. Evaluation and management in older adults. Council on Scientific Affairs, American Medical Association. JAMA 1995;274(19):1552-6.

evidence, this indicator is precise, with a raw area level rate of 139.9 per 100,000 population and a standard deviation of 103.2.

The signal ratio (i.e., the proportion of the total variation across areas that is truly related to systematic differences in area performance rather than random variation) is high, at 88.5%, indicating that the observed differences in age-sex adjusted rates likely represent true differences across areas.

Minimum bias: Is there either little effect on the indicator of variations in patient disease severity and comorbidities, or is it possible to apply risk adjustment and statistical methods to remove most or all bias?

The age structure of the population may affect admission rates for this condition, as the elderly and very young are more susceptible to dehydration. Socioeconomic factors may also affect admission rates. Differences in thresholds for admission of patients with dehydration may contribute to area rate differences. Empirical results show that area rankings are not affected by age-sex risk adjustment.

Construct validity: Does the indicator perform well in identifying true (or actual) quality-of-care problems?

Billings et al. found that low-income ZIP codes in New York City had 2 times more dehydration hospitalizations per capita than high-income ZIP codes.[87] Household income explained 42% of this variation. In addition, Millman et al.[88] reported that low-income ZIP codes had 2 times more dehydration hospitalizations per capita than high-income ZIP codes.

Based on empirical results of this study, areas with high rates of dehydration admissions also tend to have high rates of admission for other ACSCs.

Fosters true quality improvement: Is the indicator insulated from perverse incentives for providers to improve their reported performance by avoiding difficult or complex cases, or by other responses that do not improve quality of care?

Use of this indicator might lead to higher thresholds of admission for patients with dehydration, potentially denying needed care to some patients. Because some dehydration can be managed on an outpatient basis, a shift to outpatient care may occur.

Prior use: Has the measure been used effectively in practice? Does it have potential for working well with other indicators?

This indicator was originally developed by Billings et al. in conjunction with the United Hospital Fund of New York.

[87]Billings J, Zeital L, Lukomnik J, et al. Analysis of variation in hospital admission rates associated with area income in New York City. Unpublished report.
[88]Millman M, editor. Committee on Monitoring Access to Personal Health Care Services. Washington, DC: National Academy Press. 1993.

Bacterial Pneumonia Admission Rate (PQI 11)

Bacterial pneumonia is a relatively common acute condition, treatable for the most part with antibiotics. If left untreated in susceptible individuals—such as the elderly—pneumonia can lead to death.

Relationship to Quality	Proper outpatient treatment may reduce admissions for bacterial pneumonia in non-susceptible individuals, and lower rates represent better quality care.
Benchmark	State, regional, or peer group average.
Definition	Admissions for bacterial pneumonia per 100,000 population.
Outcome of Interest	Discharges with ICD-9-CM principal diagnosis code for bacterial pneumonia. Exclude patients with sickle cell anemia or HB-S disease, patients transferring from another institution, MDC 14 (pregnancy, childbirth, and puerperium), or MDC 15 (newborns and neonates).
Population at Risk	Population in MSA or county.
Empirical Results and Rating	Rate (2002): 349.7 per 100,000 population Rating: 17

Summary of Evidence

Hospital admission for bacterial pneumonia is a PQI that would be of most interest to comprehensive health care delivery systems. High admission rates may reflect a large number of inappropriate admissions or low-quality treatment with antibiotics. Admission for pneumonia is relatively common, suggesting that the indicator will be measured with good precision, and most of the observed variation reflects true differences in admission rates.

This indicator is subject to some moderate bias, and risk adjustment appears to affect the areas with the highest rates the most. Age may be a particularly important factor, and the indicator should be risk-adjusted for this factor. Areas may wish to examine the outpatient care for pneumonia and pneumococcal vaccination rates to identify potential processes of care that may reduce admission rates. The patient populations served by hospitals that contribute the most to the overall area rate for pneumonia may be a starting point for interventions.

Limitations on Use

As a PQI, admission for bacterial pneumonia is not a measure of hospital quality, but rather one measure of outpatient and other health care.

This indicator has unclear construct validity, because it has not been validated except as part of a set of indicators. Providers may reduce admission rates without actually improving quality by shifting care to an outpatient setting. Because some pneumonia care takes place in an emergency room setting, combining inpatient and emergency room data may give a more accurate picture of this indicator.

Details

Face validity: Does the indicator capture an aspect of quality that is widely regarded as important and subject to provider or public health system control?

Vaccination for pneumococcal pneumonia in the elderly and early management of bacterial respiratory infections on an ambulatory basis may reduce admissions with pneumonia. A vaccine developed for the elderly has been shown to be 45% effective in preventing hospitalizations during peak seasons.[89]

Precision: Is there a substantial amount of provider or community level variation that is not attributable to random variation?

Little evidence exists in the literature on the precision or variation in pneumonia admission rates. Based on empirical evidence, this indicator is precise, with a raw area level rate of 395.6 per 100,000 population and a standard deviation of 208.5.

[89] Foster DA, Talsma A, Furumoto-Dawson A, et al. Influenza vaccine effectiveness in preventing hospitalization for pneumonia in the elderly. Am J Epidemiol 1992;136(3):296-307.

The signal ratio (i.e., the proportion of the total variation across areas that is truly related to systematic differences in area performance rather than random variation) is very high, at 92.9%, indicating that the observed differences in age-sex adjusted rates likely represent true differences across areas. Using multivariate signal extraction techniques appears to have little additional impact on estimating true differences across areas.

Minimum bias: Is there either little effect on the indicator of variations in patient disease severity and comorbidities, or is it possible to apply risk adjustment and statistical methods to remove most or all bias?

A review of the literature suggests that comorbidities or other risk factors that may vary systematically by area do not significantly affect the incidence of hospitalization for pneumonia. Differences in thresholds for admission of patients with bacterial pneumonia may contribute to area rate differences. Empirical results show that area rankings and absolute performance are somewhat affected by age-sex risk adjustment.

Construct validity: Does the indicator perform well in identifying true (or actual) quality-of-care problems?

Billings et al. found that low-income ZIP codes in New York City had 5.4 times more pneumonia admissions per capita than high-income ZIP codes.[90] Household income explained 53% of this variation. In addition, Millman et al.[91] reported that low-income ZIP codes had 5.4 times more pneumonia hospitalizations per capita than high-income ZIP codes.

Based on empirical results, areas with high rates of bacterial pneumonia admissions also tend to have high rates of admissions for other ACSCs.

Fosters true quality improvement: Is the indicator insulated from perverse incentives for providers to improve their reported performance by avoiding difficult or complex cases, or by other responses that do not improve quality of care?

Use of this indicator might lead to higher thresholds of admission for pneumonia patients. Because pneumonia can be managed on an outpatient basis, a shift to outpatient care may occur, which might be inappropriate for more severely ill patients.

Prior use: Has the measure been used effectively in practice? Does it have potential for working well with other indicators?

This indicator was included in Weissman's set of avoidable hospitalizations.[92]

[90]Billings J, Zeital L, Lukomnik J, et al. Analysis of variation in hospital admission rates associated with area income in New York City. Unpublished report.
[91]Millman M, editor. Committee on Monitoring Access to Personal Health Care Services. Washington, DC: National Academy Press. 1993.

[92]Weissman JS, Gatsonis C, Epstein AM. Rates of avoidable hospitalization by insurance status in Massachusetts and Maryland JAMA 1992;268(17)2388-94.

Urinary Tract Infection Admission Rate (PQI 12)

Urinary tract infection is a common acute condition that can, for the most part, be treated with antibiotics in an outpatient setting. However, this condition can progress to more clinically significant infections, such as pyelonephritis, in vulnerable individuals with inadequate treatment.

Relationship to Quality	Proper outpatient treatment may reduce admissions for urinary infection, and lower rates represent better quality care.
Benchmark	State, regional, or peer group average.
Definition	Admissions for urinary tract infection per 100,000 population.
Outcome of Interest	Discharges with ICD-9-CM principal diagnosis code for urinary tract infection. Exclude patients transferring from another institution, MDC 14 (pregnancy, childbirth, and puerperium), or MDC 15 (newborns and neonates).
Population at Risk	Population in MSA or county.
Empirical Results and Rating	Rate (2002): 137.9 per 100,000 population Rating: 11

Summary of Evidence

Hospital admission for urinary tract infection is a PQI that would be of most interest to comprehensive health care delivery systems. Admission for urinary tract infection is uncommon, but the observed variation is likely to reflect true differences across areas.

Risk adjustment appears to affect the areas with the highest rates the most, and using this indicator without risk adjustment may result in the misidentification of some areas as outliers. This indicator is subject to some moderate bias and should be adjusted for age and sex. The patient populations served by hospitals that contribute the most to the overall area rate for urinary tract infection may be a starting point for interventions.

Limitations on Use

As a PQI, admission for urinary tract infection is not a measure of hospital quality, but rather one measure of outpatient and other health care. This indicator has unclear construct validity, because it has not been validated except as part of a set of indicators. Providers may reduce admission rates without actually improving quality by shifting care to an outpatient setting. Some urinary tract infection care takes place in emergency rooms. As such, combining inpatient and emergency room data may give a more accurate picture of this indicator.

Details

Face validity: Does the indicator capture an aspect of quality that is widely regarded as important and subject to provider or public health system control?

Uncomplicated urinary tract infections can be treated with antibiotics in the ambulatory setting; however, inappropriate treatment can lead to more serious complications. Admission for urinary tract infection among children, which is rare, is associated with physiological abnormalities.

Precision: Is there a substantial amount of provider or community level variation that is not attributable to random variation?

Little evidence exists in the literature on the precision and variation associated with this indicator. Based on empirical evidence, this indicator is precise, with a raw area level rate of 145.1 per 100,000 population and a standard deviation of 89.5. The signal ratio (i.e., the proportion of the total variation across areas that is truly related to systematic differences in area performance rather than random variation) is high, at 84.9%, indicating that the observed differences in age-sex adjusted rates likely represent true differences across areas. Using multivariate signal extraction techniques appears to have little additional impact on estimating true differences across areas.

Minimum bias: Is there either little effect on the indicator of variations in patient disease severity and comorbidities, or is it possible to apply risk adjustment and statistical methods to remove most or all bias?

Differences in thresholds for admission of patients with urinary tract infection may contribute to area rate differences. Empirical results show that area rankings and absolute performance are somewhat affected by age-sex risk adjustment.

Construct validity: Does the indicator perform well in identifying true (or actual) quality-of-care problems?

Billings et al. found that low-income ZIP codes in New York City had 2.2 times more urinary tract infection admissions than high-income ZIP codes.[93] Household income explained 28% of this variation. In addition, Millman et al.[94] reported that low-income ZIP codes had 2.8 times more urinary tract infection hospitalizations per capita than high-income ZIP codes.

Based on empirical results, areas with high admission rates for urinary tract infections also tend to have high admission rates for other ACSCs.

Fosters true quality improvement: Is the indicator insulated from perverse incentives for providers to improve their reported performance by avoiding difficult or complex cases, or by other responses that do not improve quality of care?

Use of this indicator might lead to higher thresholds of admission for patients with urinary tract infections.

Prior use: Has the measure been used effectively in practice? Does it have potential for working well with other indicators?

This indicator was originally developed by Billings et al. in conjunction with the United Hospital Fund of New York. It is included in Weissman's set of avoidable hospitalizations.[95]

[93]Billings J, Zeital L, Lukomnik J, et al. Analysis of variation in hospital admission rates associated with area income in New York City. Unpublished report.
[94]Millman M, editor. Committee on Monitoring Access to Personal Health Care Services. Washington, DC: National Academy Press. 1993.

[95]Weissman JS, Gatsonis C, Epstein AM. Rates of avoidable hospitalization by insurance status in Massachusetts and Maryland. JAMA 1992;268(17)2388-94.

Angina without Procedure Admission Rate (PQI 13)

Both stable and unstable angina are symptoms of potential coronary artery disease. Effective management of coronary disease reduces the occurrence of major cardiac events such as heart attacks, and may also reduce admission rates for angina.

Relationship to Quality	Proper outpatient treatment may reduce admissions for angina (without procedures), and lower rates represent better quality care.
Benchmark	State, regional, or peer group average.
Definition	Admissions for angina (without procedures) per 100,000 population.
Outcome of Interest	Discharges with ICD-9-CM principal diagnosis codes for angina. Age 18 years and older. Exclude discharges with a procedure code for cardiac procedure, patients transferring from another institution, MDC 14 (pregnancy, childbirth, and puerperium), or MDC 15 (newborns and neonates).
Population at Risk	Population in MSA or county, age 18 years and older.
Empirical Results and Rating	Rate (2002): 55.1 per 100,000 population Rating: 19

Summary of Evidence

Hospital admission for angina is a PQI that would be of most interest to comprehensive health care delivery systems. Admission for angina is relatively common, suggesting that the indicator will be measured with good precision. The observed variation likely reflects true differences in area performance.

Age-sex adjustment has a moderate impact. Other risk factors for consideration include smoking, hyperlipidemia, hypertension, diabetes, and socioeconomic status. The patient populations served by hospitals that contribute the most to the overall area rate for angina may be a starting point for interventions.

Limitations on Use

As a PQI, angina without procedure is not a measure of hospital quality, but rather one measure of outpatient and other health care. This indicator has unclear construct validity, because it has not been validated except as part of a set of indicators. Providers may reduce admission rates without actually improving quality of care by shifting care to an outpatient setting. Some angina care takes place in emergency rooms. Combining inpatient and emergency room data may give a more accurate picture.

Details

Face validity: Does the indicator capture an aspect of quality that is widely regarded as important and subject to provider or public health system control?

Stable angina can be managed in an outpatient setting using drugs such as aspirin and beta blockers, as well as advice to change diet and exercise habits.[96] Effective treatments for coronary artery disease reduce admissions for serious complications of ischemic heart disease, including unstable angina.

Precision: Is there a substantial amount of provider or community level variation that is not attributable to random variation?

Reasonably precise estimates of area angina rates should be feasible, as one study shows that unstable angina accounts for 16.3% of total

[96] Gibbons RJ, Chatterjee K, Daley J, et al. ACC/AHA/ACP-ASIM guidelines for the management of patients with chronic stable angina: a report of the American College of Cardiology/American Heart Association Task force on Practice Guidelines (Committee on Management of Patients with Chronic Stable Angina) [published erratum appears in J Am Coll Cardiol 1999 Jul;34(1):314]. J Am Coll Cardiol 1999;33(7):2092-197.

admissions for ACSCs.[97] Based on empirical evidence, this indicator is adequately precise, with a raw area level rate of 166.0 per 100,000 population and a standard deviation of 135.7.

The signal ratio (i.e., the proportion of the total variation across areas that is truly related to systematic differences in area performance rather than random variation) is very high, at 91.6%, indicating that the observed differences in age-sex adjusted rates likely represent true differences across areas. Using multivariate signal extraction techniques appears to have little additional impact on estimating true differences across areas.

Minimum bias: Is there either little effect on the indicator of variations in patient disease severity and comorbidities, or is it possible to apply risk adjustment and statistical methods to remove most or all bias?

No evidence exists in the literature on the potential bias of this indicator. The incidence of angina is related to age structure and risk factors (smoking, hyperlipidemia, hypertension, diabetes) in a population. Elderly age (over 70), diabetes, and hypertension have also been associated with being at higher risk for angina.[98]

Construct validity: Does the indicator perform well in identifying true (or actual) quality-of-care problems?

Billings et al. found that low-income ZIP codes in New York City had 2.3 times more angina hospitalizations than high-income ZIP codes.[99] Household income explained 13% of this variation. In addition, Millman et al.[100] reported that low-income ZIP codes had 2.7 times more angina hospitalizations per capita than high-income ZIP codes.

Based on empirical study, areas with high rates of angina admissions tend to have higher rates of other ACSC admissions.

Fosters true quality improvement: Is the indicator insulated from perverse incentives for providers to improve their reported performance by avoiding difficult or complex cases, or by other responses that do not improve quality of care?

Use of this quality indicator might raise the threshold for admission of angina patients. Because some angina can be managed on an outpatient basis, a shift to outpatient care may occur but is unlikely for severe angina.

Prior use: Has the measure been used effectively in practice? Does it have potential for working well with other indicators?

This indicator was originally developed by Billings et al. in conjunction with the United Hospital Fund of New York.

---

[97] Blustein J, Hanson K, Shea S. Preventable hospitalizations and socioeconomic status. Health Aff (Millwood) 1998;17(2):177-89.

[98] Brunwald E, Antman EM, Beasley JW et al. ACC/AHA guidelines for the management of patients with unstable angina and non-ST-segment elevation myocardial infarction. A report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee on the Management of Patients with Unstable Angina). J Am Coll Cardiol 2000;36(3):970-1062.

[99] Billings J, Zeital L, Lukomnik J, et al. Analysis of variation in hospital admission rates associated with area income in New York City. Unpublished report.

[100] Millman M, editor. Committee on Monitoring Access to Personal Health Care Services. Washington, DC: National Academy Press. 1993.

Uncontrolled Diabetes Admission Rate (PQI 14)

Uncontrolled diabetes should be used in conjunction with short-term complications of diabetes, which include diabetic ketoacidosis, hyperosmolarity, and coma.*

Relationship to Quality	Proper outpatient treatment and adherence to care may reduce the incidence of uncontrolled diabetes, and lower rates represent better quality care.
Benchmark	State, regional, or peer group average.
Definition	Admissions for uncontrolled diabetes per 100,000 population.
Outcome of Interest	Discharges with ICD-9-CM principal diagnosis codes for uncontrolled diabetes, without mention of a short-term or long-term complication. Age 18 years and older. Exclude patients transferring from another institution, MDC 14 (pregnancy, childbirth, and puerperium), or MDC 15 (newborns and neonates).
Population at Risk	Population in MSA or county, age 18 years and older.
Empirical Results and Rating	Rate (2002): 23.0 per 100,000 population Rating: 14

* This indicator is designed to be combined with "Short Term Diabetes Complication Admission Rate" to create the Healthy People 2010 indicator. To do so, users may simply add the rates of the two indicators together.

Summary of Evidence

Hospital admission for uncontrolled diabetes is a PQI that would be of most interest to comprehensive health care delivery systems. Healthy People 2010 has established a goal to reduce the hospitalization rate for uncontrolled diabetes in persons 18-64 years of age from 7.2 per 10,000 population to 5.4 per 10,000 population.[101] Combining this indicator with the short-term diabetes indicator will result in the Healthy People 2010 measure, except that this QI excludes transfers from another institution to reduce double counting of cases. As a result the rate for the AHRQ QI may be minimally lower than the Healthy People 2010 indicator.

This indicator is moderately precise. The observed differences across areas likely reflect true differences in area performance. Age-sex adjustment slightly changes area rankings.

Limitations on Use

As a PQI, uncontrolled diabetes is not a measure of hospital quality, but rather one measure of outpatient and other health care. Rates of diabetes may vary systematically by area, creating bias for this indicator.

Details

Face validity: Does the indicator capture an aspect of quality that is widely regarded as important and subject to provider or public health system control?

High-quality outpatient management of diabetic patients has been shown to lead to reductions in almost all types of serious avoidable hospitalizations. However, tight control may be associated with more episodes of hypoglycemia that lead to more admissions.

Precision: Is there a substantial amount of provider or community level variation that is not attributable to random variation?

Based on empirical evidence, this indicator is moderately precise, with a raw area level rate of 34.7 per 100,000 population and a standard deviation of 28.1.

The signal ratio (i.e., the proportion of the total variation across areas that is truly related to systematic differences in area performance

[101] Healthy People 2010, Office of Disease Prevention and Health Promotion. U.S. Department of Health and Human Services.

rather than random variation) is high, at 72.6%, indicating that the observed differences in age-sex adjusted rates likely represent true differences in area performance. Using multivariate signal extraction techniques appears to have little additional impact on estimating true differences across areas.

Minimum bias: Is there either little effect on the indicator of variations in patient disease severity and comorbidities, or is it possible to apply risk adjustment and statistical methods to remove most or all bias?

Minorities have higher rates of diabetes, and higher hospitalization rates may result in areas with higher minority concentrations. Empirical results show that area rankings in the highest and lowest deciles are slightly affected by age-sex adjustment.

Construct validity: Does the indicator perform well in identifying true (or actual) quality-of-care problems?

Based on empirical results, areas with high rates of uncontrolled diabetes also tend to have high rates of admission for other ACSCs.

Fosters true quality improvement: Is the indicator insulated from perverse incentives for providers to improve their reported performance by avoiding difficult or complex cases, or by other responses that do not improve quality of care?

Because diabetic emergencies are potentially life-threatening, hospitals are unlikely to fail to admit patients requiring hospitalization.

Prior use: Has the measure been used effectively in practice? Does it have potential for working well with other indicators?

This measure corresponds closely with the measure of short-term diabetes that was developed by Billings et al. and described in this document.[102] The key exception is the ICD-9-CM codes 25002 and 25003, which are the only codes included for uncontrolled diabetes.

[102]Billings J, Zeital L, Lukomnik J, et al. Analysis of variation in hospital admission rates associated with area income in New York City. Unpublished report.

Adult Asthma Admission Rate (PQI 15)

Asthma is one of the most common reasons for hospital admission and emergency room care. Most cases of asthma can be managed with proper ongoing therapy on an outpatient basis. Most published studies combine admission rates for children and adults; therefore, areas may wish to examine this indicator together with pediatric asthma.

Relationship to Quality	Proper outpatient treatment may reduce the incidence or exacerbation of asthma requiring hospitalization, and lower rates represent better quality care.
Benchmark	State, regional, or peer group average.
Definition	Admissions for adult asthma per 100,000 population.
Outcome of Interest	Discharges with ICD-9-CM principal diagnosis codes for asthma. Age 18 years and older. Exclude patients transferring from another institution, MDC 14 (pregnancy, childbirth, and puerperium), or MDC 15 (newborns and neonates).
Population at Risk	Population in MSA or county, age 18 years and older.
Empirical Results and Rating	Rate (2002): 110.9 per 100,000 population Rating: 16

Summary of Evidence

Hospital admission for asthma is a PQI that would be of most interest to comprehensive health care delivery systems.

Environmental factors such as air pollution, occupational exposure to irritants, or other exposure to allergens have been shown to increase hospitalization rates or exacerbate asthma symptoms. While race has been shown to be associated with differences in admission rates, it is unclear whether this is due to differences in severity of disease or inadequate access to care. Adjustment for race is recommended.

Admission rates have been associated with lower socioeconomic status. Areas may wish to identify hospitals that contribute the most to the overall area rate for this indicator. The patient populations served by these hospitals may be a starting point for interventions.

Limitations on Use

As a PQI, adult asthma is not a measure of hospital quality, but rather one measure of outpatient and other health care. Providers may reduce admission rates without actually improving quality by shifting care to an outpatient setting.

Admission rates that are drastically below or above the average or recommended rates should be further examined.

Details

Face validity: Does the indicator capture an aspect of quality that is widely regarded as important and subject to provider or public health system control?

According to the National Asthma Education Program, asthma is a readily treatable chronic disease that can be managed effectively in the outpatient setting.[103] Observational studies offer some evidence that inhaled steroids may decrease risk of admission by up to 50%.[104] [105]

[103] National Heart, Lung, and Blood Institute/National Asthma Education and Prevention Program. Expert Panel Report 2: Guidelines for the diagnosis and management of asthma. In: National Institutes of Health pub. no. 97-4051. Bethesda, MD; 1997.
[104] Blais L, Ernst P, Boivin JF, et al. Inhaled corticosteroids and the prevention of readmission to hospital for asthma. Am J Respir Crit Care Med 1998; 158(1):126-32.
[105] Donahue JG, Weiss ST, Livingston JM, et al. Inhaled steroids and the risk of hospitalization for asthma. JAMA 1997;277(11):887-91.

Precision: Is there a substantial amount of provider or community level variation that is not attributable to random variation?

Asthma is a common cause of admission for adults, and as such this measure is likely to have adequate precision. Based on empirical evidence, this indicator is adequately precise, with a raw area level rate of 107.9 per 100,000 population and a standard deviation of 81.7. The signal ratio (i.e., the proportion of the total variation across areas that is truly related to systematic differences in area performance rather than random variation) is high, at 83.6%, indicating that the observed differences in age-sex adjusted rates likely represent true differences across areas.

