Quality Indicators for High Acuity Pediatric Conditions

abstract

OBJECTIVE: Identifying gaps in care and improving outcomes for severely ill children requires the development of evidence-based performance measures. We used a systematic process involving multiple stakeholders to identify and develop evidence-based quality indicators for high acuity pediatric conditions relevant to any emergency department (ED) setting where children are seen.

METHODS: A prioritized list of clinical conditions was selected by an advisory panel. A systematic review of the literature was conducted to identify existing indicators, as well as guidelines and evidence that could be used to inform the creation of new indicators. A multiphase, Rand-modified Delphi method consisting of anonymous questionnaires and a face-to-face meeting of an expert panel was used for indicator selection. Measure specifications and evidence grading were created for each indicator, and the feasibility and reliability of measurement was assessed in a tertiary care pediatric ED.

RESULTS: The conditions selected for indicator development were diabetic ketoacidosis, status asthmaticus, anaphylaxis, status epilepticus, severe head injury, and sepsis. The majority of the 62 selected indicators reflect ED processes (84%) with few indicators reflecting structures (11%) or outcomes (5%). Thirty-seven percent ($n = 23$) of the selected indicators are based on moderate or high quality evidence. Data were available and interrater reliability acceptable for the majority of indicators.

CONCLUSIONS: A systematic process involving multiple stakeholders was used to develop evidence-based quality indicators for high acuity pediatric conditions. Future work will test the reliability and feasibility of data collection on these indicators across the spectrum of ED settings that provide care for children. *Pediatrics* 2013;132:752–762

AUTHORS: Antonia S. Stang, MDCM, MBA, MSc,[a,b] Sharon E. Straus, MD, MSc, FRCPC,[c] Jennifer Crotts, RN,[a,d] David W. Johnson, MD,[a,e] and Astrid Guttmann, MDCM, MSc[f]

[a]Division of Emergency Medicine, Alberta Children's Hospital, Alberta Children's Hospital Research Institute, Departments of [b]Pediatrics and Community Health Sciences, [d]Pediatrics, and [e]Pediatrics, Physiology and Pharmacology, University of Calgary, Calgary, Alberta, Canada; [c]Departments of Medicine and Geriatric Medicine, University of Toronto, and Knowledge Translation Program, Li Ka Shing Knowledge Institute, St. Michael's Hospital and University of Toronto, Toronto, Ontario; and [f]Division of Pediatric Medicine, Hospital for Sick Children, Department of Pediatrics and Health Policy, Management and Evaluation, University of Toronto and Institute for Clinical Evaluative Sciences, Toronto, Ontario, Canada

Assessing the quality of health care is an international priority.[1–4] Research has revealed that performance measurement improves health care outcomes.[5,6] According to the commonly referenced Donabedian framework, quality indicators are explicitly defined and measurable items referring to the structures (staff, equipment, and facilities), processes (prescribing, investigations, interactions between professionals and patients), or outcomes (mortality, morbidity, or patient satisfaction) of care.[4,7,8] Quality indicators have been developed for a number of health care settings, including emergency departments (EDs).[9–13]

However, despite the fact that children are frequent users of emergency care,[14] there is a lack of research on indicators specific to the pediatric population. For example, <5% of children are affected by the 3 conditions most frequently addressed in adult outcomes research (diabetes, heart disease, and arthritis).[15] Similarly, quality measures that are part of pediatric emergency practice have not been systematically developed or validated.[15] Performance measures specific to pediatrics and pediatric emergency medicine have been identified as a research priority.[16–18]

Evidence indicates that there is substantial practice variation for pediatric patients among emergency care providers, and that many providers do not optimally manage seriously injured or ill children.[19–21] Most of the recent work on practice variation and lack of adherence to practice guidelines in the pediatric ED setting has been done on common, often lower acuity conditions,[21–26] despite evidence of a similar gap between knowledge and practice in severely ill and injured children.[20] Identifying gaps in care for high acuity conditions, where improvement is likely to have the largest impact on quality of life and longevity,[19] requires valid and reliable quality indicators. The objective of this project was to use a systematic process involving multiple stakeholders to review existing indicators and develop new indicators for high acuity pediatric conditions relevant to any ED setting where children are seen.

METHODS

We used a systematic, multiphase, Rand-modified Delphi method congruent with the process for quality indicator development as outlined by the Agency for Healthcare Research and Quality (AHRQ)[27] and used in previous indicator work.[4,7,10,11,13,28] Ethics approval for this study was obtained by the Conjoint Health Research Ethics Board of the University of Calgary.

Phase 1: Selection of Target Conditions

We convened a 32-member advisory panel to select target conditions. The panel included representatives from stakeholder organizations, emergency medicine clinicians, administrators, and decision-makers from the United States and Canada. Panel members were identified by contacting stakeholder organizations and the ED directors of all pediatric EDs, and a sample of rural and general EDs across Canada, and asking them to provide names of individuals with expertise in pediatrics, ED care, or quality improvement. We analyzed health administrative data from 2006 to 2008 on the main diagnosis for high acuity pediatric patients, defined as patients age 0 to 19 years old prioritized as resuscitation and emergent by using the Canadian Triage and Acuity Scale[29] (Supplemental Information 3). We used these data on the most frequent main diagnoses seen in all EDs in Ontario and Alberta, representing 50% of Canada's population, to construct an initial list of potential conditions. We provided panelists with frequency data on the initial list and invited them to suggest additional conditions. In an e-mail survey, panelists were asked to use a scale from 1 (strongly disagree) to 9 (strongly agree)[27,30] based on National Quality Forum (NQF) measure evaluation criteria[31] to rate the final list of conditions on the following: importance (potential for morbidity or mortality associated with the condition), impact (potential to address gap between current and best practice), and validity (adequacy of scientific evidence linking performance of care to patient outcome). The survey was tested for face validity before dissemination and e-mail reminders were sent at weeks 2 and 3 to optimize response. It was decided a priori that conditions with mean scores ≥7 across all 3 criteria by ≥70% of respondents would be retained.

