Assessing the Quality of Healthcare Provided to Children

Rita Mangione-Smith
Elizabeth A. McGlynn

Reprinted from
Health Services Research
A profile of RAND Health, abstracts of its publications (including all other reprints), and ordering information can be found on the RAND Health home page on the World Wide Web at http://www.rand.org/organization/health/.

RAND is a nonprofit institution that helps improve policy and decisionmaking through research and analysis. The RAND reprint series contains offprints of RAND research originally published in journals or books. The text is reproduced here with permission of the publisher. RAND® is a registered trademark. RAND’s publications do not necessarily reflect the opinions or policies of the sponsors of RAND research.

Published 2000 by RAND
1700 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138
1333 H St., N.W., Washington, D.C. 20005-4707
RAND URL: http://www.rand.org/

To order RAND documents or to obtain additional information, contact Distribution Services: Telephone: (310) 451-7002; Fax: (310) 451-6915; Internet: order@rand.org
Assessing the Quality of Healthcare Provided to Children

Rita Mangione-Smith and Elizabeth A. McGlynn

Objective. To present a conceptual framework for evaluating quality of care for children and adolescents, summarize the key issues related to developing measures to assess pediatric quality of care, examine some existing measures, and present evidence about their current level of performance.

Principal Findings. Assessing the quality of care for children poses many challenges not encountered when making these measurements in the adult population. Children and adolescents (from this point forward referred to collectively as children unless differentiation is necessary) differ from adults in two clinically important ways (Jameson and Wehr 1993): (1) their normal developmental trajectory is characterized by change, and (2) they have differential morbidity. These factors contribute to the limitations encountered when developing measures to assess the quality of care for children.

The movement of a child through the various stages of development makes it difficult to establish what constitutes a “normal” outcome and by extension what constitutes a poor outcome. Additionally, salient developmental outcomes that result from poor quality of care may not be observed for several years. This implies that poor outcomes may be observed when the child is receiving care from a delivery system other than the one that provided the low-quality care. Attributing the suboptimal outcome to the new delivery system would be inappropriate.

Differential morbidity refers to the fact that the type, prevalence, and severity of illness experienced by children is measurably different from that observed in adults. Most children experience numerous self-limited illnesses of mild severity. A minority of children suffer from markedly more severe diseases. Thus, condition-specific measures in children are problematic to implement for routine assessments because of the extremely low incidence and prevalence of most severe pediatric diseases (Halton 1996). However, children with these conditions are potentially the segment of the pediatric population that can be most affected by variations in the quality of care. Improving the care provided to these children is likely to have the largest impact on quality of life and longevity. The low prevalence of most severe pediatric diseases also makes it difficult to evaluate the effectiveness of new treatment modalities; multi-center trials or long enrollment periods are usually required to obtain a large enough patient sample to conduct the necessary randomized controlled trials or cohort studies.

Another challenge encountered when measuring quality of care for children is that, in most cases, they depend on adults to both obtain care and to report on the outcomes of that care. Parents and their children may have different perceptions...
of what defines health or have different levels of satisfaction with the care they receive. Children, particularly those with special needs, also depend on a broad range of services including the medical system, community intervention programs, social programs, and school-based services. Dependency on these various services adds to the difficulty of measuring and appropriately attributing health outcomes observed in children to a particular service delivery entity.

Adolescents also depend on adults for access to some of their care; however, they have special needs related to confidentiality and parent-child information sharing. Adolescents commonly seek care at facilities, such as school-based clinics, that allow them to obtain confidential care. These facilities usually provide out-of-health plan care for these children, which raises special issues related to information availability for quality assessments and for assessing utilization patterns in this population. If the source of poor health outcomes is not known, quality improvement is not possible.

The many challenges faced when constructing pediatric (this term will be used to refer to both children and adolescents) quality of care measures have resulted in few of these instruments being developed specifically for children. Most of the measures developed to date have either a very limited pediatric component or still require the process or outcome validation step. Although several practice guidelines and indicators of quality have been constructed, a conceptual framework to guide the development of such tools for quality assessment in the pediatric population is lacking.

Conclusions. Pediatric health services researchers and the organizations that fund this work need to focus on developing a set of quality assessment tools that will address several challenging issues. Working within the context of the conceptual framework presented, we draw several conclusions related to issues that should be considered in developing quality of care measures for children.

Key Words. Pediatrics, quality measurement, quality of care

---

WHY MEASURE QUALITY OF CARE?

The healthcare system has changed dramatically during the past decade in response to a perceived need to contain rising costs. The resulting changes, including the shift to managed care systems, raise concerns that quality of care may be compromised. To better evaluate the cost-quality trade-off,

Address correspondence to Rita M. Mangione-Smith, M.D., Assistant Professor of Pediatrics, UCLA Department of Pediatrics, 12-467 MDCC, 10833 Le Conte Avenue, Los Angeles, CA 90095-1752. Elizabeth A. McGlynn, Ph.D. is Director, Center for Research on Quality in Health Care, RAND. This article, submitted to Health Services Research on July 8, 1997, was revised and approved for publication on December 23, 1997.
healthcare consumers, purchasers, federal and state agencies, and healthcare plans and providers are demanding better information on quality (Palmer et al. 1996).

Consumers might use quality measures to guide enrollment and provider choices. However, one recent study found that, while a majority of employees said cost and benefit information was essential, few indicated that customer satisfaction results or ratings by independent experts were essential (Tumlinson et al. 1997). A second qualitative study found that a number of consumers have difficulty interpreting the results of many quality indicators (Hibbard 1997). These early findings have led some to suggest that the investment in developing and distributing plan performance information to consumers warrants more research (Tumlinson et al. 1997). One such effort that is currently under way is the Consumer Assessment of Health Plans Study (or CAHPS). This study will in part assess the impact of report card measures on consumers' choice of health plans.

Purchasers, federal and state agencies, healthcare plans, and providers each have their own set of priorities and reasons for wanting to measure clinical performance. Private and public purchasers may want to evaluate whether the plans offered employees or beneficiaries meet adequate performance standards. Healthcare plans need to maintain and improve the quality of care delivered. With the move toward enrolling children with special healthcare needs, and receiving Medicaid, into managed care plans, many pediatric healthcare professionals want the ability to evaluate the effect of such delivery system changes on quality and health status (Fox, Wicks, and Newacheck 1993).

Quality of care measurement is also required to obtain accreditation from both the National Committee for Quality Assurance (NCQA) and the Joint Commission for Accreditation of Healthcare Organizations (JCAHO). In addition, many purchasers are requiring routine reporting of performance measures as a condition of offering plans to employees and beneficiaries. Employers use performance measures such as the Health Plan Employer Data and Information System (HEDIS) to evaluate health plan contracts (e.g., Ford Motor Company), and/or to set employee contributions to premiums (e.g., General Motors, GTE). Measures like HEDIS are designed for making valid quality comparisons among health plans, providers, and health systems. However, HEDIS is also used by health plans to identify and monitor continuous quality improvement (CQI) efforts. Little is known about how individual providers or groups of providers are using performance measures to enhance the quality of care provided.
Quality measures can be used in many different types of research studies. The most appropriate measure to use in a particular study is determined by what the investigators are attempting to understand. A measure that is appropriate for one purpose (e.g., measuring the quality of care received by a general pediatric population in a given health plan) may not be appropriate for another (e.g., measuring the effectiveness of a guideline or course of treatment). Although quality of care measurement is clearly important in internal assessment and improvement, in this article we focus on issues related to the development of external quality measures for assessing healthcare plans, providers, and systems.

