Inpatient Nursing Care and Early Warning Scores
A Workflow Mismatch

Anne Watson, PhD, RN; Chantel Skipper, MSN, RN, CPNP-AC; Rachel Steury, MSN, RN, C-NP; Heather Walsh, MSN, RN, PCNS-BC; Amanda Levin, MD

Early warning scores calculated by registered nurses (RNs) are used in hospitals to enhance the recognition of and communication about patient deterioration. This study evaluated workflow variables surrounding calculation and documentation of one pediatric hospital’s use of an early warning score. Results indicated that there were significant delays in documentation of early warning scores by RNs and inconsistencies between the early warning scores and vital signs collected and documented by non-RN personnel. These findings reflected information obtained from the RNs about how they prioritize tasks and use workarounds to specific systems issues regarding assessment and documentation in the electronic medical record. Key words: child, documentation, pediatric clinical deterioration, pediatric warning score, triage (early warning score), vital signs, workflow

A CUTE CARE REGISTERED NURSES (RNs) are often the first line of defense in the recognition of a deteriorating patient. Many hospitals use early warning scores calculated by RNs to enhance the recognition of and communication about patient deterioration.1–3 These scores, such as the Pediatric Early Warning Score (PEWS),4 incorporate patient vital signs (VS) and an RN physical assessment into a cumulative objective score. Points are given for patient behavioral state, cardiovascular status, and respiratory status and can also be given for other patient factors such as continuous nebulizer treatments, the presence of vomiting following surgery, or a tracheostomy. Scores are meant to be determined concurrently with routine VS and physical condition assessments for a stable patient. Escalation algorithms direct that bedside assessments should be increased in frequency as the score increases, and many institutions, including our own, mandate the activation of a rapid response team when the score reaches a certain level.

Early warning score tools have been validated retrospectively through medical records audits performed by individuals specifically trained for the task,5–7 although recently published systematic reviews of
normal parameters for VS and VS change parameters used in early warning score tools have found that these parameters are ill-defined and inconsistent. While one study reported that the PEWS tool was easy to use, "with minimal to no time delay in calculating the score during patient assessment," the feasibility and timeliness of score assessment and documentation under in situ hospital conditions have not been robustly examined.

A substantial body of evidence confirms that, in general, RNs work under time pressure and in environments with many interruptions and the need for multitasking. This increases the importance of understanding systems issues and the workflow around many activities such as the detection of patient deterioration. Nursing experience may not overcome the impediments of time pressure—for example, when presented with clinical vignettes using the Medical Early Warning Score (an adult early warning score), more experienced RNs were better than inexperienced RNs at detecting the signs and symptoms of a deteriorating patient. However, under conditions of time pressure, the more experienced RNs were no better than the less experienced ones.

Other barriers specific to early warning score assessment and documentation have been proposed. The increased use of VS monitors on acute care floors and use of patient care technicians (PCTs) to obtain VS may distance the "eyes" of nurses from patients and their current condition. For example, the use of continuous pulse oximetry monitoring may lead to fewer bedside assessments of other aspects of respiratory status. Second, the translation of early warning scores from paper medical records to an electronic medical record (EMR) may not be seamless. Electronic documentation takes longer than paper charting due to log-in times and delays accessing individual medical records. This has increased the use of the "batching" of data entry and charting away from patient rooms and may have an influence on timeliness and accuracy. Concurrent collection and documentation of both VS and early warning scores is important, as documentation is not simply a record of patient status but also a directive for future patient care actions.

Finally, Ludikhuize et al. found that caregivers involved in cases of patient deterioration believed that these events were not avoidable, there was no delay in recognition of deterioration, and the team taking care of the patient worked effectively and efficiently. Experts evaluating these events in retrospect disagreed. Thus, cognitive biases may lead to misinterpretation of deterioration events and inaccurate reflection on responses to deterioration. This may negatively influence how caregivers prioritize their activities with respect to the collection, documentation, and communication about patient stability.

**PURPOSE**

At our institution, preliminary auditing of our early warning score tool and feedback by RNs suggested that multiple workflow constraints existed for routine nursing activities and raised concern that scoring was more difficult and variable than previously thought. We needed to assess the degree to which RNs were able to document the early warning score and patient assessments in real time. In addition, many institutions, including our own, use non-RN personnel for the collection of VS, and we wondered about the degree to which the early warning scores that RNs entered were consistent with the VS collected and documented by other staff members.

**METHODS**

**Study design**

A mixed-methods study was conducted using (1) retrospective medical record "observations" of documentation practices and consistency related to VS and early warning scores, (2) behavioral observations of RN activities related to patient assessment, (3) and collection of RN feedback to assess perceptions about our hospital’s use of the tool. Our hospital institutional review board approved this study.
Setting

The study was conducted in a 303-bed, freestanding academic pediatric medical center in a large metropolitan area. All 7 non-intensive care unit inpatient units (196 beds) were involved in the study. Descriptive features of the 7 units are presented in the Supplemental Digital Content, Table (available at http://links.lww.com/JNCQ/A74).

