Approach to Improving Quality: the Role of Quality Measurement and a Case Study of the Agency for Healthcare Research and Quality Pediatric Quality Indicators

Kathryn M. McDonald, MM

KEYWORDS

• Quality assurance • Quality indicators • Health care
• Pediatrics • Comparative reporting

A 1-year-old undergoes abdominal surgery, and her parents are relieved it is over, start to relax, and make plans for returning home. Three days later, their baby returns to the operating room for reclosure of postoperative disruption of the abdominal wall. A 7-year-old complains loudly to his mother about sharp belly pains in his lower right side; she calls the pediatrician's office and is told to come right in. Lacking transportation, she does not make it in with the boy, and no one calls to see why. That night, he is writhing and they take a cab to the emergency room from where he is rushed to the operating room to remove his rupturing appendix.

These two stories hit us in the gut, quite literally. Anecdotes of suboptimal care make us ask what could have been done to avoid the bad outcome for that child. But how do we know how well the health care system is doing on a larger scale? Data and well-constructed measures can quantify the scale of the problem and

The author is funded by the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD, for the continued support and refinement of the AHRQ Pediatric Quality Indicators. The author reports no conflict of interest.
Center for Primary Care and Outcomes Research, Stanford University, Stanford, CA 94305-6019, USA
E-mail address: kathy.mcdonald@stanford.edu

Pediatr Clin N Am 56 (2009) 815–829
doi:10.1016/j.pcl.2009.05.009 **pediatric.theclinics.com**
0031-3955/09/$ – see front matter

thereby direct efforts to make improvements. From 2000 to 2003, the rate of wound dehiscence requiring a return to the operating room during the same hospitalization remained steady at 0.76 cases per 1000 abdominopelvic surgeries, and the rate of perforated appendix was also unchanged, but much higher, at 31% of appendectomies.[1]

A clarion call has sounded for children's health care[2] to develop interventions to improve outcomes of care, develop comparative quality reports for decision making and accountability, and create incentives that pay for performance. Each of these three goals requires data and measures. This article describes the context for measurement in pediatrics, provides a case study of the development of an indicator set using routinely collected hospital discharge data, and addresses considerations for selecting and using measures in various circumstances.

THE CURRENT MEASUREMENT LANDSCAPE AND PEDIATRICS

Many researchers, governmental agencies at the federal and state levels, provider organizations, and health care payers have instituted quality measurement aimed at performance improvement. In 2006, the Institute of Medicine published ''Performance Measurement: Accelerating Improvement,'' a report from their Committee on Redesigning Health Insurance Performance Measures, Payment, and Performance Improvement Programs.[3] This report highlighted the incredible proliferation of measures and consequent need for standardization to mitigate the potential burdens of excessive measurement and reporting requirements. Because the report resulted from a congressional mandate oriented toward the Medicare program, there was scant coverage of children's health, but there was a clear message that measurement is important to quality improvement.

Measure development in children's health is at an earlier stage compared with adult medicine, with a need for more measures across all settings of care.[4,5] At some stage in the future, though, the same problem of a plethora of measures and insufficient standardization could emerge for children's health. An obstacle to rational measurement in pediatrics—development of the right number of measures, covering the right areas, available at the right moment—may be the fragmented payment terrain and, thus, a lack of a single entity or program (eg, Medicare) that can provide the necessary incentives to make health care quality measurement and underlying information infrastructures as crucial as they are becoming for adult care.[6,7]

Nevertheless, motivation for quality measurement in pediatrics is increasing because reports show that quality gaps exist in pediatric ambulatory, emergency room, and hospital care that are similar to those demonstrated in adult settings.[8–11] Simply applying adult indicators to younger age ranges is insufficient in many situations. Specific challenges arise from the ''four Ds'' that distinguish children from ''little adults:'' differential epidemiology, dependency, demographics, and development.[4,12] Whether developing child health indicators from scratch or based on adult analogs, measure developers must consider the implications of each of these factors to produce robust indicators and comprehensive measure sets (**Table 1**). Although there are special challenges in pediatrics, many of the approaches to indicator development, assessment, and application are similar regardless of age of the patient population of interest.

Numerous frameworks exist for assessing indicators and measure sets, and vary somewhat with regard to the purpose of measurement (eg, comparative reporting among countries,[13] national consensus measures for use in public reporting and quality improvement,[14] and inclusion in a clearinghouse of measures[15]). In addition,

Table 1
Special considerations for pediatric measures

Children's Health Consideration	Description	Implication for Measurement
Differential epidemiology (versus adult care)	Relatively healthy Seldom have multiple illnesses Special populations	Measure preventive care Risk adjustment is simpler Targeted indicators are needed
Dependency	Parents or other adults involved in financing, decisions, care	Care evaluation depends on information derived from more sources
Demographics	From neonate to adolescent More poverty More ethnically/racially diverse	Sample sizes vary by group Risk adjustment potentially more complex
Development	Constantly changing physically, emotionally, and cognitively	Different measures by age group or development stage

composite measures that combine similar measures are becoming more prevalent in health care. Through their consensus process, the National Quality Forum (NQF) is currently reviewing a composite pediatric patient safety measure[16] and developing criteria to assess such composites.[17]