A Conceptual Framework for Evaluating the Quality of Care for Children

A conceptual framework is needed to guide the development of quality assessment measures for children. Although Ireys et al. recently developed a conceptual framework for evaluating quality of care for children with special healthcare needs (Ireys, Grason, and Guyer 1996), no such framework exists for the general pediatric population. Our conceptual framework for assessing the quality of care for children is based on dimensions underlying the Institute of Medicine’s (IOM’s) definition of quality; our interpretation of these dimensions is somewhat different from the one presented by Ireys et al.

The IOM defines quality as “the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge” (Institute of Medicine 1990). The definition suggests that

1. quality performance and outcomes occur on a continuum, theoretically ranging from unacceptable to excellent;
2. the scope of inquiry is limited to the structure, process, and outcomes of care provided by the healthcare delivery system;
3. quality may be assessed at multiple different levels;
4. the link between process and outcomes should be established; and
5. research evidence must be used to identify the services that improve health outcomes and in the absence of scientific evidence regarding effectiveness, professional consensus can be used to develop criteria.

The Quality Continuum

The first component of the conceptual framework suggests that quality of care should be measured in a standard and quantifiable way so that fair
comparisons can be made across individual providers, health plans, and populations. Failure to standardize the methods of quality assessment can lead to inaccurate conclusions being drawn and potentially harmful decisions being made.

Although health status is measured on a continuum that ranges from poor to excellent, health outcomes in children will be skewed toward the excellent end of the scale. In general, children are healthy and experience low rates of morbidity and mortality. For example, the prevalence of diabetes mellitus is 56 per 1,000 population in persons 45–64 years old and 110 per 1,000 individuals 65 years or older (U.S. Department of Health and Human Services [DHHS] 1992). In children, the prevalence of this disease is only 1 per 1,000 (DHHS 1992). Thus, in a health plan with 100,000 enrollees, approximately 30 percent would likely be children under age 18 years (30,000), and about 30 of those children would likely have a diagnosis of diabetes. Even common pediatric diseases do not occur with frequencies comparable to those of common adult conditions. In adults, hypertension is a common disease and occurs in 226 per 1,000 individuals ages 45–64 (DHHS 1992). In children, asthma would be considered a common disease and only occurs in 63 per 1,000 individuals (DHHS 1992).

A key issue for quality of care measurement for children, therefore, is the emphasis on preventive care in pediatrics. Guidelines related to both the periodicity and content of well-child and adolescent visits have been developed (Green 1994). Scientific evidence supporting the effectiveness of prevention is limited. Although it is possible to measure adherence to visit periodicity guidelines, measuring content is much more difficult. Parent surveys may be the best method for determining whether health professionals have adequately transmitted anticipatory guidance recommendations. Neither claims data nor medical records are likely to capture the nature and adequacy of such educational interventions.

The fact that children enjoy good to excellent health in most cases affects the utility of mortality and morbidity as outcome measures. Extremely large sample sizes are required to have enough statistical power to detect differences in mortality and morbidity rates resulting from variability in the quality of care provided for specific conditions. Additionally, the condition-specific outcomes approach to quality assessment includes only those children with diseases easy to identify and diseases that occur with moderate frequency (Perrin et al. 1993). This approach excludes children with rare or difficult-to-define conditions who may be most likely to be affected by variations in the quality of care delivered. Improving the care provided to these children
is likely to have the largest effect on quality of life and longevity. Using a condition-specific approach to quality assessment requires the development of measures across a wide range of diseases to assess the overall quality of care delivered in a health plan (McGlynn and Brook 1996).

An alternative approach to quality assessment is to assess the frequency of adverse events (e.g., emergency department use by children with poorly controlled asthma or very low birthweight infants born to mothers who lack access to adequate prenatal care). This method of assessing quality requires that adjustments be made for differences in severity and case mix. If severity adjustments are not made, an increased rate of adverse events for one health plan over another may represent the presence of a sicker patient population in the plan with the higher rate. Unfortunately, no reliable methods for making severity and case-mix adjustments currently exist (McGlynn and Brook 1996). Further, events that occur rarely (commonly the case in pediatric patient populations) are less useful for quality monitoring because it is more difficult to determine whether the differences are statistically significant.

A refinement of this approach involves establishing a set of criteria to assess whether or not an adverse outcome was preventable. Kramer et al. applied this methodology to a case-control study design (Kramer, Arsenault, and Pless 1984). They matched patients with preventable adverse outcomes (defined as preventable complications or delayed diagnosis, treatment, or referral) to patients with acceptable outcomes (defined as uncomplicated illness or prompt diagnosis, treatment, or referral). They then calculated odds ratios to measure the effect of specific structural attributes (e.g., provider type) on these outcomes. The authors argue that measuring outcome of care in children requires a methodology particularly suited to detecting rare events. The advantages of using a case control design are that (1) fewer cases are needed because outcomes are the basis for case selection, and (2) the need for long-term follow-up of a cohort of subjects to detect outcome events of interest is not needed because one starts with outcomes and looks back at process.

The Scope of Inquiry

The second component of the framework sets boundaries for the scope of inquiry for quality assessment. Quality of care can be measured by assessing the structure, process, or outcomes of care (Donabedian 1980). Donabedian makes the argument that the most direct route to an assessment of the quality of care is an examination of the actual process of that care; measuring structure or outcomes is a less direct approach to assessment. Structural
measures assess how system attributes (e.g., access to care for vulnerable populations) and characteristics of healthcare providers and hospitals affect quality (e.g., comparing the outcomes of surgical procedures in children when performed by pediatric surgeons versus general surgeons) (Snow, Cartwright, and Young 1996). Process measures collect data on what happens during the healthcare professional–patient encounter (e.g., assessments within an individual hospital of adherence to established indicators of high-quality care (McCulloch and Vidyasagar 1993). Outcome measures assess the results of health service delivery (e.g., patients’ health status or the patient/parent satisfaction with care).

Structural Measures. The first dimension of the scope of inquiry is structure. For quality assessment, structural measures that predict variations in processes or outcomes of care have the greatest utility. Assessments of pediatric quality of care involving structural measures have focused primarily on hospital characteristics, provider characteristics, and the effect of various types of health insurance coverage on access to care. Previous work in adult quality of care showed that hospital characteristics associated with outcomes of care included the type of hospital, that is, allopathic versus osteopathic (Hartz et al. 1989); size of the facility (Keeler et al. 1992; Flood, Scott, and Ewy 1984); teaching status (Keeler et al. 1992; Brennan et al. 1991; Hartz et al. 1989; Flood, Scott, and Ewy 1984); volume of patients with similar conditions (Burns and Wholey 1991; Farley and Ozminkowski 1992; Flood, Scott, and Ewy 1984); for-profit versus not-for-profit status of the institution (Brennan et al. 1991; Hartz et al. 1989); public versus private ownership of the institution (Hartz et al. 1989); percentage of minority patients served (Brennan et al. 1991); and the geographic location of the institution (Keeler et al. 1992; Brennan et al. 1991). Although the results of these studies were contradictory in some cases, the one consistent finding was that hospitals caring for higher volumes of patients with similar conditions have better adjusted mortality rates (Burns and Wholey 1991; Farley and Ozminkowski 1992; Flood, Scott, and Ewy 1984).