Phase 2: Indicator Identification and Development

We conducted a systematic review of the literature to identify existing indicators and high quality national and international guidelines, systematic reviews (SRs), and randomized controlled trials that could be used to generate new indicators for the selected conditions. The search strategy (Supplemental Table 5) was developed by a medical research librarian in consultation with the research team. We searched the following bibliographic databases: PubMed, Cumulative Index to Nursing and Allied Health Literature, Embase, the Cochrane Collaboration and Evidence-Based Emergency Medicine, Database of Abstracts of Reviews of Effects for SRs, the National Guideline Clearinghouse, the Canadian Medical Association InfoBase, the National Electronic Library for Health, Turning Research into Practice, and Best Bets from 1980 to September 2010. Targeted hand searches of relevant journals and conference proceedings (Supplemental Table 6) were conducted for a 3-year period (2007–2010). Due to resource constraints, only articles in English were included. Guidelines and existing

indicators were also identified by searching specialty society Web sites, internationally recognized guidelines such as Pediatric Advanced Life Support and Advanced Trauma Life Support, and Web sites that focus on quality and performance improvement (Supplemental Table 7).

Two research team members (Dr Stang and Ms Crotts) independently screened all titles, abstracts, and guidelines. The reviewers included for full text review any articles or guidelines that either reviewer thought might provide existing indicators for the target conditions or relevant clinical recommendations that could be used to guide indicator development. Two reviewers (Dr Stang and Ms Crotts) then independently reviewed all full text articles and selected for final inclusion any articles that reported on quality indicators for the identified conditions. The Appraisal of Guidelines Research and Evaluation instrument was applied independently by 2 reviewers (Dr Stang and Ms Crotts) to assess the quality of the guidelines.[32] We developed new indicators from high quality national or international guidelines. We defined high quality as recommended or strongly recommended by both reviewers by using the Appraisal of Guidelines Research and Evaluation instrument. Criteria for indicator development included the following: (1) the strength of the recommendation with only strong recommendations considered[33]; (2) the consistency of the recommendation between guidelines; and (3) the strength of the evidence linking ED structure or care process to patient outcome.[8] We developed new indicators based on consensus between 2 researchers (Dr Stang and Ms Crotts), and review by remaining authors (Drs Guttmann, Straus, and Johnson). Two reviewers (Dr Stang and Ms Crotts) independently assessed the quality of the evidence upon which each indicator is based by using the Grading of Recommendations Assessment, Development and Evaluation (GRADE)[34] system with 1 = very low quality (expert opinion), 2 = low quality, 3 = moderate quality, and 4 = high quality.[34] We used SRs and randomized controlled trials identified in the literature search to grade the strength of the evidence supporting a link between the performance of care specified by the indicator and patient outcomes.

Phase 3: Indicator Selection

We convened an expert panel of 14 individuals, consisting of general and pediatric emergency medicine clinicians, a nurse manager, a pediatric intensivist, quality improvement and safety researchers, and ED administrators (Supplemental Information 2). The panelists were selected based on recommendations by members of the advisory panel and represented the full spectrum of ED settings where children are seen.

We used a modified Delphi technique consisting of 2 rounds of anonymous questionnaires and a face-to-face meeting of the expert panel to generate a final list of indicators. Before completing the first questionnaire, panelists were e-mailed a description of the goals of the research project, the full list of existing and newly developed indicators classified according to Donabedian's framework (structure, process, and outcome),[8] and the grading of the quality of evidence. The first e-mailed questionnaire, sent in December 2010, was pilot tested for face validity, and e-mail reminders were sent at weeks 3 and 4 to optimize response. Panelists were asked to rate the identified indicators on the criteria of (1) relevance to the care of high acuity pediatric patients seen in any ED setting and (2) the degree to which measurement of the indicator would impact the quality of care provided. Panelists were asked to rate each indicator on both criteria by using a Likert scale from 1 (strongly disagree) to 9 (strongly agree).[27,30] We used a predetermined decision rule that any indicators rated ≤3 on both criteria by all panelists would be discarded from further consideration.

The expert panel met in-person in January 2011 after completion of the first survey. At the meeting, the panelists reviewed anonymized ratings for each indicator. Panelists were also provided with their individual ratings from the first survey and given the opportunity to suggest additional indicators. At the end of the meeting, panelists were asked to independently re-rate each indicator by using the same criteria of relevance and impact and prioritize the top 5 indicators for each condition considered the most important to measure to improve quality of care and patient outcomes. Based on previous indicator development work[28] and consensus by the expert panel, we used a predetermined decision rule that indicators rated ≥7 (moderately agree) across both criteria by ≥70% of panelists would be included in the final list. Transcripts of the meeting were documented.

Phase 4: Feasibility of Indicator Measurement

Data from January 2009 to December 2010 on the selected indicators were collected retrospectively from a tertiary care, pediatric ED with 65 000 annual visits. The goal of this phase was to determine the feasibility, described by the NQF as the extent to which the data are readily available, retrievable without undue burden, and can be implemented for performance measurement, and reliability of data collection.[31] Based on the NQF description and previous pediatric indicator development work,[35] we defined feasibility as measures that could be generated by using existing data sources including chart review, physician order entry, and ED patient tracking systems. Performance measurement in a pediatric ED also provided an initial estimate of practice variation and provider compliance with

care processes or structures specified by the indicators. We created a standard profile of measure specifications, the methods by which the target population is identified and the data actually collected.[36] This included the *International Classification of Diseases, 10th Revision* (ICD-10) diagnostic codes and inclusion criteria for each condition, and the specific data elements, such as numerator, denominator, and exclusions, for each indicator (Supplemental Information 3). Data abstractors (3 experienced chart reviewers) used a standardized database (Access 2010) that was piloted for accuracy and clarity on a sample of 10 patient visits for each condition. Interrater reliability was calculated for a random sample (10%) of charts by using intraclass correlation (ICC), Cohen's unweighted κ, or proportion agreement depending on data elements abstracted.[37–40]

RESULTS

Figure 1 summarizes the process of indicator selection.

Phase 1: Selection of Target Conditions

Ninety-one percent (32/35) of invited advisory panel members agreed to participate and identified 13 potential conditions. The number of high acuity pediatric ED visits for these conditions over a 2-year period ranged from 24 for meningitis to 1458 for burns (Table 1). Eighty-one percent (26/32) of the advisory panel members completed the condition prioritization survey, and 6 conditions had a mean score of ≥ 7 by $\geq 70\%$ of respondents (Table 1).