The American Academy of Pediatrics' Steering Committee on Quality Improvement and Management and Committee on Practice and Ambulatory Medicine published a policy statement in early 2008 on "Principles for the Development and Use of Quality Measures." They noted, "measures are an important component of improving quality," believing that "the primary purpose of quality measurement should be to identify opportunities to improve patient care and outcomes, including health status and satisfaction."[18] **Box 1** summarizes their recommendations for measure development and compares these to the NQF framework.[19] The NQF currently leads a national effort aimed at improving American health care through the endorsement of consensus-based standards. In pediatrics, diverse stakeholders have centralized around these efforts, including national, state, and regional groups representing clinicians and health care professionals, consumers, private and public employers, hospitals, and research organizations and institutions (eg, the Agency for Healthcare Research and Quality [AHRQ], the National Association of Children's Hospitals and Related Institutions [NACHRI], Child Health Corporation of America, the Alliance for Pediatric Quality, and the Hospital Quality Alliance).

Despite increasing measurement-related activities, there remain a limited number of indicators available in pediatrics, particularly for care that crosses settings and for ambulatory care. Additional indicator development is needed to feed into endorsement and reporting processes aimed at catalyzing performance improvement. Development efforts geared to specific areas provide the greatest opportunity to provide meaningful measures. For example, in 2003, the Joint Commission spurred activity on inpatient asthma measures for inclusion in their ORYX performance measurement initiative. As a result, the Pediatric Data Quality Systems Collaborative Measures Workgroup formed and focused first on asthma

Box 1
Criteria for measure development (American Academy of Pediatrics) versus Endorsement (National Quality Forum)

American Academy of Pediatrics Recommendations for Quality Measurement

Important issues for children

- Addresses discrepancy between current and ideal practice and the potential for substantial impact
- Enables assessment of disparities for vulnerable populations

Appropriateness

- Takes into consideration the unique characteristics of pediatric populations

Scientific validity

- Evidence-based, transparent, and accurately and reliably assesses what it is intended to measure
- Risk adjustment or stratification as appropriate

Feasibility

- Should not cause any undue burden on clinician, patient, or family, with attention to clear specification instructions, minimal resource use for data collection, adequate sampling
- Tailored to use in proposed practice setting
- Easily interpretable by users—includes use of methods to allow a focus on causes of variation and benchmarks

Improved quality of care

- Focuses on issues that clinicians and health systems can influence

NQF Evaluation Criteria for Endorsement Process

Importance of measuring and reporting

- Targets making gains in quality (safety, timeliness, effectiveness, efficiency, efficiency, equity, patient centeredness) and outcomes where performance varies or is poor

Requirements for consideration

- Available publicly and maintained by a responsible entity
- Fully developed and tested against NQF criteria

Scientific acceptability

- Produces reliable, valid results about quality of care

Feasibility

- Required data are available (retrievable without undue burden) and can be implemented for performance measurement

Usability

- Intended audiences can understand the results of the measure and are likely to find them useful for decision making

care and more recently on the pediatric ICU environment, vetting numerous potential indicators and ultimately seeking and receiving endorsement for several measures for each clinical situation.[20] Two of the asthma measures are already available in ''Hospital Compare''[21] for consumer decision making in hospital selection.

CASE STUDY: DEVELOPMENT OF THE AGENCY FOR HEALTHCARE RESEARCH AND QUALITY PEDIATRIC QUALITY INDICATORS

Another major recent effort, the AHRQ Pediatric Quality Indicators (PDIs), offers an example of measure development that builds on previous efforts and uses a standardized methodology to develop a full set of indicators.[1,22] The approach to the development of the AHRQ PDIs mirrored that used for earlier AHRQ indicator modules, originally developed by the University of California San Francisco–Stanford Evidence-Based Practice Center, with collaborators from other institutions including the University of California Davis and Battelle Memorial Institute.[23] Indicators from previous AHRQ sets provided the candidate list for the initial pediatric module to keep the task manageable within the resources available and to limit the sets to those that could be implemented with routinely collected hospital discharge data. This data source is ubiquitous, providing the ability to make comparisons between individual hospitals or among similar types of hospitals (eg, nonspecialized community hospitals, children's specialty hospitals). On the other hand, administrative discharge data lack important clinical details, making risk adjustment for accurate quality assessment and fair comparisons more challenging.