Minimum bias: Is there either little effect on the indicator of variations in patient disease severity and comorbidities, or is it possible to apply risk adjustment and statistical methods to remove most or all bias?

Numerous environmental risk factors for asthma have been identified, including allergens, tobacco smoke, and outdoor air pollution. Race represents one of the most complex potentially biasing factors for this indicator. Black patients have consistently been shown to have higher asthma admission rates, even when stratifying for income and age.[106] Adjustment for race is recommended. Empirical results show that area rankings are not affected by age-sex risk adjustment.

Construct validity: Does the indicator perform well in identifying true (or actual) quality-of-care problems?

Billings et al. found that low-income ZIP codes in New York City had 6.4 times more asthma hospitalizations than high-income ZIP codes.[107] Household income explained 70% of this variation. In addition, Millman et al.[108] reported that low-income ZIP codes had 5.8 times more asthma hospitalizations per capita than high-income ZIP codes.

Fosters true quality improvement: Is the indicator insulated from perverse incentives for providers to improve their reported performance by avoiding difficult or complex cases, or by other responses that do not improve quality of care?

There is little evidence to suggest that asthmatics are being inappropriately denied admission to the hospital. However, because some asthma can be managed on an outpatient basis, a shift to outpatient care may occur.

Prior use: Has the measure been used effectively in practice? Does it have potential for working well with other indicators?

This indicator was originally developed by Billings et al. in conjunction with the United Hospital Fund of New York, and is included in Weissman's set of avoidable hospitalizations.[109]

[106]Ray NF, Thamer M, Fadillioglu B, et al. Race, income, urbanicity, and asthma hospitalization in California: a small area analysis. Chest 1998;113(5):1277-84.
[107]Billings J, Zeital L, Lukomnik J, et al. Analysis of variation in hospital admission rates associated with area income in New York City. Unpublished report.
[108]Millman M, editor. Committee on Monitoring Access to Personal Health Care Services. Washington, DC: National Academy Press. 1993.

[109]Weissman JS, Gatsonis C, Epstein AM. Rates of avoidable hospitalization by insurance status in Massachusetts and Maryland. JAMA 1992;268(17)2388-94.

Rate of Lower-extremity Amputation among Patients with Diabetes (PQI 16)

Diabetes is a major risk factor for lower-extremity amputation, which can be caused by infection, neuropathy, and microvascular disease.

Relationship to Quality	Proper and continued treatment and glucose control may reduce the incidence of lower-extremity amputation, and lower rates represent better quality care.
Benchmark	State, regional, or peer group average.
Definition	Admissions for lower-extremity amputation in patients with diabetes per 100,000 population.
Outcome of Interest	Discharges with ICD-9-CM procedure codes for lower-extremity amputation in any field and diagnosis code for diabetes in any field. Age 18 years and older. Exclude discharges with trauma, patients transferring from another institution, MDC 14 (pregnancy, childbirth, and puerperium), or MDC 15 (newborns and neonates).
Population at Risk	Population in MSA or county, age 18 years and older.
Empirical Results and Rating	Rate (2002): 36.8 per 100,000 population Rating:10 (Smoothing recommended)

Summary of Evidence

Hospital admissions for lower-extremity amputation among patients with diabetes is a PQI that would be of most interest to comprehensive health care delivery systems.

Lower-extremity amputation (LEA) affects up to 15% of all patients with diabetes in their lifetimes.[110] A combination of factors may lead to this high rate of amputation, including minor trauma to the feet, which is caused by loss of sensation and may lead to gangrene.[111] Proper long-term glucose control, diabetes education, and foot care are some of the interventions that can reduce the incidence of infection, neuropathy, and microvascular diseases. Healthy People 2010 has set a goal of reducing the number of LEAs to 1.8 per 1,000 persons with diabetes.[112]

Studies have shown that LEA varies by age and sex, and age-sex risk adjustment affects moderately the relative performance of areas. Race may bias the indicator, since the rates of diabetes and poor glycemic control are higher among Native Americans and Hispanic Americans. However, results must be interpreted with care when adjusting for race, because poor quality care may also vary systematically with racial composition.

Limitations on Use

As a PQI, lower-extremity amputations among patients with diabetes is not a measure of hospital quality, but rather one measure of outpatient and other health care. PQIs are correlated with each other and may be used in conjunction as an overall examination of outpatient care.

Details

Face validity: Does the indicator capture an aspect of quality that is widely regarded as important and subject to provider or public health system control?

In the United States, diabetes is the leading cause of nontraumatic amputations (approximately 57,000 per year).[113] Possible

[110]Mayfield JA, Reiber GE, Sanders LJ, et al. Preventive foot care in people with diabetes. Diabetes Care 1998;21(12):2161-77.
[111]Pecoraro RE, Reiber BE, Burgess EM. Pathways to diabetic limb amputation. Basis of prevention. Diabetes Care 1990;13(5):513-21.
[112]Healthy People 2010, Office of Disease Prevention and Health Promotion. U.S. Department of Health and Human Services.
[113]Centers for Disease Control and Prevention (CDC). National Diabetes Fact Sheet: National Estimates and

interventions include foot clinics, wearing proper footwear, and proper care of feet and foot ulcers.[114]

Precision: Is there a substantial amount of provider or community level variation that is not attributable to random variation?

Based on empirical evidence, this indicator is moderately precise, with a raw area level rate of 30.5 per 100,000 population and a substantial standard deviation of 42.7.

The signal ratio (i.e., the proportion of the total variation across areas that is truly related to systematic differences in area performance rather than random variation) is moderate, at 68.5%, indicating that some of the observed differences in age-sex adjusted rates likely do not represent true differences in area performance. Using multivariate signal extraction techniques appears to have little additional impact on estimating true differences across areas.

*Minimum bias: Is there either little effect on the indicator of variations in patient disease severity and comorbidities, or is it possible to apply risk adjustment and statisti*cal methods to remove most or all bias?

Several sociodemographic variables are associated with the risk of lower-extremity amputation, including age, duration of diabetes, and sex.[115] [116] Empirical results show that age-sex adjustment affects the relative performance of areas.

Construct validity: Does the indicator perform well in identifying true (or actual) quality-of-care problems?

Several studies of intervention programs have noted a decrease in amputation risk. One recent study noted a 1-year post-intervention decrease of 79% in amputations in a low-income African American population. Interventions included foot care education, assistance in finding properly fitting footwear, and prescription footwear.[117] One observational study found that patients who receive no outpatient diabetes education have a three-fold higher risk of amputation than those receiving care.[118]

*Fosters true quality improvement: Is the indicator insulated from perverse incentives for providers to improve their reported performance by avoiding difficult or complex cases, or by other respons*es *that do not improve quality of care?*

Given the severity of conditions requiring lower-extremity amputation, hospitals are unlikely to fail to admit patients requiring hospitalization.

Prior use: Has the measure been used effectively in practice? Does it have potential for working well with other indicators?

This indicator is not widely used; however, it is included in the DEMPAQ measure set for outpatient care.

General Information on Diabetes in the United States. Atlanta, GA: U.S. Department of Health and Human Services, 1999.

[114]Pecoraro et al. 1990.

[115]Mayfield et al. 1998.

[116]Selby JV, Zhang D. Risk factors for lower extremity amputation in persons with diabetes. Diabetes Care 1995;18(4):509-16.

[117]Patout CA, Jr., Birke JA, Horswell R, et al. Effectiveness of a comprehensive diabetes lower-extremity amputation prevention program in a predominantly low-income African-American population. Diabetes Care 2000;23(9):1339-42.

[118]Reiber GE, Pecoraro RE, Koepsell TD. Risk factors for amputation in patients with diabetes mellitus. A case-control study. Ann Intern Med 1992;117(2):97-105.

References

Access to Health Care in America. Washington, DC: National Academy Press; 1993.

Bagg W, Sathu A, Streat S, et al. Diabetic ketoacidosis in adults at Auckland Hospital, 1988-1996. Aust NZ J Med 1998;28(5):604-8.

Billings J, Zeital L, Lukomnik J, et al. Analysis of variation in hospital admission rates associated with area income in New York City. Unpublished report.

Billings J, Zeitel L, Lukomnik J, et al. Impact of socioeconomic status on hospital use in New York City, Health Aff (Millwood) 1993;12(1):162-73.

Bindman AB, Grumbach K, Osmond D, et al. Preventable hospitalizations and access to health care. JAMA 1995;274(4):305-11.

Blais L, Ernst P, Boivin JF, et al. Inhaled corticosteroids and the prevention of readmission to hospital for asthma. Am J Respir Crit Care Med 1998; 158(1):126-32.

Blumberg MS, Juhn PI. Insurance and the risk of ruptured appendix [letter; comment]. N Engl J Med 1995;332(6):395-6; discussion 397-8.

Blustein J, Hanson K, Shea S. Preventable hospitalizations and socioeconomic status. Health Aff (Millwood) 1998;17(2):177-89.

Bratton SL, Haberkern CM, Waldhausen JH. Acute appendicitis risks of complications: age and Medicaid insurance. Pediatrics 2000;106(1 Pt 1):75-8.

Braveman P, Schaaf VM, Egerter S, et al. Insurance-related differences in the risk of ruptured appendix [see comments]. N Engl J Med 1994;331(7):444-9.

Brunwald E, Antman EM, Beasley JW et al. ACC/AHA guidelines for the management of patients with unstable angina and non-ST-segment elevation myocardial infarction. A report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee on the Management of Patients with Unstable Angina). J Am Coll Cardiol 2000;36(3):970-1062.

Burkhart DM. Management of acute gastroenteritis in children. American Family Physician 1999;60(9):2555-63, 2565-6.

Centers for Disease Control and Prevention (CDC). National Diabetes Fact Sheet: National Estimates and General Information on Diabetes in the United States. Atlanta, GA: U.S. Department of Health and Human Services, 1999.

Centers for Disease Control and Prevention (CDC). Asthma mortality and hospitalization among children and young adults—United States, 1980-1993. MMWR Morb Mortal Wkly Rep 1996;45(17):350-3.

Chassin MR, Galvin RW, The urgent need to improve health care quality. Institute of Medicine National Roundtable on Health Care Quality. JAMA 1998;280(11):1000-5.

Donahue JG, Weiss ST, Livingston JM, et al. Inhaled steroids and the risk of hospitalization for asthma. JAMA 1997;277(11):887-91.

Edep ME, Shah NB, Tateo IM, et al. Differences between primary care physicians and cardiologists in management of congestive heart failure: relation to practice guidelines. J Am Coll Cardiol 1997;30(2):518-26.

Feinleib M, Rosenberg HM, Collins JG, et al. Trends in COPD morbidity and mortality in the United States. Am Rev Respir Dis 1989;140(3 pt 2):S9-18.

Foster DA, Talsma A, Furumoto-Dawson A, et al. Influenza vaccine effectiveness in preventing hospitalization for pneumonia in the elderly. Am J Epidemiol 1992;136(3):296-307.

Gaster B, Hirsch IB. The effects of improved glycemic control on complications in type 2 diabetes. Arch Intern Med 1998;158(2):134-40.

Gibbons RJ, Chatterjee K, Daley J, et al. ACC/AHA/ACP-ASIM guidelines for the management of patients with chronic stable angina: a report of the American College of Cardiology/American Heart Association Task force on Practice Guidelines (Committee on Management of Patients with Chronic Stable Angina) [published erratum appears in J Am Coll Cardiol 1999 Jul;34(1):314]. J Am Coll Cardiol 1999;33(7):2092-197.

Hackner D, Tu G, Weingarten S, et al. Guidelines in pulmonary medicine: a 25-year profile. Chest 1999;116(4):1046-62.

Harris MI. Diabetes in America: epidemiology and scope of the problem. Diabetes Care 1998;21 Suppl 3:C11-4.

Healthy People 2010. Office of Disease Prevention and Health Promotion, U.S. Department of Health and Human Services.

Hessol NA, Fuentes-Afflick E, Bacchetti P. Risk of low birth weight infants among black and white parents. Obstet Gynecol 1998;92(5):814-22.

Hiss RG. Barriers to care in non-insulin-dependent diabetes mellitus. The Michigan Experience. Ann Intern Med 1996;124(1 Pt 2):146-8.

Humphrey LL, Palumbo PJ, Butters MA, et al. The contribution of non-insulin-dependent diabetes to lower-extremity amputation in the community. Arch Intern Med 1994;154(8):885-92.

Institute of Medicine Division of Health Care Services. Medicare: a strategy for quality assurance. Washington, DC: National Academy Press; 1990.

Lin, S, Fitzgerald E, Hwang SA, et al. Asthma hospitalization rates and socioeconomic status in New York State (1987-1993) J Asthma 1999;36(3):239-51.

Mayfield JA, Reiber GE, Sanders LJ, et al. Preventive foot care in people with diabetes. Diabetes Care 1998;21(12):2161-77.

McConnochie KM, Russo MJ, McBride JT, et al. Socioeconomic variation in asthma hospitalization: excess utilization or greater need? Pediatrics 1999;103(6):375.

Millman M, editor. Committee on Monitoring Access to Personal Health Care Services. Washington, DC: National Academy Press; 1993.

Murray JL, Bernfield M. The differential effect of prenatal care on the incidence of low birth weight among blacks and whites in a prepaid health care plan. N Engl J Med 1988;319(21):1385-91.

Musey VC, Lee JK, Crawford R, et al. Diabetes in urban African-Americans. I. Cessation of insulin therapy is the major precipitating cause of diabetic ketoacidosis. Diabetes Care 1995;18(4):483-9.

National Heart, Lung, and Blood Institute/National Asthma Education and Prevention Program. Expert Panel Report 2: Guidelines for the diagnosis and management of asthma. In: National Institutes of Health pub. no. 97-4051. Bethesda, MD; 1997.
O'Campo P, Xue X, Wang MC, et al. Neighborhood risk factors for low birthweight in Baltimore: a multilevel analysis. Am J Public Health 1997;87(7):1113-8.

Patout CA, Jr., Birke JA, Horswell R, et al. Effectiveness of a comprehensive diabetes lower-extremity amputation prevention program in a predominantly low-income African-American population. Diabetes Care 2000;23(9):1339-42.

Pecoraro RE, Reiber BE, Burgess EM. Pathways to diabetic limb amputation. Basis of prevention. Diabetes Care 1990;13(5):513-21.

Philbin EF, Andreaou C, Rocco TA, et al. Patterns of angiotensin-converting enzyme inhibitor use in congestive heart failure in two community hospitals. Am J Cardio. 1996;77(1):832-8.

Ray NF, Thamer M, Fadillioglu B, et al. Race, income, urbanicity, and asthma hospitalization in California: a small area analysis. Chest 1998;113(5):1277-84.

Reiber GE, Pecoraro RE, Koepsell TD. Risk factors for amputation in patients with diabetes mellitus. A case-control study. Ann Intern Med 1992;117(2):97-105.

Reis, SE, Holubkov R, Edmundowicz D, et al. Treatment of patients admitted to the hospital with congestive heart failure: specialty-related disparities in practice patterns and outcomes. J Am Coll Cardiol 1997;30(3):733-8.

Rosenthal GE, Harper DL, Shah A, et al. A regional evaluation of variation in low-severity hospital admissions. J Gen Intern Med 1997;12(7):416-22.

Selby JV, Zhang D. Risk factors for lower extremity amputation in persons with diabetes. Diabetes Care 1995;18(4):509-16.

Silver MP, Babitz ME, Magill MK. Ambulatory care sensitive hospitalization rates in the aged Medicare population in Utah, 1990 to 1994: a rural-urban comparison. J Rural Health 1997;13(4):285-94.

Trudeau ME, Solano-McGuire SM. Evaluating the quality of COPD care. Am J. Nurs 1999;99(3):47-50.

Weinberg AD, Minaker KL. Dehydration. Evaluation and management in older adults. Council on Scientific Affairs, American Medical Association. JAMA 1995;274(19):1552-6.

Weinberger M, Oddone EZ, Henderson WG. Does increased access to primary care reduce hospital readmissions? VA Cooperative Study Group on Primary Care and Hospital readmission. N Engl J Med 1996;334(22):1441-7.

Weissman JS, Gatsonis C, Epstein AM. Rates of avoidable hospitalization by insurance status in Massachusetts and Maryland. JAMA 1992;268(17)2388-94.

Zoorob RJ, Hagen MD. Guidelines on the care of diabetic nephropathy, retinopathy and foot disease. Am Fam Physician 1997;56(8):2021-8, 2033-4.

Appendix A: Prevention Quality Indicator Definitions

Definitions use ICD-9-CM codes valid from October 1, 1994 to September 30, 2005. For ICD-9-CM codes introduced after October 1995, the date of introduction is indicated after the code label. For example, "OCT96-" indicates the ICD-9-CM code was introduced in October 1996.

Diabetes Short-term Complications Admission Rate (PQI 1)

Numerator:

Discharges with ICD-9-CM principal diagnosis code for short-term complications (ketoacidosis, hyperosmolarity, coma) (see below).

All non-maternal/non-neonatal discharges of age 18 years and older.

Exclude:

Transfer from other institution, MDC 14 (pregnancy, childbirth, and puerperium) and MDC 15 (newborns and other neonates).

Include ICD-9-CM diagnosis codes:

25010	DM KETO T2, DM CONT	25022	DM W/ HYPROSM T2, DM UNCNT
25011	DM KETO T1, DM CONT	25023	DM W/ HYPROSM T1, DM UNCNT
25012	DM KETO T2, DM UNCONT	25030	DM COMA NEC TYP II, DM CNT
25013	DM KETO T1, DM UNCONT	25031	DM COMA NEC T1, DM CONT
25020	DM W/ HYPROSM T2, DM CONT	25032	DM COMA NEC T2, DM UNCONT
25021	DM W/ HYPROSM T1, DM CONT	25033	DM COMA NEC T1, DM UNCONT

Denominator:

Population in MSA or county, age 18 years and older.

Perforated Appendix Admission Rate (PQI 2)

Numerator:

Discharges with ICD-9-CM diagnosis code for perforations or abscesses of appendix (see below) in any field.

Exclude:

Transfer from other institution, MDC 14 (pregnancy, childbirth, and puerperium) and MDC 15 (newborns and other neonates).

Include ICD-9-CM diagnosis codes (outcome of interest):

5400	AC APPEND W PERITONITIS
5401	ABSCESS OF APPENDIX

Perforated Appendix Admission Rate (PQI 2)

Denominator:

Number of discharges with diagnosis code for appendicitis in any field in MSA or county.

Include ICD-9-CM diagnosis codes (population at risk):

5400	AC APPEND W PERITONITIS	5409	ACUTE APPENDICITIS NOS
5401	ABSCESS OF APPENDIX	541	APPENDICITIS NOS

Diabetes Long-term Complications Admission Rate (PQI 3)

Numerator:

Discharges with ICD-9-CM principal diagnosis code for long-term complications (renal, eye, neurological, circulatory, or complications not otherwise specified) (see below).

All non-maternal/non-neonatal discharges of age 18 years and older.

Exclude:

Transfer from other institution, MDC 14 (pregnancy, childbirth, and puerperium) and MDC 15 (newborns and other neonates).

Include ICD-9-CM diagnosis codes:

25040	DM RENAL COMP T2 CONT	25070	DM CIRCU DIS T2 CONT
25041	DM RENAL COMP T1 CONT	25071	DM CIRCU DIS T1 CONT
25042	DM RENAL COMP T2 UNCNT	25072	DM CIRCU DIS T2 UNCNT
25043	DM RENAL COMP T1 UNCNT	25073	DM CIRCU DIS T1 UNCNT
25050	DM EYE COMP T2 CONT	25080	DM W COMP NEC T2 CONT
25051	DM EYE COMP T1 CONT	25081	DM W COMP NEC T1 CONT
25052	DM EYE COMP T2 UNCNT	25082	DM W COMP NEC T2 UNCNT
25053	DM EYE COMP T1 UNCNT	25083	DM W COMP NEC T1 UNCNT
25060	DM NEURO COMP T2 CONT	25090	DM W COMPL NOS T2 CONT
25061	DM NEURO COMP T1 CONT	25091	DM W COMPL NOS T1 CONT
25062	DM NEURO COMP T2 UNCNT	25092	DM W COMPL NOS T2 UNCNT
25063	DM NEURO COMP T1 UNCNT	25093	DM W COMPL NOS T1 UNCNT

Denominator:

Population in MSA or county, age 18 years and older.

Pediatric Asthma Admission Rate (PQI 4)

Numerator:

Discharges with ICD-9-CM principal diagnosis code of asthma (see below).

All non-maternal/non-neonatal discharges of under age 18.

Exclude:

Transfer from other institution, MDC 14 (pregnancy, childbirth, and puerperium) and MDC 15 (newborns and other neonates).

Include ICD-9-CM diagnosis codes:

49300	EXT ASTHMA W/O STAT ASTH	49321	CH OB ASTHMA W STAT ASTH
49301	EXT ASTHMA W STATUS ASTH	49322	CH OBS ASTH W ACUTE EXAC OCT00-
49302	EXT ASTHMA W ACUTE EXAC OCT00-	49381	EXERCSE IND BRONCHOSPASM OCT03-
49310	INT ASTHMA W/O STAT ASTH	49382	COUGH VARIANT ASTHMA OCT03-
49311	INT ASTHMA W STATUS ASTH	49390	ASTHMA W/O STATUS ASTHM
49312	INT ASTHMA W ACUTE EXAC OCT00-	49391	ASTHMA W STATUS ASTHMAT
49320	CH OB ASTH W/O STAT ASTH	49392	ASTHMA W ACUTE EXACERBTN OCT00-

Denominator:

Population in MSA or county, under age 18.

Chronic Obstructive Pulmonary Disease (COPD) Admission Rate (PQI 5)

Numerator:

Discharges with ICD-9-CM principal diagnosis code for COPD (see below).

All non-maternal/non-neonatal discharges of age 18 years and older.

Exclude:

Transfer from other institution, MDC 14 (pregnancy, childbirth, and puerperium) and MDC 15 (newborns and other neonates).

Include ICD-9-CM diagnosis codes:

490	BRONCHITIS NOS*	4919	CHRONIC BRONCHITIS NOS
4660	AC BRONCHITIS*	4920	EMPHYSEMATOUS BLEB
4910	SIMPLE CHR BRONCHITIS	4928	EMPHYSEMA NEC
4911	MUCOPURUL CHR BRONCHITIS	494	BRONCHIECTASIS OCT00-
49120	OBS CHR BRNC W/O ACT EXA	4940	BRONCHIECTAS W/O AC EXAC OCT00-
49121	OBS CHR BRNC W ACT EXA	4941	BRONCHIECTASIS W AC EXAC OCT00-
4918	CHRONIC BRONCHITIS NEC	496	CHR AIRWAY OBSTRUCT NEC

*Qualifies only if accompanied by secondary diagnosis of 491.xx, 492.x, 494.x or 496 (i.e., any other code on this list).

Denominator:

Population in MSA or county, age 18 years and older.

Pediatric Gastroenteritis Admission Rate (PQI 6)
Numerator: Discharges with ICD-9-CM principal diagnosis code for gastroenteritis (see below). All non-maternal/non-neonatal discharges under age 18. Exclude: Transfer from other institution, MDC 14 (pregnancy, childbirth, and puerperium) and MDC 15 (newborns and other neonates). Include ICD-9-CM diagnosis codes: 00861 ENTERITIS ROTAVIRUS 00862 ENTERITIS ADENOVIRUS 00863 ENTERITIS NORWALK VIRUS 00864 ENTERITIS OTH SML RND VIRUS 00865 ENTERITIS CALICIVIRUS 00866 ENTERITIS ASTROVIRUS 00867 ENTERITIS ENTEROVIRUS NEC 00869 ENTERITIS NOS 0088 VIRAL ENTERITIS NOS 0090 INFECTIOUS ENTERITIS NOS 0091 ENTERITIS OF INFECT ORIG 0092 INFECTIOUS DIARRHEA 0093 DIARRHEA OF PRESU INFECT ORIG 5589 NONINF GASTROENTERIT NEC
Denominator: Population in MSA or county, under age 18.

Hypertension Admission Rate (PQI 7)

Numerator:

Discharges with ICD-9-CM principal diagnosis code for hypertension (see below).

All non-maternal/non-neonatal discharges of age 18 years and older.

Exclude:

Transfer from other institution, MDC 14 (pregnancy, childbirth, and puerperium), MDC 15 (newborns and other neonates), and discharges with cardiac procedure codes (see below) in any field.