Phase 2: Indicator Identification and Development

Table 2 reveals the results of the search of the literature and quality improvement Web sites. We identified 46 existing indicators for the 6 targeted conditions. We derived 51 new indicators from rec-
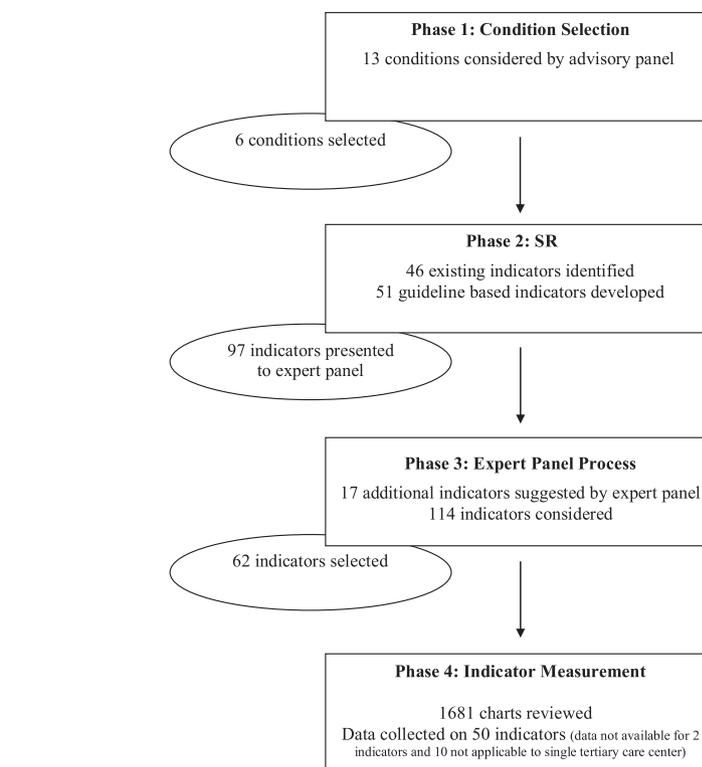

Phase 1: Condition Selection
13 conditions considered by advisory panel

6 conditions selected

Phase 2: SR
46 existing indicators identified
51 guideline based indicators developed

97 indicators presented to expert panel

Phase 3: Expert Panel Process
17 additional indicators suggested by expert panel
114 indicators considered

62 indicators selected

Phase 4: Indicator Measurement
1681 charts reviewed
Data collected on 50 indicators (data not available for 2 indicators and 10 not applicable to single tertiary care center)

FIGURE 1
Process of indicator selection.

ommendations contained in national and international guidelines. The interrater reliability for determining the level of evidence upon which each indicator was based was acceptable ($\kappa > 0.6$)[39] for all conditions except status epilepticus ($\kappa = 0.15$). The overall interrater reliability for the GRADE rating was $\kappa = 0.68$.

Phase 3: Indicator Selection

We presented 97 indicators to the expert panel for initial rating, and none of the indicators were discarded after the first survey. The expert panel suggested an additional 17 indicators for discussion at the face-to-face meeting. The expert panel selected 62 quality indicators. In addition to the indicators for each condition, the panel selected 2 general measures relevant for all high acuity pediatric patients (Table 3). The majority of the indicators reflect ED processes (84%, $n = 52$), with few indicators reflecting structures (11%, $n = 7$) or outcomes (5%, $n = 3$).[8]

Thirty-seven percent ($n = 23$) of the indicators selected are based on moderate or high quality evidence.

Phase 4: Feasibility of Indicator Measurement

Table 4 reveals the age and proportion of patients who met the inclusion criteria for each of the conditions. A total of 1681 unique visits were identified based on age, acuity, and ICD-10 codes. The proportion of patients meeting the inclusion criteria for each condition ranged from 22% for severe head injury to 84% for anaphylaxis. The interrater reliability for determining which patients met the inclusion criteria was acceptable ($\kappa > 0.6$)[39] for all conditions except for severe sepsis ($\kappa = 0.23$).

Results of indicator measurement in a pediatric ED are shown in Table 3. For the indicators reflecting timeliness of care, ED arrival (first time recorded) was used as time zero based on consensus of the expert panel. For diabetic

TABLE 1 Volume of Pediatric Patients (Age 0–19 Years) With Identified Conditions Seen in Alberta and Ontario in 2006–2008 and Results of Advisory Panel Condition Selection

Condition	Number of Visits	Rate[a] (Per Thousand ≤ Age 19)	% Panelists With Mean Score for Condition ≥7
Anaphylaxis	1334	0.34	92
Cardiac arrest	170	0.04	54
DKA	1138	0.29	100
Meningitis	24	0.01	65
Neutropenia	406	0.10	65
Septic shock	240	0.06	85
Severe head injury	941	0.24	92
Sickle cell anemia crisis	770	0.19	50
Status asthmaticus	489	0.12	96
Status epilepticus	439	0.11	88
Supraventricular tachycardia	545	0.14	58
Burn[b]	1458	0.37	62
Tylenol ingestion[b]	1390	0.35	58

[a] Denominator = population of Alberta and Ontario ≤ age 19 = 3 967 400.

[b] Data from Ontario only.

ketoacidosis (DKA), data were accessible from chart review and a physician order entry system for all of the applicable indicators with high interrater reliability. Compliance with the processes of care specified by the indicators was good with minimal practice variation identified. For example, no patients received bicarbonate ($n = 62$), and the expert panel agreed that this number should be low (ie, <1%). The majority of patients received potassium replacement (91%, $n = 65$) and were treated with the appropriate insulin dose and route (88%, $n = 59$). Required data were available for the status asthmaticus indicators, but interrater reliability was more variable. The majority of patients received a systemic corticosteroid during the visit (99%, $n = 180$), and β2-agonists and systemic steroids were provided in a median of 19 and 27 minutes, respectively. Reliability was lower for indicators that relied on information written in the chart by a clinician (compared with data from physician order entry or patient tracking systems), such as the "percentage of admitted patients with objective assessment of severity of their condition" (49%, $\kappa = 0.08$) and "patients referred to an asthma education program" (32%, $\kappa = 0.48$). From the currently available data sources, it is not possible to determine if the performance on these indicators was low due to practice variation and poor provider compliance or lack of documentation. For anaphylaxis, 68% of patients received epinephrine in the ED, and 94% of patients who received epinephrine in the ED were treated by the appropriate route. Interrater reliability was good for the anaphylaxis indicators.

