Quality Measures in Pediatric Hospital Medicine: Moneyball or Looking for Fabio?

Mark W. Shen, MD,¹ and Jack Percelay, MD²

¹Dell Children’s Medical Center, Austin, Texas
²E.L.M.O. Pediatrics, New York, New York

“What are we doing here?” Brad Pitt asks a table full of weathered scouts in a memorable scene in the movie Moneyball. Playing Oakland Athletics manager Billy Beane, he goes on to critique the scouts’ traditional, subjective ranking of players as “looking for Fabio.” The film is based on a book of the same name that chronicles the record-breaking success of the 2002 Oakland Athletics baseball team. At the heart of this achievement was the application of novel evidence-based metrics to maximize performance in the setting of a limited supply of fiscal resources.

Optimizing value has evolved into a larger movement in sports, and quality measurement in medicine would be envious of a similar trajectory. Although reporting and regulatory requirements in medicine have increased in recent years, our metrics remain crude, often with an inability to attribute value to distinct players. Within pediatrics, hospital-based measures comprise <5% of the total available measures, despite the fact that inpatient care accounts for 40% of pediatric health care spending.¹ Of the specific pediatric measures endorsed by the National Quality Forum for use in the inpatient setting (Table 1),²–⁵ few have relevance to the bulk of daily care provided by pediatric hospitalists.

The measures that are the most relevant to pediatric hospital medicine (PHM), the Joint Commission’s Children’s Asthma Care (CAC) core measures for inpatient asthma, have recently been called into question in light of a large study that revealed no immediate link between components of the measures and readmission rates for asthma.⁶ Although many hospitals, over time, increased compliance with the home management plan of care requirement, readmission rates did not budge. Thus, there was systems-level process improvement but no change in outcome. Meanwhile, compliance rates for reliever and systemic corticosteroid use were too high to analyze potential relation with outcomes; measurement reflected high performance but could not drive improvement. This study raises larger questions about the “how” and “why” of quality metrics.

There are many reasons why the quality of care should be measured:

- Improvement
- Meeting standards
- Comparisons (for consumers/purchasers of health care)
Identifying outliers for reward or punishment
Monitoring changes over time

Similarly, several frameworks may be used to organize measurement:

- Donabedian: structure, process, outcomes
- Misuse-underuse-overuse
- Institute of Medicine dimensions: safe, equitable, timely, efficient, effective, patient centered
- Other: by location, provider, and/or diagnostic category

A simplified approach is that measurement may benefit performance either through selection or improvement.8 The process of selection, whereby consumers or regulatory agencies "punish" low performers by selecting out higher quality care providers, limits choices within the system by narrowing the spectrum of quality providers. The process of improvement seeks to shift the performance distribution of the entire population to a higher range (Fig 1). The Joint Commission’s CAC core measures are an example of the former, designed and publicly reported to allow for evaluation or selection of sites based on their quality of care. Measurement for selection may stimulate providers to effect change within systems, and many hospitalist groups were likely motivated to improve rates of asthma action plan delivery. However, high performers had little incentive to improve their care based on these measures, illustrating that measurement for selection does not always lead to improvement.

Measurement for selection is perhaps best left to national and regulatory agencies. For many reasons, this process is understandably slow and less than ideal as the sole driver of widespread improvement. Measures designed for public reporting must typically meet predefined criteria (or standards) for levels of validity, reliability, and feasibility. As exemplified by the CAC core measures, even a good faith effort to define worthwhile measures may fall below expectations.6

Cautionary tales of unintended consequences have similarly emerged from the adult experience with national measures.9 Small sample sizes, low occurrence rates, and low volume at most hospitals have further limited the applicability of many measures, adult and pediatric.10,11 In addition, under careful analysis, even the validity of certain measures, such as postoperative respiratory failure, may be called into question.12 These concerns have led to calls for even more stringent requirements for national measures, including: (1) strong research foundation (based on >1 study); (2) accurate capture of the evidence base; (3) focus on a process proximate to the outcome of interest; and (4) implementation that has little or no chance of inducing unintended adverse consequences.13 Although the stakes for such measures are high, increasingly rigorous requirements will ultimately slow the pace at which measures are developed.