Include ICD-9-CM diagnosis codes:

4010	MALIGNANT HYPERTENSION	40310	BEN HYP REN W/O REN FAIL
4019	HYPERTENSION NOS	40390	HYP REN NOS W/O REN FAIL
40200	MAL HYPERTEN HRT DIS NOS	40400	MAL HY HT/REN W/O CHF/RF
40210	BEN HYPERTEN HRT DIS NOS	40410	BEN HY HT/REN W/O CHF/RF
40290	HYPERTENSIVE HRT DIS NOS	40490	HY HT/REN NOS W/O CHF/RF
40300	MAL HYP REN W/O REN FAIL		

Exclude ICD-9-CM procedure codes:

0050	IMPL CRT PACEMAKER SYS OCT02-	3601	PTCA-1 VESSEL W/O AGENT
0051	IMPL CRT DEFIBRILLAT OCT02-	3602	PTCA-1 VESSEL WITH AGNT
0052	IMP/REP LEAD LF VEN SYS OCT02-	3603	OPEN CORONRY ANGIOPLASTY
0053	IMP/REP CRT PACEMKR GEN OCT02-	3604	INTRCORONRY THROMB INFUS
0054	IMP/REP CRT DEFIB GENAT OCT02-	3605	PTCA-MULTIPLE VESSEL
3500	CLOSED VALVOTOMY NOS	3606	INSERT OF COR ART STENT OCT95-
3501	CLOSED AORTIC VALVOTOMY	3607	INS DRUG-ELUT CORONRY ST OCT02-
3502	CLOSED MITRAL VALVOTOMY	3609	REM OF COR ART OBSTR NEC
3503	CLOSED PULMON VALVOTOMY	3610	AORTOCORONARY BYPASS NOS
3504	CLOSED TRICUSP VALVOTOMY	3611	AORTOCOR BYPAS-1 COR ART
3510	OPEN VALVULOPLASTY NOS	3612	AORTOCOR BYPAS-2 COR ART
3511	OPN AORTIC VALVULOPLASTY	3613	AORTOCOR BYPAS-3 COR ART
3512	OPN MITRAL VALVULOPLASTY	3614	AORTCOR BYPAS-4+ COR ART
3513	OPN PULMON VALVULOPLASTY	3615	1 INT MAM-COR ART BYPASS
3514	OPN TRICUS VALVULOPLASTY	3616	2 INT MAM-COR ART BYPASS
3520	REPLACE HEART VALVE NOS	3617	ABD-CORON ART BYPASS OCT96-
3521	REPLACE AORT VALV-TISSUE	3619	HRT REVAS BYPS ANAS NEC
3522	REPLACE AORTIC VALVE NEC	362	ARTERIAL IMPLANT REVASC
3523	REPLACE MITR VALV-TISSUE	363	OTH HEART REVASCULAR
3524	REPLACE MITRAL VALVE NEC	3631	OPEN CHEST TRANS REVASC
3525	REPLACE PULM VALV-TISSUE	3632	OTH TRANSMYO REVASCULAR
3526	REPLACE PULMON VALVE NEC	3639	OTH HEART REVASULAR
3527	REPLACE TRIC VALV-TISSUE	3691	CORON VESS ANEURYSM REP
3528	REPLACE TRICUSP VALV NEC	3699	HEART VESSLE OP NEC
3531	PAPILLARY MUSCLE OPS	3731	PERICARDIECTOMY
3532	CHORDAE TENDINEAE OPS	3732	HEART ANEURYSM EXCISION
3533	ANNULOPLASTY	3733	EXC/DEST HRT LESION OPEN
3534	INFUNDIBULECTOMY	3734	EXC/DEST HRT LES OTHER
3535	TRABECUL CARNEAE CORD OP	3735	PARTIAL VENTRICULECTOMY
3539	TISS ADJ TO VALV OPS NEC	375	HEART TRANSPLANTATION (NOT VALID AFTER OCT 03)
3541	ENLARGE EXISTING SEP DEF	3751	HEART TRANPLANTATION OCT03-
3542	CREATE SEPTAL DEFECT	3752	IMPLANT TOT REP HRT SYS OCT03-
3550	PROSTH REP HRT SEPTA NOS	3753	REPL/REP THORAC UNIT HRT OCT03-
3551	PROS REP ATRIAL DEF-OPN	3754	REPL/REP OTH TOT HRT SYS OCT03-

Hypertension Admission Rate (PQI 7)			
3552	PROS REPAIR ATRIA DEF-CL	3770	INT INSERT PACEMAK LEAD
3553	PROST REPAIR VENTRIC DEF	3771	INT INSERT LEAD IN VENT
3554	PROS REP ENDOCAR CUSHION	3772	INT INSERT LEAD ATRI-VENT
3560	GRFT REPAIR HRT SEPT NOS	3773	INT INSER LEAD IN ATRIUM
3561	GRAFT REPAIR ATRIAL DEF	3774	INT OR REPL LEAD EPICAR
3562	GRAFT REPAIR VENTRIC DEF	3775	REVISION OF LEAD
3563	GRFT REP ENDOCAR CUSHION	3776	REPL TV ATRI-VENT LEAD
3570	HEART SEPTA REPAIR NOS	3777	REMOVAL OF LEAD W/O REPL
3571	ATRIA SEPTA DEF REP NEC	3778	INSER TEAM PACEMAKER SYS
3572	VENTR SEPTA DEF REP NEC	3779	REVIS OR RELOCATE POCKET
3573	ENDOCAR CUSHION REP NEC	3780	INT OR REPL PERM PACEMKR
3581	TOT REPAIR TETRAL FALLOT	3781	INT INSERT 1-CHAM, NON
3582	TOTAL REPAIR OF TAPVC	3782	INT INSERT 1-CHAM, RATE
3583	TOT REP TRUNCUS ARTERIOS	3783	INT INSERT DUAL-CHAM DEV
3584	TOT COR TRANSPOS GRT VES	3785	REPL PACEM W 1-CHAM, NON
3591	INTERAT VEN RETRN TRANSP	3786	REPL PACEM 1-CHAM, RATE
3592	CONDUIT RT VENT-PUL ART	3787	REPL PACEM W DUAL-CHAM
3593	CONDUIT LEFT VENTR-AORTA	3789	REVISE OR REMOVE PACEMAK
3594	CONDUIT ARTIUM-PULM ART	3794	IMPLT/REPL CARDDEFIB TOT
3595	HEART REPAIR REVISION	3795	IMPLT CARDIODEFIB LEADS
3596	PERC HEART VALVULOPLASTY	3796	IMPLT CARDIODEFIB GENATR
3598	OTHER HEART SEPTA OPS	3797	REPL CARDIODEFIB LEADS
3599	OTHER HEART VALVE OPS	3798	REPL CARDIODEFIB GENRATR

Denominator:

Population in MSA or county, age 18 years and older.

Congestive Heart Failure (CHF) Admission Rate (PQI 8)

Numerator:

Discharges with ICD-9-CM principal diagnosis code for CHF (see below).

All non-maternal/non-neonatal discharges of age 18 years and older.

Exclude:

Transfer from other institution, MDC 14 (pregnancy, childbirth, and puerperium), MDC 15 (newborns and other neonates), and discharges with cardiac procedure codes (see below) in any field.

Include ICD-9-CM diagnosis codes:

39891	RHEUMATIC HEART FAILURE
40201	MAL HYPERT HRT DIS W CHF
40211	BENIGN HYP HRT DIS W CHF
40291	HYPERTEN HEART DIS W CHF
40401	MAL HYPER HRT/REN W CHF
40403	MAL HYP HRT/REN W CHF/RF
40411	BEN HYPER HRT/REN W CHF
40413	BEN HYP HRT/REN W CHF/RF
40491	HYPER HRT/REN NOS W CHF
40493	HYP HT/REN NOS W CHF/RF
4280	CONGESTIVE HEART FAILURE
4281	LEFT HEART FAILURE
42820	SYSTOLIC HRT FAILURE NOS OCT02-
42821	AC SYSTOLIC HRT FAILURE OCT02-
42822	CHR SYSTOLIC HRT FAILURE OCT02-
42823	AC ON CHR SYST HRT FAIL OCT02-
42830	DIASTOLC HRT FAILURE NOS OCT02-
42831	AC DIASTOLIC HRT FAILURE OCT02-
42832	CHR DIASTOLIC HRT FAIL OCT02-
42833	AC ON CHR DIAST HRT FAIL OCT02-
42840	SYST/DIAST HRT FAIL NOS OCT02-
42841	AC SYST/DIASTOL HRT FAIL OCT02-
42842	CHR SYST/DIASTL HRT FAIL OCT02-
42843	AC/CHR SYST/DIA HRT FAIL OCT02-
4289	HEART FAILURE NOS

Exclude ICD-9-CM procedure codes:

0050	IMPL CRT PACEMAKER SYS OCT02-
0051	IMPL CRT DEFIBRILLAT OCT02-
0052	IMP/REP LEAD LF VEN SYS OCT02-
0053	IMP/REP CRT PACEMKR GEN OCT02-
0054	IMP/REP CRT DEFIB GENAT OCT02-
3500	CLOSED VALVOTOMY NOS
3501	CLOSED AORTIC VALVOTOMY
3502	CLOSED MITRAL VALVOTOMY
3503	CLOSED PULMON VALVOTOMY
3504	CLOSED TRICUSP VALVOTOMY
3510	OPEN VALVULOPLASTY NOS
3511	OPN AORTIC VALVULOPLASTY
3512	OPN MITRAL VALVULOPLASTY
3513	OPN PULMON VALVULOPLASTY
3514	OPN TRICUS VALVULOPLASTY
3520	REPLACE HEART VALVE NOS
3521	REPLACE AORT VALV-TISSUE
3522	REPLACE AORTIC VALVE NEC
3523	REPLACE MITR VALV-TISSUE
3524	REPLACE MITRAL VALVE NEC
3525	REPLACE PULM VALV-TISSUE
3526	REPLACE PULMON VALVE NEC
3527	REPLACE TRIC VALV-TISSUE
3528	REPLACE TRICUSP VALV NEC
3531	PAPILLARY MUSCLE OPS
3532	CHORDAE TENDINEAE OPS
3533	ANNULOPLASTY
3534	INFUNDIBULECTOMY
3601	PTCA-1 VESSEL W/O AGENT
3602	PTCA-1 VESSEL WITH AGNT
3603	OPEN CORONRY ANGIOPLASTY
3604	INTRCORONRY THROMB INFUS
3605	PTCA-MULTIPLE VESSEL
3606	INSERT OF COR ART STENT OCT95-
3607	INS DRUG-ELUT CORONRY ST OCT02-
3609	REM OF COR ART OBSTR NEC
3610	AORTOCORONARY BYPASS NOS
3611	AORTOCOR BYPAS-1 COR ART
3612	AORTOCOR BYPAS-2 COR ART
3613	AORTOCOR BYPAS-3 COR ART
3614	AORTCOR BYPAS-4+ COR ART
3615	1 INT MAM-COR ART BYPASS
3616	2 INT MAM-COR ART BYPASS
3617	ABD-CORON ART BYPASS OCT96-
3619	HRT REVAS BYPS ANAS NEC
362	ARTERIAL IMPLANT REVASC
363	OTH HEART REVASCULAR
3631	OPEN CHEST TRANS REVASC
3632	OTH TRANSMYO REVASCULAR
3639	OTH HEART REVASULAR
3691	CORON VESS ANEURYSM REP
3699	HEART VESSLE OP NEC
3731	PERICARDIECTOMY
3732	HEART ANEURYSM EXCISION
3733	EXC/DEST HRT LESION OPEN
3734	EXC/DEST HRT LES OTHER

Congestive Heart Failure (CHF) Admission Rate (PQI 8)			
3535	TRABECUL CARNEAE CORD OP	3735	PARTIAL VENTRICULECTOMY
3539	TISS ADJ TO VALV OPS NEC	375	HEART TRANSPLANTATION (NOT VALID AFTER OCT 03)
3541	ENLARGE EXISTING SEP DEF	3751	HEART TRANPLANTATION OCT03-
3542	CREATE SEPTAL DEFECT	3752	IMPLANT TOT REP HRT SYS OCT03-
3550	PROSTH REP HRT SEPTA NOS	3753	REPL/REP THORAC UNIT HRT OCT03-
3551	PROS REP ATRIAL DEF-OPN	3754	REPL/REP OTH TOT HRT SYS OCT03-
3552	PROS REPAIR ATRIA DEF-CL	3770	INT INSERT PACEMAK LEAD
3553	PROST REPAIR VENTRIC DEF	3771	INT INSERT LEAD IN VENT
3554	PROS REP ENDOCAR CUSHION	3772	INT INSERT LEAD ATRI-VENT
3560	GRFT REPAIR HRT SEPT NOS	3773	INT INSER LEAD IN ATRIUM
3561	GRAFT REPAIR ATRIAL DEF	3774	INT OR REPL LEAD EPICAR
3562	GRAFT REPAIR VENTRIC DEF	3775	REVISION OF LEAD
3563	GRFT REP ENDOCAR CUSHION	3776	REPL TV ATRI-VENT LEAD
3570	HEART SEPTA REPAIR NOS	3777	REMOVAL OF LEAD W/O REPL
3571	ATRIA SEPTA DEF REP NEC	3778	INSER TEAM PACEMAKER SYS
3572	VENTR SEPTA DEF REP NEC	3779	REVIS OR RELOCATE POCKET
3573	ENDOCAR CUSHION REP NEC	3780	INT OR REPL PERM PACEMKR
3581	TOT REPAIR TETRAL FALLOT	3781	INT INSERT 1-CHAM, NON
3582	TOTAL REPAIR OF TAPVC	3782	INT INSERT 1-CHAM, RATE
3583	TOT REP TRUNCUS ARTERIOS	3783	INT INSERT DUAL-CHAM DEV
3584	TOT COR TRANSPOS GRT VES	3785	REPL PACEM W 1-CHAM, NON
3591	INTERAT VEN RETRN TRANSP	3786	REPL PACEM 1-CHAM, RATE
3592	CONDUIT RT VENT-PUL ART	3787	REPL PACEM W DUAL-CHAM
3593	CONDUIT LEFT VENTR-AORTA	3789	REVISE OR REMOVE PACEMAK
3594	CONDUIT ARTIUM-PULM ART	3794	IMPLT/REPL CARDDEFIB TOT
3595	HEART REPAIR REVISION	3795	IMPLT CARDIODEFIB LEADS
3596	PERC HEART VALVULOPLASTY	3796	IMPLT CARDIODEFIB GENATR
3598	OTHER HEART SEPTA OPS	3797	REPL CARDIODEFIB LEADS
3599	OTHER HEART VALVE OPS	3798	REPL CARDIODEFIB GENRATR

Denominator:

Population in MSA or county, age 18 years and older.

Low Birth Weight Rate (PQI 9)

Numerator:

Number of births with ICD-9-CM diagnosis code for less than 2500 grams (see below) in any field.

Exclude:

Transfer from other institution.

Include ICD-9-CM diagnosis codes:

76400	LIGHT-FOR-DATES WTNOS	76490	FET GROWTH RETARD WTNOS
76401	LIGHT-FOR-DATES <500G	76491	FET GROWTH RETARD <500G
76402	LT-FOR-DATES 500-749G	76492	FET GROWTH RET 500-749G
76403	LT-FOR-DATES 750-999G	76493	FET GROWTH RET 750-999G
76404	LT-FOR-DATES 1000-1249G	76494	FET GRWTH RET 1000-1249G
76405	LT-FOR-DATES 1250-1499G	76495	FET GRWTH RET 1250-1499G
76406	LT-FOR-DATES 1500-1749G	76496	FET GRWTH RET 1500-1749G
76407	LT-FOR-DATES 1750-1999G	76497	FET GRWTH RET 1750-1999G
76408	LT-FOR-DATES 2000-2499G	76498	FET GRWTH RET 2000-2499G
76410	LT-FOR-DATE W/MAL WTNOS	76500	EXTREME IMMATUR WTNOS
76411	LT-FOR-DATE W/MAL <500G	76501	EXTREME IMMATUR <500G
76412	LT-DATE W/MAL 500-749G	76502	EXTREME IMMATUR 500-749G
76413	LT-DATE W/MAL 750-999G	76503	EXTREME IMMATUR 750-999G
76414	LT-DATE W/MAL 1000-1249G	76504	EXTREME IMMAT 1000-1249G
76415	LT-DATE W/MAL 1250-1499G	76505	EXTREME IMMAT 1250-1499G
76416	LT-DATE W/MAL 1500-1749G	76506	EXTREME IMMAT 1500-1749G
76417	LT-DATE W/MAL 1750-1999G	76507	EXTREME IMMAT 1750-1999G
76418	LT-DATE W/MAL 2000-2499G	76508	EXTREME IMMAT 2000-2499G
76420	FETAL MALNUTRITION WTNOS	76510	PRETERM INFANT NEC WTNOS
76421	FETAL MALNUTRITION <500G	76511	PRETERM NEC <500G
76422	FETAL MALNUTR 500-749G	76512	PRETERM NEC 500-749G
76423	FETAL MAL 750-999G	76513	PRETERM NEC 750-999G
76424	FETAL MAL 1000-1249G	76514	PRETERM NEC 1000-1249G
76425	FETAL MAL 1250-1499G	76515	PRETERM NEC 1250-1499G
76426	FETAL MAL 1500-1749G	76516	PRETERM NEC 1500-1749G
76427	FETAL MALNUTR 1750-1999G	76517	PRETERM NEC 1750-1999G
76428	FETAL MALNUTR 2000-2499G	76518	PRETERM NEC 2000-2499G

Denominator:

All births (discharges in MDC 15 – newborns and other neonates) in MSA or county.

Dehydration Admission Rate (PQI 10)

Numerator:

Discharges with ICD-9-CM principal diagnosis code for hypovolemia (see below).

Exclude:

Transfer from other institution, MDC 14 (pregnancy, childbirth, and puerperium) and MDC 15 (newborns and other neonates).

Include ICD-9-CM diagnosis code:

2765 HYPOVOLEMIA

Denominator:

Population in MSA or county.

Bacterial Pneumonia Admission Rate (PQI 11)

Numerator:

Discharges with ICD-9-CM principal diagnosis code for bacterial pneumonia (see below).

Exclude:

Transfer from other institution, MDC 14 (pregnancy, childbirth, and puerperium), MDC 15 (newborns and other neonates), and discharges with diagnosis code for sickle cell anemia or HB-S disease (see below) in any field.

Include ICD-9-CM diagnosis codes:

481	PNEUMOCOCCAL PNEUMONIA	4829	BACTERIAL PNEUMONIA NOS
4822	H.INFLUENZAE PNEUMONIA	4830	MYCOPLASMA PNEUMONIA
48230	STREP PNEUMONIA UNSPEC	4831	CHLAMYDIA PNEUMONIA OCT96-
48231	GRP A STREP PNEUMONIA	4838	OTH SPEC ORG PNEUMONIA
48232	GRP B STREP PNEUMONIA	485	BRONCOPNEUMONIA ORG NOS
48239	OTH STREP PNEUMONIA	486	PNEUMONIA, ORGANISM NOS
481	PNEUMOCOCCAL PNEUMONIA	4829	BACTERIAL PNEUMONIA NOS
4822	H.INFLUENZAE PNEUMONIA	4830	MYCOPLASMA PNEUMONIA

Exclude ICD-9-CM diagnosis codes:

28241	THLASEMA HB-S W/O CRISIS OCT03-	28263	SICKLE-CELL/HB-C DISEASE
28242	THLASSEMIA HB-S W CRISIS OCT03-	28264	HB-S/HB-C DIS W CRISIS OCT03-
28260	SICKLE-CELL ANEMIA NOS	28268	HB-S DIS W/O CRISIS NEC OCT03-
28261	HB-S DISEASE W/O CRISIS	28269	SICKLE-CELL ANEMIA NEC
28262	HB-S DISEASE WITH CRISIS		

Denominator:

Population in MSA or county.

Urinary Tract Infection Admission Rate (PQI 12)

Numerator:

Discharges with ICD-9-CM principal diagnosis code of urinary tract infection (see below).

Exclude:

Transfer from other institution, MDC 14 (pregnancy, childbirth, and puerperium) and MDC 15 (newborns and other neonates).

Include ICD-9-CM diagnosis codes:

59000	CHR PYELONEPHRITIS NOS	59080	PYELONEPHRITIS NOS
59001	CHR PYELONEPH W MED NECR	59081	PYELONEPHRIT IN OTH DIS
59010	AC PYELONEPHRITIS NOS	5909	INFECTION OF KIDNEY NOS
59011	AC PYELONEPHR W MED NECR	5950	ACUTE CYSTITIS
5902	RENAL/PERIRENAL ABSCESS	5959	CYSTITIS NOS
5903	PYELOURETERITIS CYSTICA	5990	URIN TRACT INFECTION NOS

Denominator:

Population in MSA or county.

Angina without Procedure Admission Rate (PQI 13)

Numerator:

Discharges with ICD-9-CM principal diagnosis code for angina (see below).

All non-maternal/non-neonatal discharges of age 18 years and older.

Exclude:

Transfer from other institution, MDC 14 (pregnancy, childbirth, and puerperium), MDC 15 (newborns and other neonates), and discharges with a procedure code for cardiac procedure (see below) in any field.