The team's indicator set development process for the PDIs and other AHRQ indicator modules is serial and iterative, as shown in **Fig. 1**. The first task is to identify all of the possible indicators from multiple sources. Consulting the research literature and seeking ideas from key informants about indicators in use but not reported in the literature result in a candidate list of indicators for a given purpose. The goal for the candidate list is to identify as many potential indicators as possible for the anticipated use. Sometimes there will be areas of clinical significance for which no indicators exist. In these cases, new indicators could be developed by using quality concepts from Donabedian's classic framework of structure, process, and outcome of care.[24] Is there a structural element of care, such as neonatal ICU level, that has been shown to correspond to quality? Is there a process of care, such as weight checks every 2 weeks on patients who have eating disorders, that has not been adopted in a measure but emanates from a clinical guideline or standard of care? Is there an outcome, such as disease-free survival in patients who have acute lymphocytic leukemia, that might vary from one setting to another on the basis of quality of care? Looking at clinical guidelines, practice standards, trigger tools from the patient safety literature, and other evaluations showing a connection between something that might be measured and high-quality care is a more time-intensive approach to identifying potential indicators.

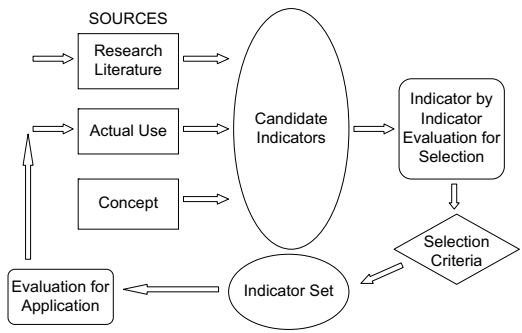

Fig. 1. Quality indicator set development process. This figure shows the major activities undertaken to develop and continually refine indicator sets.

After a candidate list of indicators is assembled, the team's process requires a thorough review of each indicator drawing from published literature, external input, and further research (**Fig. 2**). The overall review is guided by a series of questions relating to the scientific soundness and usability of the indicator (**Fig. 3**). When evidence from the readily available sources is inconclusive, the evaluation process requires a clinical panel review.

This structured review by a multidisciplinary group of clinicians establishes consensual validity, which "extends face validity from one expert to a panel of experts who examine and rate the appropriateness of each item."[25] For the PDIs review, 19 national organizations nominated 125 individuals, from which 45 panelists were selected to create four panels with applicable and diverse membership in terms of setting (eg geographic region, community hospital, children's hospital, and ambulatory care) and field of practice (eg, adolescent medicine, allergy/immunology, cardiology, cardiothoracic surgery, critical care medicine, emergency medicine, endocrinology, family medicine, general pediatrics, general surgery, hematology/oncology, hospital medicine, infectious diseases, pulmonology, neonatology, neurosurgery, nursing, radiology, and urology).

Adapting the RAND/UCLA Appropriateness Method,[26] the process for clinical panel assessment consists of sending each panelist a packet of background information and a questionnaire for each indicator, reviewing each panelist's ratings and comments, and using areas of disagreement to develop an agenda for moderated discussion by way of telephone with all panel members, followed by final independent assessment by clinician panelists using the same questionnaire. For the PDIs, a previous clinical panel review questionnaire was modified to seek input on different indicator applications, specifically comparative reporting among hospitals and quality improvement within a hospital. In earlier work on the AHRQ Patient Safety Indicators, quality indicators were mostly used for quality improvement, and comparative

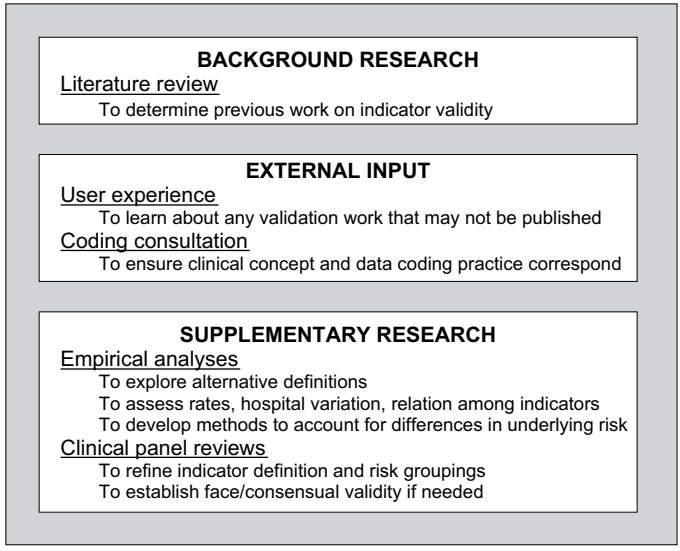

Fig. 2. Quality indicator evaluation process. The three boxes show the major activities required for evaluation of a single indicator. The data sources and goals are displayed within each box.