If national measurement for the purpose of selection is too slow to produce

---

**TABLE 1 National Quality Forum–Endorsed Inpatient Pediatric Measures**

<table>
<thead>
<tr>
<th>Agency</th>
<th>Measures</th>
</tr>
</thead>
</table>
| Agency for Healthcare Research and Quality | Hospital-level:  
  - Accidental puncture or laceration  
  - Pressure ulcer  
  - Foreign body left in during procedure  
  - Central venous catheter-related bloodstream infections  
  - Iatrogenic pneumothorax in neonates  
  - Iatrogenic pneumothorax  
  - Neonatal mortality  
  - Bloodstream infections in neonates  
  - Pediatric heart surgery mortality  
  - Pediatric heart surgery volume  
  - Postoperative hemorrhage or hematoma  
  - Postoperative respiratory failure  
  - Postoperative sepsis  
  - Postoperative wound dehiscence  
  - Transfusion reactions  
  Area-level indicators:  
  - Asthma admissions  
  - Diabetes, short-term complications  
  - Gastroenteritis admissions  
  - Perforated appendix admissions  
  - Urinary tract infection admissions |
| The Joint Commission | CAC core measures:  
  - Reliever medication  
  - Systemic corticosteroids  
  - Home management plan of care  
  Immunization:  
  - Pneumococcal polysaccharide  
  - Influenza |
| Centers for Disease Control and Prevention |  
  - Catheter-associated bloodstream infection  
  - Central line-associated bloodstream infection  |
any wide-scale, real-time change, then the priority should be measurement for improvement as a potentially more effective and efficient means of driving performance. Metrics, therefore, should be defined from the ground-level perspective of an individual hospitalist or hospital medicine group with a focus on how to best advance quality care. It should be acknowledged that attribution of quality (whether individual versus group or hospitalist versus emergency physician) will often be raised as a legitimate concern, particularly as patients may encounter a variety of providers during an uncomplicated hospitalization. However, the issue of attribution remains one more relevant to selection and less to the longitudinal journey of a patient through an integrated and accountable system. Thus, by moving away from a focus on judging groups or individuals, the path toward meaningful inpatient pediatric quality measures becomes clear. Questions of “what” or “whom” we should measure transform into “What should we improve?”

A roadmap for improvement must necessarily begin with areas that have the most potential for improvement, with evidence of variable or substandard care. Conditions or processes with high volumes and/or representing significant burdens of disease may also be used to help identify target measures. Although high-level reviews of gaps in care are underway, each group will ultimately need to define measures that best represent the potential for improvement at each site. Just as spending on players for professional sports teams is limited by local markets and salary caps, so too will each group need to individualize potential dashboards based on local value equations. Examples of such work already exist. Based on volume, morbidity, and established variability in current care, potential targets for improvement within PHM are listed in Table 2.

All of these prospective metrics may be critiqued based on a lack of validity, reliability, feasibility, risk stratification, and/or clear linkage to outcomes of interest. In addition, some are outcome measures without a clear evidence-based path toward improvement. However, it remains abundantly clear that there are unexplainable levels of unwarranted variation in those same processes. There is room for us to do better. We cannot afford to wait for perfect measures before we embark on our journey to improve the quality of the health care that we deliver to hospitalized children. Thus, returning to the example of asthma, the absence of meaningful measures should not deter our quest toward improved outcomes. As each group looks to move forward, the first step on that journey will be to identify the best initial targets for improving the value of care delivered locally.