Pediatric studies that examine the link between structural measures and outcomes have been limited because mortality is a rare outcome in children. As a result, most pediatric studies that compare hospital characteristics as structural measures of quality have focused on outcomes in intensive care unit (ICU) patients and trauma patients (Pollack et al. 1991, 1994; Knudson, Shagoury, and Lewis 1997). In order to help guide policy decisions regarding the regionalization of inpatient pediatric care to children’s hospitals, future studies examining hospital characteristics as they relate to quality will need to
examine outcomes other than mortality (e.g., functional status). This will be possible only when assessment tools for measuring such outcomes become more widely available and routinely used in research and practice.

Two studies focusing on provider characteristics as structural measures of quality were identified in the pediatric literature (Snow, Cartwright, and Young 1996; Schulman, Lucchesee, and Sullivan 1995). Both of these investigations illustrate that structural measures are feasible to use in the pediatric population and point to how these measures can be linked to outcomes of care other than mortality. The first of these two studies compared complication rates for children operated on by adult subspecialty surgeons versus pediatric subspecialty surgeons (Snow, Cartwright, and Young 1996). This study found that the pediatric subspecialty surgeons performed a higher volume of the studied surgical procedure and had a lower complication rate than the adult subspecialty surgeons in the study. This finding is consistent with the results of adult studies examining the relationship between physicians’ and hospitals’ procedure volumes and patient outcomes (Hughes, Hunt, and Luft 1987; Hannan et al. 1989).

A second pediatric study of provider characteristics compared weight-specific survival rates and adherence to neonatal intensive care unit (NICU) protocols by pediatric residents versus non-physician providers, that is, nurse practitioners and physician assistants (Schulman, Lucchesee, and Sullivan 1995). This study found equivalent weight-specific survival rates for both types of providers and significantly increased adherence to NICU protocols by non-physician providers.

Some healthcare professionals have expressed concern that the current trend toward enrolling pediatric Medicaid recipients with special healthcare needs into managed care plans will limit their access to needed subspecialty care and community services (Fox, Wicks, and Newacheck 1993; Ireys, Grason, and Guyer 1996). One randomized controlled trial that examined the content of care provided to pediatric Medicaid recipients with special needs in fee-for-service plans versus an HMO found that the health services provided to the two groups were comparable (Mauldon et al. 1994). However, no studies have examined health status outcomes as they relate to type of insurance coverage for children with special healthcare needs. Clearly, as more of this population is enrolled in managed care health plans, investigations examining these potential differences will be essential.

Access to care is limited for children with no medical insurance coverage (Stoddard, St. Peter, and Newacheck 1994; St. Peter, Newacheck, and Halfon 1992; Redlener and Karich 1994). Adolescent females are among those most
likely to be uninsured and to lack access to healthcare services (Davidson, Gibbs, and Chapin 1991). Children receiving Medicaid benefits have better access to care than do children with no insurance; however, it is unknown whether they receive care of a quality equal to the care provided to children covered by private insurance plans (St. Peter, Newacheck, and Halton 1992). Some studies suggest that Medicaid recipients are more likely to use emergency rooms as a source of non-emergent primary care than are privately insured patients (DeAngelis, Fosarelli, and Duggan 1985; Chessare 1986). As state governments continue to mandate managed care enrollment for their Medicaid-eligible populations, concerns are raised about whether changes in the use of emergency department services will affect health outcomes (Glotzer et al. 1991). This population is accustomed to a model of health service delivery that includes limited access to primary care providers in the outpatient setting. The enrollment of these patients in managed care plans necessarily imposes a new model of healthcare delivery on them and their families. In the short term, difficulty in adjusting to this new model will likely decrease parental satisfaction. However, providing this population with a regular source of ambulatory primary care may improve their health outcomes in the future.

Studies examining the impact of denied access to the emergency room by Medicaid managed care plans on health outcomes in children have had mixed results (Shaw, Selbst, and Gill 1990; Gadowski et al. 1995). These studies have limited external validity because they were conducted at single institutions. Many large managed care organizations have patient information tracking systems, which could potentially provide more representative data in future studies of this issue.

Concerns about confidentiality have been identified by both providers and adolescents as having a negative effect on the willingness to seek care. This is particularly true for reproductive services (Davidson, Gibbs, and Chapin 1991), and it represents a significant health systems problem because more than 10 percent of all 15- to 19-year-old females become pregnant annually (Green 1994). Understanding the structural elements that affect adolescents' likelihood of seeking care will facilitate better system design in the future.

Secondary data analyses have been conducted to examine the link between geographic proximity to healthcare facilities and the likelihood of receiving an invasive cardiac procedure such as catheterization or angioplasty in adults (Carlisle et al. 1995); however, no such study has been performed for children or adolescents. Primary data on a population-based sample are currently lacking for both adults and children. One effort that is currently
under way that may provide a great deal of information on the impact of access to care on outcomes is a study being conducted by the Center for Studying Health System Change, under a grant from The Robert Wood Johnson Foundation. This study is examining the effects of changes in the organization and financing of health services on the utilization of services and on selected patient outcomes. The Community Tracking Study (CTS), which is being conducted in 60 communities, provides a unique perspective on these issues because the sample is nationally representative.

Process Measures. The second dimension of the scope of inquiry is process. Donabedian pointed out that elements of the process of care do not signify quality until their relationship to desirable outcomes have been established, that is, until they are outcome-validated. However, he goes on to say that, "once it has been established that certain procedures used in specified situations are clearly associated with good results, the mere presence or absence of these procedures in these situations can be accepted as evidence of good or bad quality" (Donabedian 1980).

Proponents of the use of process measures to assess quality of care in pediatrics argue that effective measurement is necessary to identify opportunities for improvement and to guide the process of improvement (O’Leary 1994). Processes of care can be observed during the time frame of measurement (often 12 months or less), while most outcomes of care in pediatrics require a considerably longer period of observation. The target of quality improvement is usually the physician or the healthcare system. Process measures address aspects of care over which physicians are more likely to have control and to have the benefit of explicitly identifying the criteria by which quality will be evaluated. The results of process measures potentially point the way to getting a better result.