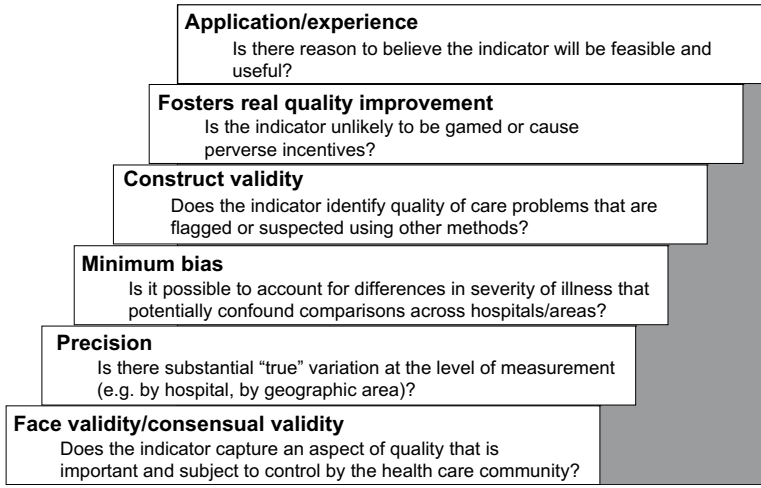

Fig. 3. Evaluation criteria and guiding questions. The figure presents the building blocks to assessing individual indicators for scientific soundness, unintended consequences of use, and potential acceptability.

reporting applications were rare. Additional indicators for the PDIs set included rates of potentially avoidable hospitalization, which required a different set of questions about threats to validity than those used for patient safety metrics.[22,27] Given the changes in the measurement environment, future indicator assessments could also evaluate usefulness for pay-for-performance applications.

Of the 30 indicators in the AHRQ Quality Indicators that were potentially applicable to pediatrics, 18 indicators were selected for inclusion in the PDIs module.[1] Some were removed from consideration prior to panel review because of issues raised by empiric analyses. For example, pneumonia mortality is too rare in children to be used as a reliable measure. Input from users and unpublished chart reviews led to elimination of the failure-to-rescue indicator. Some indicators were eliminated or changed based on clinical panel review. The clinical review results for the final set of indicators are shown in **Table 2**.

Selection is relative, in that an indicator set attempts to bring in the best indicators at the time that meet a reasonable level of credibility in response to the quality indicator validity framework (see **Fig. 3**). As indicator sets are made available, further research and actual use of the indicators provides additional data regarding validity. As shown in **Fig. 1**, this information also can feed back into the process of indicator refinement, may suggest potential revisions to the set as a whole, and expand knowledge for interpreting indicator results in practice. The AHRQ Quality Indicator program provides ongoing support and access to the research team (by way of e-mail to www.support@qualityindicators.ahrq.gov), making technical assistance and this feedback loop seamless.

APPLICATION OF THE AGENCY FOR HEALTHCARE RESEARCH AND QUALITY PEDIATRIC QUALITY INDICATORS

An important early application of the AHRQ PDIs module is research.[28–30] More in-depth study may guide appropriate use and interpretation. For example, NACHRI supported a chart review study of cases flagged by 11 indicators from the PDIs set at

Table 2
Indicators selected for inclusion in the Agency for Healthcare Research and Quality Pediatric Quality Indicators

		Panel Usefulness Rating		76-hospital Evaluation	
PDIs	Age Range[a]	Internal QI	Comparative Reporting	PPV Range[b]	Preventable Events/y
Accidental puncture or laceration	0–17 y	Yes (–)	No	32.4–68.0	286–601
Decubitus ulcer	0–17 y	Yes (–)	Yes (–)	51.4–78.9	157–241
Foreign body left in during procedure	0–17 y	Yes	Yes	44.4–80.0	14–26
Iatrogenic pneumothorax (in neonates at risk)	<28 days, birth weight of 2500 g or less	Yes	Yes (–)	10.0–20.0	2–4
Iatrogenic pneumothorax (in nonneonates)	0–17 y, excludes those who have a birth weight of 2500 g or less	Yes	Yes (–)	29.0–64.2	50–111
Pediatric heart surgery mortality rate	<18 y, excludes those <30 d with PDA closure only	Yes	Yes	n/a	n/a
Pediatric heart surgery volume	Same as above	n/a	n/a	n/a	n/a
Postoperative hemorrhage and hematoma	0–17 y	Yes	Yes	12.9–57.1	18–79
Postoperative respiratory failure	0–17 y	Yes	Yes (–)	13.9–39.8	106–305
Postoperative sepsis	0–17 y, neonates excluded	Yes (–)	No	25.6–66.9	177–464

Postoperative wound dehiscence	0–17 y	Yes	Yes	34.1–75.0	10–22
Selected infection due to medical care	0–17 y	Yes (–)	No	40.0–80.7	605–1221
Transfusion reaction	0–17 y, neonates excluded	Yes	Yes (–)	0.0–20.0	0–1
Area-level potentially preventable hospitalizations					
Asthma	2–17 y	Yes	Yes (–)	n/a	n/a
Diabetes short-term complications	6–17 y	Yes	No	n/a	n/a
Gastroenteritis	3 mo–17 y	Yes (–)	No	n/a	n/a
Perforated appendix	1–17 y	Yes (–)	Yes (–)	n/a	n/a
Urinary tract infection	3 mo–17 y	Yes (–)	No	n/a	n/a

The notation "Yes (–)" reflects a rating with less concordance among panelists than a simple "Yes" without the minus sign.
Abbreviations: n/a, not accessed; PPV, positive predictive value; QI, quality indicator.