For the status epilepticus indicators, the median time to second line anticonvulsant administration was 31 minutes (ICC = 0.89), with 87% ($\kappa = 1.0$) of patients receiving a benzodiazepine as initial therapy and 86% ($\kappa = 1.0$) of patients with rapid bedside glucose documented. Interrater reliability was not calculated for "attainment of seizure control within 30 minutes" and "receipt of an antiepileptic within 10 minutes" due to the small number of cases available for reliability comparison.[38] Four of the severe head injury indicators were specific to referring (nontrauma) centers (Table 3) and were not applicable to the center where data were collected. Data for 2 of the indicators "head computed tomography (CT) scan performed and analyzed within 1 hour of request" and "neurosurgeon response time >30 minutes" were not available by using existing data sources. Compliance was high with respect to documentation of central nervous system (CNS), blood pressure, and oxygen saturation monitoring. However, 9.5% ($n = 21$) of patients were not intubated before leaving the ED, and only 46% ($n = 13$) of intubated patients had documented end tidal CO_2 monitoring. Interrater reliability was high for the head injury

TABLE 2 Results of the Systematic Review of the Literature and Indicator Selection Process

Condition	Existing Indicators	New Guideline Based Indicators	Additional Indicators Suggested by Expert Panel	Total Indicators Considered by Expert Panel	Final Indicators Selected	Interrater Reliability (κ)[a] for GRADE[b] Rating
DKA	6	4	6	16	10	0.74
Status asthmaticus	16	11	1	28	16	0.73
Anaphylaxis	1	6	2	9	3	1.00
Status epilepticus	2	13	2	17	6	0.15
Severe head injury	14	7	4	25	16	0.66
Sepsis/septic shock	2	9	2	13	9	0.55
General measures[c]	5	1	0	6	2	1.00
Total	46	51	17	114	62	0.68

[a] κ(κ) = Cohen's unweighted κ.

[b] Grade = level of evidence upon which indicator is based (1 = very low quality [expert opinion], 2 = low quality, 3 = moderate quality, 4 = high quality).[34]

[c] Measures identified in literature search that were not condition specific.

TABLE 3 Final List of Indicators Selected by Expert Panel and Results of Indicator Measurement at a Pediatric ED

Quality Indicators (First 5 Listed for Each Category Were Rated Highest Priority by Expert Panel)	Type of Indicator	Grade[a]	Direction[b]	Results of Measurement, % (Denominator) or Median Min (IQR)	Interrater Reliability, κ,[c] ICC, or % Agreement (n)
DKA					
% of EDs with clinical guidelines for the treatment of DKA in children	S	1	↑	NA (single site)	NA
% of patients receiving IV fluids within 60 min of ED arrival	P	1	↑	36% (n = 47)	κ = 1.00 (n = 9)
Time from triage to initiation of IV therapy	P	1	NA	78 min (69)	ICC = 1.00 (n = 11)
% of patients treated with appropriate insulin dose and route	P	4	↑	88% (n = 59)	κ = 1.00 (n = 11)
Time from ED arrival to insulin administration	P	1	NA	115 min (88)	ICC = 1.00 (n = 11)
% of patients treated initially with IV normal saline	P	2	↑	72% (n = 47)	κ = 1.00 (n = 11)
% of patients who receive potassium replacement	P	4	↑	91% (n = 65)	κ = 1.00 (n = 11)
% of patients who receive bicarbonate	P	3	↓	0% (n = 65)	Agree = 1.0 (n = 9)
% of patients administered IV insulin	P	4	↑	88% (n = 65)	κ = 1.00 (n = 11)
Time from arrival to expert consultation	P	1	NA	161 min (71)	ICC = 1.00 (n = 11)
Status asthmaticus/severe asthma					
Time from arrival to systemic steroid administered	P	4	NA	27 min (23)	ICC = 0.88 (n = 68)
% of EDs with clinical guidelines for the treatment of asthma in children	S	2	↑	NA (single site)	NA
Time from arrival to first inhaled β2-agonist treatment	P	4	NA	19 min (16)	ICC = 0.90 (n = 69)
% of admitted patients with objective assessment of severity on initial presentation	P	1	↑	49% (n = 94)	κ = 0.08 (n = 14)
Unplanned return visit within 72 h for same/related asthma exacerbation	0	1	↓	5.4% (n = 95)	Agree = 0.93 (n = 68)
% of patients with asthma who received respiratory assessment score	P	3	↑	98% (n = 190)	κ = 1.0 (n = 69)
% of patients receiving systemic corticosteroid during visit	P	4	↑	99% (n = 180)	κ = 0.87 (n = 68)
% of patients who received ipratropium bromide in ED	P	4	↑	99% (n = 180)	κ = 0.80 (n = 68)
% of patients receiving combination β2-agonist/anticholinergic therapy	P	4	↑	98% (n = 180)	κ = 0.22 (n = 68)
Time from arrival to expert consult for patients not improving after conventional treatment	P	1	NA	253 min (197)	ICC = 0.66 (n = 13)
% of patients not improving who receive at least 1 second-line therapy	P	3	↑	65% (n = 49)	κ = 0.38 (n = 68)
% of patients discharged from the hospital from ED with a prescription/supply of oral steroids	P	4	↑	92% (n = 95)	κ = 0.62 (n = 53)
% of patients discharged from the hospital with a controller medication	P	4	↑	76% (n = 95)	κ = 0.50 (n = 53)
% of discharged patients referred to an asthma education program	P	2	↑	32% (n = 95)	κ = 0.48 (n = 53)
% of patients discharged from the hospital from the ED with a written action plan	P	3	↑	49% (n = 95)	κ = 0.83 (n = 53)
% of patients discharged with follow-up instructions	P	3	↑	87% (n = 95)	κ = 0.51 (n = 53)
Anaphylaxis					
% of EDs with clinical guidelines for the treatment of anaphylaxis in children	S	2	↑	NA (single site)	NA
% of patients with anaphylaxis who received epinephrine in ED	P	3	↑	68% (n = 211)	κ = 0.89 (n = 19)
% of patients treated with epinephrine in ED treated by the appropriate route	P	3	↑	94% (n = 82)	κ = 0.68 (n = 19)
Status epilepticus					
% of EDs with clinical guidelines for the treatment of status epilepticus in children	S	1	↑	NA (single site)	NA
Time from arrival to administration of second-line anticonvulsant	P	2	NA	31 min (44)	ICC = 0.89 (n = 7)
Failure to achieve seizure control within 30 min of ED arrival	0	2	↓	54% (n = 59)	NC[d] (n = 4)
% of patients receiving antiepileptic drug within 10 min of ED arrival	P	1	↑	58% (n = 59)	NC[d] (n = 4)
% of patients in whom initial drug therapy in ED included a benzodiazepine	P	3	↑	87% (n = 76)	κ = 1.00 (n = 12)
% of patients with rapid bedside glucose documented	P	2	↑	86% (n = 85)	κ = 1.00 (n = 12)