Include ICD-9-CM diagnosis codes:

4111	INTERMED CORONARY SYND	4130	ANGINA DECUBITUS
41181	CORONARY OCCLSN W/O MI	4131	PRINZMETAL ANGINA
41189	AC ISCHEMIC HRT DIS NEC	4139	ANGINA PECTORIS NEC/NOS

Exclude ICD-9-CM procedure codes:

0050	IMPL CRT PACEMAKER SYS OCT02-	3601	PTCA-1 VESSEL W/O AGENT
0051	IMPL CRT DEFIBRILLAT OCT02-	3602	PTCA-1 VESSEL WITH AGNT
0052	IMP/REP LEAD LF VEN SYS OCT02-	3603	OPEN CORONRY ANGIOPLASTY
0053	IMP/REP CRT PACEMKR GEN OCT02-	3604	INTRCORONRY THROMB INFUS
0054	IMP/REP CRT DEFIB GENAT OCT02-	3605	PTCA-MULTIPLE VESSEL
3500	CLOSED VALVOTOMY NOS	3606	INSERT OF COR ART STENT OCT95-
3501	CLOSED AORTIC VALVOTOMY	3607	INS DRUG-ELUT CORONRY ST OCT02-
3502	CLOSED MITRAL VALVOTOMY	3609	REM OF COR ART OBSTR NEC
3503	CLOSED PULMON VALVOTOMY	3610	AORTOCORONARY BYPASS NOS

Angina without Procedure Admission Rate (PQI 13)			
3504	CLOSED TRICUSP VALVOTOMY	3611	AORTOCOR BYPAS-1 COR ART
3510	OPEN VALVULOPLASTY NOS	3612	AORTOCOR BYPAS-2 COR ART
3511	OPN AORTIC VALVULOPLASTY	3613	AORTOCOR BYPAS-3 COR ART
3512	OPN MITRAL VALVULOPLASTY	3614	AORTCOR BYPAS-4+ COR ART
3513	OPN PULMON VALVULOPLASTY	3615	1 INT MAM-COR ART BYPASS
3514	OPN TRICUS VALVULOPLASTY	3616	2 INT MAM-COR ART BYPASS
3520	REPLACE HEART VALVE NOS	3617	ABD-CORON ART BYPASS OCT96-
3521	REPLACE AORT VALV-TISSUE	3619	HRT REVAS BYPS ANAS NEC
3522	REPLACE AORTIC VALVE NEC	362	ARTERIAL IMPLANT REVASC
3523	REPLACE MITR VALV-TISSUE	363	OTH HEART REVASCULAR
3524	REPLACE MITRAL VALVE NEC	3631	OPEN CHEST TRANS REVASC
3525	REPLACE PULM VALV-TISSUE	3632	OTH TRANSMYO REVASCULAR
3526	REPLACE PULMON VALVE NEC	3639	OTH HEART REVASULAR
3527	REPLACE TRIC VALV-TISSUE	3691	CORON VESS ANEURYSM REP
3528	REPLACE TRICUSP VALV NEC	3699	HEART VESSLE OP NEC
3531	PAPILLARY MUSCLE OPS	3731	PERICARDIECTOMY
3532	CHORDAE TENDINEAE OPS	3732	HEART ANEURYSM EXCISION
3533	ANNULOPLASTY	3733	EXC/DEST HRT LESION OPEN
3534	INFUNDIBULECTOMY	3734	EXC/DEST HRT LES OTHER
3535	TRABECUL CARNEAE CORD OP	3735	PARTIAL VENTRICULECTOMY
3539	TISS ADJ TO VALV OPS NEC	375	HEART TRANSPLANTATION (NOT VALID AFTER OCT 03)
3541	ENLARGE EXISTING SEP DEF	3751	HEART TRANPLANTATION OCT03-
3542	CREATE SEPTAL DEFECT	3752	IMPLANT TOT REP HRT SYS OCT03-
3550	PROSTH REP HRT SEPTA NOS	3753	REPL/REP THORAC UNIT HRT OCT03-
3551	PROS REP ATRIAL DEF-OPN	3754	REPL/REP OTH TOT HRT SYS OCT03-
3552	PROS REPAIR ATRIA DEF-CL	3770	INT INSERT PACEMAK LEAD
3553	PROST REPAIR VENTRIC DEF	3771	INT INSERT LEAD IN VENT
3554	PROS REP ENDOCAR CUSHION	3772	INT INSERT LEAD ATRI-VENT
3560	GRFT REPAIR HRT SEPT NOS	3773	INT INSER LEAD IN ATRIUM
3561	GRAFT REPAIR ATRIAL DEF	3774	INT OR REPL LEAD EPICAR
3562	GRAFT REPAIR VENTRIC DEF	3775	REVISION OF LEAD
3563	GRFT REP ENDOCAR CUSHION	3776	REPL TV ATRI-VENT LEAD
3570	HEART SEPTA REPAIR NOS	3777	REMOVAL OF LEAD W/O REPL
3571	ATRIA SEPTA DEF REP NEC	3778	INSER TEAM PACEMAKER SYS
3572	VENTR SEPTA DEF REP NEC	3779	REVIS OR RELOCATE POCKET
3573	ENDOCAR CUSHION REP NEC	3780	INT OR REPL PERM PACEMKR
3581	TOT REPAIR TETRAL FALLOT	3781	INT INSERT 1-CHAM, NON
3582	TOTAL REPAIR OF TAPVC	3782	INT INSERT 1-CHAM, RATE
3583	TOT REP TRUNCUS ARTERIOS	3783	INT INSERT DUAL-CHAM DEV
3584	TOT COR TRANSPOS GRT VES	3785	REPL PACEM W 1-CHAM, NON
3591	INTERAT VEN RETRN TRANSP	3786	REPL PACEM 1-CHAM, RATE
3592	CONDUIT RT VENT-PUL ART	3787	REPL PACEM W DUAL-CHAM
3593	CONDUIT LEFT VENTR-AORTA	3789	REVISE OR REMOVE PACEMAK
3594	CONDUIT ARTIUM-PULM ART	3794	IMPLT/REPL CARDDEFIB TOT
3595	HEART REPAIR REVISION	3795	IMPLT CARDIODEFIB LEADS
3596	PERC HEART VALVULOPLASTY	3796	IMPLT CARDIODEFIB GENATR
3598	OTHER HEART SEPTA OPS	3797	REPL CARDIODEFIB LEADS
3599	OTHER HEART VALVE OPS	3798	REPL CARDIODEFIB GENRATR

Denominator:

Population in MSA or county, age 18 years and older.

Uncontrolled Diabetes Admission Rate (PQI 14)

Numerator:

Discharges with ICD-9-CM principal diagnosis code for uncontrolled diabetes, without mention of a short-term or long-term complication (see below).

All non-maternal/non-neonatal discharges of age 18 years and older.

Exclude:

Transfer from other institution, MDC 14 (pregnancy, childbirth, and puerperium) and MDC 15 (newborns and other neonates).

Include ICD-9-CM diagnosis codes:

25002	DM, T2, UNCONT
25003	DM, T1, UNCONT

Denominator:

Population in MSA or county, age 18 years and older.

May be combined with diabetes short-term complications as a single indicator as a simple sum of the rates to form the Health People 2010 indicator (note that the AHRQ QI excludes transfers to avoid double counting cases).

Adult Asthma Admission Rate (PQI 15)

Numerator:

Discharges with ICD-9-CM principal diagnosis code of asthma (see below).

All non-maternal/non-neonatal discharges age 18 years and older.

Exclude:

Transfer from other institution, MDC 14 (pregnancy, childbirth, and puerperium) and MDC 15 (newborns and other neonates).

Include ICD-9-CM diagnosis codes:

49300	EXT ASTHMA W/O STAT ASTH	49321	CH OB ASTHMA W STAT ASTH
49301	EXT ASTHMA W STATUS ASTH	49322	CH OBS ASTH W ACUTE EXAC OCT00-
49302	EXT ASTHMA W ACUTE EXAC OCT00-	49381	EXERCSE IND BRONCHOSPASM OCT03-
49310	INT ASTHMA W/O STAT ASTH	49382	COUGH VARIANT ASTHMA OCT03-
49311	INT ASTHMA W STATUS ASTH	49390	ASTHMA W/O STATUS ASTHM
49312	INT ASTHMA W ACUTE EXAC OCT00-	49391	ASTHMA W STATUS ASTHMAT
49320	CH OB ASTH W/O STAT ASTH	49392	ASTHMA W ACUTE EXACERBTN OCT00-

Denominator:

Population in MSA or county, age 18 years and older.

Rate of Lower-extremity Amputation among Patients with Diabetes (PQI 16)

Numerator:

Discharges with ICD-9-CM procedure code for lower-extremity amputation (see below) in any field and diagnosis code of diabetes in any field (see below).

All non-maternal/non-neonatal discharges of age 18 years and older.

Exclude:

Transfer from other institution, MDC 14 (pregnancy, childbirth, and puerperium), MDC 15 (newborns and other neonates), and trauma diagnosis code (see below) in any field.

Include ICD-9-CM procedure codes:

8410	LOWER LIMB AMPUTAT NOS	8415	BELOW KNEE AMPUTAT NEC
8411	TOE AMPUTATION	8416	DISARTICULATION OF KNEE
8412	AMPUTATION THROUGH FOOT	8417	ABOVE KNEE AMPUTATION
8413	DISARTICULATION OF ANKLE	8418	DISARTICULATION OF HIP
8414	AMPUTAT THROUGH MALLEOLI	8419	HINDQUARTER AMPUTATION

ICD-9-CM diagnosis codes for diabetes:

25000	DMII WO CMP NT ST UNCNTR	25050	DMII OPHTH NT ST UNCNTRL
25001	DMI WO CMP NT ST UNCNTRL	25051	DMI OPHTH NT ST UNCNTRLD
25002	DMII WO CMP UNCNTRLD	25052	DMII OPHTH UNCNTRLD
25003	DMI WO CMP UNCNTRLD	25053	DMI OPHTH UNCNTRLD
25010	DMII KETO NT ST UNCNTRLD	25060	DMII NEURO NT ST UNCNTRL
25011	DMI KETO NT ST UNCNTRLD	25061	DMI NEURO NT ST UNCNTRLD
25012	DMII KETOACD UNCONTROLD	25062	DMII NEURO UNCNTRLD
25013	DMI KETOACD UNCONTROLD	25063	DMI NEURO UNCNTRLD
25020	DMII HPRSM NT ST UNCNTRL	25070	DMII CIRC NT ST UNCNTRLD
25021	DMI HPRSM NT ST UNCNTRLD	25071	DMI CIRC NT ST UNCNTRLD
25022	DMII HPROSMLR UNCONTROLD	25072	DMII CIRC UNCNTRLD
25023	DMI HPROSMLR UNCONTROLD	25073	DMI CIRC UNCNTRLD
25030	DMII O CM NT ST UNCNTRLD	25080	DMII OTH NT ST UNCNTRLD
25031	DMI O CM NT ST UNCNTRL	25081	DMI OTH NT ST UNCNTRLD
25032	DMII OTH COMA UNCONTROLD	25082	DMII OTH UNCNTRLD
25033	DMI OTH COMA UNCONTROLD	25083	DMI OTH UNCNTRLD
25040	DMII RENL NT ST UNCNTRLD	25090	DMII UNSPF NT ST UNCNTRL
25041	DMI RENL NT ST UNCNTRLD	25091	DMI UNSPF NT ST UNCNTRLD
25042	DMII RENAL UNCNTRLD	25092	DMII UNSPF UNCNTRLD
25043	DMI RENAL UNCNTRLD	25093	DMI UNSPF UNCNTRLD

Exclude ICD-9-CM diagnosis codes associated with trauma:

8950	AMPUTATION TOE	8971	AMPUTAT BK, UNILAT-COMPL
8951	AMPUTATION TOE-COMPLICAT	8972	AMPUT ABOVE KNEE, UNILAT
8960	AMPUTATION FOOT, UNILAT	8973	AMPUT ABV KN, UNIL-COMPL
8961	AMPUT FOOT, UNILAT-COMPL	8974	AMPUTAT LEG, UNILAT NOS
8962	AMPUTATION FOOT, BILAT	8975	AMPUT LEG, UNIL NOS-COMP
8963	AMPUTAT FOOT, BILAT-COMP	8976	AMPUTATION LEG, BILAT
8970	AMPUT BELOW KNEE, UNILAT	8977	AMPUTAT LEG, BILAT-COMPL

Denominator:

Population in MSA or county, age 18 years and older.

Appendix B: Detailed Methods

This appendix describes the methods used by the University of California-San Francisco (UCSF) Evidence-based Practice Center to refine the Healthcare Cost and Utilization Project (HCUP) quality indicators.

Semi-structured Interviews

The project team and previous developers of the HCUP Quality Indicators (HCUP QIs) developed a contact list of individuals associated with hospital associations, business coalitions, State data groups, and Federal agencies. This list was designed to include QI users and potential users from a broad spectrum of organizations in both the public and private sectors; it was not intended as a representative sample. All contacts were faxed an introductory letter and asked to participate as advisors on the project with a short telephone interview. This request was well received; only six out of 37 declined participation themselves without suggesting an alternative respondent. Overall, the 31 contacts phoned expressed interest in the study, offering many suggestions and comments. The composition of the 31 interviewees is as follows: three consultants, two Federal agency employees, one health plan medical director, five representatives of hospital associations, one international academic researcher, four representatives of private accreditation groups, two representatives of private data groups, two members of professional organizations, five representatives of provider and other private organizations, three representatives of State data groups, and three representatives of other health care organizations.

The semi-structured interviews were designed to identify potential indicators, concerns of end users, and other factors important in the development of quality indicators that may not be captured in the published literature. Thus, academic researchers, whose work is more likely to appear in peer-reviewed journals, were reserved as peer reviewers for the final document. As a result, the results of the semi-structured interviews are not intended to be a non-biased representation of the opinions regarding quality indicators, but rather a sampling of those opinions not likely to be available in the peer-reviewed literature.

The interviewers solicited information on the development and use of quality indicators by the targeted organizations, as well as other known measures and additional contacts. Interviewers used a semi-structured interview and recorded information from the interview on a data-collection form. Further, some advisors provided the project team with materials regarding quality indicators and the use of HCUP QIs.

Quality Indicators Evaluation Framework

Six areas were considered essential for evaluating the reliability and validity of a proposed quality indicator. Several sources contributed to the development of the evaluation criteria framework: (1) results of the semi-structured interviews, including the interests and concerns of HCUP QI users, (2) task order document describing the Agency for Healthcare Research and Quality's (AHRQ) interests, (3) evidence available in the policy and research literature and (4) evidence available through statistical analyses. The six criteria were quite similar to the criteria for "testing the scientific strength of a measure" proposed by McGlynn and Asch.[1] They describe a measure as reliable "if, when repeatedly applied to the same population, the same result is obtained a high proportion of the time." They propose evaluating validity in terms of face validity, criterion validity ("an objective assessment of the ability of the measure to predict a score on some other measure that serves as the evaluation criterion"), and construct validity ("whether the correlations between the measure and other measures are of the right magnitude and in the right direction"). Criterion validity was viewed as an assessment of bias (criterion #3), where the "gold standard" measure is purged of bias due to severity of illness. Face validity captures a variety of concepts discussed by McGlynn and Siu, including the importance of the condition, the efficacy of

available treatments (e.g., the ability of providers to improve outcomes), and the potential for improvement in quality of care.[2]

Evidence supporting the use of current and candidate quality indicators was assembled in terms of the following six areas.

- Face validity: Does the indicator capture an aspect of quality that is widely regarded as important and subject to provider or public health system control?
- Precision: Is there a substantial amount of provider or community level variation that is not attributable to random variation?
- Minimum bias: Is there either little effect on the indicator of variations in patient disease severity and comorbidities, or is it possible to apply risk adjustment and statistical methods to remove most or all bias?
- Construct validity: Does the indicator perform well in identifying true (or actual) quality-of-care problems?
- Fosters real quality improvement: Is the indicator insulated from perverse incentives for providers to improve their reported performance by avoiding difficult or complex cases, or by other responses that do not improve quality of care?
- Application: Has the measure been used effectively in practice? Does it have potential for working well with other indicators?

In addition to the above framework, the Donabedian paradigm of structure, process, and outcome was followed to categorize current (HCUP) and candidate QIs.[3, 4] For example, potentially inappropriate utilization falls into the category of process, while in-hospital mortality, adverse events, and complication rates represent outcome measures.

Three broad audiences for the quality measures were considered: health care providers and managers, who would use the quality measures to assist in initiatives to improve quality; public health policy-makers, who would use the information from indicators to target public health interventions; and health care purchasers and consumers, who would potentially use the measures to guide decisions about health policies and providers. Because of the limitations of quality indicators derived based on administrative data, the focus was primarily on applications oriented to "screening for potential quality problems." For the purpose of the Evaluation Framework, indicators must at least pass tests indicating that they are appropriate for the use of screening. The rest of this section provides a more detailed explanation of each part of the Evaluation Framework, considering these three audiences wherever differences have been noted in the literature.

- Face validity: Does the indicator capture an aspect of quality that is widely regarded as important and subject to provider or public health system control?

This question considers the degree to which potential users view the quality indicator as important and informative. There are two parts to this question: Does the indicator relate to an aspect of health care that users regard as important? And does performance on the measure credibly indicate high-quality care? Obviously, face validity will be influenced by how well the indicator performs in the other areas covered in the Evaluation Framework. Clinicians tend to distrust outcome measures because of concerns over the adequacy of risk adjustment and the multiple factors beyond providers' control that contribute to poor outcomes. Other critics add that outcome measures suffer from imprecision (with random noise outweighing provider differences) and important selection biases (e.g., due to variations in admitting practices). Addressing this issue at the outset serves as a point of reference for the findings of the literature review and empirical analysis.

Broadly speaking, consumers, health care payers, regulators, and public health officials are likely to be most interested in measures based on outcomes that are relatively frequent, costly, or have serious implications for an individual's health. In addition, there should be reason to believe that the outcome may be (at least somewhat) under providers' control (in other words, controlled trials or well-designed cohort studies have shown that specific diagnostic or therapeutic modalities may reduce its frequency or severity). Outcome measures might include operative mortality rates or mortality after hospitalization with serious acute illnesses such as a heart attack. These measures seem most intuitive, since they assess the main outcomes that medical treatments are intended to affect.

Perhaps surprisingly, however, reports of hospital mortality rates appear to have little effect on where patients seek their care.[5, 6] One reason may be that many patients describe difficulty in interpreting indicators involving mortality and morbidity rates, and consequently view them as unhelpful.[7] Another reason may be that providers prefer measures of process, particularly if there is reason to believe (generally from randomized controlled trials) that certain processes truly lead to better patient outcomes. Patients appear to prefer reports of other patients' satisfaction with care, and especially informal recommendations from family, friends, and their own physicians.[7] Thus, developing indicators with high face validity for patients may require active participation from patients, targeting aspects of care identified as important in patient surveys, or taking additional steps to enhance provider perceptions about the validity of outcome measures.[8-17]

Many providers view outcome-based QIs with considerable skepticism.[18] For most outcomes, the impacts of random variation and patient factors beyond providers' control often overwhelm differences attributable to provider quality.[19-24] Consequently, providers tend to support measures of quality based on processes of care that have been documented in clinical trials to lead to better health outcomes in relatively broad groups of patients — for example, the processes of acute MI care measured in the Cooperative Cardiovascular Project.[25-30] Such process measures focus precisely on the aspects of care under providers' control. As long as the process measures are based on evidence of effectiveness, they serve as useful proxies for outcome measures that would otherwise be difficult to observe or measure. For example, when using inpatient discharge data only, it is not possible to ascertain out-of-hospital mortality. In general, process measures are not as noisy as outcome measures, because they are less subject to random variation. They also suggest specific steps that providers may take to improve outcomes or reduce costs—even if such outcome improvements are difficult to document at the level of particular providers.

The relationship between some structural quality measures and important outcomes has been well-documented, although some concerns remain about the interpretation of the measures.[3, 4, 31, 32] These measures include measures of hospital volume for volume-sensitive conditions, technological capabilities (e.g., ability to perform certain intensive procedures like coronary angioplasty), and teaching status.[33-61] All of these measures have limited face validity, because they are widely acknowledged to be weak surrogates for true quality of care.[62] For example, many low-volume hospitals have been shown to achieve excellent outcomes, whereas many high-volume hospitals have surprisingly poor outcomes.

- Precision: Is there a substantial amount of provider or community level variation that is not attributable to random variation?

The impact of chance on apparent provider or community health system performance must be considered. Unobserved patient and environmental factors may result in substantial differences in performance among providers in the absence of true quality differences. Moreover, the same providers may appear to change from year to year, in the absence of changes in the care they deliver. Thus, using "raw" quality data will often result in poorly reproducible, or imprecise, measurements, giving an incorrect impression of provider quality.

An extensive literature on the importance of random variations in quality measures now exists.[19, 21-24, 63-68] In general, random variation is most problematic when there are relatively few observations per provider, when adverse outcome rates are relatively low, and when providers have little control over

patient outcomes or when variation in important processes of care is minimal. If a large number of patient factors that are difficult to observe influence whether or not a patient has an adverse outcome, it may be difficult to separate the "quality signal" from the surrounding noise. The evidence on the precision of each of the evaluated QIs was reviewed. Empirical methods can be used to assess both the importance of sample size and the importance of provider effects (versus patient and area effects) in explaining observed variation in the measure.

But this is not entirely a statistical question, and considerations of mechanisms and concerns related to face validity can also be helpful in assessing the precision of a measure. For example, if better hospitals invariably admit sicker patients, then the apparent variation in a measure at the hospital level will be significantly less than the true variation (see the discussion of unbiasedness below). In such a case, other sources of evidence suggesting that a measure is valid or that such bias exists can be helpful in assessing the quality measure. The literature review encompasses both empirical and other sources of evidence on measure precision, and the empirical analysis presents systematic evidence on the extent of provider-level or area-level variation in each quality measure.

Statistical techniques can account for random variations in provider performance by estimating the extent to which variation across providers appears to be clustered at the provider level, versus the extent to which it can be explained by patient and area effects.[68-71] Under reasonable statistical assumptions, the resulting estimates of the extent to which quality truly varies at the provider or area level can be used to "smooth" or "shrink" estimates of the quality of specific providers or areas. The methods are Bayesian: the data used to construct the quality measures are used to update a "prior" distribution of provider quality estimates, so that the "posterior" or smoothed estimate of a provider's (or area's) quality is a best guess, reflecting the apparent patient- and provider-level (or area-level) variance of measure performance.

- Minimum bias: Is there either little effect on the indicator of variations in patient disease severity and comorbidities, or is it possible to apply risk adjustment and statistical methods to remove most or all bias?

A QI may exhibit precision, but nonetheless yield inaccurate results due to systematic measurement biases. Extensive research has documented the importance of *selection problems* in interpreting many quality measures, especially measures related to mortality.[72-76] Such biases may have two basic forms: differences in admitting practices between two hospitals produce non-random samples from the same underlying patient population (selection biases) or the patient populations may in fact contain different case-mixes. Selection effects presumably exert a greater influence on measures involving elective admissions and procedures, for which physician admission and treatment practice styles show marked variation.[56, 57] Nonetheless, selection problems exist even for conditions involving urgent "non-discretionary" admissions, likely due to modest practice variation, and non-random distribution of patient characteristics across hospital catchment areas.[59, 77] The attention of researchers and quality analysts has focused on developing valid models to adjust for patient factors, especially when comparing hospital mortality.[72, 74]

The principal statistical approach to address concerns about bias is risk adjustment.[78, 79, 60, 61, 80-86] Numerous risk adjustment instruments currently exist, but current methods are far from perfect.[79, 87] In general, risk adjustment methods are based on data drawn from administrative data and medical chart reviews.[78] Previous studies suggest that administrative data have at least two major limitations. First, coding errors and variations are common; some diagnoses are frequently entered with errors and with some inconsistency across hospitals.[88-90] Factors affecting the accuracy of these codes include restrictions on the number of secondary diagnoses permitted, as well as systematic biases in documentation and coding practices introduced by awareness that risk-adjustment and reimbursement are related to the presence of particular complications.[91-96]

Second, most administrative data sources do not distinguish disorders that can be in-hospital complications from pre-existing comorbidities.[78, 97] To the extent that diagnoses such as shock and

pulmonary edema may result from poor quality of care, their incorporation in prediction models may bias estimates of expected mortality, and even favor hospitals whose care results in more complications. One proprietary risk-adjustment system has been shown to be significantly biased by its inclusion of conditions that actually developed after admission, but this study was limited to one condition (acute MI) and its conclusions are somewhat controversial.[98, 99] In another study, estimates of mortality differences between municipal and voluntary hospitals in New York City were substantially affected by whether potential complications were excluded from risk-adjustment.[61] New York and California have recently added a "6th digit" to ICD-9-CM codes to distinguish secondary diagnoses present at admission from those that developed during hospitalization. This refinement may allow valid comparisons of risk-adjusted mortality using administrative data for certain conditions, although the accuracy of the "6th digit" has not been established.[100]

Clinically based risk adjustment systems supplement hospital discharge data with information available from medical records. Because exact clinical criteria can be specified for determining whether a diagnosis is present, coding errors are diminished. In addition, complications can be distinguished from comorbidities focusing on whether the diagnosis was present at admission.[79] Because the number of clinical variables that may potentially influence outcomes is small, and because these factors differ to some extent across diseases and procedures, progress in risk-adjustment has generally occurred by focusing on patients with specific conditions. Thus, sophisticated chart-based risk adjustment methods have been developed and applied for interpreting mortality rates for patients undergoing cardiac surgery and interventional cardiology procedures; critically ill patients; patients undergoing general surgery; and medical patients with acute myocardial infarction, community-acquired pneumonia, and upper gastrointestinal hemorrhage.[29, 36, 85, 101-107]

However, chart-based risk adjustment methods are not without their own limitations. First, especially for severely ill patients and those who die soon after admission—some of the most important patients for computing many quality measures—complete diagnosis information may not have been ascertained prior to death, and therefore would not be in the patient's medical record. Important observations might be missing for such patients, resulting in biased estimates in the risk-adjusted model. Second, medical chart reviews are very costly, and so routine collection of detailed risk information is not always feasible. As a result, the impact of chart-based risk adjustment may vary across measures. For some measures, its impact is modest and does not substantially alter relative rankings of providers.[113-116] For others, it is much more important.[79, 97, 108-112] Of course, because all risk adjustment methods generally leave a substantial amount of outcome variation unexplained, it is possible that unmeasured differences in patient mix are important even in the most detailed chart-based measures.

For each quality measure, this report reviews the evidence on whether important systematic differences in patient mix exist at the provider and community level, and whether various risk adjustments significantly alter the quality measure for particular providers. A distinction is made between risk adjustment methods that rely only on administrative data and have been validated with clinical data, and those that are not validated. Risk adjustment methods requiring clinical data cannot be applied to the HCUP data, and therefore are not covered in this report. The empirical analysis then assesses whether a common approach to risk adjustment using administrative data—the 3M™ All Patient Refined Diagnosis Related Groups (APR-DRG) system—significantly alters the quality measure for specific providers. Emphasis is placed on the impact on *relative* measures of performance (whether risk adjustment affects which hospitals are regarded as high- or low-quality) rather than *absolute* measures of performance (whether risk adjustment affects a hospital's quantitative performance on the quality measure). As noted above, this system is not ideal, because it provides only four severity levels within each base APR-DRG, omits important physiologic and functional predictors, and potentially misadjusts for iatrogenic complications. Information on the 3M™ APR-DRG system is available at http://www.3m.com/us/healthcare/his/products/coding/refined_drg.jhtml.

