Beyond Statistical Prediction: Qualitative Evaluation of the Mechanisms by Which Pediatric Early Warning Scores Impact Patient Safety

Christopher P. Bonafide, MD, MSCE, Kathryn E. Roberts, RN, MSN, CNS, CCRN, CCNS, Christine M. Weirich, MPH, Breah Paciotti, MPH, Kathleen M. Tibbetts, MS, Ron Keren, MD, MPH, Frances K. Barg, PhD, John H. Holmes, PhD


BACKGROUND: Early warning scores (EWSs) assign points to clinical observations and generate scores to help clinicians identify deteriorating patients. Despite marginal predictive accuracy in retrospective datasets and a paucity of studies prospectively evaluating their clinical effectiveness, pediatric EWSs are commonly used.

OBJECTIVE: To identify mechanisms beyond their statistical ability to predict deterioration by which physicians and nurses use EWSs to support their decision making.

DESIGN: Qualitative study.

SETTING: A children’s hospital with a rapid response system.

PARTICIPANTS: Physicians and nurses who recently cared for patients with false-positive and false-negative EWSs (score failures).

INTERVENTION: Semistructured interviews.

MEASUREMENTS: Themes identified through grounded theory analysis.

RESULTS: Four themes emerged among the 57 subjects interviewed: (1) The EWS facilitates safety by alerting physicians and nurses to concerning changes and prompting them to think critically about deterioration. (2) The EWS provides less-experienced nurses with vital sign reference ranges. (3) The EWS serves as evidence that empowers nurses to overcome barriers to escalating care. (4) In stable patients, those with baseline abnormal physiology, and those experiencing neurologic deterioration, the EWS may not be helpful.

CONCLUSIONS: Although pediatric EWSs have marginal performance when applied to datasets, clinicians who recently experienced score failures still considered them valuable to identify deterioration and transcend hierarchical barriers. Combining an EWS with a clinician’s judgment may result in a system better equipped to respond to deterioration than retrospective data analyses alone would suggest. Future research should seek to evaluate the clinical effectiveness of EWSs in real-world settings.

ORIGINAL RESEARCH

Thousands of hospitals have recently implemented rapid response systems (RRSs), attempting to reduce mortality outside of intensive care units (ICUs). These systems have 2 clinical components, a response (efferent) arm and an identification (afferent) arm. The response arm is usually composed of a medical emergency team (MET) that responds to calls for urgent assistance. The identification arm includes tools to help clinicians recognize patients who require assistance from the MET. In many hospitals, the identification arm includes an early warning score (EWS). In pediatric patients, EWSs assign point values to vital signs that fall outside of age-based ranges, among other clinical observations. They then generate a total score intended to help clinicians identify patients exhibiting early signs of deterioration. When experimentally applied to vital sign datasets, the test characteristics of pediatric EWSs in detecting clinical deterioration are highly variable across studies, with major tradeoffs between sensitivity, specificity, and predictive values that differ by outcome, score, and cut-point (Table 1). This reflects the difficulty of identifying deteriorating patients using only objective measures. However, in real-world settings, EWSs are used by clinicians in conjunction with their clinical judgment. We hypothesized that EWSs have benefits that extend beyond their ability to predict deterioration, and thus have value not demonstrated by test characteristics alone. In order to further explore this issue, we aimed to qualitatively evaluate mechanisms beyond their statistical ability to predict deterioration by which physicians and nurses use EWSs to support their decision making.
METHODS
Overview
As 1 component of a larger study, we conducted semi-structured interviews with nurses and physicians at The Children’s Hospital of Philadelphia (CHOP) between May and October 2011. In separate subprojects using the same participants, the larger study also aimed to identify residual barriers to calling for urgent assistance and assess the role of families in the recognition of deterioration and MET activation.

Setting
The Children’s Hospital of Philadelphia is an urban, tertiary-care pediatric hospital with 504 beds. Surgical patients hospitalized outside of ICUs are cared for by surgeons and surgical nurses without pediatrician co-management. Implementation of a RRS was prompted by serious safety events in which clinical deterioration was either not recognized or was recognized and not escalated. Prior to RRS implementation, a code blue team could be activated for patients in immediate need of resuscitation, or, for less-urgent needs, a pediatric ICU fellow could be paged by physicians for informal consults.

A multidisciplinary team developed and pilot-tested the RRS, then implemented it hospital-wide in February 2010. Representing an aspect of a multipronged approach to improve safety culture, the RRS consisted of (1) an EWS based upon Parshuram’s Bedside Paediatric Early Warning System,8–10 calculated by hand on a paper form (see online supplementary content) at the same frequency as vital signs (usually every 4 hours), and (2) a 30-minute response MET available for activation by any clinician for any concern, 24 hours per day, 7 days per week. Escalation guidelines included a prompt to activate the MET for a score that increased to the red zone (≥9). For concerns that could not wait 30 minutes, any hospital employee could activate the immediate-response code blue team.