TABLE 3 Continued

Quality Indicators (First 5 Listed for Each Category Were Rated Highest Priority by Expert Panel)	Type of Indicator	Grade[a]	Direction[b]	Results of Measurement, % (Denominator) or Median Min (IQR)	Interrater Reliability, κ,[c] ICC, or % Agreement (n)
Severe head injury					
% of EDs with clinical guidelines for the treatment of severe head injury in children	S	2	↑	NA (single site)	NA
Time from arrival to definitive airway management/rapid sequence intubation	P	1	NA	20 min (19)	0.99 (n = 6)
% of patients with documented, untreated hypotension	P	3	↓	5% (n = 21)	κ = 0.77 (n = 9)
Time from ED arrival to definitive care/trauma center[e]	P	1	NA	NA	NA
Local guidelines for referral of patients to definitive care/trauma center[e]	S	1	↑	NA	NA
Time from ED arrival to initiation of transfer to definitive care/trauma center[e]	P	1	↓	NA	NA
Time from ED arrival to arrival of specialized transport[e]	P	1	↓	NA	NA
% of patients with hourly CNS monitoring/charting	P	1	↑	90% (n = 21)	κ = 0.41 (n = 21)
% of patients without sequential blood pressure monitoring	P	3	↓	0% (n = 21)	κ = 1.00 (n = 21)
% of patients without sequential oxygen saturation and respiratory rate monitoring	P	2	↓	0% (n = 21)	κ = 1.00 (n = 21)
% of patients who are not intubated before leaving ED (ie, for admission, imaging)	P	1	↓	9.5% (n = 21)	κ = 0.63 (n = 13)
% of patients with CT scan analyzed within 1 h of request	P	1	↑	NC[f]	NC[f]
% of patients who do not have a CT scan performed within 1 h of ED arrival	P	1	↓	24% (n = 21)	Agree = 0.44 (n = 9)
Neurosurgeon response time >30 min from time of call-in request	P	1	↓	NC[f]	NC[f]
% of patients without spinal cord immobilization	P	2	↓	14% (n = 21)	κ = 1.00 (n = 21)
% of intubated patients with end tidal CO_2 monitoring	P	1	↑	46% (n = 13)	κ = 1.00 (n = 9)
Severe sepsis/septic shock					
Time from ED arrival to isotonic fluid bolus	P	3	NA	68 min (104)	ICC = 0.99 (n = 12)
% of EDs with clinical guidelines for the treatment of severe sepsis/septic shock in children	S	1	↑	NA (single site)	NA
Time from ED arrival to antibiotic administration in suspected severe sepsis/septic shock	P	1	NA	189 min (212)	ICC = 1.0 (n = 12)
% of patients treated with appropriate volume of isotonic fluid in first hour	P	3	↑	57% (n = 28)	κ = 0.85 (n = 16)
% of patients with fluid refractory shock treated with appropriate vasopressor/inotrope	P	2	↑	43% (n = 7)	NC[d] (n = 3)
% of patients with dopamine resistant shock treated with epinephrine/norepinephrine	P	2	↑	0% (n = 3)	NC[d] (n = 3)
% of patients treated with pressors who have not received 60 cc/kg fluid	P	2	↓	25% (n = 8)	NC[d] (n = 3)
Time from ED arrival to IV/IO insertion	P	3	NA	63 min (78)	ICC = 0.98 (n = 9)
Glucose measured in first hour of resuscitation	P	1	↑	47% (n = 30)	Agree = 0.59 (n = 17)
General measures					
Time from ED arrival to first provider for all urgent and emergency patients	P	1	NA	34 min (32)	NC[f]
Return visits within 48 h resulting in admission for all urgent and emergency patients	O	1	↓	1.3% (n = 12 636)	NC[f]

IQR, interquartile range; NA, results not reported as denominator requires multiple sites or indicator refers only to referral centers; O, outcome; P, process; S, structure.

[a] Grade denotes level of evidence upon which indicator is based (1 = very low quality [expert opinion], 2 = low quality, 3 = moderate quality, 4 = high quality).[34]

[b] Direction denotes directionality of proportion-based indicators with a higher (↑) or lower (↓) number indicative of better performance.

[c] κ(κ) = Cohen's unweighted κ.

[d] NC denotes reliability not calculated due to sample size (<5 observations).

[e] This indicator was specific to referring (nontrauma) centers and was not applicable to the center where data were collected.

[f] NC denotes data not available from existing data sources.

indicators with the exception of "hourly CNS monitoring" (κ = 0.41) and "CT within 1 hour of arrival" (Agreement = 0.44). For severe sepsis/septic shock, the median time from ED arrival to isotonic fluid bolus was 68 minutes, 63 minutes to intravenous (IV)/intraosseous (IO) insertion, and 189 minutes to antibiotic administration. Sample size was also small, and interrater reliability not calculated, for the severe sepsis indicators that measured fluid refractory shock (n = 7), dopamine resistant shock (n = 3), and patients treated with pressors who had not received 60 cc/kg

TABLE 4 Demographic Data and Proportion of Patients Who Met Inclusion Criteria for High Acuity Conditions

Condition	Age, Median (Range), y	Charts Reviewed, n	Meet Criteria, % (n)[a]	Reliability, κ (n)[b]
DKA	11.7 (1.3–17.9)	78	83% (n = 65)	1.0 (n = 11)
Severe asthma/status asthmaticus	5.25 (1.75–18)	1058	18% (n = 190)	1.0 (n = 69)
Anaphylaxis	6.9 (0.3–17.9)	250	84% (n = 210)	1.0 (n = 23)
Status epilepticus	3.75 (0.5–18)	127	67% (n = 85)	0.67 (n = 12)
Severe head injury	6.4 (0–16.9)	94	22% (n = 21)	0.83 (n = 13)
Severe sepsis/septic shock	2.2 (0.3–17.2)	74	41% (n = 30)	0.23 (n = 17)