A remaining methodological issue concerns the appropriateness of adjusting for certain "risk factors."[117-126] For example, "Do Not Resuscitate" status may be associated with differences in care that not only reflect patient preferences (e.g., less use of intensive treatments) but also true differences in quality of care (e.g., inadequate physician visits), resulting in increased complications that would result in

a “Do Not Resuscitate” order, and increased mortality.[127] Importantly, the prevalence of patients with DNR status may vary nonrandomly between hospitals, with large referral centers having greater percentages of patients seeking (and receiving) aggressive medical care.[128]

Adjusting for race implies that patients of different races respond differently to the same treatments, when patients of different races may actually receive different treatments. A substantial literature documents systematic differences in the care delivered to patients by race and gender.[116, 129-135] For example, African-American diabetics undergo limb amputations more often than do diabetics of other races.[136] Thus, wherever possible it is noted if review of the literature indicates particularly large differences in a quality measure by race or gender. Some gender or race differences may be due to either patient preference or physiological differences that would be appropriate to include in a risk adjustment model. In other cases, differences denote lower quality care, and in this case race and gender should not be included in the risk adjustment model. Where applicable, this is noted in the literature review.

- Construct validity: Does the indicator perform well in identifying true (or actual) quality-of-care problems?

Ideally, a hospital will perform well on a quality measure if and only if it does not have a significant quality problem, and will perform poorly if and only if it does. In practice, of course, no measure performs that well. The analyses of noise and bias problems with each measure are intended to assess two of the principal reasons why a hospital might appear relatively good or bad (or not appear so) when it really is not (or really is). Detecting quality problems is further complicated by the fact that adverse outcomes are often the result of the course of an illness, rather than an indication of a quality problem at a hospital. Formally, one would like to know the sensitivity and specificity of a quality measure, or at least the positive predictive value (PPV) of a quality measure for detecting a true hospital quality problem.[119]

When available, for each measure, any existing literature was reviewed on its sensitivity or PPV for true provider quality problems. In most cases, however, no true gold standard, or ideal measure of quality, was found. Therefore, construct validity was tested, i.e., the construct is that different measures of quality, on the same patients, should be related to each other at the provider level, even if it is not always clear which measure is better. It may be easier to ask “is the indicator correlated with other, accepted measures of quality at the provider level?” rather than “does the indicator perform well in identifying providers with quality problems?” For example, studies have validated survey rankings of “best” hospitals by examining the relation with actual process and outcome measures for AMI, and peer review failure rates with HCFA risk-adjusted mortality rates.[137, 138]

- Fosters real quality improvement: Is the indicator insulated from perverse incentives for providers to improve their reported performance by avoiding difficult or complex cases, or by other responses that do not improve quality of care?

Ideally, when quality measures are used to guide quality improvement initiatives or reward good providers, the best way for a provider to perform well on the measure is to provide high-quality care. Unfortunately, many quality indicators appear to at least leave open the possibility of improving *measured* performance without improving *true* quality of care.

In measures that are risk-adjusted, measured performance can be improved by “upcoding” — including more comorbid diagnoses in order to increase apparent severity of illness.[68. 96] Systematic biases in diagnostic codes were observed after the introduction of the Prospective Payment System and may also explain much of the apparent reduction in adjusted mortality attributed to the Cardiac Surgery Reporting System in New York.[93-96] The extent to which upcoding is a problem probably increases with

[119]The PPV represents that the chance that a positive test result reflects a “true positive.” It combines the properties of the test itself (e.g., sensitivity and specificity for detecting quality problems) with the prevalence of true quality problems in the target population.

the ambiguity of the specific data element, and decreases when auditing programs maximize the reliability and validity of submitted data. In recent years, an aggressive auditing program has significantly reduced the extent to which comorbidities not substantiated by the medical chart are recorded for Medicare patients, leading some analysts to conclude that "upcoding" is no longer as substantial of a problem for Medicare patients.[139] However, such audit standards have generally not been imposed on the State discharge databases used in the HCUP project. In this review, indicators for which risk adjustment appears to be important are noted, and thus upcoding is a potentially important problem.

Indicators capturing patient morbidity, such as adverse events and complications, must overcome a reporting bias in the reverse direction (i.e., toward under-reporting). With some exceptions, most hospitals in most States rely on voluntary incident reporting for adverse events. Such methods are known to detect only a fraction of true adverse drug events (ADEs).[140] The Institute of Medicine has recently recommended mandatory reporting systems for adverse events emanating from certain egregious errors.[141] However, the JCAHO's sentinel reporting system tracks many of these same errors (e.g., operating on the wrong patient or body part, suicide or rape of an inpatient), and it was received very negatively by hospitals, despite being a voluntary system. Thus, the degree to which mandatory reporting requirements alleviate or exacerbate reporting bias for adverse events remains to be seen. In addition, high-quality hospitals with sophisticated error detection systems may report errors more frequently, leading to high apparent complication rates in hospitals that may have superior quality in other dimensions.[142-144]

Perverse incentives may arise from the criteria used to define or identify the target patient population. For instance, restricting mortality measures to inpatient deaths potentially allows hospitals to lower their mortality rates simply by discharging patients to die at home or in other institutions.[91, 100, 145, 146] Measures of surgical site infections and other complications of hospital care that only capture in-hospital events will similarly reward hospitals that merely reduce length of stay by discharging or transferring high-risk cases.[147-149] Early concerns that surgeons in New York avoided operating on high-risk patients may have proved unfounded, though this issue remains unsettled.[150-153] In general, the incentive for providers to avoid treating sicker patients remains a significant concern for outcome-based quality measures.[68]

The available evidence on each of these possible undesirable responses to the use of each quality measure was reviewed. For the most part, evidence was lacking on responses to indicators, particularly since many of the proposed indicators have not been subjected to public reporting. Potential responses were noted when appropriate.

- Application: Has the measure been used effectively in practice? Does it have potential for working well with other indicators?

While important problems exist with many specific applications of HCUP QIs and other quality indicators, they have been applied in a range of settings. As noted in the section on face validity, these applications broadly include initiatives to improve provider quality and initiatives to provide quality-related information to providers and consumers. Studies describing its use in these activities were reviewed for each quality indicator. However, a thorough review of the non-peer reviewed literature was not conducted. Therefore, indicators may have been adopted, and may continue to be used, by many provider organizations or Government agencies.

A recent systematic review more comprehensively summarizes the literature on the impact of performance reports on consumers, providers, and purchasers.[154] Useful and accurate information on quality remains a desirable goal for consumers and providers alike. The interest in quality and the resulting data and research has had some impact on the field of health services research. For instance, the HCUP project has provided a valuable resource for a number of studies in health services research.[124-126, 155-169]

Literature Review of Quality Indicators

A literature review was conducted to identify quality indicators reported as such and potential quality measures. The result of this first stage was a comprehensive list of measures that could be defined based on routinely collected hospital discharge data. In the second phase, the literature was searched for further evidence on these indicators to provide information on their suitability for the new QI set. This second phase resulted in a comprehensive bibliography for each indicator. In addition, a sub-set of the entire indicator list was selected for detailed review using specific evaluation criteria. The entire process for this systematic review of the literature is described in the following sections.

Phase 1: Identification of Indicators

Step 1: Selecting the articles. To locate literature pertaining to quality indicators, a strategic literature search was conducted using the Medline database. Over 30 search strategies were compared using Medical Subject Headings (MeSH) based on their ability to retrieve a set of key articles known to the project team. Successful combinations of MeSH term searches returned all the key articles. The final MeSH terms used were "hospital, statistic and methods" and "quality indicators." Articles were also limited to those published in 1994 or later. Articles prior to 1994 had been reviewed for the original QI development. This search returned approximately 2,600 articles—the highest number of known key articles in the most concise manner.

Articles were screened using the titles and abstracts for preliminary abstraction. To qualify for preliminary abstraction, the articles must have described a potential indicator or quality relationship that could be adequately defined using administrative data, and be generalizable to a national data set. For the purpose of this study, a quality indicator was defined as an explicit measure (defined by the developer) of some aspect of health care quality. Some literature defines only a quality relationship, in that the article expounds on a process or structural aspect of a health care provider that is related to better outcomes. However, the author does not specifically define or recommend that the relationship be used as a quality measure. In this case, the article only describes a quality relationship, not a quality indicator. Only 181 articles met the criteria for preliminary abstraction. This reflects the small number of quality indicators with published formal peer-reviewed evaluations.

Step 2: Preliminary abstraction. The preliminary round was designed to screen articles for applicability and quality, to obtain and assess the clinical rationale of the indicators, and to identify those articles with enough detail for a more comprehensive abstraction. Nine abstractors participated in this phase. Five of these abstractors were medical doctors with health services research training. The remaining four abstractors were familiar with the project and the literature, and included a project manager, the research coordinator, and two undergraduate research assistants.

The articles were sorted into clinical groupings. The research coordinator rated these clinical groupings according to the amount of clinical knowledge required to abstract the articles. Those requiring the most clinical knowledge were assigned to physicians, while those requiring the least clinical knowledge were assigned to the undergraduate research assistants. Abstractors selected clinical groupings that were of interest or that corresponded to their clinical specialties.

Abstractors recorded information about each article on a one-page abstraction form. Information coded included:

- Indicator type (i.e. mortality, readmission, potentially overused procedures)
- Clinical domain (i.e. medical, surgical, obstetric, pediatric, and psychiatric)
- Measure category (i.e. structure, process, proxy-outcome, and outcome)
- Clinical rationale for the indicators
- Use of longitudinal data

- Use of data beyond hospital discharge data
- Strengths and weaknesses identified by the author
- Strengths and weaknesses not identified by the author.

Each abstraction form was reviewed by the research coordinator for quality of the abstraction and for accuracy of the coding. All data were then entered into a Microsoft Access database.

Step 3: Full abstraction. The purpose of the full abstraction phase was to identify potential indicators for the new QI set, and to assess the evidence for validity of existing indicators. To accomplish this, only articles that described an indicator in conjunction with specific and comprehensive information on its validity were fully abstracted. Four of the original abstractors participated in this phase of the abstraction. Three of these abstractors were medical doctors, the fourth a master's level research coordinator.

Each of the articles for preliminary abstraction and the corresponding abstraction form was reviewed by both the research coordinator and the project manager independently. To qualify for full abstraction, the articles needed to meet the previously noted criteria and the following criteria:

- Define a quality indicator, as opposed to only a relationship that was not formulated or explicitly proposed as a measurement tool
- Discuss a novel indicator, as opposed to indicators defined elsewhere and used in the article only to discuss its relationship with another variable (i.e., socioeconomic status, race, urbanization)
- Define an indicator based on administrative data only.

Only 27 articles met these formal criteria. This highlights an important aspect of the literature on quality indicators: most indicators are based on published clinical literature to identify important patient and provider characteristics and processes of care for specific clinical conditions; there is also a substantial literature on technical aspects such as severity adjustment, coding, and data collection. It should be noted that, while only 27 articles qualified for formal abstraction, these are not the only useful articles. Many articles provide important information about quality measurement. However, few quality indicators are specifically defined, evaluated, and reported in the literature besides descriptive information on the process of development. (The Complication Screening Program is a noteworthy and laudable exception that has been extensively validated in the published literature, mostly by the developers). This evidence report will be an important contribution to the paucity of literature on indicator validation.

An abstraction form was filled out for each indicator defined in an article. The abstraction form coded the following information:

- All the information coded in the preliminary abstraction form
- Measure administrative information (i.e. developer, measure set name, year published)
- Level of care (primary (prevention), secondary (screening or early detection) or tertiary (treatment to prevent mortality/morbidity))
- Scoring method (i.e. rate, ratio, mean, proportion)
- A priori suggested quality standard (i.e. accepted benchmark, external comparison, and internal comparison)
- Indicator definition (numerator, denominator statements, inclusions, and exclusions)
- Extent of prior use
- Current status (i.e. measure defined, pilot tested, implemented, discontinued)

- Scientific support for measure (i.e. published guidelines, clinician panel, literature review, revision of pre-existing instruments, theory only)
- Other essential references for the measure
- Validity testing
- Risk adjustment.

If the measure included risk adjustment, a separate form for the risk adjustment method was filled out. This included:

- Method administrative information
- Adjustment rationale
- Classification or analytic approach (i.e. stratification, logistic or linear regression)
- System development method (i.e. logistic regression, score based on empirical model, a priori/clinical judgment)
- Published performance for discrimination and calibration
- Use of comorbidities, severity of illness, or patients demographics
- Use of longitudinal data, or additional data sources beyond discharge data
- Extent of current use
- Other essential references for the method
- Abstractor comments.

The abstraction forms were reviewed by the research coordinator and entered into a Microsoft Access database.

Parallel Step: Supplementing literature review using other sources. Because the literature in this area is not the primary source for reporting the use of quality indicators, a list of suitable indicators was compiled from a variety of sources. As previously noted, the phone interviews with project advisors led to information on some indicators. In addition, the Internet sites of known organizations using quality indicators; the CONQUEST database; National Library of Healthcare Indicators (NLHI), developed by the Joint Commission on Accreditation of Healthcare Organizations (JCAHO); and a list of ORYX-approved indicators provided by the JCAHO were searched. Indicators that could be defined using administrative data were recorded in an indicator database.

Breakdown of indicators by primary source. During Phase 1, no one literature search was sufficiently sensitive for the purpose of identifying either quality indicators or quality relationships. In addition, there was relatively little literature defining quality indicators. Web sites, organizations, and additional literature describing quality indicators were searched to be confident that a large percentage of the quality indicators in use were identified. In general, most volume, utilization, and ACSC indicators have been described primarily in the literature. On the other hand, the primary sources for most mortality and length of stay indicators were current users or databases of indicators. However, many indicators found in the literature were also reported by organizations, and vice versa. Thus, it is difficult to delineate which indicators were derived only from the literature and which were derived from the parallel step described above.

Phase 2: Evaluation of Indicators

The result of Phase 1 was a list of potential indicators with varied information on each depending on the source. Since each indicator relates to an area that potentially screens for quality issues, a

structured evaluation framework was developed to determine measurement performance. A series of literature searches were then conducted to assemble the available scientific evidence on the quality relationship each indicator purported to measure. Due to limited resources, not all of the indicators identified in Phase 1 could be reviewed, and therefore some were selected for detailed review using the evaluation framework. The criteria used to select these indicators are described later.

Step 1. Development of evaluation framework. As described previously, a structured evaluation of each indicator was developed and applied to assess indicator performance in six areas:

- Face validity
- Precision
- Minimum bias
- Construct validity
- Fosters real quality improvement
- Prior use.

Step 2. Identification of the evidence. The literature was searched for evidence in each of the six areas of indicator performance described above, and in the clinical areas addressed by the indicators. The search strategy used for Phase 2 began with extensive electronic searching of MEDLINE, PsycINFO, and the Cochrane Library.[170-172] (A decision was made not to search EMBASE on the grounds that the studies of quality measurement necessarily must take into account the particular health care system involved.[173]) In contrast to conducting systematic reviews of purely clinical topics, it was reasoned that the European literature not captured in the Medline database or Cochrane Library would almost certainly represent studies of questionable relevance to the U.S. health system.

The extensive electronic search strategy involved combinations of MeSH terms and keywords pertaining to clinical conditions, study methodology, and quality measurement (Figure B-1).

Additional literature searches were conducted using specific measure sets as “keywords”. These included “Maryland Quality Indicators Project,” “HEDIS and low birth weight, or cesarean section, or frequency, or inpatient utilization,” “IMSystem,” “DEMPAQ,” and “Complications Screening Program.”

The bibliographies of key articles were searched, and the Tables of Contents of general medical journals were hand searched, as well as journals focusing in health services research or in quality measurement. This list of journals included *Medical Care*, *Health Services Research*, *Health Affairs*, *Milbank Quarterly*, *Inquiry*, *International Journal for Quality in Healthcare*, and *the Joint Commission Journal on Quality Improvement.* These literature searches and on-line screening for relevancy retrieved over 2,000 additional articles, which were added to the project database. These articles were used for evaluations of individual indicators.

The use of medical literature databases likely eliminated much of the “gray literature” that may be applicable to this study. Given the limitations and scope of this study, a formal search of the “gray literature” was not completed beyond that which was previously known by the project team or resulted from telephone interviews.

Figure B-1: Example Search

Mortality Following Stroke	
Medline Search String	**Number of References Retrieved**
1. Cerebrovascular disorders [MeSH terms]	47,264
2. Epidemiologic studies [MeSH terms] OR clinical trials [MeSH terms]	32,630
3. Search mortality [MeSH Terms] OR prognosis [MeSH terms]	18,460
4. #1 AND #2 AND #3	2,410
5. #4 AND stroke [title]	524
6. Quality of health care [MeSH term]	852,714
7. #1 AND #2 AND (#3 OR #6)	1.988
8. Reproducibility of results [MeSH terms] OR sensitivity and specificity [MeSH terms]	110,384
9. Records [MeSH terms] OR hospitalization [MeSH terms]	55,739
10. #8 AND #9	3,835
11. #1 AND #10	106
Note: The results of searches 5 and 11 were scanned (titles and abstracts) to pull relevant studies, and the bibliographies of these studies were hand-searched for additional references. All searches included limits: Publication date from 1990 to 2000 and language English.	

Step 3. Selection of a sub-set of indicators. Since there were too many indicators identified in Phase 1 (literature search and parallel steps) for detailed evaluation using the Evaluation Framework, criteria were developed to select a group for further evaluation. These criteria were intended to be top-level evaluations of the face validity and precision of the indicators. A subset of indicators was selected for preliminary empirical evaluation. To do this, first the indicators related to complications were disqualified for this particular report, since they will be included in an expansion to the report that will include patient safety indicators. Second, all of the current HCUP QIs (except those related to complications of care) were selected for empirical evaluation. Third, the priority of clinical areas well covered by the current HCUP indicator set was lowered (for example, obstetrical indicators). Finally, a set of criteria for selection was applied to the remaining indicators.

The following were specific criteria for evaluation for all indicators:

Indicator must be definable with HCUP data (i.e., uses only administrative data available in HCUP data set).

- Conditions that affect at least 1% of hospitalized patients or 20% of providers, as tested using the Nationwide Inpatient Sample data set
- Conditions that are the subject of public reporting, previous use, or large dollar volume
- Clear relationship to quality apparent as evaluated by clinical judgment of health services researchers and medical doctors.

In addition, several specific criteria were noted for the indicator types:

- Volume:
  - < Widely documented volume-outcome relationship
  - < Recent evidence regarding volume-outcome relationship

- Utilization rates:
  - < Condition must have an alternative surgical or medical therapy with lower/higher morbidity or mortality.
- Ambulatory care sensitive conditions:
  - < Differences in patient management practices for that condition
  - < Existence of treatment guidelines, and evidence of failure to comply.
- In-hospital mortality
  - < Relatively homogenous group.

When selecting between competing alternatives that met all the above criteria, the choice was made to evaluate clinical areas in depth rather than evaluating a large breadth of indicators. To do this, multiple aspects in one clinical domain were evaluated (i.e., evaluations of CABG, PTCA, and AMI; stroke and carotid endarterectomy). In these clinical areas, at least two different types of indicators were evaluated (i.e., mortality and utilization).

The selected indicators were then evaluated empirically, using preliminary tests of precision. Those demonstrating adequate precision were then evaluated by a literature review (Phase 2), as well as further empirical analysis.

Step 4. Evaluation of evidence. The abstracts from relevant articles for each indicator were reviewed and selected according to the following criteria:

- The article addressed some aspect of the six areas of indicator performance.
- The article was relevant to a national sample, rather than a local population.

Based on this literature, a team member or clinician developed a draft write-up of the indicator following the evaluation framework. The literature review strategy is depicted in the flow diagram in Figure B-2.

Risk Adjustment of HCUP Quality Indicators

"Raw" unadjusted measures of hospital or area performance for each indicator are simple means constructed from the HCUP discharge data and census population counts. Obviously, simple means do not account for differences in the indicators that are attributable to differences in patient mix across hospitals that are measured in the discharge data, or demographic differences across areas. In general, risk adjustment involves conducting a multivariate regression to adjust expected performance for these measured patient and population characteristics. Although complex, multivariate regression methods are the standard technique for risk-adjustment because they permit the simultaneous consideration of multiple patient characteristics and interaction among those characteristics. The interpretation of the risk-adjusted estimate is straightforward: it is the value of the indicator expected at that hospital if the hospital had an "average" patient case-mix.

This section contains the methods for the evaluation of risk adjustment systems, leading to the decision to use APR-DRGs. The purpose of this evaluation is to briefly outline the evidence regarding the use of risk adjustment systems for evaluating potential bias in indicators and for risk adjusting established indicators to compare provider performance. The first section discusses criteria used to evaluate the risk adjustment systems. Such criteria arise from the literature-based evidence on risk adjustment systems, as well as user criteria obtained through the semi-structured telephone interviews. Second, the methods used to implement APR-DRGs empirically in the new QI set are outlined. The methods for risk-adjustment of the hospital level quality indicators are described. An analogous method was used for the

area level quality indicators. However, the area level indicators account only for demographic differences.

Risk Adjustment Literature Review Methods

The literature review for risk adjustment of the HCUP QIs combined evaluation criteria common to evidence studies on the performance of risk adjustment systems with additional considerations of importance to the potential HCUP QI users. These considerations were determined through semi-structured interviews with users, discussed earlier in this report. In general, users viewed risk adjustment as an important component of the HCUP QIs' refinement. State data organizations and agencies involved in reporting of hospital performance measures especially tended to view risk-adjustment as essential for the validity of the results and acceptance by participating hospitals. Concerns that patient severity differed systematically among providers, and that this difference might drive the performance results, was frequently mentioned as a reason for limited reporting and public release of the HCUP QIs to date, especially for outcome-oriented measures like mortality following common elective procedures.

Literature-based Criteria for Evaluating Risk Adjustment Systems

HCUP QI users were concerned about the validity or performance of possible risk adjustment systems. Evidence was assessed on the performance of risk-adjustment systems from published reports using the following commonly applied criteria.[79, 87, 174]

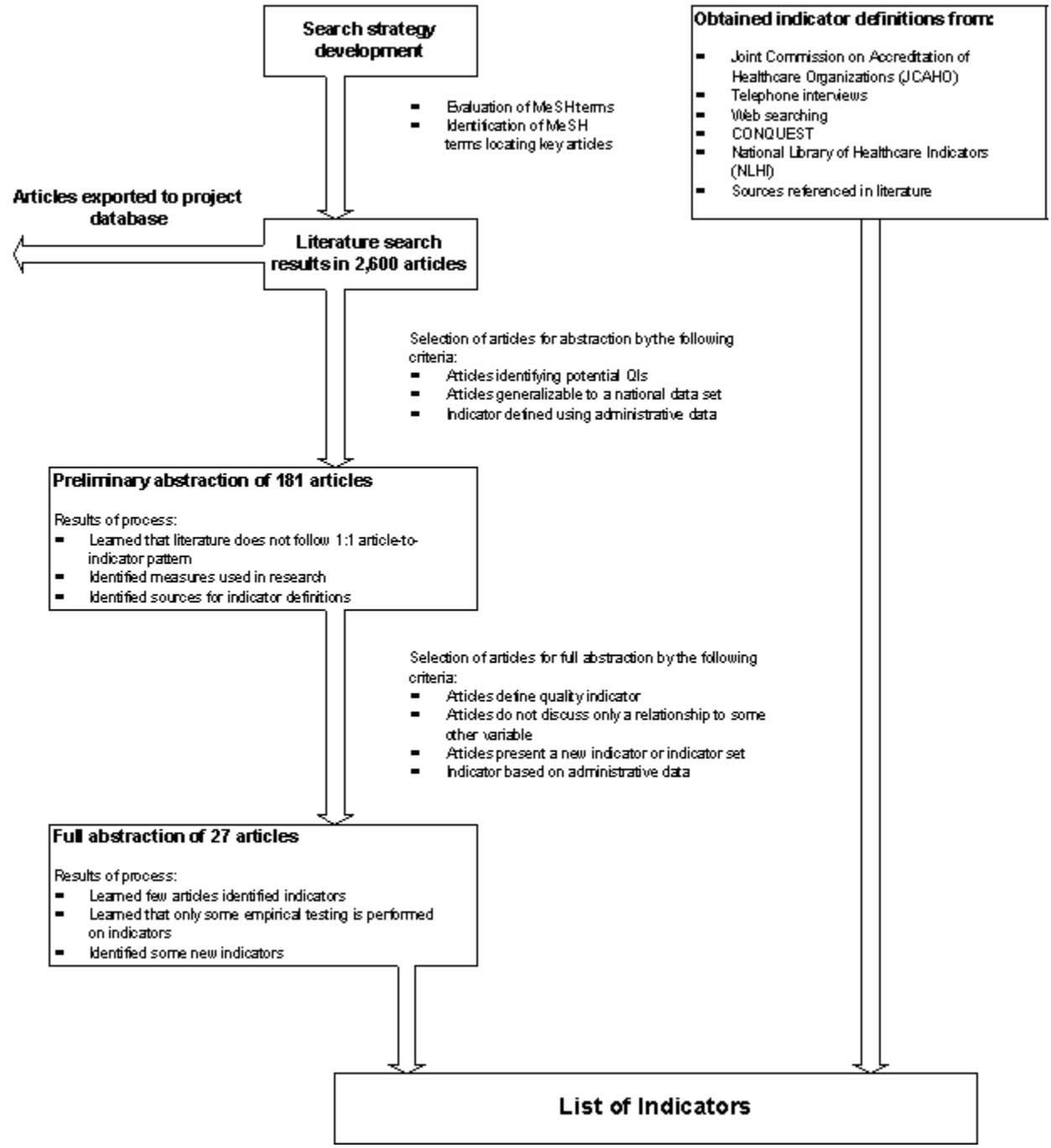
- Classification and analytic approach. Risk adjustment systems have been developed to predict complications, resource use, and mortality. Alternative analytic approaches included stratification (assigning individuals to mutually exclusive cells), logistic regression, or linear regression (calculating an expected level of medical utilization based on a statistical model). Methods based on logistic or linear statistical models are generally able to consider more dimensions of patient characteristics than stratification. Even more effective approaches might involve combining multivariate adjustment and stratification through propensity score methods and accounting for the relationship between aspects of disease severity that are measured and those that are not.[175, 176] However, no currently available risk adjustment systems are based on these analytic methods.