Utilization of the RRS at CHOP is high, with 2–3 calls to the MET per day and a combined MET/code-blue team call rate of 27.8 per 1000 admissions.12 Previously reported pediatric call rates range from 2.8 to 44.0, with a median of 9.6 per 1000 admissions across 6 studies.13–19 Since implementation, there has been a statistically significant net reduction in critical deterioration events (unpublished data).

Participants
We recruited nurses and physicians who had recently cared for children age ≤18 years on general medical or surgical wards with false-negative or false-positive EWSs (instances when the score failed to predict deterioration). Recruitment ceased when we reached thematic data saturation (a qualitative research term for the point at which no new themes emerge with additional interviews).20

Data Collection
Through a detailed review of the relevant literature and consultation with experts, we developed a semi-structured interview guide (see online supplementary content) to elicit nurses’ and physicians’ viewpoints regarding the mechanisms by which they use EWSs to support their decision making.

Experienced qualitative research scientists (F.K.B. and J.H.H.) trained 2 study interviewers (B.P. and K.M.T.). In order to minimize social-desirability bias, the interviewers were not clinicians and were not involved in RRS operations. Each interview was recorded, professionally transcribed, and imported into NVivo 8.0 software for analysis (QSR International, Melbourne, Australia).

Data Analysis
We coded the interviews inductively, without using a predetermined set of themes. This approach is known as grounded theory methodology.21 Two team members coded each interview independently. They then reviewed their coding together and discussed discrepancies until reaching consensus. In weekly meetings while the interviews were ongoing, we compared newly collected data with themes that had previously emerged in order to guide further thematic development and refinement (the constant comparative method).22 After all of the interviews were completed and consensus had been reached for each individual interview, the study team convened a series of additional meetings to further refine and finalize the themes.

### TABLE 1. Test Characteristics of Early Warning Scores

<table>
<thead>
<tr>
<th>Score and Citation</th>
<th>Outcome Measure</th>
<th>Score</th>
<th>Sens</th>
<th>Spec</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brighton Paediatric Early Warning Score&lt;sup&gt;5&lt;/sup&gt;</td>
<td>RRT or code blue call</td>
<td>≥4</td>
<td>86%</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Bristol Paediatric Early Warning Tool&lt;sup&gt;6,11&lt;/sup&gt;</td>
<td>Escalation to higher level of care</td>
<td>≥1</td>
<td>89%</td>
<td>64%</td>
<td>2%</td>
<td>&gt;99%</td>
</tr>
<tr>
<td>Cardiff and Vale Paediatric Early Warning System&lt;sup&gt;7&lt;/sup&gt;</td>
<td>Respiratory or cardiac arrest, HDU/ICU admission, or death</td>
<td>≥1</td>
<td>78%</td>
<td>95%</td>
<td>4%</td>
<td>NR</td>
</tr>
<tr>
<td>Bedside Paediatric Early Warning System score, original version&lt;sup&gt;8&lt;/sup&gt;</td>
<td>Code blue call</td>
<td>≥5</td>
<td>64%</td>
<td>91%</td>
<td>9%</td>
<td>NR</td>
</tr>
<tr>
<td>Bedside Paediatric Early Warning System score, simplified version&lt;sup&gt;9&lt;/sup&gt;</td>
<td>Urgent ICU admission without a code blue call</td>
<td>≥7</td>
<td>64%</td>
<td>91%</td>
<td>9%</td>
<td>NR</td>
</tr>
</tbody>
</table>

Note: Abbreviations: ER, erroneously reported; HDU, high dependency unit; ICU, intensive care unit; NPV, negative predictive value; NR, not reported; PPV, positive predictive value; RRT, rapid response team; Sens, sensitivity; Spec, specificity.
TABLE 2. Characteristics of Physician and Nurse Participants