Increasing the use of process of care measures in pediatrics requires the development of explicit indicators of excellent processes across the range of preventive, acute, and chronic care. Much of the existing work on quality for children focuses on preventive care (e.g., immunizations, early initiation of prenatal care). Greater attention to acute and chronic care is warranted. One such effort is currently under way at RAND under funding from the Health Care Financing Administration. This project has developed almost 500 process of care indicators for children across 21 health conditions; the indicators are distributed relatively equally across preventive, acute, and chronic care (Schuster et al. 1997). This system of indicators, designed to be aggregated into summary scores, offers one potential solution to the challenge of assessing quality for diseases of rare incidence and prevalence in children.
Each indicator in the system is rated by the level of scientific evidence supporting it. A national expert panel of physicians was convened to review the entire set of indicators using the RAND/UCLA modified Delphi process. Development of the system should be completed in 1998.

Administrative data sets, medical chart reviews, and/or survey data (patient or clinician) are the usual sources for assessing process of care. These data are used to determine the level of adherence to practice guidelines or sets of quality indicators. However, each of these sources of data has certain limitations and may miss some important processes of care. Using administrative data requires that a claim be filed; therefore, only billable events can be studied. Results associated with laboratory and radiology studies are rarely included in these administrative data sets. Use of medical records requires that the process be documented in the record. Certain processes, such as history and physical examination, may be documented only when the patient has an abnormal finding. Other processes such as counseling or education may lack adequate detail in documentation. Survey data depend on the accuracy and recall of events by respondents. Patients may know they had blood drawn, for example, but not know what tests were run. For many dimensions of process, one source of data may be optimal for conducting assessments. These considerations should be taken into account when developing grading systems to assess performance.

Outcome Measures. The third dimension of the scope of inquiry is outcome. Donabedian stated that outcomes are “the ultimate validation of the effectiveness and quality of medical care” (Donabedian 1966). However, he also noted some limitations to using outcomes when assessing quality. These include the fact that some outcomes are difficult to measure, outcomes can be affected by many factors outside of the medical care system, and relevant outcomes may not be apparent for years. In pediatrics, for example, using brain damage as a criterion of the effectiveness of neonatal resuscitation is complicated by all of these limitations. Brain damage is difficult to define and quantify, it can have many causes other than inadequate resuscitation in the birthing area, and it may not manifest itself until a child reaches school age (McCulloch and Vidyasagar 1993).

The low rates of disease and death in the pediatric population require that future outcome assessments of quality of care for children focus on alternative measures to mortality and morbidity rates. One of these alternatives is to measure health status as an outcome of care. Many investigators have developed measures of pediatric health status. Some of these instruments have focused on a single dimension of health status (Stein and Jessop 1990;
Walker and Greene 1990), while others have included multiple dimensions (Eisen et al. 1979; Lewis, Pantell, and Keickhefer 1989; Starfried et al. 1993; Landgraf, Abetz, and Ware 1996; Starfield 1995). To our knowledge, these measures have not been validated as quality of care measures; that is, we do not know if variations on these measures are reliably and specifically associated with differences in the quality of care delivered.

All but one of these measures (Lewis, Pantell, and Keickhefer 1989) were developed for use across different disease categories. Because of the low prevalence of pediatric disease, the use of generic measures across various disease categories is necessary to achieve adequate sample sizes for comparative analyses (McGlynn, Hafon, and Leibowitz 1995).

Assessing health status as an outcome measure of child healthcare is challenging for several reasons. These include differences between child, parent, and physician perceptions of the value associated with different dimensions of health and variability in what constitutes a "normal" outcome in children of different ages and states of development. Previous work has shown that children and their parents may differ in how they assess the various dimensions of child health (Pantell and Lewis 1987). Most instruments used to assess child health status examine the parent's perspective (Stein and Jessop 1990; Eisen et al. 1979; Lewis, Pantell, and Keickhefer 1989). However, Starfield's Child Health and Illness Profile is one of the few available measures that assesses health status from the child's perspective (Starfield 1995). Another health status measure called the Child Health Questionnaire was recently developed by Landgraf et al. and includes both child- and parent-completed surveys (Landgraf, Abetz, and Ware 1996). Clearly, age, developmental stage, and reading ability may affect the feasibility of using children as respondents. One alternative for younger children that avoids the use of written survey instruments involves having children respond to pictorial representations of various dimensions of health status. The Dartmouth COOP charts have been used for this purpose (Nelson et al. 1990). Future measures should attempt, where possible, to assess health status outcomes from multiple perspectives—the child's, parent's, and physician's.

Because children are still developing, disease can manifest itself as a deceleration of normal development rather than an observable abnormality (Starfield et al. 1993). Future measures will either need to be age- and developmental stage-specific or be amenable to standardization of scale scores to correct for these differences. However, in order to develop standardized scale scores we must have normative data on the population to be assessed. These data are frequently not available for disabled or chronically ill children
(Young and Wright 1995), and may not be available across the range of sociodemographic groups.

The delayed expression of many salient outcomes requires that future research in this area define meaningful short- and intermediate-term outcomes. Some examples of such outcomes might include lost school days, activity limitation days, or bed days. These outcomes will also need to be examined across multiple disease categories to achieve large enough sample sizes to conduct meaningful statistical analyses.

Another frequently used outcome measure is satisfaction with care, which assesses whether respondents’ expectations of the medical care system were met. As with all outcome measures, however, there is little and often conflicting information on the structure and processes of care that produce satisfaction (Korsch, Gozzi, and Francis 1968; Like and Zywnaski 1986; Bertakis, Roter, and Putnam 1991; Brody 1980; Hall and Dornan 1988; Hall et al. 1994; Linder-Pelz and Stewart 1986; Rubin et al. 1993). Establishing this link is critical to the design of quality improvement programs. Furthermore, low levels of satisfaction are relatively uncommon (Hall and Dornan 1988). Finally, while we know that dissatisfaction can predict the increased likelihood of changes in physicians or health plans in adults (Rubin et al. 1993), there is little information on the contribution of parental assessments of satisfaction with children’s healthcare on decisions to change doctors or health plans. For children themselves, there may be little opportunity for their assessments to influence household decision making.

Currently, a collaborative effort among Harvard University, RAND, Research Triangle Institute, and the Agency for Health Care Policy and Research is under way to develop a set of consumer satisfaction survey instruments for national use in making healthcare comparisons. This project is called the Consumer Assessments of Health Plans Study (CAHPS). The goals of this project are to (1) develop and test questionnaires that assess health plans and services, (2) produce easily understandable reports for communicating survey information to consumers, and (3) evaluate the usefulness of these reports for consumers selecting healthcare plans and services. In addition to a core set of items designed for use with all respondents, supplemental questions were designed for use with certain subgroups including families with children and Medicaid recipients. The information from CAHPS questionnaires and reports is intended to help consumers and group purchasers compare health plans and make more informed choices.
The Levels of Quality Assessment

The third component of the proposed theoretical framework refers to the concept that quality of care can be assessed at the individual or population levels. Each of these units of analysis provides different information regarding strategies for quality improvement based on the results of quality measurement. For example, consider adherence to the American Academy of Pediatrics' (AAP's) recommended immunization schedule for children from birth to two years old as a measure of high-quality care. At the individual level, delays or failure to immunize might be improved by encouraging providers to educate parents in their practice on the importance of completing the full immunization series to prevent their child from contracting serious illnesses. At the community level, a large, school-based vaccination program could be implemented to enhance herd immunity in the community.