[a] All indicators have specific inclusion, exclusion, and risk adjustment defined in detail elsewhere.[22] All exclude neonates who had birth weight less than 500 g, obstetric patients, and normal newborns (if applicable).

[b] PPV for preventable cases relative to all cases flagged by indicator, after excluding cases present on admission. Low estimates more certain and high estimates possible, depending on clinical interpretation.

Data from McDonald K, Romano P, Davies S, et al. Measures of pediatric health care quality based on hospital administrative data: the pediatric quality indicators. Rockville (MD): Agency for Healthcare Research and Quality (AHRQ); 2007; and Scanlon MC, Harris JM, Levy F, et al. Evaluation of the Agency for Healthcare Research and Quality Pediatric Quality Indicators. Pediatrics 2008;121(6):e1723–31.

76 children's hospitals. The study quantified the effect of missing information from the administrative data source about whether a complication of care was present on admission.[27] Three indicators had high rates of flagged cases present prior to the hospital admission: decubitis ulcer (40%), postoperative sepsis (40%), and selected infections caused by medical care (43%). With new requirements for hospital administrative data sets to distinguish whether a complication occurred during the hospital stay, and with modifications to the PDI software to use this information, the accuracy of finding cases of potentially preventable complications will improve. The study also examined the preventability of complications experienced by patients, calculating positive predictive values for finding preventable complications relative to all complications uncovered by the PDI (see **Table 2**). The assessment of preventability depends on the individual clinician evaluating the chart and the state of medical science at the time. A complication that is not preventable today may be amenable to strategies developed in the future to reduce its occurrence.

Similarly to the NACHRI study, hospitals can run (or have another organization analyze) the entire PDIs set with their data to efficiently produce a subset of cases, and then drill down with chart review and additional data capture to guide quality improvement efforts. In the NACHRI review that started with 1.7 million hospitalizations, approximately 25,000 discharges (1.5%) were flagged for further study. Alternatively, a hospital quality program officer might use specific indicators that line up with particular quality improvement initiatives and monitor indicator results over time. These and other applications to quality improvement may benefit from comparison to national data on the PDIs, which are available by query from the Healthcare Cost and Utilization Project (HCUP) Web site (www.HCUPnet.ahrq.gov).[31] Any comparative use of a quality indicator necessitates appropriate risk adjustment to the extent feasible with the given data source. The AHRQ PDIs have this capability built into the software.

Figs. 4 and **5** are based on the HCUP Nationwide Inpatient Sample extracted from HCUPnet and show national estimates for adjusted rates for 3 years—2001, 2003, and 2005—for the four area-level indicators with population denominators (see **Fig. 4**) and for the hospital-level indicators that netted at least 100 potentially preventable complications per year among the 76 hospitals in the NACHRI study (see **Fig. 5**). Rates have not changed much over the period, indicating opportunity for improvements in the future.

In addition to uses by researchers and hospitals, the PDIs are under review or already implemented in several state reports.[32,33] Some reports are not officially mandated and are made available by the research community, such as an analysis of 2005 state data for Tennessee of the five area-level PDIs that identify potentially avoidable pediatric hospitalizations (PAPH).[34] This report's application of the PDIs found 6725 PAPH discharges that cost hospitals a total of $17.6 million in 2005. Of note, the diabetes short-term complication rate was much higher than the national average, indicating a useful starting point for delving into potential barriers to high-quality primary care for those diabetic children who were hospitalized. Texas, like some other states, has ongoing mandates for comparative public reporting of hospital data. With the availability of the AHRQ PDIs, the Texas data organization considered potential inclusion of the indicators in its annual reports. A collaborative developed between hospitals and the state's representatives to anticipate and address each other's needs in the short- and long-range.[35] Selected AHRQ PDIs were incorporated and released at the end of 2007 and are available for comparison of hospitals.[36] Meanwhile, the collaborative is working toward incorporating other sources of data and indicators into public reports.[35] The ongoing challenge of these efforts is finding publicly available, clinically rich data collected in a standardized fashion in all relevant settings.