<table>
<thead>
<tr>
<th></th>
<th>Physicians (n = 30)</th>
<th>Nurses (n = 27)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Race</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Black</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>White</td>
<td>26</td>
<td>22</td>
</tr>
<tr>
<td>Prefer not to say</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>&gt;1 race</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic/Latino</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Not Hispanic/Latino</td>
<td>23</td>
<td>25</td>
</tr>
<tr>
<td>Prefer not to say</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>16</td>
<td>25</td>
</tr>
<tr>
<td>M</td>
<td>14</td>
<td>2</td>
</tr>
<tr>
<td>Practice setting</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medical</td>
<td>21</td>
<td>22</td>
</tr>
<tr>
<td>Surgical</td>
<td>9</td>
<td>18.5</td>
</tr>
<tr>
<td>Among physicians only, experience level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intern</td>
<td>7</td>
<td>23.3</td>
</tr>
<tr>
<td>Senior resident</td>
<td>7</td>
<td>23.3</td>
</tr>
<tr>
<td>Attending physician</td>
<td>16</td>
<td>53.3</td>
</tr>
<tr>
<td>Among attending physicians only, no. of years practicing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;5</td>
<td>8</td>
<td>50.0</td>
</tr>
<tr>
<td>5–&lt;10</td>
<td>3</td>
<td>18.8</td>
</tr>
<tr>
<td>≥10</td>
<td>5</td>
<td>31.3</td>
</tr>
<tr>
<td>Among nurses only, no. of years practicing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;1</td>
<td>5</td>
<td>18.5</td>
</tr>
<tr>
<td>1–&lt;2</td>
<td>5</td>
<td>18.5</td>
</tr>
<tr>
<td>2–&lt;5</td>
<td>8</td>
<td>33.3</td>
</tr>
<tr>
<td>5–&lt;10</td>
<td>4</td>
<td>14.6</td>
</tr>
<tr>
<td>10–&lt;20</td>
<td>1</td>
<td>3.7</td>
</tr>
<tr>
<td>≥20</td>
<td>3</td>
<td>11.1</td>
</tr>
<tr>
<td>Recruitment method</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cared for patient with false-positive score</td>
<td>10</td>
<td>33.3</td>
</tr>
<tr>
<td>Cared for patient with false-negative score</td>
<td>13</td>
<td>43.3</td>
</tr>
<tr>
<td>Randomly selected to ensure data saturation for surgical settings</td>
<td>7</td>
<td>23.3</td>
</tr>
</tbody>
</table>

NOTE: Abbreviations: F, female; M, male. Due to rounding of percentages, some totals do not equal 100.0%.

Human Subjects
The CHOP Institutional Review Board approved this study. All participants provided written informed consent.

RESULTS
Participants
We recruited 27 nurses and 30 physicians before reaching thematic data saturation. Because surgical patients are underrepresented relative to medical patients among the population with false-positive and false-negative scores in our hospital, this included 3 randomly selected surgical nurses and 7 randomly selected surgical physicians recruited to ensure thematic data saturation for surgical settings. Characteristics of the participants are displayed in Table 2.

Thematic Analysis
We provide the final themes, associated subthemes, and representative quotations below, with additional supporting quotations in Table 3. Because CHOP’s MET is named the Critical Assessment Team, the term CAT appears in some quotations.

Of note, after interviewing 9 surgeons, we found that they were not very familiar with the EWS and had little to say either positively or negatively about the system. For example, when asked what they thought about the EWS, a surgical intern said, “I have no idea. I don’t have enough experience with it. This is probably the first time that I ever had anybody telling me that the system is in place.” Therefore, surgeons did not contribute meaningfully to the themes below.

Theme 1: The EWS facilitates patient safety by alerting nurses and physicians to concerning vital sign changes and prompting them to think critically about the possibility of deterioration.
Nurses and physicians frequently discussed the direct role of the EWS in revealing changes consistent with early signs of deterioration. A medical nurse with <1 year of experience said, “The higher the number gets, the more it sets off a red flag to you to kind of keep an eye on certain things. They are just as important as taking a set of vitals.” When asked if the EWS had ever helped to identify deterioration, a medical attending physician in practice for 5 years said, “I think sometimes we will blow off, so to speak, certain things, but when you look at the numbers and you see a big [EWS] change versus if you were [just] looking at individual vital signs, then yeah, I think it has made a difference.”

Nurses and physicians also discussed the role of the EWS in prompting them to closely examine individual vital signs and think critically about whether or not a patient is exhibiting early signs of deterioration. A surgical nurse with <1 year of experience said, “Sometimes I feel like if you want things to be okay you can kind of write them off, but when you have to write [the EWS] down . . . it kind of jogs you to think, maybe something is going on or maybe someone else needs to know about this.” A medical senior resident commented, “I think it has alerted me earlier to changes in vital signs that I might not necessarily have known. I think there are nurses that use it and they see that there is an elevation and they call you about it. Then it makes me go back and look through and see what their vital signs are and if it happens in time—we only go through and look at everyone’s vital signs about twice a day—it can be very helpful.”