One key difference between individual and population-based assessments of quality is their relative ability to detect potential underuse of services. Typically, population-based assessments are designed to identify everyone who would be eligible to receive a particular intervention (children under age two for childhood immunizations, for instance, or children with asthma for evaluation of medication use). Failure to receive needed care by some of these children can be identified if the eligible population is known. These evaluations are easiest to conduct when the entire population can be identified and the entity responsible for delivering needed services can be linked to the population. This is one of the reasons for the high level of quality assessment activity in managed care plans. By contrast, individual assessments tend to focus on the quality of care received by those who present for services. While underuse is possible to detect in this group, it will in many cases represent an underestimate.

The level at which to report quality assessments is a subject of considerable debate. Four of the most common recommendations are at the health plan, medical group, facility, or individual physician level. For population-based assessments, the health plan is often the most straightforward level at which to report. However, consumers are likely to find the medical group or physician levels more salient. Three key challenges to quality reporting at the medical group or individual physician level are attribution of responsibility, number of eligible patients, and costs associated with assessing quality. The attribution issue revolves around the need to identify the physician or set of physicians who are responsible for each patient; in a gatekeeper system this might be the primary care physician. In systems without formal assignment it
could be the physician seen most often. Another issue to consider is the group or individual physician's scope of responsibility. Should the pediatrician be held accountable for all care, including that provided by subspecialists? Does responsibility end with making a referral, or must the primary care physician ensure that referral appointments are kept? Once such responsibilities are outlined, the next challenge is whether any one physician has enough patients with a particular condition to make assessments statistically meaningful. For most preventive care and some acute care this is not likely to pose a problem, but for chronic conditions greater difficulties arise. Finally, the burden of data collection greatly increases when one moves from the health plan level to the individual physician level. Large numbers of assessments are most feasible when the measure relies on administrative data and are considerably less feasible when surveys or medical record abstractions are required.

*Establishing the Link Between Process and Outcomes*

The fourth component of the conceptual framework examines the importance of linking process and outcomes. Donabedian stated that "causal validity resides neither in process nor in outcome, but in the link that joins them" (Donabedian 1980). In some cases good outcomes occur despite rather than because of the care we provide. This is particularly the case in pediatrics where most children are naturally resilient and a majority of their illnesses are self-limited. The opposite is also true: bad outcomes may result despite the fact that excellent processes of care were provided. Health outcomes, particularly in vulnerable populations, can result from a complex interaction of medical care, genetic, environmental, and behavioral factors. If we desire to make comparisons among healthcare plans or providers, we can only hold them accountable for aspects of care that are under their control to change. Outcome measures must be adjusted for factors outside the health system if fair comparisons are to be made. Therefore, we may either limit the measures we use to those that assess processes or outcomes under the control of health plans and providers, or we must adjust for significant confounding factors so that variation attributable to health services delivery can be isolated.

Some pediatric investigators have measured adherence to practice guidelines to assess process quality (Cockington, Drew, and Eberhard 1989; McCulloch and Vidyasagar 1993; Zenni and Robinson 1996). However, whether or not adhering to practice guidelines actually improves a child's health outcomes has not been investigated extensively. Portions of some pediatric practice guidelines are based on strong scientific evidence of improved
health outcomes (e.g., the indications for use of penicillin prophylaxis in children included in the AHCPR sickle cell anemia practice guideline) (Agency for Health Care Policy and Research [AHCPR] 1993). However, only one pediatric study has attempted to link adherence to a practice guideline in its entirety with improved health outcomes (Cockington, Drew, and Eberhard 1989). This study examined long-term outcomes in infants treated according to a practice guideline for neonatal jaundice. Infants in the study were assessed at three years of age for physical, neurological, ophthalmological, audiological, and psychological abnormalities. The authors concluded that long-term follow-up of jaundiced infants treated according to the practice guideline revealed a satisfactory outcome. However, this study does not truly present evidence of outcome-validated processes of care because none of the infants who were managed conservatively without phototherapy were assessed for health status outcomes. Fiser recently conducted a multi-center study to validate a scale that measures neurological outcomes in pediatric intensive care patients (Fiser 1992). Ultimately, this scale could be used to outcome-validate the effectiveness of recommended practice guidelines for pediatric advanced life support (Zaritsky et al. 1995).

Scientific Evidence and Professional Consensus

The final component of the conceptual framework addresses the importance of using scientific evidence in the development of quality assessment measures and where scientific evidence is lacking, the need to use expert consensus. Although pediatric investigations have typically used explicit process measures to assess quality of care (Nakayama et al. 1989; Zimmerman and Applebaum 1992; Fisher and Vinci 1995; Sapien et al. 1995; Gloor, Kissoon, and Joubert 1993; Kemper 1988), most of these studies fail to describe how the explicit criteria were selected (Nakayama et al. 1989; Zimmerman and Applebaum 1992; Fisher and Vinci 1995; Sapien et al. 1995). It is not clear whether the criteria used in these studies are based on scientific evidence, determined through a process of professional consensus, or are instead based primarily on the opinions of the investigators, that is, implicit assessments.

In most cases no research exists to support the efficacy or effectiveness of medical or surgical interventions used in pediatric practice (O’Leary 1987). This lack of research evidence is the result of several factors. First, randomized controlled trials (RCTs) are particularly difficult to conduct in pediatrics because of ethical concerns related to carrying out this type of research with children. Children are dependent on their parents for the medical care they receive. Their dependence calls into question their ability to truly give
informed consent regarding receiving an experimental drug or technology. Second, technologies and treatments are frequently accepted as the "standard of care" before they undergo rigorous scientific evaluation. As Cunningham points out, "The fact that we cannot explain many of our patients' symptoms is a frustrating reality, and technology offers a kind of release in the form of a number, a tracing, or an image" (Cunningham 1989). Once a technology or treatment is accepted as standard care, even if this is not supported by research, it becomes unethical to withhold such therapies in RCTs or cohort studies. Third, the prevalence of most pediatric diseases is extremely low making it difficult to collect enough cases in a given RCT or cohort study to have the statistical power to make assessments about efficacy or effectiveness of an intervention. This type of research requires extremely expensive and labor-intensive multi-center trials. For example, clinical studies of treatment effectiveness for sickle cell anemia have frequently involved enrolling patients at multiple different tertiary care centers around the country to acquire a large enough sample size to perform comparative analyses (Bjornson et al. 1996; Charache et al. 1995).

The evidence problem has led to the development of several expert consensus practice guidelines. As defined by Bergman, a practice guideline is an explicit statement of optimal care for the average patient. These guidelines provide a representation of the most current medical knowledge in an action-oriented format that is easily used by physicians (Bergman 1995).