- System development method. Risk adjustment classifications may be based either on an empirical model clinical judgment or some combination. For example, an assessment of whether two heart attack patients are expected to have similar outcomes can be based on statistical tests or clinical expertise or both.[79]

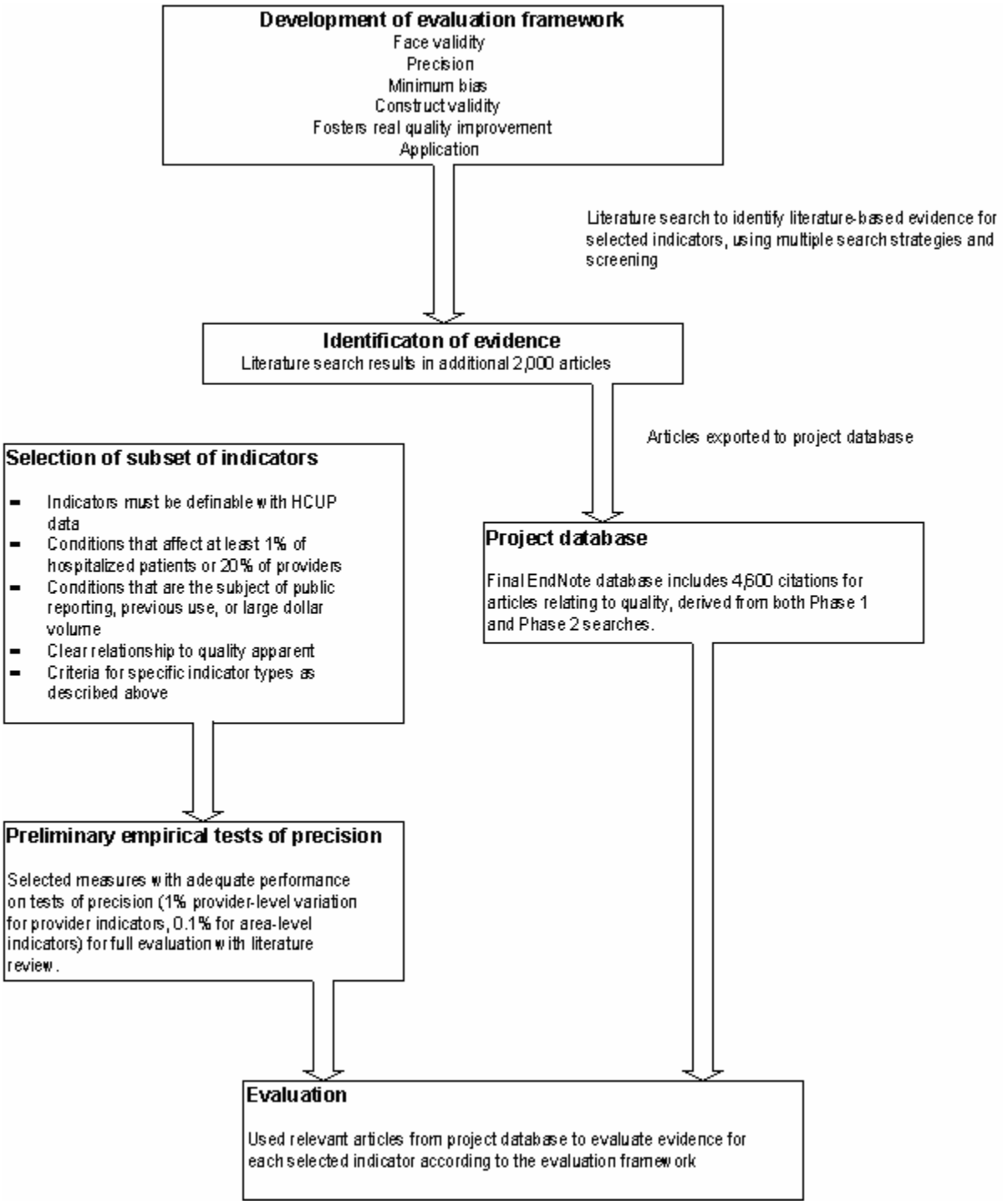
- Feasibility. Feasibility is largely determined by the data requirements of the risk-adjustment method. We reviewed whether a system required hospital data elements other than those found on the discharge abstract (e.g., data from medical charts or laboratory data) or non-hospital data (e.g., outpatient hospital or physician data). We also evaluated whether the method was likely to be enhanced with discharge data that included a unique patient identifier, so that risk adjusters could be developed based on data from multiple hospitalizations or encounters. Because only a subset of the States participating in HCUP collect supplementary data beyond discharge abstracts or unique patient identifiers for use in longitudinal analyses, a risk adjustment system was selected that did not depend on such information.

Figure B-2: Literature Review Strategy

Phase 1. Identification of Indicators



Phase 2. Evaluation of Indicators



- Empirical performance: Discrimination. A critical aspect of the performance of a risk-adjustment model is the extent to which the model predicts a higher probability of an event for patients who actually experience the event. The statistical test of discrimination is generally expressed as a C-statistic or R2 (how much of the variation in the patient level data the model explains). In general, systems that discriminate more have the potential to influence QI measures more substantially. Many severity-adjustment systems were designed primarily to predict in subsequent periods (e.g., resource consumption next year). However, for purposes of evaluating QI performance, the estimation of concurrent risk is more important (i.e., differences in the likelihood of experiencing an outcome in the current time period). Ideally, discrimination would be assessed using an R2 or other statistic of predicted variation that is computed on a separate data source from the one used to develop the model, to avoid "over-fitting" (i.e., the model might appear do well in part because it explains nonsystematic variations in the data used to develop it).

- Empirical performance: Calibration. Calibration is a measure of whether the mean of the predicted outcomes equals the mean of the actual outcomes for the entire population and for population subgroups. The statistical test is often expressed as a Chi-square or "goodness-of-fit" for the equivalence of means of population subgroups. Even if the severity-adjustment system does not predict well at the level of individuals, it may predict well at the aggregate (group) level of, say, women, 70-74 years of age. Over-fitting will be an issue here as well, unless a different data source is used to validate the model than was used to estimate the model.

Not many risk-adjustment systems have been evaluated in published reports using all of these criteria, nor have they been evaluated using consistent data sources. These limitations of the literature on risk adjustment complicate comparisons of risk adjustment systems based on performance criteria. In the end, the user-specified criteria determined a narrow set of potential risk adjustment systems to consider. The performance criteria delineated between these potential systems and informed the empirical evaluation of the impact of risk adjustment on the assessment of provider and area quality.

User-specified Criteria for Evaluating Risk Adjustment Systems

Evidence on the performance of a risk adjustment system is a primary consideration for HCUP QI users, and is essential to the validity of reported performance measures. However, users also cited other factors as potentially important determinants of the acceptance of HCUP QIs reporting by hospitals, State regulators and State legislatures, and other potential consumers of hospital performance data. These factors included the following:

- "Open" systems preferable to "black box" systems. Although there was no specific prohibition against using proprietary systems vs. systems in the public domain, there was a preference for using "open" systems where the risk adjustment logic was published and available for scrutiny by interested parties.

- Data collection costs minimized and well-justified. The widespread recognition that data collection was costly for hospitals meant that any risk-adjustment system that would be imposed on hospitals had to justify the cost of data collection by documenting that the additional information led to substantially different and more accurate inferences about performance. At least one State had stopped using a risk adjustment system that required medical chart review because the high cost of implementation was not considered worth the efficiency gained from improved accuracy.

- Multiple-use coding system. Some risk adjustment systems were designed to categorize patients according to expected resource use, defined either as charges or length of stay, while others were designed to categorize patients according to expected health outcomes, including mortality and complications. For example, several States calculated and reported mortality rates by

diagnosis-related group (DRG). These users generally believed that a risk-adjustment system for health outcomes based on discharge records that relied on the same diagnostic groups used for reimbursement was more likely to be accurate than a system that relied on codes used for quality and health outcome comparisons only, since there would be less financial and audit incentives to record codes accurately for the latter. Thus, coding systems that affected reimbursement for at least some patients were likely to capture diagnoses and procedures reported in medical charts.

One potentially important limitation of relying on codes that are also used for payment is that changes in reimbursement-related coding practices (e.g., as a result of tighter Medicare rules implemented in 1996) may alter apparent severity. However, because of the financial implications of changes in coding practices, any significant changes are likely to be identified and reported by payers, and so can be considered in interpreting variations and trends in reported quality measures.

- Official recognition. Many users indicated that systems that had been supported or otherwise recognized by Government agencies such as AHRQ were preferable to other systems, because such support facilitated acceptance by legislative and hospital groups. Adoption of the HCUP QIs themselves was often justified in part by their sponsorship by AHRQ. State agencies, especially those from smaller States, often cited the lack of staff resources and expertise needed to make independent evaluations of competing indicator sets and risk adjustment methods.

Risk Adjustment Empirical Methods

The APR-DRG system, with severity and risk of mortality classifications, was used in two ways:

- To evaluate the impact of measured differences in patient severity on the relative performance of hospitals and areas, by comparing QI measures with and without risk adjustment
- To risk-adjust the hospital- and area-specific measures.

The available literature on the impact of risk adjustment on indicator performance is limited, but suggests that at least in some cases different systems may give different results. Problems of incomplete or inconsistent coding across institutions are probably important contributing factors to the differences in results. Thus, definitive risk adjustment for some indicators may require detailed reviews of medical charts and additional data sources (charts may also be incomplete), just as definitive quality measures for many indicators may require additional sources of information. However, the importance of random variations in patients means that whatever risk adjustment and quality measurement system is chosen should be used in conjunction with statistical methods that seek to minimize other sources of noise and bias.

The empirical analysis is intended to illustrate the approach of combining risk adjustment with smoothing techniques, including suggestive evidence on the importance of risk adjustment for potential new QIs, using a risk adjustment system that can be implemented on discharge data by most HCUP QI users. The empirical analysis is supplemented with a review of the clinical literature to identify additional clinical information that is important to consider for certain indicators. In particular, the literature review highlights a few indicators where risk adjustment with additional clinical data has been shown to be particularly important, and where important differences in case mix seem less likely to be related to the secondary diagnoses used to risk-adjust discharge data.

This section describes how risk-adjustment is implemented using patient demographics (age and sex) along with the APR-DRG classification system. The next section describes statistical methods used to account for additional sources of noise and bias not accounted for by observed patient characteristics. By applying these methods to all of the potential new QIs, the relative importance of both risk adjustment and smoothing can be evaluated in terms of the relative performance of hospitals (or areas) compared to the “raw” unadjusted QIs based on simple means from NIS discharge data. The simple means fail to account both for differences in the indicators that are attributable to systematic differences in measured

and unmeasured patient mix across hospitals/areas that are measured in the discharge data, and for random variations in patient mix. A multivariate regression approach was adopted to adjust performance measures for measured differences in patient mix, which permits the inclusion of multiple patient demographic and severity characteristics.

Specifically, if it is denoted whether or not the event associated with a particular indicator $\mathbf{Y^k}$ (**k**=1,…,K) was observed for a particular patient **i** at hospital/area **j** (**j**=1,…,J) in year **t** (**t**=1,…,T), then the regression to construct a risk-adjusted "raw" estimate of a hospital or area's performance on each indicator can be written as:

(1) $$Y^k_{ijt} = M^k_{jt} + Z_{ijt}\Pi^k_t + \varepsilon^k_{ijt} \quad \text{where}$$

Y^k_{ijt} is the $\mathbf{k^{th}}$ quality indicator for patient **i** discharged from hospital/area **j** in year **t** (i.e., whether or not the event associated with the indicator occurred on that discharge);

M^k_{jt} is the "raw" adjusted measure for indicator **k** for hospital/area **j** in year **t** (i.e., the hospital/area "fixed effect" in the patient-level regression);

Z_{ijt} is a vector of patient covariates for patient **i** discharged from hospital/area **j** in year **t** (i.e., the patient-level measures used as risk adjusters);

Π^k_t is a vector of parameters in each year **t**, giving the effect of each patient risk adjuster on indicator **k** (i.e., the magnitude of the risk adjustment associated with each patient measure); and

ε^k_{ijt} is the unexplained residual in this patient-level model.

The hospital or area specific intercept $\mathbf{M^k_{jt}}$ is the "raw" adjusted measure of a hospital or area's performance on the indicator, holding patient covariates constant. In most of the empirical analysis that follows, the patient-level analysis is conducted using data from all hospitals and areas. (The model shown implies that each hospital or area has data for all years, and with each year has data on all outcomes; however, this is not essential to apply risk adjustment methods.)

These patient-level regressions were estimated by linear ordinary least-squares (OLS). In general, the dependent variables in the regressions are dichotomous, which raises the question of whether a method for binary dependent variables such as logit or probit estimation might be more appropriate. However, previous work by McClellan and Staiger has successfully used OLS regression for similar analyses of hospital/area differences in outcomes. In addition, estimating logit or probit models with hospital or area fixed effects cannot be done with standard methods; it requires computationally intensive conditional maximum likelihood methods that are not easily extended to multiple years and multiple measures.[177]

A commonly used "solution" to this problem is to estimate a logit model without hospital or area effects, and then to use the resulting predictions as estimates of the expected outcome. However, this method yields biased estimates and predictions of hospital performance. In contrast, it is easy to incorporate hospital or area fixed effects into OLS regression analysis, the resulting estimates are not biased, and the hospital or area fixed effects provide direct and easily-interpretable estimates of the outcome rate for a particular hospital or area measure in a particular year, holding constant all observed patient characteristics.

Of course, it is possible that a linear probability model is not the correct functional form. However, as in earlier work, a very flexible functional form is specified, including full interactions among age and sex covariates as well as a full set of APR-DRG risk adjusters. In the sensitivity analyses for selected quality measures, this flexible linear probability model produced estimates of the effects of the risk adjusters that did not differ substantially from nonlinear (logit and probit) models. Another potential limitation of the OLS approach is that it may yield biased estimates of confidence intervals, because the

errors of a linear probability model are necessarily heteroskedastic. Given the large sample sizes for the parameters estimated from these regressions (most indicators involve thousands of "denominator" discharges per year), such efficiency is not likely to be an important concern. Nevertheless, models were estimated using Weighted Least Squares to account for heteroskedasticity, to see if estimates were affected [178]. Very similar estimates of adjusted indicator performance were obtained.

Specifically, in addition to age, sex, and age*sex interactions as adjusters, the model also included the APR-DRG category for the admission and the APR-DRG constructed severity subclass (or risk-of-mortality subclass for mortality measures). APR-DRGs are a refinement of the DRGs used by the Centers for Medicare and Medicaid Services (formerly the Health Care Financing Administration), with additional classifications for non-Medicare cases (e.g., neonates). The severity subclass evaluates the episode of care on a scale of 1 (minor) to 4 (extreme). In the APR-DRG Version 12, Severity of Illness is defined as the "extent of physiologic de-compensation or organ system loss of function." The APR-DRG severity of illness subclass was designed principally to predict resource use, particularly length-of-stay. As such, because this risk-adjustment system was not designed to predict utilization rates, for example, the evaluation of each indicator does not consider lack of impact of risk-adjustment to be evidence of lack of real bias. However, impact of risk-adjustment is considered to be evidence of problems of potential bias. The literature review further informs potential sources of bias, and the prior use of each indicator may require collection of supplemental data for confounding clinical conditions.

For each indicator, the APR-DRG groupings in the Major Diagnostic Category (MDC) related to that indicator were excluded from the risk adjustment model. The groupings are either medical (based on diagnoses) or surgical (based on procedures), and groupings in the MDC of the same type were excluded. For example, for the Coronary Artery Bypass Graft rate indicator, all surgical APR-DRGs in MDC '05' ('Diseases and Disorders of the Circulatory System') were excluded. For GI Hemorrhage mortality, all medical APR-DRGs in MDC '06' ('Diseases and Disorders of the Digestive System') were excluded. Some of the indicators fall into only a few DRG categories. All discharges with carotid endarterectomy, for example, were within DRG '005', ('Extracranial Vascular Procedures'). These indicators relied primarily on the severity subclass, which is independent of the DRG.

Actual implementation of the model involves running a regression with potentially a few thousand variables (each DRG divided into four severity subclasses) on millions of observations, straining the capacity of most statistical software and computer systems. In order to limit the number of covariates (DRG groups) in the model, the total number was restricted to 165 categories (DRG by severity), which was for all indicators sufficient to include 80% of discharges. All severity or risk-of-mortality subgroups were maintained for each APR-DRG included in the model in the construction of the raw adjusted estimates. The adjusted estimates of hospital performance are reported and used to compute descriptive statistics for each indicator in each year. They are also used to construct smoothed estimates of each indicator.

The risk-adjusted estimates of hospital performance (age, gender, APR-DRG) and area performance (age, gender only) were used to construct descriptive statistics and smoothed estimates for each QI.

Empirical Methods

Analysis Approach

Data sources. The data sources used in the empirical evaluation were the 1995-97 Nationwide Inpatient Sample (NIS), which has been used for previous HCUP QI development, and the complete State Inpatient Data (SID) for five HCUP participating States (California, Florida, Illinois, New York, and Pennsylvania). The annual NIS consists of about 6 million discharges and over 900 hospitals. The NIS contains all-payer data on hospital inpatient stays from selected States (Arizona, California, Colorado, Connecticut, Florida, Georgia, Hawaii, Illinois, Iowa, Kansas, Maryland, Massachusetts, Missouri, New Jersey, New York, Oregon, Pennsylvania, South Carolina, Tennessee, Utah, Washington, and Wisconsin). All discharges from sampled hospitals are included in the NIS database. The NIS is

designed to approximate a 20% sample of U.S. community hospitals, defined as all non-Federal, short-term, general, and other specialty hospitals, excluding hospital units of institutions. Included among community hospitals are specialty hospitals such as obstetrics-gynecology, ear-nose-throat, short-term rehabilitation, orthopedic, and pediatric. Excluded are long-term hospitals, psychiatric hospitals, and alcoholism/chemical dependency treatment facilities. A complete description of the content of the NIS, including details of the participating States discharge abstracts, can be found on the Agency for Healthcare Research and Quality Web site (http://www.hcup-us.ahrq.gov/nisoverview.jsp).

The SID sample consisted of 10 million discharges and over 1,300 hospitals in over 200 metropolitan areas. Only the SID empirical results are reported, because the provider-level results were similar in both data sources, and the SID data were needed for the direct construction of area measures. All of the quality indicators can be constructed from the NIS with two caveats: First, the area measures are based on a weighted sample of discharges and are less precise than if complete State discharge data are used; and second, even though hospital sampling for the NIS was supposed to allow construction of a representative sample at the State level, it is possible that the Metropolitan Service Area (MSA)-level samples are not representative (i.e., biased). These limitations are not applicable when using the software on the full data from the SID to construct measures based on complete data from area hospitals.

Reported quality indicators. All potential indicators were assessed empirically by developing and conducting statistical tests for evaluation framework criteria of precision, bias, and construct validity. For each statistical test, we calculated up to four different estimates of indicator performance. First, the raw indicator was the simple observed value (e.g., the rate or volume) for each provider or area. Second, the adjusted indicator was based on the use of multivariate regression to account for differences among providers in demographics and comorbidities (defined using the 3M™ APR-DRG) of patients, and among areas in demographics of the population. Third, univariate smoothing techniques were applied to estimate the amount of random error relative to the true difference in performance (the "reliability") for each indicator.[68] Fourth, new multivariate signal extraction methods were applied by combining information from multiple indicators over several years to extract more quality signal from each individual indicator than is possible with the univariate methods.[179]

Overview of empirical analysis. The approach included several stages and generated a series of analyses on potential quality indicators that sequentially assessed some of the problems identified in the literature review. For reference, the "raw" or minimally adjusted indicator was constructed, based on the discharge data for each hospital and census data for each area. This measure was then "risk-adjusted" through a discharge-level regression that included controls for patient mix. The hospital-level and area-level fixed effects in these regressions are the estimates of quality indicators that are typically reported for particular hospitals and areas, and they typically reflect substantial noise. In the second stage of the analysis, these estimates were then "smoothed" using a Bayesian procedure to yield a best-guess estimate of true hospital or area performance on the indicator—the "signal" in the observed noisy measure. This was done in two ways. First, a univariate approach was used, in which the distribution of the indicator itself is used to construct the best guess. This is the smoothing or shrinkage approach most widely used in the literature on provider quality.[69-71] Second, a multivariate approach was used, in which the joint distribution of a large number of indicators (and the indicator of interest in previous time periods) is used to construct the best-guess estimate of performance. In general, the covariation among different indicators and within each indicator over time implies that much more precise estimates of true hospital or area quality can be generated using this multivariate signal extraction approach. All of the estimates of factor loadings and correlations are based on smoothed estimates, which helps to improve the ability to detect correlations, thereby addressing the multidimensionality of quality. Finally, summary statistics are reported describing the performance of the indicator in terms of the principal domains described in the literature review: precision, bias, and construct validity.

Intuition Behind Univariate and Multivariate Methods

An important limitation of many quality indicators is their imprecision, which complicates the reliable identification of persistent differences among providers in performance. The imprecision in quality indicators arises from two sources. The first is sampling variation, which is a particular problem for

indicators based on small numbers of patients per provider (where the particular patients treated by the provider in a given year are considered a "sample" of the entire population who might have been treated or will be treated in the near future). The amount of variation due to the particular sample of patients is often large relative to the total amount of provider-level variation that is observed in any given quality indicator. A second source of imprecision arises from non-persistent factors that are not sensitive to the size of the sample; for example, a severe winter results in higher than usual rates of pneumonia mortality. Both small samples and other one-time factors that are not sensitive to sample size can add considerable volatility to quality indicators. Also, it is not the absolute amount of imprecision that matters, but rather the amount of imprecision relative to the underlying signal (i.e., true provider-level variation) that dictates the reliability of any particular indicator. Even indicators based on relatively large samples with no non-persistent factors at work can be imprecise if the true level of variation among providers is negligible.