Theme 2: The EWS provides less-experienced nurses with helpful age-based reference ranges for vital signs that they use when caring for hospitalized children.
Although this theme did not appear among physicians, nurses frequently noted that they referred to the
TABLE 3. Additional Representative Quotations Identified in Semistructured Interviews

Theme 1: The EWS facilitates patient safety by alerting nurses and physicians to concerning vital sign changes and prompting them to think critically about the possibility of deterioration.
- “I think [the EWS] helps us to be focused and gives us definite criteria to look for if there is an issue or change. It hopefully gives us a head start if there is going to be a change. They have a way of tracking it with the different color-coding system they use. … Like, ‘Oh geez, the heart rate is a little bit higher,’ that changes the color from yellow to orange, then I have to let the charge nurse know because that is a change from where they were earlier … it kind of organizes it, I feel like, from where it was before.” (medical nurse with 23 years of experience)
- “I think for myself, as a new clinician, one of our main goals is to help judge sick versus not sick. So to have a concrete system for thinking about that is helpful.” (medical intern)
- “I think [the EWS] can help put things together for us. When you are really busy, you don’t always get to focus on a lot of details. It is like another red flag to say you might have not realized that the child’s heart rates went up further, but now here’s some evidence that they did.” (medical resident)
- “I think that the ability to use the EWS to watch the progression of a patient over time is really helpful. I’ve had a few patients that have gotten sicker from a respiratory standpoint. We can have multiple on the floor at the same time, and what’s nice is that sometimes nurses have been able to come to me and we can really see through the score that we are at the point where a higher level of care is needed, whereas, in the old system, without that, we would have had to essentially wait for true clinical deceleration before the ICU would have been involved. I think that does help to deliver better care.” (medical resident)

Theme 2: The EWS provides less-experienced nurses with helpful age-based reference ranges for vital signs that they use when caring for hospitalized children.
- “Sometimes you just write down the vitals and maybe you are not really thinking, and then when you go to do the EWS you looked at the score and it’s really off in their age range. It kind of gives you 1 more step to recognize that there’s a problem.” (medical nurse with <1 year of experience)
- “I see the role [of the EWS] more broadly as a guide of where your patient should fall with their vital signs according to their age. I think that has been the biggest help for me, to be able to visualize, ‘I have a 3-year-old; this is where they should be for their respiratory rate or heart rate.’ I think it has been good to be able to see that they are falling within the range appropriate for their age.” (surgical nurse with 9 years of experience)

Theme 3: The EWS provides concrete evidence of clinical changes in the form of a score. This empowers nurses to overcome escalation barriers and communicate their concerns, helping them take action to rescue their deteriorating patients.
- “Sometimes you just write down the vitals and maybe you are not really thinking, and then when you go to do the EWS you looked at the score and it’s really off in their age range. It kind of gives you 1 more step to recognize that there’s a problem.” (medical nurse with <1 year of experience)
- “I think that since it has been instituted nursing is coming to residents more than they did beforehand … Can you reassees this patient? Do you think that we should call CAT?” (medical nurse with 1 year of experience)
- “I think [the EWS] helps us to be focused and gives us definite criteria to look for if there is an issue or change. It hopefully gives us a head start if there is going to be a change.” (medical attending physician in practice for 2 years)

Theme 4: In some patients, the EWS may not help with decision-making. These include patients who are very stable and have a low likelihood of deterioration, and patients with abnormal physiology at baseline who consistently have very high EWSs.
- “The patient I took care of in this situation was a really sick kid to begin with, and it wasn’t so much they were concerned about his EWS because, unless there was a really serious event, he would probably be staying on our floor anyway … in some cases we just have some really sick kids whose scores may constantly be high all the time, so it wouldn’t be helpful for the doctors or us to really bring it up.” (medical nurse with 1 year of experience)

NOTE: Abbreviations: CAT, critical assessment team; EWS, early warning score; ICU, intensive care unit.

scoring sheet as a reference for vital signs appropriate for hospitalized children. A surgical nurse with <1 year of experience said, “In nursing school, I mostly dealt with adults. So, to figure out the different ranges for normal vital signs, it helps to have it listed on paper so I can see, ‘Oh, I didn’t realize that this 10-year-old’s heart rate is higher than it should be.’”

A medical nurse with 14 years of experience cited the benefits for less-experienced nurses, noting, “[The EWS helps] newer nurses who don’t know the ranges. Where it’s ‘Oh, my kid’s blood pressure is 81 [mm Hg] over something,’ then they can look at their age and say, ‘Oh, that is completely normal for a 2-month-old.’ But [before the EWS] there was nowhere to look to see the ranges. Unless you were [Pediatric Advanced Life Support] certified where you would know that stuff, there was a lot of anxiety related to vital signs.”