Many practice guidelines addressing common conditions are based on expert consensus rather than scientific evidence (Bergman 1995). In addition, the content of practice guidelines can vary widely depending on the membership of the expert panel and the standards used to judge evidence (Wagner, Herdman, and Alberts 1989). This is illustrated by the different recommendations for well-child care up to age six issued by the American Academy of Pediatrics and the U.S. Preventive Services Task Force (USPSTF). The AAP recommends 13 visits between birth and six years of age and the USPSTF recommends six or seven visits for children without high-risk conditions during this same time period (American Academy of Pediatrics 1985; U.S. Preventive Services Task Force 1989). If a quality of care measure was structured to reflect the AAP guidelines, physicians following the USPSTF guidelines might unfairly be found to be deficient in their delivery of care. This could also result in a bias favoring one specialty over another. For example, if pediatricians followed the AAP guidelines and family practitioners used the USPSTF guidelines, family practitioners would inappropriately be downgraded for the quality of care they were providing. Neither set of standards is
based on evidence demonstrating differential outcomes associated with one or the other schedule. Additionally, periodicity of visits is irrelevant if the content of services delivered is suboptimal or if the content is ineffective in modifying outcomes. Guidelines that are based on frequency and periodicity often suffer from lack of supporting evidence (Wagner, Herdman, and Alberts 1989). Most well-child care practice guidelines are of this type.

In addition to their dependence on expert opinion, for practice guidelines to remain useful they must be routinely updated to reflect advances in science. Quality assessment tools in general must be flexible enough to integrate new scientific evidence as appropriate. One challenge for the future will be determining when new findings are adequate to overturn prior opinion or research.

Some Existing Measures and Their Current Level of Performance

It is beyond the scope of this article to enumerate all of the measures currently in use that apply to the pediatric population. Therefore, we have chosen to highlight a recently developed index of measures, CONQUEST 1.1 (COmputerized Needs-oriented QUality Measurement EVaulation SysTemS), and one quality assessment system that is in widespread use, HEDIS 3.0 (Health Plan Employer Data and Information System), as an illustration of some of the pediatric measures that are available.

CONQUEST 1.1 is a recently developed system commissioned by the Agency for Health Care Policy and Research for cataloging and evaluating clinical performance measures. This system functions as a tool to identify and compare selected performance measures (Palmer et al. 1996). The system provides information about whether psychometric testing of measures has been conducted and whether or not each performance measure has been risk-adjusted. The database does not contain results from the application of the cataloged measures.

Clinical performance measures as defined by the developers of CONQUEST 1.1 are tools that assess the delivery of clinical services. Good clinical performance requires providing services that are appropriate for each patient's condition, providing them safely, competently, and in an appropriate time frame, and achieving desired outcomes in terms of those aspects of patient health and patient satisfaction that can be affected by clinical services (Palmer et al. 1996).

These measures are often condition-specific and focus on the relationship between the processes of care and patient outcomes. This definition of
clinical performance measures excludes (1) individual case reviews because they do not sum across a sample of patients; (2) utilization statistics unless the rates are compared to a standard of quality; (3) unadjusted health status or patient outcome measures (outcome measures that do not adjust for differences in patient characteristics will mainly be comparing case mix rather than differences in quality); and (4) measures of satisfaction.

Currently, the CONQUEST 1.1 database contains 1,185 measures derived from 53 measure sets. Eleven of the measure sets contain 58 measures that apply to pediatric clinical performance. (See the Appendix for a full list of all pediatric measures and their performance assessments.) The majority of these measures (38) were developed for and apply predominantly to the inpatient hospital setting. Only 15 measures were designed primarily for the ambulatory setting and these represent only two aspects of care (childhood immunizations and timing of prenatal care). The remaining five apply to emergency department or ambulatory surgery centers.

Only 50 percent of these pediatric measures have undergone test-retest, interrater, or internal consistency reliability testing. Face validity testing was performed on all of these measures, and convergent validity testing was performed on one-third of them. Criterion validity was not examined for any of these measures; however, this type of validation is difficult because it requires a gold standard measure against which the newly developed measure’s performance can be compared. Currently, there is no such criterion available for comparison. Of the 58 pediatric clinical performance measures cataloged in the CONQUEST 1.1 database, only 23 have been risk adjusted.

Since the release of CONQUEST 1.1, the original HEDIS has been updated to HEDIS 3.0. This new version includes the following effectiveness of care measures that apply to children and/or adolescents: childhood immunizations, adolescent immunizations, prenatal care in the first trimester, low-birthweight babies, and checkups after delivery. Some HEDIS measures are currently reported only in the adult population but could be extended in the future to include adolescents (e.g., cervical cancer screening, advice to quit smoking).

CONCLUSIONS

Few tools exist currently for assessing the quality of pediatric care. This article has provided a review of the current status in the field of quality assessment
in pediatrics and has identified some of the methodological and practical challenges likely to be encountered in developing new tools. A targeted set of research activities in this area could expedite advancements in this field and expand the dimensions of care for pediatrics that could be included in future quality-monitoring activities. The research activities should be organized around a set of principles that we discuss here.

First, quality of care measures for children should be developed in a scientifically rigorous manner. Quality indicators should be explicit statements of desirable (or undesirable) structural, process, or outcome dimensions. These statements should be supported either by research that establishes the efficacy or effectiveness of the indicator or by a formal process of obtaining expert consensus. The actual tools for measurement should be tested in multiple settings and evaluated for reliability, validity, and feasibility. Results should be reported to different audiences in a format that maximizes the likelihood that the information can be interpreted and used in appropriate decision contexts.

Second, although it is important to develop a full complement of measures across the dimensions of structure, process, and outcomes, greater emphasis should be placed on the development of process measures because of their greater utility in both internal and external evaluations of performance. Although preventive services are critical for children, substantial gaps exist in the tools available to assess care for acute and chronic conditions. Relatively greater investments should be made in these areas.

Third, improved tools for assessing child health outcomes will facilitate the development of a strong scientific foundation for quality measures. In particular, work to develop measures of child functional status will expand the areas for which links between the processes and outcomes of care can be evaluated. A particular issue for pediatrics is the design of functional status tools that elicit information directly from younger children who might otherwise be represented only through less desirable proxy measures. Promising approaches in this area include the use of graphics and computer interactive tools.

Effective methods for assessing quality of care delivered to children and adolescents is critical to ensure continued improvements in the functioning of the healthcare system. Failure to develop an adequate toolbox for evaluating quality for children may place them at increased risk for receiving suboptimal care and experiencing avoidable adverse outcomes. Investments in a coordinated research strategy with adequate opportunities to test the implementation and application of these tools are needed.
## Appendix: CONQUEST 1.1 Quality Measure Grading System