The approach to account for the imprecision or lack of reliability is a generalization of the idea of applying a "shrinkage factor" to each provider's estimate so that less reliable estimates are shrunk toward the national average. These "reliability-adjusted" estimates are sometimes referred to as "smoothed" estimates (because provider performance is less volatile over time) or "filtered" estimates (because the methods filter out the non-systematic noise, much like a radio filters our background noise to improve the radio signal). If the observed provider variation = signal variation + noise variation, then the shrinkage factor would be signal variation ÷ (signal variation + noise variation). For example, suppose that the observed variation among providers in the in-hospital pneumonia mortality rate was a standard deviation of 10.2 percentage points, and the signal variation was a standard deviation of 5.0 percentage points. Then the shrinkage factor for this indicator is 0.240 = (0.050^2) ÷ (0.102^2). The generalization of this approach seeks to extract additional signal using information on the relationship among multiple indicators over time.

Many of the key ideas behind the reliability-adjusted or filtered estimates are illustrated through a simple example. Suppose that one wants to evaluate a particular provider's performance based on in-hospital mortality rates among patients admitted with pneumonia, and data are available for the most recent 2 years. Consider the following three possible approaches: (1) use only the most recent mortality rate, (2) construct a simple average of the mortality rates from the 2 recent years, or (3) ignore the provider's mortality rate and assume that mortality is equal to the national average. The best choice among these three approaches depends on two important considerations: the signal-to-noise ratio in the provider's data and how strongly correlated performance is from one year to the next.

For example, suppose that the mortality rate for the provider was based on only a few patients, and one had reason to believe that mortality did not vary much across providers. Then one would be tempted to choose the last option and ignore the provider's own data because of its low reliability (e.g., low signal-to-noise ratio). This is the idea of simple shrinkage estimators, in which less reliable estimates are shrunk toward the average for all providers. Alternatively, if one had reason to believe that provider mortality changed very slowly over time, one might choose the second option in hopes that averaging the data over 2 years would reduce the noise in the estimates by effectively increasing the sample size in the provider. Even with large numbers of patients, one might want to average over years if idiosyncratic factors (such as a bad winter) affected mortality rates from any single year. Finally, one would tend to choose the first option, and rely solely on mortality from the most recent year, if such idiosyncratic factors were unimportant, if the provider admitted a large number of patients each year, and if mortality was likely to have changed from the previous year.

The method of creating filtered estimates formalizes the intuition from this simple example. The filtered estimates are a combination of the provider's own quality indicator, the national average, and the provider's quality indicators from past years or other patient outcomes. As suggested by the example, to form the optimal combination, one must know the amount of noise and signal variance in each indicator, as well as the correlation across indicators in the noise and signal variance.

The noise variance (and covariance) is estimated in a straightforward manner for each provider, based on the number of patients on which each indicator is based. To estimate the signal variance (and covariance) for each quality indicator, the noise variance is subtracted from the total variance observed in

each indicator across providers (which reflects both signal and noise variance). In other words, the observed variation in quality indicators is sure to overstate the amount of actual variation across providers (because of the noise in the indicators). Therefore, the amount of true variation in performance is estimated based on how much the observed variation exceeded what would have been expected due to sampling error. Importantly, this method does not *assume* that provider performance is correlated from one year to the next (or that performance is correlated across indicators). Instead, these correlations are estimated directly from the data, and information from past years or other indicators is incorporated only to the extent that these empirically estimated correlations are large.

Smoothed Estimates of Hospital Performance

For each hospital, a vector of **K** adjusted indicator estimates was observed over **T** years from estimating the patient-level regressions (1) run separately by year for each indicator as described in the preceding section. Each indicator is a noisy estimate of true hospital quality; in other words, it is likely that hospitals that performed especially well or badly on the measure did so at least in part due to chance. This fact is incorporated in Bayesian methods for constructing best-guess "posterior" estimates of true provider performance based on observed performance and the within-provider noise in the measures.

In particular, let $\mathbf{M_j}$ be the **1xTK** vector of estimated indicator performance for hospital **j**. Then:

(2) $$M_j = \mu_{.j} + \varepsilon_j$$

Where $\boldsymbol{\mu_j}$ is a **1xTK** vector of the true hospital intercepts for hospital **j**, and $\boldsymbol{\varepsilon_j}$ is the estimation error (which has a mean zero and is uncorrelated with $\boldsymbol{\mu_j}$). Note that the variance of $\boldsymbol{\varepsilon_j}$ can be estimated from the patient-level regressions, since this is simply the variance of the regression estimates $\mathbf{M_j}$. In particular, $\mathbf{E(\varepsilon_{jt}' \varepsilon_{jt}) = \Omega_{jt}}$ and $\mathbf{E(\varepsilon_{jt}' \varepsilon_{js}) = 0}$ for $\mathbf{t \neq s}$, where $\boldsymbol{\Omega_{jt}}$ is the covariance matrix of the intercept estimates for hospital **j** in year **t**.

A linear combination of each hospital's observed indicators must be created in such a way that it minimizes the mean-squared prediction error. In other words, the following hypothetical regression should be run:

(3) $$\mu_{jt}^{k} = M_j \beta_{jt}^{k} + v_{jt}^{k}$$

but cannot be run directly, since μ is unobserved and the optimal β varies by hospital and year. While equation (3) cannot be directly estimated, it is possible to estimate the parameters for this hypothetical regression. In general, the minimum mean squared error linear predictor of μ is given by $\mathbf{M_j\beta}$, where $\boldsymbol{\beta} = \mathbf{[E(M_j'M_j)]^{-1}\ E(M_j'\mu_j)}$. This best linear predictor depends on two moment matrices:

(4.1) $$E(M_j'M_j) = E(\mu_j'\mu_j) + E(\varepsilon_j'\varepsilon_j)$$

(4.2) $$E(M_j'\mu_j) = E(\mu_j'\mu_j)$$

The required moment matrices are estimated directly as follows:

- Estimate $\mathbf{E(\varepsilon_j'\varepsilon_j)}$ with the patient-level OLS estimate of the covariance matrix for the parameter estimates $\mathbf{M_j}$. Call this estimate $\mathbf{S_j}$. Note that $\mathbf{S_j}$ varies across hospitals.

- Estimate $\mathbf{E(\mu_j'\mu_j)}$ by noting that $\mathbf{E(M_j'M_j\text{-}S_j) = E(\mu_j'\mu_j)}$. If assumed that $\mathbf{E(\mu_j'\mu_j)}$ is the same for all hospitals, then it can be estimated by the sample average of $\mathbf{M_j'M_j\text{-}S_j}$. Note that it is easy

to relax the assumption that $\mathbf{E(\mu_j'\mu_j)}$ is the same for all hospitals by calculating $\mathbf{M_j'M_j-S_j}$ for subgroups of hospitals.

With estimates of $\mathbf{E(\mu_j'\mu_j)}$and $\mathbf{E(\varepsilon_j'\varepsilon_j)}$, one can form least squares estimates of the parameters in equation 3 which minimize the mean squared error. Analogous to simple regression, the prediction of a hospital's true intercept is given by:

$$\hat{\mu}_j = M_j E\left(M_j' M_j\right)^{-1} E\left(M_j' \mu_j\right) = M_j \left[E\left(\mu_j' \mu_j\right) + E\left(\varepsilon_j' \varepsilon_j\right)\right]^{-1} E\left(\mu_j' \mu_j\right) \tag{5}$$

using estimates of $\mathbf{E(\mu_j'\mu_j)}$and $\mathbf{E(\varepsilon_j'\varepsilon_j)}$ in place of their true values. One can use the estimated moments to calculate other statistics of interest as well, such as the standard error of the prediction and the r-squared for equation 3, based on the usual least squares formulas. Estimates based on equation (5) are referred to as "filtered" estimates, since the key advantage of such estimates is that they optimally filter out the estimation error in the raw quality indicators.

Equation 5 in combination with estimates of the required moment matrices provides the basis for estimates of hospital quality. Such estimates of hospital quality have a number of attractive properties. First, they incorporate information in a systematic way from many outcome indicators and many years into the predictions of any one outcome. Moreover, if the moment matrices were known, the estimates of hospital quality represent the optimal linear predictors, based on a mean squared error criterion. Finally, these estimates maintain many of the attractive aspects of existing Bayesian approaches, while dramatically simplifying the complexity of the estimation.[69] It is possible to construct univariate smoothed estimates of hospital quality, based only on empirical estimates for particular measures, using the models just described but restricting the dimension of $\mathbf{M_j}$ to only a particular indicator **k** and time period **t**. Of course, to the extent that the provider indicators are correlated with each other and over time, this will result in a less precise (efficient) estimate.

With many years of data, it helps to impose some structure on $\mathbf{E(\mu_j'\mu_j)}$ for two reasons. First, this improves the precision of the estimated moments by limiting the number of parameters that need to be estimated. Second, a time series structure allows for out-of-sample forecasts. A non-stationary, first-order Vector Autoregression structure (VAR) is used. The VAR model is a generalization of the usual autoregressive model, and assumes that each hospital's quality indicators in a given year depend on the hospital's quality indicators in past years plus a contemporaneous shock that may be correlated across quality indicators. In most of what follows, a non-stationary first-order VAR is assumed for μ_{jt} **(1xK)**, where:

$$\mu_{jt} = \mu_{j,t-1}\Phi + u_{jt}, \text{ with } V\left(u_{jt}\right) = \Sigma \text{ and } V\left(\mu_{j1}\right) = \Gamma \tag{6}$$

Thus, estimates are needed of the lag coefficient (**Φ**), the variance matrix of the innovations (**Σ**), and the initial variance condition (**Γ**), where **Σ** and **Γ** are symmetric **K**x**K** matrices of parameters and **Φ** is a general **K**x**K**matrix of parameters, for a total of $\mathbf{2K^2+K}$ parameters. For example, 10 parameters must be estimated for a VAR model with two outcomes (**K=2**).

The VAR structure implies that $\mathbf{E(M_j'M_j-S_j) = E(\mu_j'\mu_j) = f(\Phi,\Sigma,\Gamma)}$. Thus, the VAR parameters can be estimated by Optimal Minimum Distance (OMD) methods, i.e., by choosing the VAR parameters so that the theoretical moment matrix, $\mathbf{f(\Phi,\Sigma,\Gamma)}$, is as close as possible to the corresponding sample moments from the sample average of $\mathbf{M_j'M_j-S_j}$. More specifically, let $\mathbf{d_j}$ be a vector of the non-redundant (lower triangular) elements of $\mathbf{M_j'M_j-S_j}$ and let **δ** be a vector of the corresponding moments from the true moment matrix, so that $\mathbf{\delta=g(\Phi,\Sigma,\Gamma)}$.[177] Then the OMD estimates of **(Φ,Σ,Γ)** minimize the following OMD objective function:

(7) $$q = M\left[\bar{d} - g(\Phi, \Sigma, \Gamma)\right]' V^{-1}\left[\bar{d} - g(\Phi, \Sigma, \Gamma)\right]$$

where **V** is the sample estimate of the covariance matrix for **d**, and $\bar{\mathbf{d}}$ is the sample average of **d**. If the VAR model is correct, the value of the objective function, **g**, will be distributed $\mathbf{X^2}$ **(p)** where **p** is the degree of over-identification (the difference between the number of elements in **d** and the number of parameters being estimated). Thus, **q** provides a goodness of fit statistic that indicates how well the VAR model fits the actual covariances in the data.

Finally, estimated $\mathbf{R^2}$ statistics are used to evaluate the filtered estimates' ability to predict (in sample) and forecast (out-of-sample) variation in the true intercepts, and to compare methods used to conventional methods (e.g., simple averages, or univariate shrinkage estimators). If true hospital intercepts **(μ)** were observed, a natural metric for evaluating the predictions would be the sample **R-squared:**

(8) $$R^2 = 1 - \left(\sum_{j=1}^{N} \hat{u}_j^2\right) \Big/ \left(\sum_{j=1}^{N} \mu_j^2\right)$$

where $$\hat{u}_j = \mu_j - \hat{\mu}_j$$

is the prediction error. Of course **μ** is not observed. Therefore, an estimate is constructed using the estimate of $\mathbf{E(\mu_j'\mu_j)}$ for the denominator, and the estimate of

$$E\left[(\mu_j - \hat{\mu}_j)'(\mu_j - \hat{\mu}_j)\right]$$

for the terms in the numerator (where this can be constructed from the estimated moment matrices in equations 4.1 and 4.2). Finally, a weighted **R-squared** is reported (weighting by the number of patients treated by each hospital).

As in earlier work using this method for cardiac care in the adult population, the indicators are validated using out-of-sample performance, based on forecasts (e.g., using the first 2 years of data to predict in subsequent year) and based on split-sample prediction (e.g., using one-half of the patient sample to predict outcome indicators in the other half of the sample). For evaluating out-of-sample forecasts, a modified **R-squared** of the forecast is constructed that estimates the fraction of the systematic (true) hospital variation in the outcome measure (**M**) that was explained:

(9) $$\tilde{R}^2 = 1 - \left(\sum_{j=1}^{N} (\hat{\nu}_j^2 - S_j)\right) \Big/ \left(\sum_{j=1}^{N} (M_j^2 - S_j)\right)$$

where $$\hat{\nu}_j = M_j - \hat{\mu}_j$$

is the forecast error and **Sj** is the OLS estimate of the variance of the estimate **M**j. This modified **R-squared** estimates the amount of variance in the true hospital effects that has been forecast. Note that because these are out-of-sample forecasts, the **R-squared** can be negative, indicating that the forecast performed worse than a naive forecast in which one assumed that quality was equal to the national average at all hospitals.

Empirical Analysis Statistics

Using the methods just described, a set of statistical tests was constructed to evaluate precision, bias, and construct validity. Each of the key statistical test results for these evaluation criteria was summarized and explained in the beginning of this appendix. Tables B1-B3 provides a summary of the statistical analyses and their interpretation. Indicators were tested for precision first, and ones that performed poorly were eliminated from further consideration. Bias and construct validity were assessed for all recommended indicators.

TableB-1: Precision Tests

Measure	Statistic		Interpretation
Precision. Is most of the variation in an indicator at the level of the provider? Do smoothed estimates of quality lead to more precise measures?			
a. Raw variation in indicator	Provider Standard Deviation Signal Standard Deviation Provider/Area Share	Unadjusted Age-sex adjusted Age-sex+APR-DRG adjusted	Provider variation is signal variation + noise variation. What percentage of the total variation (patient + provider) is between-provider variation (a measure of how much variation is subject to provider control). Risk adjustment can either increase or decrease true variation.
b. Univariate smoothing	Signal/Signal-to-noise ratio: Unadjusted Age-sex adjusted Age-sex + APR-DRG adjusted		Estimates what percentage of the observed variation between providers reflects “true” quality differences versus random noise. Risk adjustment can increase or decrease estimates of “true” quality differences.
c. MSX methods	In-sample R-squared: Unadjusted Age-sex adjusted Age-sex + APR-DRG adjusted		To the extent that indicators are correlated with each other and over time, MSX methods can extract more “signal” (a higher percentage of observed variation between providers that reflects “true” quality).

Table B-2: Bias Tests

Measure	Statistic	Interpretation
Bias. Does risk-adjustment change the assessment of relative provider performance, after accounting for reliability? Is the impact greatest among the best or worst performers, or overall? What is the magnitude of the change in performance?		
a. MSX methods: unadjusted vs. age, sex, APR-DRG risk adjustment	Rank correlation coefficient (Spearman)	Risk-adjustment matters to the extent that it alters the assessment of relative provider performance. This test determines the impact overall.
	Average absolute value of change relative to mean	This test determines whether the absolute change in performance was large or small relative to the overall mean.
	Percentage of the top 10% of providers that remains the same	This test measures the impact at the highest rates (in general, the worse performers, except for measures like VBAC).
	Percentage of the bottom 10% of providers that remains the same	This tems measure the impact at the lowest rates (in general, the best performers, except for measures like VBAC).
	Percentage of providers that move more than two deciles in rank (up or down)	This test determines the magnitude of the relative changes.

Table B-3: Construct Validity Tests

Measure	Statistic	Interpretation
Construct validity. Is the indicator related to other indicators in a way that makes clinical sense? Do methods that remove noise and bias make the relationship clearer?		
a. Correlation of indicator with other indicators	Pearson correlation coefficient	Are indicators correlated with other indicators in the direction one might expect?
b. Factor loadings of indicator with other indicators	Factor loadings	Do indicators load on factors with other indicators that one might expect?

References for Appendix B

1. McGlynn EA, Asch SM. Developing a clinical performance measure. Am J Prev Med 1998;14(3 Suppl):14-21.

2. Siu AL, McGlynn EA, Morgenstern H, et al. Choosing quality of care measures based on the expected impact of improved care on health. Health Serv Res 1992;27(5):619-50.

3. Donabedian A. Explorations in Quality Assessment and Monitoring. The definition of quality and approaches to its assessment. Ann Arbor, MI: Health Administration Press; 1980.

4. Donabedian A. The quality of care: how can it be assessed? JAMA 1988;260(12):1743-1748.

5. Schneider EC, Epstein AM. Influence of cardiac-surgery performance reports on referral practices and access to care. A survey of cardiovascular specialists. N Engl J Med 1996;335(4):251-6.

6. Mennemeyer ST, Morrisey MA, Howard LZ. Death and reputation: how consumers acted upon HCFA mortality information. Inquiry 1997;34(2):117-28.

7. Hibbard JH, Jewett JJ. Will quality report cards help consumers? Health Aff (Millwood) 1997;16(3):218-28.

8. Normand SL, McNeil BJ, Peterson LE, et al. Eliciting expert opinion using the Delphi technique: identifying performance indicators for cardiovascular disease. Int J Qual Health Care 1998;10(3):247-60.

9. Veroff DR, Gallagher PM, Wilson V, et al. Effective reports for health care quality data: lessons from a CAHPS demonstration in Washington State. Int J Qual Health Care 1998;10(6):555-60.

10. Delbanco TL, Stokes DM, Cleary PD, et al. Medical patients' assessments of their care during hospitalization: insights for internists. J Gen Intern Med 1995;10(12):679-85.

11. Laine C, Davidoff F. Patient-centered medicine. A professional evolution. JAMA 1996;275(2):152-6.

12. Allen HM, Jr., Rogers WH. Consumer surveys of health plan performance: a comparison of content and approach and a look to the future. Jt Comm J Qual Improv 1996;22(12):775-94.

13. Cleary PD, Edgman-Levitan S. Health care quality. Incorporating consumer perspectives. JAMA 1997;278(19):1608-12.

14. Eye on patients: excerpts from a report on patients' concerns and experiences about the health care system. American Hospital Association and the Picker Institute. J Health Care Finance 1997;23(4):2-11.

15. Calnan MW. The patient's perspective. Int J Technol Assess Health Care 1998;14(1):24-34.

16. Tye L. Patient surveys show how Massachusetts hospitals stack up. Boston Globe 1998 November 13, 1998;Sect. A1, A34.

17. Finkelstein BS, Harper DL, Rosenthal GE. Patient assessments of hospital maternity care: a useful tool for consumers? Health Serv Res 1999;34(2):623-40.

18. Berwick DM, Wald DL. Hospital leaders' opinions of the HCFA mortality data. JAMA 1990;263(2):247-9.

19. Jencks SF, Daley J, Draper D, et al. Interpreting hospital mortality data. The role of clinical risk adjustment. JAMA 1988;260(24):3611-6.

20. Park RE, Brook RH, Kosecoff J, et al. Explaining variations in hospital death rates. Randomness, severity of illness, quality of care. JAMA 1990;264(4):484-90.

21. Localio AR, Hamory BH, Sharp TJ, et al. Comparing hospital mortality in adult patients with pneumonia. A case study of statistical methods in a managed care program. Ann Intern Med 1995;122(2):125-32.

22. Localio AR, Hamory BH, Fisher AC, et al. The public release of hospital and physician mortality data in Pennsylvania. A case study. Med Care 1997;35(3):272-286.

23. Hofer TP, Hayward RA. Identifying poor-quality hospitals. Can hospital mortality rates detect quality problems for medical diagnoses? Med Care 1996;34(8):737-53.

24. Thomas JW, Hofer TP. Accuracy of risk-adjusted mortality rate as a measure of hospital quality of care. Med Care 1999;37(1):83-92.

25. Mant J, Hicks N. Detecting differences in quality of care: the sensitivity of measures of process and outcome in treating acute myocardial infarction. BMJ 1995;311(7008):793-6.

26. Palmer RH. Process-based measures of quality: the need for detailed clinical data in large health care databases. Ann Intern Med 1997;127(8 Pt 2):733-8.

27. Eddy DM. Performance measurement: problems and solutions . Health Aff (Millwood) 1998;17(4):7-25.

28. Harr DS, Balas EA, Mitchell J. Developing quality indicators as educational tools to measure the implementation of clinical practice guidelines. Am J Med Qual 1996;11(4):179-85.

29. Ellerbeck EF, Jencks SF, Radford MJ, et al. Quality of care for Medicare patients with acute myocardial infarction. A four-state pilot study from the Cooperative Cardiovascular Project . JAMA 1995;273(19):1509-14.

30. Marciniak TA, Ellerbeck EF, Radford MJ, et al. Improving the quality of care for Medicare patients with acute myocardial infarction: results from the Cooperative Cardiovascular Project . JAMA 1998;279(17):1351-7.

31. Donabedian A. Evaluating the quality of medical care. Milbank Mem Fund Q 1966;44(3):Suppl:166-206.

32. Brook RH, McGlynn EA, Cleary PD. Quality of health care. Part 2: measuring quality of care. N Engl J Med 1996;335(13):966-70.

33. Luft HS, Bunker JP, Enthoven AC. Should operations be regionalized? The empirical relation between surgical volume and mortality. N Engl J Med 1979;301(25):1364-9.

34. Hughes RG, Garnick DW, Luft HS, et al. Hospital volume and patient outcomes. The case of hip fracture patients. Med Care 1988;26(11):1057-67.

35. Hannan EL, JF OD, Kilburn H, Jr., et al. Investigation of the relationship between volume and mortality for surgical procedures performed in New York State hospitals. JAMA 1989;262(4):503-10.

36. Hannan EL, Kilburn H, Jr., Bernard H, et al. Coronary artery bypass surgery: the relationship between inhospital mortality rate and surgical volume after controlling for clinical risk factors. Med Care 1991;29(11):1094-107.

37. Stone VE, Seage GRd, Hertz T, et al. The relation between hospital experience and mortality for patients with AIDS. JAMA 1992;268(19):2655-61.

38. Hosenpud JD, Breen TJ, Edwards EB, et al. The effect of transplant center volume on cardiac transplant outcome. A report of the United Network for Organ Sharing Scientific Registry. JAMA 1994;271(23):1844-9.

39. Jones A, O'Driscoll K, Luke LC. Head injuries and the observation ward [letter; comment]. J Accid Emerg Med 1995;12(2):160-1.

40. Phibbs CS, Bronstein JM, Buxton E, et al. The effects of patient volume and level of care at the hospital of birth on neonatal mortality. JAMA 1996;276(13):1054-9.

41. Ellis SG, Weintraub W, Holmes D, et al. Relation of operator volume and experience to procedural outcome of percutaneous coronary revascularization at hospitals with high interventional volumes. Circulation 1997;95(11):2479-84.

42. Jollis JG, Peterson ED, Nelson CL, et al. Relationship between physician and hospital coronary angioplasty volume and outcome in elderly patients. Circulation 1997;95(11):2485-91.

43. Hannan EL, Racz M, Ryan TJ, et al. Coronary angioplasty volume-outcome relationships for hospitals and cardiologists. JAMA 1997;277(11):892-8.

44. Dardik A, Burleyson GP, Bowman H, et al. Surgical repair of ruptured abdominal aortic aneurysms in the state of Maryland: factors influencing outcome among 527 recent cases. J Vasc Surg 1998;28(3):413-20.

45. Rosenthal GE, Shah A, Way LE, et al. Variations in standardized hospital mortality rates for six common medical diagnoses: implications for profiling hospital quality. Med Care 1998;36(7):955-64.

46. Cebul RD, Snow RJ, Pine R, et al. Indications, outcomes, and provider volumes for carotid endarterectomy. JAMA 1998;279(16):1282-7.

47. Begg CB, Cramer LD, Hoskins WJ, et al. Impact of hospital volume on operative mortality for major cancer surgery. JAMA 1998;280(20):1747-51.

48. Thiemann DR, Coresh J, Oetgen WJ, et al. The association between hospital volume and survival after acute myocardial infarction in elderly patients . N Engl J Med 1999;340(21):1640-8.