The definition of DKA was based on biochemical criteria[51]; severe asthma/status asthmaticus: score of ≥ 9 on a validated clinical pediatric asthma score[52]; anaphylaxis criteria were based on the National Institute of Allergy and Infectious Disease and Food Allergy and Anaphylaxis Network consensus definition for anaphylaxis[53]; status epilepticus based on ≥ 1 of the following: (1) a 30-min duration of seizures; (2) 2 or more sequential seizures without full recovery of consciousness between seizures; or (3) ongoing seizure for >10 min in the ED[54]; severe head injury: Glasgow Coma Scale ≥ 9[55]; severe sepsis/septic shock criteria were based on the international consensus definitions for sepsis and organ dysfunction in pediatrics.[46]
[a] Inclusion criteria outlined in Supplemental Information 3.
[b] κ = Cohen's unweighted κ for the number of patients who met inclusion criteria for each condition.

of fluid (n = 8). The median time to first provider for all resuscitation and emergent patients was 34 minutes. In addition, 1.3% of 12 636 resuscitation and emergent patients discharged from the hospital returned within 48 hours and were admitted. Data on time to provider and return visits were from an administrative database, and interreliability data could not be assessed.

DISCUSSION

This rigorous process provides 62 evidence and expert consensus based quality indicators for high acuity conditions relevant for any ED setting where children are seen. Previous work on indicators for pediatric ED patients has focused on administrative and clinical measures, such as length of stay in the ED after admission, that are not presentation or condition specific[41,42]; common conditions seen in any ED setting[28]; and creation of a balanced scorecard to reflect all facets of pediatric emergency care.[43] None of the previous work targets high acuity conditions. A recently published analysis of existing pediatric measures relevant to emergency care revealed that most disease specific measures address a few common pediatric conditions and suggested that future measures should consider illness severity.[44]

The 4 phases of this study followed the process for indicator development and assessment as outlined by the AHRQ.[27] These phases included the following: expert engagement of an advisory panel to identify conditions for indicator development, identification of candidate indicators including literature review and summary of evidence (using GRADE), expert panel review and selection of indicators by using a modified Delphi process, and assessment of feasibility of candidate indicators including empirical analyses.

One of the strengths of this project was the comprehensive search for existing indicators and high quality guidelines for the development of new indicators, and the systematic application of GRADE in assessing the level of evidence upon which each indicator is based. Although the GRADE system was a useful means of summarizing the evidence for the expert panel, we identified a number of challenges with its use, including the need for significant time, resources, and research methodology expertise. Even for raters with clinical and research backgrounds (physician with masters level epidemiology training and an experienced research nurse), the interrater reliability for GRADE assignment was variable. Not surprisingly, the κ was lower for conditions with a less developed evidence base, such as status epilepticus, as compared with asthma or DKA.

Given the large number of indicators considered and the variable quality of evidence available, the opportunity for the expert panel to discuss the indicators in person was an integral part of the indicator selection process. Previous work has also emphasized the importance of a face-to-face meeting of the expert panel.[28] Another similarity with previous work on pediatric quality measurement was that the majority of indicators selected by the expert panel reflected ED processes.[28,44] The only structural indicators retained by the expert panel assessed the presence of clinical guidelines for each of the conditions, despite a paucity of evidence linking guidelines to patient outcome. These findings illustrate the need for further work developing outcome indicators and establishing links between structure and process indicators and patient outcome.

A final strength of the project was the inclusion of a data collection phase to assess the feasibility and reliability of indicator measurement. Previous work on quality measures for the pediatric population has emphasized the importance of testing measures in the real world settings where care will be assessed.[45] The measurement stage of this study highlighted a number of issues that are relevant to the interpretation and application of the indicators. For example, a challenge identified in the data collection phase was the difficulty assigning time zero for complex conditions such as severe sepsis, DKA, and status epilepticus. We decided a priori to use ED arrival as time zero but recognize that this may not be accurate as a septic child may decompensate while in the ED and not have met the criteria for severe sepsis at presentation. Similarly, a child

may start seizing while in the ED such that the time from ED arrival to first anticonvulsant treatment may not accurately describe the timeliness of seizure treatment. Difficulty in assigning a time zero may account in part for the relatively long length of time from ED arrival to isotonic fluid bolus; IV/IO insertion; and antibiotic administration (Table 3) for patients with severe sepsis. These results highlight the need for data collection across multiple centers to establish reasonable benchmarks for these indicators, especially for conditions such as severe sepsis where identifying the denominator is a challenge, as illustrated by our poor interrater reliability ($\kappa = 0.23$) in applying an operational definition based on an international consensus definition.[46]

Another challenge we encountered was the small sample size for a number of the indicators. The combination of low event rates and small numbers of eligible patients is a recognized issue in performance measurement,[47] particularly in pediatrics.[48] The conventional minimum sample size is ≥ 30 eligible patients.[48] Our experience collecting 2 years of feasibility data suggests that even tertiary care pediatric centers may not be able to accrue sufficient numbers to adequately measure performance for conditions such as severe head injury or severe sepsis (Table 4).

A number of methods have been suggested to address the issue of small sample size in indicator reporting. One approach is to only report on institutions with adequate numbers of eligible patients (≥ 30).[48] However, applying measures only to institutions with a particular volume of high acuity cases would miss a significant portion of patients who are seen in smaller centers, and it is in these centers where practice variation and the potential for improvement may be greatest.[21,24–26,49] Many of the indicators developed here would be useful even for smaller volume centers for local quality improvement initiatives, such as measuring the impact of a new clinical pathway. Another solution already in use by the US Department of Health and Human Services is to aggregate data over 3 years.[48] In addition to changing the time frame of reporting, other methods that could be used to adapt certain indicators for public reporting and accountability for smaller volume centers include the following: changing reporting conventions to reflect uncertainty when it exists[48]; using a composite measure of multiple outcomes[48]; application of statistical methods such as indirect estimation and hierarchical modeling[47]; and utilizing selected measures on a regional rather than an institutional level.