<table>
<thead>
<tr>
<th>Measure Name</th>
<th>Measure Set</th>
<th>Validity Testing</th>
<th>Type of Test Used</th>
<th>Reliability Testing</th>
<th>Type of Test Used</th>
<th>Denominator Data Source</th>
<th>Risk Adjustment?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asthma admission rate</td>
<td>HEDA 2.5†</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Data Accuracy</td>
<td>A*</td>
<td>Proposed</td>
</tr>
<tr>
<td>Bloodstream infection</td>
<td>JCAHO IMSysystem‡</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Test-Retest, Interrater Internal Consistency</td>
<td>A,R**</td>
<td>Yes</td>
</tr>
<tr>
<td>Cancellation of ambulatory procedure on day of procedure</td>
<td>MDQIQ§</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Not Specified</td>
<td>A,R,O***</td>
<td>Yes</td>
</tr>
<tr>
<td>Cesarean section</td>
<td>JCAHO IMSysystem</td>
<td>Yes</td>
<td>Face</td>
<td>Convergent</td>
<td>Data Accuracy</td>
<td>A,R</td>
<td>Yes</td>
</tr>
<tr>
<td>Childhood immunization</td>
<td>HEDR 2.5††</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Data Accuracy</td>
<td>A</td>
<td>Proposed</td>
</tr>
<tr>
<td>Childhood immunization</td>
<td>HEDA 2.5</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Data Accuracy</td>
<td>A</td>
<td>Proposed</td>
</tr>
<tr>
<td>Childhood immunization</td>
<td>HEDH 2.5++</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Data Accuracy</td>
<td>A</td>
<td>Proposed</td>
</tr>
<tr>
<td>Childhood immunization</td>
<td>PBGH§§</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Data Accuracy</td>
<td>A</td>
<td>Proposed</td>
</tr>
<tr>
<td>Childhood immunization</td>
<td>PBGH</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Data Accuracy</td>
<td>A</td>
<td>Proposed</td>
</tr>
<tr>
<td>Childhood immunization</td>
<td>PBGH</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Data Accuracy</td>
<td>A</td>
<td>Proposed</td>
</tr>
<tr>
<td>Diagnosis of congestive heart failure</td>
<td>JCAHO IMSysystem</td>
<td>Yes</td>
<td>Face</td>
<td>Convergent</td>
<td>Data Accuracy</td>
<td>A</td>
<td>Yes</td>
</tr>
</tbody>
</table>

*continued*
<table>
<thead>
<tr>
<th>Measure Name</th>
<th>Measure Set</th>
<th>Validity Testing</th>
<th>Type of Test Used</th>
<th>Reliability Testing</th>
<th>Type of Test Used</th>
<th>Denominator Data Source</th>
<th>Risk Adjustment?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digoxin level</td>
<td>JCAHO IMSystem Medication Use</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Interrater</td>
<td>Data Accuracy Internal Consistency</td>
<td>A</td>
</tr>
<tr>
<td>Discrepancy between initial and final x-ray reports requiring change in patient management</td>
<td>JCAHO IMSystem MDQI</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Not Specified</td>
<td>A,R,O</td>
<td>No</td>
</tr>
<tr>
<td>Glasgow coma scale for patients with head injuries</td>
<td>JCAHO IMSystem Trauma</td>
<td>Yes</td>
<td>Face</td>
<td>Convergent</td>
<td>Test-retest</td>
<td>Interrater Data Accuracy</td>
<td>A,R</td>
</tr>
<tr>
<td>Inpatient mortality</td>
<td>MDQI</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Not Specified</td>
<td>A,R,O</td>
<td>Yes</td>
</tr>
<tr>
<td>Laparotomy or thoracotomy for selected trauma patients who die</td>
<td>JCAHO IMSystem Trauma</td>
<td>Yes</td>
<td>Face</td>
<td>Convergent</td>
<td>Test-retest</td>
<td>Interrater Data Accuracy</td>
<td>A,R</td>
</tr>
<tr>
<td>Lithium level</td>
<td>JCAHO IMSystem Medication Use</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Interrater</td>
<td>Data Accuracy Internal Consistency</td>
<td>A</td>
</tr>
<tr>
<td>Low Apgar score in low-birthweight babies</td>
<td>JCAHO IMSystem Prenatal Care</td>
<td>Yes</td>
<td>Face</td>
<td>Convergent</td>
<td>Test-retest</td>
<td>Interrater Data Accuracy</td>
<td>A,R</td>
</tr>
<tr>
<td>Low birthweight</td>
<td>HEDA 2.5</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Data Accuracy</td>
<td>A</td>
<td>Proposed</td>
</tr>
<tr>
<td>Measure Name</td>
<td>Measure Set</td>
<td>Type of Test Used</td>
<td>Type of Reliability Testing</td>
<td>Validity Testing</td>
<td>Type of Validity Testing</td>
<td>Denominator</td>
<td>Data Source</td>
</tr>
<tr>
<td>------------------------------</td>
<td>-------------------</td>
<td>-------------------</td>
<td>----------------------------</td>
<td>-----------------</td>
<td>-------------------------</td>
<td>-------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Low-birthweight</td>
<td>JCAHO IMSytem</td>
<td>Face</td>
<td>Face</td>
<td>Yes</td>
<td>Convergent</td>
<td>A.R</td>
<td>Data Accuracy</td>
</tr>
<tr>
<td>Management of coma in trauma patients</td>
<td>JCAHO IMSytem</td>
<td>Face</td>
<td>Face</td>
<td>Yes</td>
<td>Convergent</td>
<td>A.R</td>
<td>Data Accuracy</td>
</tr>
<tr>
<td>Neonatal mortality</td>
<td>JCAHO IMSytem</td>
<td>Face</td>
<td>Face</td>
<td>Yes</td>
<td>Face</td>
<td>A.R</td>
<td>Data Accuracy</td>
</tr>
<tr>
<td>Number of drugs at discharge</td>
<td>JCAHO IMSytem</td>
<td>Face</td>
<td>Face</td>
<td>Yes</td>
<td>Face</td>
<td>A.R</td>
<td>Internal Consistency</td>
</tr>
<tr>
<td>Patient education regarding insulin</td>
<td>JCAHO IMSytem</td>
<td>Face</td>
<td>Face</td>
<td>Yes</td>
<td>Face</td>
<td>A.R</td>
<td>Data Accuracy</td>
</tr>
<tr>
<td>Perinatal adverse event</td>
<td>JCAHO IMSytem</td>
<td>Face</td>
<td>Face</td>
<td>Yes</td>
<td>Face</td>
<td>A.R</td>
<td>Data Accuracy</td>
</tr>
<tr>
<td>Perioperative mortality</td>
<td>JCAHO IMSytem</td>
<td>Face</td>
<td>Face</td>
<td>Yes</td>
<td>Face</td>
<td>A.R</td>
<td>Data Accuracy</td>
</tr>
<tr>
<td>Phenyl level</td>
<td>JCAHO IMSytem</td>
<td>Face</td>
<td>Face</td>
<td>Yes</td>
<td>Face</td>
<td>A.R</td>
<td>Data Accuracy</td>
</tr>
</tbody>
</table>

*Quality of Healthcare for Children*
Appendix: (Continued)