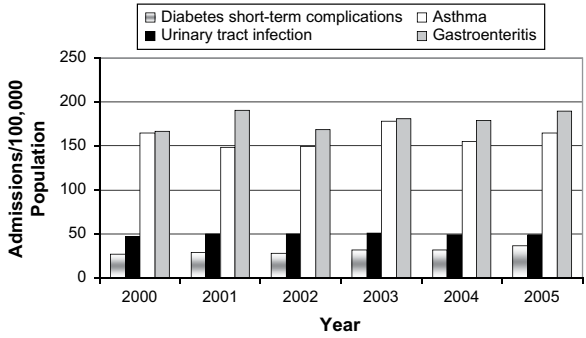

Fig. 4. Area-level PDI rates. The graph presents the risk-adjusted rates of four indicators from the AHRQ PDIs. Diabetes short-term complications admissions include ketoacidosis, hyperosmolarity, or coma in children aged 6 to 17 years. Urinary tract infection admissions include children aged 3 months to 17 years, and exclude patients who have disorders of the kidneys or urinary tract and patients in an immunocompromised state. Asthma admissions for children aged 2 to 17 years exclude patients who have cystic fibrosis or respiratory system anomalies. Gastroenteritis admissions for children 3 months to 17 years exclude patients who have gastrointestinal abnormalities or bacterial gastroenteritis. All area-level indicators exclude patients transferred from another institution. Each indicator is calculated and risk adjusted according to detailed specifications available at: www.qualityindicators.ahrq.gov, and is derived from the HCUP Nationwide Inpatient Sample for the given year and Version 3.1 of the AHRQ PDIs software.

CHALLENGES IN INDICATOR SELECTION AND USE

As with the Texas experience and that of others, real stumbling blocks exist in selecting and using measures. If we wait for ideal data sources, then progress will be slow on the path toward quality improvement; however, in using data sources that are missing some of the information that would make assessments more meaningful, we risk misinterpretation. This risk can be mitigated if we consider the use of a particular measure and apply it appropriately, benefiting from the indicator's strengths and carefully accommodating its deficiencies.

Quality measurements are part of the diagnostic toolbox for understanding what ails the health care system at a given point in time. As with patient care, interpreting a diagnostic test result is a function of the likelihood of the patient having a disease and the performance characteristics of the test. We have mounting evidence about gaps in quality, and so measures that are "good enough" and not necessarily as high performing on sensitivity and specificity as we might ultimately want can move us toward directing quality improvement activities to good effect.

But thinking about any adverse consequences to treating prematurely or empirically in a given context merits careful consideration. If quality indicators are applied directly to quality improvement, then adverse consequences or overinspection of the indicators' shortcomings may distract providers from investigating and treating potential clinical and organizational sources of suboptimal care that could relate to the indicator finding, if true. Thus, the assumption that the indicator result is pointing to a deficiency could be followed up with an assessment about which interventions might reduce the likelihood of poor performance on the metric. Empiric therapy would result in implementing the interventions and seeing whether improvement in the indicator follows.

If the quality indicators are used for public comparative reporting, then adverse consequences of a false positive include patients choosing a hospital they think is

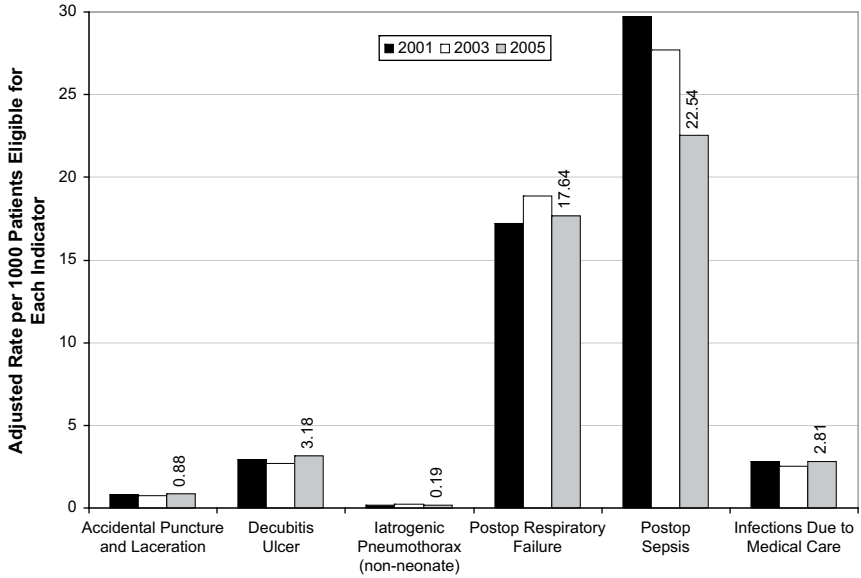

Fig. 5. Selected hospital-level AHRQ PDIs. The graph presents the risk-adjusted rates for six indicators from the AHRQ PDIs. All indicators rely on secondary diagnosis coding to count a complication of care for the numerator (after excluding patients in whom the complication is unlikely to be preventable to the extent feasible with administrative data) and use patients at risk for the complication as the denominator. For example, postoperative sepsis is defined as the number of patients who have sepsis (secondary diagnosis for septicemia, septic shock, inflammatory response attributable to infectious process, and so forth) per 1000 eligible admissions (all surgical patients aged 0–17 years, except with a principal diagnosis of sepsis or infection, neonates, and stays under 4 days). Each indicator is calculated and risk adjusted according to detailed specifications available at: www.qualityindicators.ahrq.gov, and is derived from the HCUP Nationwide Inpatient Sample for the given year and Version 3.1 of the AHRQ PDI software. Postop, postoperative.