In addition to the challenges identified above, the results of this study are subject to a few limitations. First, despite our efforts to incorporate the best possible evidence, less than half of the final indicators selected by the expert panel are based on moderate or high quality evidence. Unfortunately, this is likely a reflection of gaps in the overall quantity and quality of child health research[50] and reinforces the need for further high quality research in pediatrics. Second, information on the volume of patients seen and the availability of data may not be generalizable to other pediatric institutions and is certainly not applicable to smaller, nonpediatric hospitals. However, the measure specifications and reliability data should be applicable to most settings. A final limitation is the dependence on written documentation for some of the indicators. In general, interrater reliability and compliance with the process or structure specified by the indicator was lower for indicators that required clinician documentation, as compared with indicators where data were available from a patient tracking or physician order entry system. This limitation needs to be taken into account in future applications of the indicators, either at the data collection phase through the use of alternate data such as physician billing, electronic health records, and pharmacy data or in the benchmarking/reporting phase.

CONCLUSIONS

This evidence and expert consensus based process provides indicators for high acuity pediatric conditions potentially suitable for a range of applications from local quality improvement initiatives to public reporting. The results of this study contribute significantly to the existing body of quality indicators for the emergency care of pediatric patients. Future work will focus on multicenter benchmarking and data collection to test the validity and feasibility of these indicators across the spectrum of ED settings that provide care for children. This research provides clinicians, researchers, and policy makers with tools to improve the quality of pediatric care for severely ill children seen in any ED setting.

REFERENCES

1. Institute of Medicine. *Statement on Quality of Care: National Roundtable on Health Care Quality—The Urgent Need to Improve Health Care Quality.* Washington, DC: National Academies Press; 1998

2. Institute of Medicine. *Crossing the Quality Chasm: A New Health System for the 21st Century.* Washington, DC: National Academies Press; 2001

3. Beal AC, Co JP, Dougherty D, et al. Quality measures for children's health

care. *Pediatrics.* 2004;113(1 pt 2):199–209

4. Campbell SM, Cantrill JA, Roberts D. Prescribing indicators for UK general practice: Delphi consultation study. *BMJ.* 2000;321 (7258):425–428

5. Lee TH. Eulogy for a quality measure. *N Engl J Med.* 2007;357(12):1175–1177

6. Bradley EH, Holmboe ES, Mattera JA, Roumanis SA, Radford MJ, Krumholz HM. A qualitative study of increasing beta-blocker use after myocardial infarction: why do some hospitals succeed? *JAMA.* 2001;285(20): 2604–2611

7. Campbell SM, Braspenning J, Hutchinson A, Marshall MN. Research methods used in developing and applying quality indicators in primary care. *BMJ.* 2003;326(7393):816–819

8. Donabedian A. The quality of care. How can it be assessed? *JAMA.* 1988;260(12):1743–1748

9. Graff L, Stevens C, Spaite D, Foody J. Measuring and improving quality in emergency medicine. *Acad Emerg Med.* 2002;9(11): 1091–1107

10. Lindsay P, Schull M, Bronskill S, Anderson G. The development of indicators to measure the quality of clinical care in emergency departments following a modified-delphi approach. *Acad Emerg Med.* 2002;9(11): 1131–1139

11. Mourad SM, Hermens RP, Nelen WL, Braat DD, Grol RP, Kremer JA. Guideline-based development of quality indicators for subfertility care. *Hum Reprod.* 2007;22(10): 2665–2672

12. McGory ML, Shekelle PG, Ko CY. Development of quality indicators for patients undergoing colorectal cancer surgery. *J Natl Cancer Inst.* 2006;98(22):1623–1633

13. Guru V, Anderson GM, Fremes SE, O'Connor GT, Grover FL, Tu JV; Canadian CABG Surgery Quality Indicator Consensus Panel. The identification and development of Canadian coronary artery bypass graft surgery quality indicators. *J Thorac Cardiovasc Surg.* 2005;130(5):1257

14. National Center for Health Statistics. *Health, United States, 2012: With Special Feature on Emergency Care.* Hyattsville, MD: National Center for Health Statistics; 2013

15. Bordley WC. Outcomes research and emergency medical services for children: domains, challenges, and opportunities. *Ambul Pediatr.* 2002;2(suppl 4):306–310

16. Clancy CM, Dougherty D, Walker E. The importance of outcomes research in pediatric emergency medicine. *Ambul Pediatr.* 2002;2(suppl 4):293–300

17. Moody-Williams JD, Krug S, O'Connor R, Shook JE, Athey JL, Holleran RS. Practice guidelines and performance measures in emergency medical services for children. *Ann Emerg Med.* 2002;39(4):404–412

18. Dougherty D, Simpson LA. Measuring the quality of children's health care: a prerequisite to action. *Pediatrics.* 2004;113(1 pt 2):185–198

19. Institute of Medicine, Committee on the Future of Emergency Care in the United States Health System. *Emergency Care for Children: Growing Pains.* Washington, DC: National Academies Press; 2007

20. Zebrack M, Dandoy C, Hansen K, Scaife E, Mann NC, Bratton SL. Early resuscitation of children with moderate-to-severe traumatic brain injury. *Pediatrics.* 2009;124(1): 56–64

21. Hampers LC, Trainor JL, Listernick R, et al. Setting-based practice variation in the management of simple febrile seizure. *Acad Emerg Med.* 2000;7(1):21–27

22. 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):1165–1170

23. Schweich PJ, Smith KM, Dowd MD, Walkley EI. Pediatric emergency medicine practice patterns: a comparison of pediatric and general emergency physicians. *Pediatr Emerg Care.* 1998;14(2):89–94

24. Petrack EM, Christopher NC, Kriwinsky J. Pain management in the emergency department: patterns of analgesic utilization. *Pediatrics.* 1997;99(5):711–714

25. Plint AC, Johnson DW, Wiebe N, et al. Practice variation among pediatric emergency departments in the treatment of bronchiolitis. *Acad Emerg Med.* 2004;11(4):353–360

26. Isaacman DJ, Kaminer K, Veligeti H, Jones M, Davis P, Mason JD. Comparative practice patterns of emergency medicine physicians and pediatric emergency medicine physicians managing fever in young children. *Pediatrics.* 2001;108(2):354–358

27. Fabian LA, Geppert J. Quality Indicator Measure Development, Implementation, Maintenance, and Retirement Summary (Prepared by Battelle, under Contract No. 290-04-0020). Rockville, MD: Agency for Healthcare Research and Quality; May 2011

28. Guttmann A, Razzaq A, Lindsay P, Zagorski B, Anderson GM. Development of measures of the quality of emergency department care for children using a structured panel process. *Pediatrics.* 2006;118(1):114–123