Theme 3: The EWS provides concrete evidence of clinical changes in the form of a score. This empowers nurses to overcome escalation barriers and communicate their concerns, helping them take action to rescue their deteriorating patients.

Nurses and physicians often described the role of the EWS as a source of objective evidence that a patient was exhibiting a concerning change. They shared the ways in which the EWS was used to convey concerns, noting most commonly that this was used as a communication tool by nurses to raise their concerns with physicians. A medical nurse with 23 years of experience said, “[With the EWS] you feel like you have concrete evidence. It’s not just a feeling [that] they are not looking as well as they were … it feels scientific.” Building upon this concept, a medical attending physician in practice for 2 years said, “The EWS is a number that certainly gives people a sense of ‘Here’s the data behind why I am really coming to you and insisting on this.’ It is not calling and saying, ‘I just have a bad feeling,’ it is, ‘I have a bad feeling and his EWS has gone to a 9.’”

Theme 4: In some patients, the EWS may not help with decision-making. These include patients who are very stable and have a low likelihood of deterioration, patients with abnormal physiology at baseline who consistently have very high EWSs, and patients experiencing neurologic deterioration.

Nurses and physicians described some patient scenarios in which the EWS may not help with decision making. Discussing postoperative patients, a surgical nurse with 1 year of experience said, “I love doing [the EWS] for some patients. I think it makes perfect sense. Then there are some patients [for whom] I am doing it just to do it because they are only here for 24 hours. They are completely stable. They never had
I vital sign that was even a little bit off. It’s kind of like we are just filling it out to fill it out.” Commenting on patients at the other end of the spectrum, a medical attending physician in practice for 2 years said, “[The EWS] can be a useful composite tool, but for specialty patients with abnormal baselines, I think it is much more a question of making sure you pay attention to the specific changes, whether it is the EWS or heart rate or vital signs or pain score or any of those things.” A final area in which nurses and physicians identified weaknesses in the EWS surrounded neurologic deterioration. Specifically, nurses and physicians described experiences when the EWS increased minimally or not at all in patients with sudden seizures or concerning mental status changes that warranted escalated care.

**DISCUSSION**

This study is the first to analyze viewpoints on the mechanisms by which EWSs impact decision making among physicians and nurses who had recently experienced score failures. Our study, performed in a children’s hospital, builds upon the findings of related studies performed in hospitals that care primarily for adults.23-28 Andrews and Waterman found that nurses consider the utility of EWSs to extend beyond detecting deterioration by providing quantifiable evidence, “packaged” in the form of a score that improves communication between nurses and physicians.23 Mackintosh and colleagues found that a RRS that included an EWS helped to formalize the way nurses and physicians understand deterioration, enable them to overcome hierarchical boundaries through structured discussions, and empower them to call for help.24 In a quasi-experimental study, McDonnell and colleagues found that an EWS improved self-assessed knowledge, skills, and confidence of nursing staff to detect and manage deteriorating patients.25 In addition, we describe novel findings, including the use of EWS parameters as reference ranges independent of the score, and specific situations when the EWS fails to support decision making. The weaknesses we identified could be used to drive EWS optimization for low-risk patients who are stable as well as higher-risk patients with abnormal baseline physiology and those at risk of neurologic deterioration.

This study has several limitations. Although the interviewees were not involved in RRS operations, it is possible that social desirability bias influenced responses. Next, we identified a knowledge gap among surgeons, and they contributed minimally to our findings. This is most likely because (1) surgical patients deteriorate on the wards less often than medical patients in our hospital, so surgeons are rarely presented with EWSs; (2) surgeons spend less time on the wards compared with medical physicians; and (3) surgical residents rotate in short blocks interspersed with rotations at other hospitals and may be less engaged in hospital safety initiatives.

**CONCLUSIONS**

Although EWSs perform only marginally well as statistical tools to predict clinical deterioration, nurses and physicians who recently experienced score failures described substantial benefits in using them to help identify deteriorating patients and transcend barriers to escalation of care by serving as objective communication tools. Combining an EWS with a clinician’s judgment may result in a system better equipped to respond to deterioration than previous EWS studies focused on their test characteristics alone suggest. Future research should seek to compare and prospectively evaluate the clinical effectiveness of EWSs in real-world settings.

Disclosures: This project was funded by the Pennsylvania Health Research Formula Fund Award (awarded to Keren and Bonafide) and the CHOP Nursing Research and Evidence-Based Practice Award (awarded to Roberts). The funders did not influence the study design; the collection, analysis, or interpretation of data; the writing of the report; or the decision to submit the article for publication. The authors have no other conflicts to report.

**References**