better, when in fact it is not, and having a worse outcome from their care. In addition, the hospitals and care providers measured will expend lots of time and energy to support their credible claim that the measure is giving an incorrect reading on the actual quality delivered. Thus, public exposure argues for more robust measures and a good system of conveying the confidence interval reflecting the measurement error. If, however, comparative reporting is used in hospital consortiums and away from the public's eye, then the situation and thresholds for "good enough" measurement are more akin to those in the quality improvement application.

Finally, measure use for reimbursement is becoming commonplace. This movement offers the possibility of rewarding systems and providers who offer the best quality of care. This newer territory, however, makes the need for high-fidelity measures more pressing. Potential adverse consequences of using less refined measures depend on the specific deficiencies in measurement and the particular circumstance but could include (1) changing practice behavior to avoid taking care of sicker patients (eg, "cherry-picking") if risk adjustment is inadequate, (2) giving less attention to that which is not measured but may be vital clinically, and (3) moving the focus to the measures as opposed to the actual care of the children and support of their families. As with other applications, the issue is not only the measurement's performance characteristics but

also what steps are taken on the basis of results from the measure. Even pay-for-performance could use less developed measures if the initial results were followed by careful audits and efforts to supplement the data and analysis to assure fairness.

SUMMARY

With a large pipeline of measures for potential NQF endorsement, with the recent developments to not pay for some iatrogenic complications (for Medicare patients), and with public knowledge about quality variation, measures are here to stay. They can cause some distress, given the exposure of both accurate and inaccurate findings about quality of care. The pediatric field has some advantages compared with the adult counterpart, as it enters more deeply into measurement territory. The smaller number of patients makes the world of pediatrics cozier and potentially more collaborative, and more lessons are available from the intense scrutiny in the adult-care arena. These factors offer an opportunity to the pediatrics community to view the development and application of measures from the perspective of an evolutionary process of continuous quality improvement. Past history has shown that criticism of measures will not obviate the need for this vital tool for quality improvement. Instead, critiques can make the measures better. Within this context, measures and their uses are best approached the same way we think about quality itself—through a continuous cycle of improvement.

ACKNOWLEDGEMENTS

I thank Sheryl Davies, Corinna Haberland, Jeffrey Geppert, Patrick Romano, and Olga Saynina for their ongoing teamwork and contributions to quality measure research and continuous improvement of the AHRQ PDIs. Thanks also to Kristin Cox for administrative assistance and help with background research.

REFERENCES

1. McDonald KM, Davies SM, Haberland CA, et al. Preliminary assessment of pediatric health care quality and patient safety in the United States using readily available administrative data. Pediatrics 2008;122(2):e416–25.
2. Perrin JM, Homer CJ. The quality of children's health care matters—time to pay attention. N Engl J Med 2007;357(15):1549–51.
3. Committee on Redesigning Health Insurance Performance Measures, Payment, and Performance Improvement Programs. Performance measurement: accelerating improvement. Washington DC: National Academies Press; 2005.
4. Beal AC, Co JP, Dougherty D, et al. Quality measures for children's health care. Pediatrics 2004;113(1 Pt 2):199–209.
5. Dougherty D, Simpson LA. Measuring the quality of children's health care: a prerequisite to action. Pediatrics 2004;113(1 Pt 2):185–98.
6. Shaller D. Implementing and using quality measures for children's health care: perspectives on the state of the practice. Pediatrics 2004;113(1 Pt 2):217–27.
7. Menachemi N, Struchen-Shellhorn W, Brooks RG, et al. Influence of pay-for-performance programs on information technology use among child health providers: the devil is in the details. Pediatrics 2009;123(Suppl 2):S92–6.
8. Mangione-Smith R, DeCristofaro AH, Setodji CM, et al. The quality of ambulatory care delivered to children in the United States. N Engl J Med 2007;357(15):1515–23.