<table>
<thead>
<tr>
<th>Measure Name</th>
<th>Measure Set</th>
<th>Validity Testing</th>
<th>Type of Test Used</th>
<th>Reliability Testing</th>
<th>Type of Test Used</th>
<th>Denominator Data Source</th>
<th>Risk Adjustment?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Postoperative peripheral neurological deficit</td>
<td>JCAHO IMSysterm Peri-operative Care</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Test-retest Interrater Data Accuracy</td>
<td>A,R</td>
<td>Yes</td>
</tr>
<tr>
<td>Postoperative cardiac arrest</td>
<td>JCAHO IMSysterm Peri-operative Care</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Test-retest Interrater Data Accuracy</td>
<td>A,R</td>
<td>Yes</td>
</tr>
<tr>
<td>Postoperative intrahospital mortality</td>
<td>JCAHO IMSysterm Peri-operative Care</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Test-retest Interrater Data Accuracy</td>
<td>A,R</td>
<td>Yes</td>
</tr>
<tr>
<td>Postoperative CNS complications</td>
<td>JCAHO IMSysterm Peri-operative Care</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Test-retest Interrater Data Accuracy</td>
<td>A,R</td>
<td>Yes</td>
</tr>
<tr>
<td>Prenatal care in the first trimester</td>
<td>HEDR 2.5</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Data Accuracy</td>
<td>A</td>
<td>Proposed</td>
</tr>
<tr>
<td>Prenatal care in the first trimester</td>
<td>HEDH 2.5</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Data Accuracy</td>
<td>A</td>
<td>Proposed</td>
</tr>
<tr>
<td>Prenatal care in the first trimester</td>
<td>HEDA 2.5</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Data Accuracy</td>
<td>A</td>
<td>Proposed</td>
</tr>
<tr>
<td>Prenatal care in the first trimester</td>
<td>PBGH</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Data Accuracy</td>
<td>A</td>
<td>Proposed</td>
</tr>
</tbody>
</table>
| Prenatal care timing             | PBGH                      | Yes              | Face              | Yes                 | Data Accuracy      | S****        | Proposed     | continued
### Appendix: (Continued)

<table>
<thead>
<tr>
<th>Measure Name</th>
<th>Measure Set</th>
<th>Validity Testing</th>
<th>Type of Test Used</th>
<th>Reliability Testing</th>
<th>Type of Test Used</th>
<th>Denominator Data Source</th>
<th>Risk Adjustment?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Registered patients in the emergency room more than 6 hours</td>
<td>MDQI</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Not Specified</td>
<td>A,R,O</td>
<td>Yes</td>
</tr>
<tr>
<td>Registered patients who leave the emergency room prior to completion of treatment</td>
<td>MDQI</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Not Specified</td>
<td>A</td>
<td>No</td>
</tr>
<tr>
<td>Surgical site infection</td>
<td>JCAHO IMS System Infection Control</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Not Specified</td>
<td>Int rater Internal Consistency</td>
<td>Yes</td>
</tr>
<tr>
<td>Surgical wound infection</td>
<td>MDQI</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Not Specified</td>
<td>A,O</td>
<td>Yes</td>
</tr>
<tr>
<td>Theophylline level</td>
<td>JCAHO IMS System Medication Use</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Not Specified</td>
<td>Int rater Internal Consistency</td>
<td>A</td>
</tr>
<tr>
<td>Thoracostomy or thoracotomy for selected trauma patients</td>
<td>JCAHO IMS System Trauma</td>
<td>Yes</td>
<td>Face</td>
<td>Convergent</td>
<td>Test-retest</td>
<td>Interrater Data Accuracy</td>
<td>No</td>
</tr>
<tr>
<td>Timeliness of abdominal surgical intervention for trauma patients</td>
<td>JCAHO IMS System Trauma</td>
<td>Yes</td>
<td>Face</td>
<td>Convergent</td>
<td>Test-retest</td>
<td>Interrater Data Accuracy</td>
<td>No</td>
</tr>
</tbody>
</table>

continued
<table>
<thead>
<tr>
<th>Measure Name</th>
<th>Measure Set</th>
<th>Validity Testing</th>
<th>Type of Test Used</th>
<th>Reliability Testing</th>
<th>Type of Test Used</th>
<th>Denominator Data Source</th>
<th>Risk Adjustment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Timeliness of head CT for trauma patients</td>
<td>JCAHO IMSysytem</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Test-retest Interrater</td>
<td>A,R</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Trauma</td>
<td></td>
<td>Convergent</td>
<td></td>
<td>Data Accuracy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Timeliness of neurosurgical intervention for trauma patients</td>
<td>JCAHO IMSysytem</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Test-retest Interrater</td>
<td>A,R</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Trauma</td>
<td></td>
<td>Convergent</td>
<td></td>
<td>Data Accuracy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Timeliness of orthopedic surgical intervention for trauma patients</td>
<td>JCAHO IMSysytem</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Test-retest Interrater</td>
<td>A,R</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Trauma</td>
<td></td>
<td>Convergent</td>
<td></td>
<td>Data Accuracy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Timing of prophylactic antibiotic administration</td>
<td>JCAHO IMSysytem</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Interrer</td>
<td>A,R</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Medication Use</td>
<td></td>
<td></td>
<td></td>
<td>Data Accuracy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unscheduled admission following an ambulatory procedure</td>
<td>MDQI</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Not Specified</td>
<td>A,R,O</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unscheduled readmission</td>
<td>MDQI</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Not Specified</td>
<td>A,R,O</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unscheduled return to the operating room</td>
<td>MDQI</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Not Specified</td>
<td>A,R,O</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unscheduled return to a special care unit</td>
<td>MDQI</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Not Specified</td>
<td>A,R,O</td>
<td>Yes</td>
</tr>
</tbody>
</table>

continued
## Appendix: (Continued)

<table>
<thead>
<tr>
<th>Measure Name</th>
<th>Measure Set</th>
<th>Validity Testing</th>
<th>Type of Test Used</th>
<th>Reliability Testing</th>
<th>Type of Test Used</th>
<th>Denominator Data Source</th>
<th>Risk Adjustment?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unscheduled return to the emergency room within 72 hours</td>
<td>MDQI</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Not Specified</td>
<td>A,R,O</td>
<td>Yes</td>
</tr>
<tr>
<td>Vaginal birth after cesarean section</td>
<td>JCAHO IMSYstem Prenatal Care</td>
<td>Yes</td>
<td>Face</td>
<td>Convergent</td>
<td>Yes</td>
<td>Data Accuracy</td>
<td>A,R</td>
</tr>
<tr>
<td>Ventilator pneumonia</td>
<td>JCAHO IMSYstem Infection Control</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Internal Consistency</td>
<td>A,R,O</td>
<td>Yes</td>
</tr>
<tr>
<td>Very low birthweight</td>
<td>HEDA 2.5</td>
<td>Yes</td>
<td>Face</td>
<td>Yes</td>
<td>Data Accuracy</td>
<td>A</td>
<td>Proposed</td>
</tr>
<tr>
<td>Vital signs monitoring</td>
<td>JCAHO IMSYstem Trauma</td>
<td>Yes</td>
<td>Face</td>
<td>Convergent</td>
<td>Yes</td>
<td>Data Accuracy</td>
<td>A</td>
</tr>
</tbody>
</table>

*Administrative Data; **Medical Record Data; ***Other Data Source Not Specified; ****Survey Data. 
†Health Plan Employer Data and Information System (HEDIS) Administrative version 2.5. 
‡Joint Commission on Accreditation of Healthcare Organizations IMSYstem. 
§Maryland Hospital Association Quality Improvement Project. 
††HEDIS 2.5 Medical Records. 
‡‡HEDIS 2.5 Hybrid. 
§§Pacific Business Group on Health.
REFERENCES