<table>
<thead>
<tr>
<th>Category</th>
<th>Disease/Condition</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volume</td>
<td>Bronchiolitis</td>
<td>↓ Steroid, bronchodilator, radiograph utilization</td>
</tr>
<tr>
<td></td>
<td>Pneumonia</td>
<td>↓ Broad-spectrum antibiotic usage</td>
</tr>
<tr>
<td></td>
<td>Asthma</td>
<td>↓ Readmission rates</td>
</tr>
<tr>
<td></td>
<td>Gastroenteritis</td>
<td>↑ Enteral rehydration rates</td>
</tr>
<tr>
<td></td>
<td>Skin and soft tissue infections</td>
<td>↓ Vancomycin usage</td>
</tr>
<tr>
<td></td>
<td>Codes on the floor</td>
<td>↓ Preventable codes on the floor</td>
</tr>
<tr>
<td></td>
<td>Serious safety events (definition may be institution-specific)</td>
<td>↑ Days between events</td>
</tr>
<tr>
<td>Morbidity</td>
<td>Sepsis</td>
<td>↓ Failure to rescue rate</td>
</tr>
<tr>
<td></td>
<td>Osteomyelitis/septic arthritis</td>
<td>↓ Discharge with central catheter</td>
</tr>
<tr>
<td>Universal processes</td>
<td>Hand-washing</td>
<td>↑ Rates of compliance</td>
</tr>
<tr>
<td></td>
<td>Communication to medical home</td>
<td>↑ Rates on discharge</td>
</tr>
<tr>
<td></td>
<td>Patient satisfaction</td>
<td>↑ Scores related to physician communication</td>
</tr>
<tr>
<td></td>
<td>Readmission and risk-adjusted length of stay for acute conditions</td>
<td>↓ Length of stay and readmission rate (linked together)</td>
</tr>
</tbody>
</table>
As pediatric hospital medicine groups mobilize to measure across similar domains, crude data will begin to be available for benchmarking that may facilitate collaborative learning across sites. In an environment in which the primary goal is improved care for children, and not simply the selection of the “best” providers of care, open sharing of rates and lessons learned will increase the pace at which change occurs. The importance of transparent measurement of performance has been advocated nationally within PHM and is not a new concept in pediatrics. A moving story has been that of cystic fibrosis programs over the past half century. The catalyst for change was 1 center publicizing survival rates that were unthinkably better than the norm, prompting incredulous responses but subsequent sharing of “secrets” across sites. Survival rates increased dramatically over this time period, and these early lessons have prompted the creation of formal learning collaboratives grounded in improvement science that aim to further mobilize busy clinicians toward a common aim.

Home runs such as increased survival will be hard to come by in PHM, which is dominated by acute, self-limited conditions. However, with the hospital as the most costly and complex site of care, reducing inpatient waste becomes a burning platform for change in our inefficient system. Just as we can no longer afford to wait on national measures to guide us, so too must we act now to sustain our health care system, if not our national economy. Prompted by an ethical call to arms, several professional societies have responded by organizing grassroots campaigns to move toward reducing unnecessary care in medicine. Highlighting clinical decisions such as early imaging for low back pain in adults and routine use of antibiotics for sinusitis, overintervention is clearly in the strike zone.

We propose that PHM join these movements by measuring and reducing waste. Pediatric hospitalists have a unique perch overlooking the intersection of inefficient outpatient, emergency, and intensive care. Opportunities for value abound, including reducing chest radiographs (and subsequent antibiotic use) in lower respiratory tract disease, eliminating use of central intravenous catheters for routine osteomyelitis/septic arthritis, eliminating overuse in bronchiolitis, and reducing rates of unnecessary intravenous hydration. An effective campaign will require thoughtful consideration of measures balanced with prioritized and motivational goals.

CONCLUSIONS
Waiting for national agencies to define a measurement agenda will not result in timely, meaningful change. The fundamental goal of measurement should be to drive improvement. This change may then be accelerated through open sharing and collaboration. Although a variety of frameworks may be used to define prospective metrics, high-impact goals should be identified on the basis of their ability to improve the value of health care delivered. Reducing waste is a moral imperative, and an easy target to get us on base. Now is our time to move beyond subjective searches for Fabio and mobilize as a field toward playing real Moneyball.

REFERENCES