9. Knapp JF, Simon SD, Sharma V. Quality of care for common pediatric respiratory illnesses in United States emergency departments: analysis of 2005 National Hospital Ambulatory Medical Care Survey Data. Pediatrics 2008;122(6): 1165–70.
10. Miller MR, Zhan C. Pediatric patient safety in hospitals: a national picture in 2000. Pediatrics 2004;113(6):1741–6.
11. Woods D, Thomas E, Holl J, et al. Adverse events and preventable adverse events in children. Pediatrics 2005;115(1):155–60.
12. Simpson L, Dougherty D, Krause D, et al. Measuring children's health care quality. Am J Med Qual 2007;22(2):80–4.
13. Kelley E, Hurst J. Health care quality indicators project conceptual framework paper. OECD Health Working Paper. 2006. Available at: http://www.oecd.org/dataoecd/1/36/36262363.pdf. Accessed January 29, 2009.
14. NQF Measure Evaluation Criteria. Available at: http://www.qualityforum.org/about/leadership/measure_evaluation.asp. Accessed January 29, 2009.
15. National Quality Measures Clearinghouse Inclusion Criteria. Available at: http://www.qualitymeasures.ahrq.gov/submit/inclusion.aspx. Accessed January 29, 2009.
16. NQF Candidate Consensus Standards as of October 29, 2008. Available at: http://www.qualityforum.org/pdf/tblcandidatemeasurescurrent.xls. Accessed January 29, 2009.
17. Kaplan S, Normand S. Conceptual and analytical issues in creating composite measures of ambulatory care performance. Standardizing Ambulatory Care Performance Measures Project (Phase 3). Available at: http://www.qualityforum.org/pdf/projects/CEF/5%2024%2007%20Kaplan-Normand.pdf. Accessed January 29, 2009.
18. Hodgson ES, Simpson L, Lannon CM. Principles for the development and use of quality measures. Pediatrics 2008;121(2):411–8.
19. The National Quality Forum Measure Evaluation Criteria. Available at: http://www.qualityforum.org/about/leadership/tbEvalCriteria2008-08-28Final.pdf. Accessed February 1, 2009.
20. Pediatric Data Quality Systems (Pedi-QS) Collaborative Measures Workgroup. Available at: http://www.pediqs.com/home.html. Accessed January 29, 2009.
21. Hospital Compare. A quality tool provided by medicare. Available at: http://www.hospitalcompare.hhs.gov/Hospital/Static/About-HospQuality.asp?dest=NAV%7CHome%7CAbout%7CQualityMeasures. Accessed January 29, 2009.
22. McDonald K, Romano P, Davies S, et al. Measures of pediatric health care quality based on hospital administrative data: the pediatric quality indicators. Rockville (MD): Agency for Healthcare Research and Quality (AHRQ); 2007.
23. AHRQ quality indicators. Available at: http://qualityindicators.ahrq.gov/general_faq.htm. Accessed January 29, 2009.
24. Donabedian A. Evaluating the quality of medical care. Milbank Mem Fund Q 1966;44(3 Suppl):166–206.
25. Green L, Lewis F. Measurement and evaluation in health education and heath promotion. Mountain View (CA): Mayfield Publishing Company; 1998.
26. Fitch K, Bernstein S, Aguilar M, et al. The RAND/UCLA Appropriateness Method user's manual. Santa Monica (CA): RAND; 2001.
27. McDonald K, Romano P, Geppert J, et al. Measures of patient safety based on hospital administrative data—the patient safety indicators. Technical Review 5 (Prepared by the University of California San Francisco-Stanford Evidence-based Practice Center under Contract No. 290-97-0013). AHRQ Publication

No. 02-0038. Rockville (MD): Agency for Healthcare Research and Quality; 2002.
28. Scanlon MC, Harris JM, Levy F, et al. Evaluation of the agency for healthcare research and quality pediatric quality indicators. Pediatrics 2008;121(6): e1723–31.
29. Herrod HG, Chang CF. Potentially avoidable pediatric hospitalizations as defined by the Agency for Healthcare Research and Quality: what do they tell us about disparities in child health? Clin Pediatr (Phila) 2008;47(2):128–36.
30. Smith RB, Cheung R, Owens P, et al. Medicaid markets and pediatric patient safety in hospitals. Health Serv Res 2007;42(5):1981–98.
31. HCUPnet, Healthcare Cost and Utilization Project. Agency for Healthcare Research and Quality. Available at: http://hcupnet.ahrq.gov. Accessed January 29, 2009.
32. State of New Jersey. Department of Health and Senior Services: Office of Health Care Quality Assessment. Available at: http://www.state.nj.us/health/healthcarequality/index.shtml. Accessed January 24, 2009.
33. Community engagement. Iowa Healthcare Collaborative. 2008 Iowa Report. Available at: http://www.ihconline.org/iowareport/2008/documents/IowaReport2008.pdf. Accessed January 24, 2009.
34. Chang C, Herrod H, Steinberg S. Potentially avoidable pediatric hospitalizations in Tennessee, 2005. Memphis (TN): Methodist LeBonheur Center for Healthcare Economics, The University of Memphis; 2008.
35. Levy FH, Henion JS, Harris JM. Journey toward meaningful pediatric quality metric reporting: the Texas experience. J Healthc Qual 2008;30(3):36–42, 50.
36. Quality of Children's Care in Texas Hospitals, 2006. Center for Health Statistics, Texas Health Care Information Collection. Available at: http://www.dshs.state.tx.us/THCIC/Publications/Hospitals/PDIReport/PDIReport.shtm. Accessed January 29, 2009.