29. Canadian Paediatric Triage and Acuity Scale: implementation guidelines for emergency

departments. *Canadian Journal of Emergency Medicine.* 2001;3(4)

30. Fitch K, Bernstein SJ, Aguilar MD, et al. *The RAND/UCLA Appropriateness Method User's Manual.* Santa Monica CA; RAND, Santa Monica CA: 2001

31. National Quality Forum. Measure evaluation criteria. December 2009. Available at: www.qualityforum.org/docs/measure_evaluation _criteria.aspx. Accessed August 13, 2013

32. Appraisal of Guidelines for Research and Evaluation (AGREE) Instrument. Available at: www.agreetrust.org/wp-content/uploads/2013/06/AGREE_II_Users_Manual_and_23-item_Instrument_ENGLISH.pdf. Accessed August 22, 2013

33. Guyatt GH, Oxman AD, Kunz R, et al; GRADE Working Group. Going from evidence to recommendations. *BMJ.* 2008;336(7652): 1049–1051

34. Atkins D, Best D, Briss PA, et al; GRADE Working Group. Grading quality of evidence and strength of recommendations. *BMJ.* 2004;328(7454):1490

35. Nkoy FL, Fassl BA, Simon TD, et al. Quality of care for children hospitalized with asthma. *Pediatrics.* 2008;122(5):1055–1063

36. Rubin HR, Pronovost P, Diette GB. From a process of care to a measure: the development and testing of a quality indicator. *Int J Qual Health Care.* 2001;13 (6):489–496

37. Reeves MJ, Mullard AJ, Wehner S. Inter-rater reliability of data elements from a prototype of the Paul Coverdell National Acute Stroke Registry. *BMC Neurol.* 2008;8:19

38. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. *Stat Med.* 1998;17(1):101–110

39. Worster A, Haines T. Advanced statistics: understanding medical record review (MRR) studies. *Acad Emerg Med.* 2004;11 (2):187–192

40. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas.* 1960;20 (1):37–46

41. Hung GR, Chalut D. A consensus-established set of important indicators of pediatric emergency department performance. *Pediatr Emerg Care.* 2008;24(1):9–15

42. Shaw KN, Ruddy RM, Gorelick MH. Pediatric emergency department directors' benchmarking survey: fiscal year 2001. *Pediatr Emerg Care.* 2003;19(3):143–147

43. Alessandrini E, Alpern E, Varadarajan K, et al. Developing quality performance measures for pediatric emergency care (Abstract). *E-PAS.* 2009;2725.1

44. Alessandrini E, Varadarajan K, Alpern ER, et al; Pediatric Emergency Care Applied

Research Network. Emergency department quality: an analysis of existing pediatric measures. *Acad Emerg Med*. 2011;18(5):519–526

45. Lannon C, Peterson LE, Goudie A. Quality measures for the care of children with otitis media with effusion. *Pediatrics*. 2011; 127(6). Available at: www.pediatrics.org/cgi/content/full/127/6/e1490

46. Goldstein B, Giroir B, Randolph A; International Consensus Conference on Pediatric Sepsis. International pediatric sepsis consensus conference: definitions for sepsis and organ dysfunction in pediatrics. *Pediatr Crit Care Med*. 2005;6(1):2–8

47. Zaslavsky AM. Statistical issues in reporting quality data: small samples and case-mix variation. *Int J Qual Health Care*. 2001; 13(6):481–488

48. Bardach NS, Chien AT, Dudley RA. Small numbers limit the use of the inpatient pediatric quality indicators for hospital comparison. *Acad Pediatr*. 2010;10(4):266–273

49. Grol R. Successes and failures in the implementation of evidence-based guidelines for clinical practice. *Med Care*. 2001; 39(8 suppl 2):II46–II54

50. Zylke JW, Rivara FP, Bauchner H. Challenges to excellence in child health research: call for papers. *JAMA*. 2012;308(10):1040–1041

51. Dunger DB, Sperling MA, Acerini CL, et al; ESPE; LWPES. ESPE/LWPES consensus statement on diabetic ketoacidosis in children and adolescents. *Arch Dis Child*. 2004;89(2): 188–194

52. Chalut DS, Ducharme FM, Davis GM. The Preschool Respiratory Assessment Measure (PRAM): a responsive index of acute asthma severity. *J Pediatr*. 2000;137 (6):762–768

53. Sampson HA, Muñoz-Furlong A, Campbell RL, et al. Second symposium on the definition and management of anaphylaxis: summary report—Second National Institute of Allergy and Infectious Disease/Food Allergy and Anaphylaxis Network symposium. *J Allergy Clin Immunol*. 2006; 117(2):391–397

54. Epilepsy Foundation. Prolonged or serial seizures (status epilepticus). Available at: www.epilepsyfoundation.org/about/types/types/statusepilepticus.cfm. Accessed August 13, 2013

55. Ghajar J. Traumatic brain injury. *Lancet*. 2000;356(9233):923–929

Updated Information & Services	including high resolution figures, can be found at: http://pediatrics.aappublications.org/content/132/4/752
References	This article cites 44 articles, 13 of which you can access for free at: http://pediatrics.aappublications.org/content/132/4/752#BIBL
Subspecialty Collections	This article, along with others on similar topics, appears in the following collection(s): **Emergency Medicine** http://www.aappublications.org/cgi/collection/emergency_medicine_sub **Administration/Practice Management** http://www.aappublications.org/cgi/collection/administration:practice_management_sub **Quality Improvement** http://www.aappublications.org/cgi/collection/quality_improvement_sub
Permissions & Licensing	Information about reproducing this article in parts (figures, tables) or in its entirety can be found online at: http://www.aappublications.org/site/misc/Permissions.xhtml
Reprints	Information about ordering reprints can be found online: http://www.aappublications.org/site/misc/reprints.xhtml

PEDIATRICS®

OFFICIAL JOURNAL OF THE AMERICAN ACADEMY OF PEDIATRICS

Quality Indicators for High Acuity Pediatric Conditions

Antonia S. Stang, Sharon E. Straus, Jennifer Crotts, David W. Johnson and Astrid Guttmann

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://pediatrics.aappublications.org/content/132/4/752

Data Supplement at:
http://pediatrics.aappublications.org/content/suppl/2013/09/18/peds.2013-0854.DCSupplemental

American Academy of Pediatrics

DEDICATED TO THE HEALTH OF ALL CHILDREN®