49. Pratt R, Burr G, Leelarthaepin B, et al. The effects of All-RN and RN-EN staffing on the quality and cost of patient care. Aust J Adv Nurs 1993;10(3):27-39.

50. Archibald LK, Manning ML, Bell LM, et al. Patient density, nurse-to-patient ratio and nosocomial infection risk in a pediatric cardiac intensive care unit. Pediatr Infect Dis J 1997;16(11):1045-8.

51. Blegen MA, Goode CJ, Reed L. Nurse staffing and patient outcomes. Nurs Res 1998;47(1):43-50.

52. Czaplinski C, Diers D. The effect of staff nursing on length of stay and mortality. Med Care 1998;36(12):1626-38.

53. Kovner C, Gergen PJ. Nurse staffing levels and adverse events following surgery in U.S. hospitals. Image J Nurs Sch 1998;30(4):315-21.

54. McCloskey JM. Nurse staffing and patient outcomes. Nurs Outlook 1998;46(5):199-200.

55. Bond CA, Raehl CL, Pitterle ME, et al. Health care professional staffing, hospital characteristics, and hospital mortality rates. Pharmacotherapy 1999;19(2):130-8.

56. Wennberg J, Gittelsohn. Small area variations in health care delivery. Science 1973;182(117):1102-8.

57. Wennberg J, Gittelsohn A. Variations in medical care among small areas. Sci Am 1982;246(4):120-34.

58. Markowitz JS, Pashko S, Gutterman EM, et al. Death rates among patients hospitalized with community-acquired pneumonia: a reexamination with data from three states. Am J Public Health 1996;86(8 Pt 1):1152-4.

59. Hofer TP, Wolfe RA, Tedeschi PJ, et al. Use of community versus individual socioeconomic data in predicting variation in hospital use. Health Serv Res 1998;33(2 Pt 1):243-59.

60. Jencks SF, Dobson A. Refining case-mix adjustment. The research evidence. N Engl J Med 1987;317(11):679-86.

61. Shapiro MF, Park RE, Keesey J, et al. The effect of alternative case-mix adjustments on mortality differences between municipal and voluntary hospitals in New York City. Health Serv Res 1994;29(1):95-112.

62. Halm EA, Lee C, Chassin MR. How is volume related to quality in health care? A systematic review of the research literature: Institute of Medicine, National Academy of Sciences, Division of Health Care Services, Committee on Quality of Care in America; 2000 May 1.

63. Thomas J, Holloway J, Guire K. Validating risk-adjusted mortality as an indicator for quality of care. Inquiry 1993;30(1):6-22.

64. Hofer TP, Hayward RA. Can early re-admission rates accurately detect poor-quality hospitals? Med Care 1995;33(3):234-45.

65. Thomas JW. Does risk-adjusted readmission rate provide valid information on hospital quality? Inquiry 1996;33(3):258-70.

66. Normand ST, Glickman ME, Sharma RG, et al. Using admission characteristics to predict short-term mortality from myocardial infarction in elderly patients. Results from the Cooperative Cardiovascular Project . JAMA 1996;275(17):1322-8.

67. Thomas JW, Hofer TP. Research evidence on the validity of risk-adjusted mortality rate as a measure of hospital quality of care. Med Care Res Rev 1998;55(4):371-404.

68. Hofer TP, Hayward RA, Greenfield S, et al. The unreliability of individual physician "report cards" for assessing the costs and quality of care of a chronic disease . JAMA 1999;281(22):2098-105.

69. Normand S, Glickman M, Gastonis C. Statistical methods for profiling providers of medical care: Issues and applications. JASA 1997;92(439):803-814.

70. O'Hagan A. Bayesian Inference. In: al. GPe, editor. Kendall's Advanced Theory of Statistics. New York: Halstead Press; 1994.

71. Goldstein H. Multilevel Statistical Models. 2nd ed. New York: Halstead Press; 1995.

72. Wennberg JE, Freeman JL, Shelton RM, et al. Hospital use and mortality among Medicare beneficiaries in Boston and New Haven. N Engl J Med 1989;321(17):1168-73.

73. Fisher ES, Wennberg JE, Stukel TA, et al. Hospital readmission rates for cohorts of Medicare beneficiaries in Boston and New Haven. N Engl J Med 1994;331(15):989-95.

74. Miller MG, Miller LS, Fireman B, et al. Variation in practice for discretionary admissions. Impact on estimates of quality of hospital care . JAMA 1994;271(19):1493-8.

75. Rosenthal GE, Harper DL, Shah A, et al. A regional evaluation of variation in low-severity hospital admissions. J Gen Intern Med 1997;12(7):416-22.

76. Fisher ES, Wennberg JE, Stukel TA, et al. Associations among hospital capacity, utilization, and mortality of US Medicare beneficiaries, controlling for sociodemographic factors. Health Serv Res 2000;34(6):1351-62.

77. McClellan M, McNeil BJ, Newhouse JP. Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality? Analysis using instrumental variables. JAMA 1994;272(11):859-66.

78. Iezzoni LI, Ash AS, Shwartz M, et al. Judging hospitals by severity-adjusted mortality rates: the influence of the severity-adjustment method. Am J Public Health 1996;86(10):1379-87.

79. Iezzoni LI. The risks of risk adjustment. JAMA 1997;278(19):1600-7.

80. Iezzoni LI, Heeren T, Foley SM, et al. Chronic conditions and risk of in-hospital death. Health Serv Res 1994;29(4):435-60.

81. Jones RH, Hannan EL, Hammermeister KE, et al. Identification of preoperative variables needed for risk adjustment of short-term mortality after coronary artery bypass graft surgery. The Working Group Panel on the Cooperative CABG Database Project. J Am Coll Cardiol 1996;28(6):1478-87.

82. Khuri SF, Daley J, Henderson W, et al. Risk adjustment of the postoperative mortality rate for the comparative assessment of the quality of surgical care: results of the National Veterans Affairs Surgical Risk Study. J Am Coll Surg 1997;185(4):315-27.

83. Wray NP, Hollingsworth JC, Peterson NJ, et al. Case-mix adjustment using administrative databases: a paradigm to guide future research. Med Care Res Rev 1997;54(3):326-56.

84. Kuttner R. The risk-adjustment debate. N Engl J Med 1998;339(26):1952-6.

85. Block PC, Peterson EC, Krone R, et al. Identification of variables needed to risk adjust outcomes of coronary interventions: evidence-based guidelines for efficient data collection. J Am Coll Cardiol 1998;32(1):275-82.

86. Richardson D, Tarnow-Mordi WO, Lee SK. Risk adjustment for quality improvement. Pediatrics 1999;103(1 Suppl E):255-65.

87. Iezzoni LI. Risk adjustment for measuring healthcare outcomes. 2nd ed. Chicago, Ill.: Health Administration Press; 1997.

88. Iezzoni LI, Foley SM, Daley J, et al. Comorbidities, complications, and coding bias. Does the number of diagnosis codes matter in predicting in-hospital mortality? JAMA 1992;267(16):2197-203.

89. Green J, Wintfeld N. How accurate are hospital discharge data for evaluating effectiveness of care? Med Care 1993;31(8):719-31.

90. Malenka DJ, McLerran D, Roos N, et al. Using administrative data to describe casemix: a comparison with the medical record. J Clin Epidemiol 1994;47(9):1027-32.

91. Jencks SF, Williams DK, Kay TL. Assessing hospital-associated deaths from discharge data. The role of length of stay and comorbidities. JAMA 1988;260(15):2240-6.

92. Romano PS, Mark DH. Bias in the coding of hospital discharge data and its implications for quality assessment. Med Care 1994;32(1):81-90.

93. Simborg DW. DRG creep: a new hospital-acquired disease. N Engl J Med 1981;304(26):1602-4.

94. Keeler EB, Kahn KL, Draper D, et al. Changes in sickness at admission following the introduction of the prospective payment system. JAMA 1990;264(15):1962-8.

95. Hsia DC, Krushat WM, Fagan AB, et al. Accuracy of diagnostic coding for Medicare patients under the prospective-payment system. N Engl J Med 1988;318(6):352-355.

96. Green J, Wintfeld N. Report cards on cardiac surgeons. Assessing New York State's approach. N Engl J Med 1995;332(18):1229-32.

97. Hannan EL, Racz MJ, Jollis JG, et al. Using Medicare claims data to assess provider quality for CABG surgery: does it work well enough? Health Serv Res 1997;31(6):659-78.

98. Romano PS, Chan BK. Risk-adjusting acute myocardial infarction mortality: are APR-DRGs the right tool? Health Serv Res 2000;34(7):1469-89.

99. Goldfield N, Averill R. On "risk-adjusting acute myocardial infarction mortality: are APR-DRGs the right tool"? [comment]. Health Serv Res 2000;34(7):1491-5; discussion 1495-8.

100. Jollis JG, Romano PS. Pennsylvania's Focus on Heart Attack--grading the scorecard. N Engl J Med 1998;338(14):983-7.

101. O'Connor GT, Plume SK, Olmstead EM, et al. Multivariate prediction of in-hospital mortality associated with coronary artery bypass graft surgery. Northern New England Cardiovascular Disease Study Group. Circulation 1992;85(6):2110-8.

102. O'Connor GT, Plume SK, Olmstead EM, et al. A regional intervention to improve the hospital mortality associated with coronary artery bypass graft surgery. The Northern New England Cardiovascular Disease Study Group. JAMA 1996;275(11):841-6.

103. Hannan EL, Kilburn H, Jr., Racz M, et al. Improving the outcomes of coronary artery bypass surgery in New York State. JAMA 1994;271(10):761-6.

104. Knaus WA, Draper EA, Wagner DP, et al. An evaluation of outcome from intensive care in major medical centers. Ann Intern Med 1986;104(3):410-8.

105. Knaus WA, Wagner DP, Zimmerman JE, et al. Variations in mortality and length of stay in intensive care units. Ann Intern Med 1993;118(10):753-61.

106. Fine MJ, Auble TE, Yealy DM, et al. A prediction rule to identify low-risk patients with community-acquired pneumonia. N Engl J Med 1997;336(4):243-50.

107. Cooper GS, Chak A, Harper DL, et al. Care of patients with upper gastrointestinal hemorrhage in academic medical centers: a community-based comparison. Gastroenterology 1996;111(2):385-90.

108. Iezzoni LI, Ash AS, Shwartz M, et al. Predicting who dies depends on how severity is measured: implications for evaluating patient outcomes. Ann Intern Med 1995;123(10):763-70.

109. Iezzoni LI, Shwartz M, Ash AS, et al. Predicting in-hospital mortality for stroke patients: results differ across severity-measurement methods. Med Decis Making 1996;16(4):348-56.

110. Iezzoni LI, Shwartz M, Ash AS, et al. Using severity-adjusted stroke mortality rates to judge hospitals. Int J Qual Health Care 1995;7(2):81-94.

111. Iezzoni LI, Shwartz M, Ash AS, et al. Severity measurement methods and judging hospital death rates for pneumonia. Med Care 1996;34(1):11-28.

112. Iezzoni LI, Shwartz M, Ash AS, et al. Using severity measures to predict the likelihood of death for pneumonia inpatients. J Gen Intern Med 1996;11(1):23-31.

113. Pine M, Norusis M, Jones B, et al. Predictions of hospital mortality rates: a comparison of data sources. Ann Intern Med 1997;126(5):347-54.

114. Krumholz HM, Chen J, Wang Y, et al. Comparing AMI mortality among hospitals in patients 65 years of age and older: evaluating methods of risk adjustment. Circulation 1999;99(23):2986-92.

115. Luft H, Romano P, Remy L, et al. Second Report of the California Hospital Outcomes Project: Acute Myocardial Infarction. ,Office of Statewide Health Planning and Development.

116. Iezzoni LI, Ash AS, Shwartz M, et al. Differences in procedure use, in-hospital mortality, and illness severity by gender for acute myocardial infarction patients: are answers affected by data source and severity measure? Med Care 1997;35(2):158-71.

117. Hibbard JH, Jewett JJ, Legnini MW, et al. Choosing a health plan: do large employers use the data? Health Aff (Millwood) 1997;16(6):172-80.

118. Hibbard JH, Jewett JJ, Engelmann S, et al. Can Medicare beneficiaries make informed choices? Health Aff (Millwood) 1998;17(6):181-93.

119. Schneider EC, Epstein AM. Use of public performance reports: a survey of patients undergoing cardiac surgery. JAMA 1998;279(20):1638-42.

120. Booske BC, Sainfort F, Hundt AS. Eliciting consumer preferences for health plans. Health Serv Res 1999;34(4):839-54.

121. Rainwater JA, Romano PS, Antonius DM. The California Hospital Outcomes Project: how useful is California's report card for quality improvement? Jt Comm J Qual Improv 1998;24(1):31-9.

122. Romano PS, Rainwater JA, Antonius D. Grading the graders: how hospitals in California and New York perceive and interpret their report cards. Med Care 1999;37(3):295-305.

123. Palmer RH, Louis TA, Peterson HF, et al. What makes quality assurance effective? Results from a randomized, controlled trial in 16 primary care group practices. Med Care 1996;34(9 Suppl):SS29-39.

124. Duffy SQ, Farley DE. Patterns of decline among inpatient procedures. Public Health Rep 1995;110(6):674-81.

125. Rutledge R. Can medical school-affiliated hospitals compete with private hospitals in the age of managed care? An 11-state, population-based analysis of 351,201 patients undergoing cholecystectomy. J Am Coll Surg 1997;185(3):207-17.

126. Maynard C, Chapko MK, Every NR, et al. Coronary angioplasty outcomes in the Healthcare Cost and Utilization Project, 1993-1994. Am J Cardiol 1998;81(7):848-52.

127. Shepardson LB, Youngner SJ, Speroff T, et al. Increased risk of death in patients with do-not-resuscitate orders. Med Care 1999;37(8):727-37.

128. Layde PM, Broste SK, Desbiens N, et al. Generalizability of clinical studies conducted at tertiary care medical centers: a population-based analysis. J Clin Epidemiol 1996;49(8):835-41.

129. Hannan EL, Kilburn H, Jr., O'Donnell JF, et al. Interracial access to selected cardiac procedures for patients hospitalized with coronary artery disease in New York State. Med Care 1991;29(5):430-41.

130. Buckle JM, Horn SD, Oates VM, et al. Severity of illness and resource use differences among white and black hospitalized elderly. Arch Intern Med 1992;152(8):1596-603.

131. McBean AM, Gornick M. Differences by race in the rates of procedures performed in hospitals for Medicare beneficiaries. Health Care Financ Rev 1994;15(4):77-90.

132. McBean AM, Warren JL, Babish JD. Continuing differences in the rates of percutaneous transluminal coronary angioplasty and coronary artery bypass graft surgery between elderly black and white Medicare beneficiaries. Am Heart J 1994;127(2):287-95.

133. Williams JF, Zimmerman JE, Wagner DP, et al. African-American and white patients admitted to the intensive care unit: is there a difference in therapy and outcome? Crit Care Med 1995;23(4):626-36.

134. Phillips RS, Hamel MB, Teno JM, et al. Race, resource use, and survival in seriously ill hospitalized adults. The SUPPORT Investigators. J Gen Intern Med 1996;11(7):387-96.

135. Romano PS, Campa DR, Rainwater JA. Elective cervical discectomy in California: postoperative in-hospital complications and their risk factors. Spine 1997;22(22):2677-92.

136. Huber TS, Wang JG, Wheeler KG, et al. Impact of race on the treatment for peripheral arterial occlusive disease. J Vasc Surg 1999;30(3):417-25.

137. Chen J, Radford MJ, Wang Y, et al. Do "America's Best Hospitals" perform better for acute myocardial infarction? N Engl J Med 1999;340:286-92.

138. Hartz AJ, Gottlieb MS, Kuhn EM, et al. The relationship between adjusted hospital mortality and the results of peer review. Health Serv Res 1993;27(6):765-77.

139. Hsia DC, Ahern CA, Ritchie BP, et al. Medicare reimbursement accuracy under the prospective payment system, 1985 to 1988. JAMA 1992;268(7):896-899.

140. Cullen DJ, Bates DW, Small SD, et al. The incident reporting system does not detect adverse drug events: a problem for quality improvement. Jt Comm J Qual Improv 1995;21(10):541-8.

141. Kohn L, Corrigan J, Donaldson M, et al., editors. To Err Is Human: Building a Safer Health System. Washington, D.C.: National Academy Press; 2000.

142. Silber JH, Rosenbaum PR, Schwartz JS, et al. Evaluation of the complication rate as a measure of quality of care in coronary artery bypass graft surgery. JAMA 1995;274(4):317-23.

143. Silber JH, Rosenbaum PR, Ross RN. Comparing the contribution of predictors: which outcomes vary with hospital rather than patient characteristics? J Am Stat Assoc 1995;90:7-18.

144. Silber JH, Rosenbaum PR, Williams SV, et al. The relationship between choice of outcome measure and hospital rank in general surgical procedures: implications for quality assessment. Int J Qual Health Care 1997;9:193-200.

145. Kahn KL, Brook RH, Draper D, et al. Interpreting hospital mortality data. How can we proceed? JAMA 1988;260(24):3625-8.

146. Mullins RJ, Mann NC, Hedges JR, et al. Adequacy of hospital discharge status as a measure of outcome among injured patients. JAMA 1998;279(21):1727-31.

147. Sands K, Vineyard G, Platt R. Surgical site infections occurring after hospital discharge. J Infect Dis 1996;173(4):963-70.

148. Sands K, Vineyard G, Livingston J, et al. Efficient identification of postdischarge surgical site infections: use of automated pharmacy dispensing information, administrative data, and medical record information. J Infect Dis 1999;179(2):434-41.

149. Iezzoni LI, Mackiernan YD, Cahalane MJ, et al. Screening inpatient quality using post-discharge events. Med Care 1999;37(4):384-98.

150. Omoigui NA, Miller DP, Brown KJ, et al. Outmigration for coronary bypass surgery in an era of public dissemination of clinical outcomes. Circulation 1996;93(1):27-33.

151. Hannan EL, Siu AL, Kumar D, et al. Assessment of coronary artery bypass graft surgery performance in New York. Is there a bias against taking high-risk patients? Med Care 1997;35(1):49-56.

152. Petersen LA, Orav EJ, Teich JM, et al. Using a computerized sign-out program to improve continuity of inpatient care and prevent adverse events. Jt Comm J Qual Improv 1998;24(2):77-87.

153. Dranove. Information is good except when its bad. ,Stanford University Working Paper.

154. Marshall MN, Shekelle PG, Leatherman S, et al. The public release of performance data: what do we expect to gain? A review of the evidence. JAMA 2000;283(14):1866-74.

155. Johantgen M, Elixhauser A, Bali JK, et al. Quality indicators using hospital discharge data: state and national applications. Jt Comm J Qual Improv 1998;24(2):88-105.

156. Olfson M, Marcus S, Sackeim HA, et al. Use of ECT for the inpatient treatment of recurrent major depression. Am J Psychiatry 1998;155(1):22-9.

157. Meurer JR, Kuhn EM, George V, et al. Charges for childhood asthma by hospital characteristics. Pediatrics 1998;102(6):E70.

158. Lanska DJ, Hartz AJ. Measurement of quality in health care. Neurology 1998;50(3):584-7.

159. Lanska DJ, Kryscio RJ. In-hospital mortality following carotid endarterectomy. Neurology 1998;51(2):440-7.

160. Schnitzler MA, Lambert DL, Mundy LM, et al. Variations in healthcare measures by insurance status for patients receiving ventilator support. Clin Perform Qual Health Care 1998;6(1):17-22.

161. Niederman MS, McCombs JS, Unger AN, et al. The cost of treating community-acquired pneumonia. Clin Ther 1998;20(4):820-37.

162. Zhao SZ, Wong JM, Davis MB, et al. The cost of inpatient endometriosis treatment: an analysis based on the Healthcare Cost and Utilization Project Nationwide Inpatient Sample. Am J Manag Care 1998;4(8):1127-34.

163. Rentz AM, Halpern MT, Bowden R. The impact of candidemia on length of hospital stay, outcome, and overall cost of illness. Clin Infect Dis 1998;27(4):781-8.

164. Ritchie JL, Maynard C, Chapko MK, et al. Association between percutaneous transluminal coronary angioplasty volumes and outcomes in the Healthcare Cost and Utilization Project 1993-1994. Am J Cardiol 1999;83(4):493-7.

165. Best AE. Secondary data bases and their use in outcomes research: a review of the area resource file and the Healthcare Cost and Utilization Project. J Med Syst 1999;23(3):175-81.

166. Krumholz HM, Chen YT, Bradford WD, et al. Variations in and correlates of length of stay in academic hospitals among patients with heart failure resulting from systolic dysfunction. Am J Manag Care 1999;5(6):715-23.

167. Seifeldin R, Hantsch JJ. The economic burden associated with colon cancer in the United States. Clin Ther 1999;21(8):1370-9.

168. Williams GR, Jiang JG, Matchar DB, et al. Incidence and occurrence of total (first-ever and recurrent) stroke. Stroke 1999;30(12):2523-8.

169. Kirsner RS, Yang DG, Kerdel FA. Dermatologic disease accounts for a large number of hospital admissions annually. J Am Acad Dermatol 1999;41(6):970-3.

170. MEDLINE [database online]. Bethesda (MD),National Library of Medicine.1985-Updated weekly.Available from: National Library of Medicine; BRS Information Technologies, McLean, VA; DIALOG information Services, Inc., Palo Alto, CA.

171. PsycINFO. Washington (DC),American Psychological Association.1887-Updated monthly (journals), 1987-current (books and chapters).Available from: American Psychological Association, Washington DC; Ovid Technologies, New York, NY; Silverplatter Information, Norwood, MA.

172. The Cochrane Library. Oxford, England, Update Software, Ltd. Available from: Update Software, San Diego, CA.

173. EMBASE. The Netherlands, Elsevier Science Publishers B.V.1974-Updated weekly. Available from: Elsevier Science, Secondary Publishing Division, New York, NY.

174. Ingber MJ. The current state of risk adjustment technology for capitation. J Ambulatory Care Manage 1998;21(4):1-28.

175. Rosenbaum P, Rubin D. Reducing bias in observational studies using subclassification on the propensity score. J Am Stat Assoc 1894;79:516-24.

176. Glazer J, McGuire T. Minimum variance optimal risk adjustment. American Economic Review 1999; In Press.

177. Chamberlain G. Panel Data. In: Griliches Z, Intriligator M, editors. Handbook of Economics. New York: Elsevier Science; 1984. p. 1247-1318.

178. Green. Economic Analysis. 2nd ed. New Jersey: Prentice Hall; 1997.

179. McClellan M, Staigner D. The quality of health care providers. National Bureau of Economic Research Working Paper #7327.

Appendix C: Log of Revisions to PQI Documentation and Software Version 2.1, Revision 4

The following table summarizes the revisions made to the PQI software, software documentation and the Guide to Patient Safety Indicators (Guide) document in release version 2.1, revision 4. The table lists the component(s) affected by the change and a short summary of the changes that were made.

Component	***Changes***
Software (SAS and SPSS), Software Documentation, and Guide	1. The years for which the ICD-9-CM and DRG codes defining PQIs are valid was amended to be through FY 2005 instead of FY 2004, that is, the codes in the software are effective through September 30, 2005. 2. Added new module that calculates condition-specific rates for the diabetes PQIs across stratifiers.
Software Documentation (SAS and SPSS)	Table 3 was amended to include the 2003 census data (i.e., QICTY03.TXT and QICTYA03.TXT) and condition-specific module files (PQSASC2 and QICTYC03.TXT).
Software (SAS and SPSS)	Added the 2003 census data (i.e., QICTY03.TXT and QICTYA03.TXT) and condition-specific module files (PQSASC2 and QICTYC03.TXT).
Guide	Rearranged the sequence of PQIs to place in numerical order.
Software (SAS)	Inserted “PQ” in format names for age aggregations in SAS programs to distinguish these formats from similarly named formats used by other indicator software.

Appendix D: ICD-9-CM and DRG Coding Updates in PQI Release Version 2.1, Revision 4

For Fiscal Year 2005 (effective 10-1-2004) there were no ICD-9-CM or DRG coding changes that affected indicator definitions. Refer to the Prevention Quality Indicators Archive (http://www.qualityindicators.ahrq.gov/pqi_archive.htm) for changes to indicator definitions made in previous PQI revisions.