Data collection

Medical record observations

Medical records were examined by 6 RNs over 2 time periods approximately 6 months apart. Only one time period was initially planned, but subsequently, our hospital embarked on an educational campaign to improve detection of patient deterioration and our escalation culture. Targeted clinician groups received supplemental training on our hospital’s early warning score and its associated escalation algorithm and received a badge card with scoring and escalation algorithm information. Enhancements also were made to the EMR to increase the ease of score entry for RNs and the visibility of the score for physicians. The campaign did not target the timeliness of score documentation, nor did it address any aspects of the way in which the PCT-collected VS were used in score calculation. However, we thought it important to add another collection period after the campaign to assess any effects on our data.

At each time period, four 12-hour day and four 12-hour night shifts were audited for all inpatients in acute care beds at the end of each shift that had been on the unit for the entire previous 12 hours. Heart rate (HR), respiratory rate (RR), early warning cardiovascular subscore, and respiratory subscore within 1 hour before or 1 hour after hospital-standardized VS documentation times (12 AM/12 PM, 4 AM/4 PM, and 8 AM/8 PM) were extracted. Any of these time blocks in the EMR without documented HR, RR, or early warning cardiovascular or respiratory subscores were marked as missing. Only early warning score data were collected during the second data collection period. We did not suspect that VS collection and documentation practices had changed as a result of the early warning score educational campaign.

Workflow documentation practices were examined in the EMR by abstracting the “posting” time for each piece of information entered, which is chosen by the provider, and the actual time of data entry, which is recorded automatically. That is, an 8 AM VS or early warning score may be posted for 8 AM, but the actual entry time might be 10:30 AM, reflecting either delayed documentation or delayed acquisition of the VS or early warning score. In addition, because our EMR has separate screens for our early warning score and VS documentation, it is difficult to visualize the most recent PCT-collected VS when an RN is entering the early warning score. An RN has the option of documenting new vitals should the early warning score not match the previously documented PCT vitals, but RNs rarely enter a second set of VS since this is a task assigned to the PCTs.

We decided that no cardiovascular or respiratory subscore should be lower than that required of the VS that most closely corresponded in time. It was possible that these subscores could be higher and still be consistent with the recorded VS depending on other aspects of the RN’s physical assessment. For example, the respiratory subscore can be elevated if the patient is breathing with retractions, even if the RR is within normal parameters. For our retrospective review, for all patients with an elevated RR or HR, we defined “inconsistency” as any documented early warning respiratory or cardiovascular subscore that was too low compared with the most closely previously documented VS.

The Wilcoxon test was used to analyze delay-in-posting time differences between the PCTs and RNs for VS and pre- and postedocumentation for early warning scores. The Fisher exact test was used to evaluate the early warning score and VS consistency.

Behavioral observations

Nurse managers of 5 of the acute care units agreed to allow behavioral observations of bedside RNs. A graduate student in a Human
Factors and Applied Cognition program at a local university was the observer. Every minute, for 90 minutes, the observer noted the occurrence of bedside physical assessment, documentation (either engagement with the EMR or use of a personal paper note system), and the location of this engagement (bedside or other—computer on wheels, nurse station in hallway, medication room, or team room). In addition, communication with other providers was noted and all other behaviors were noted as “other.” Pilot work was conducted to create the data collection instrument. Because the tasks of interest were relatively few in number, a paper tool was decided as the best way to document behaviors and a stopwatch was used for timing.

Prior to the observations, a nurse was randomly selected from the daily assignment sheet located at the desk of the unit clerk. The observer provided an information sheet about the observations, and the RN was asked for permission to observe, with a waiver of written informed consent approved by the institutional review board. Participants were told that they could suspend or stop the observation at any time.

**Nurse feedback**

Feedback sessions conducted by quality improvement staff members had indicated significant dissatisfaction with our current early warning score tool. Therefore, after each observation, participants were asked “Tell me your satisfaction level with using the PEWS on your unit.”

**RESULTS**

**Medical record observations**

Vital signs and early warning scores were missing for 11% and 15% of the data retrospectively, perhaps because patients were unavailable (eg, off the unit), routine VS assessments were not ordered, or there was an omission by the provider. We also excluded from the analysis the assessments posted after the time they were actually entered.

The final number of VS instances in the first data collection was 2583, 73% documented by non-RNs (n = 1878) and 27% documented by RNs (n = 675). The Figure shows the average delay in minutes in charting of VS by type of personnel for 4 time blocks: 0 to 9 minutes, 10 to 59 minutes, 60 to 119 minutes, and more than 120 minutes. The mean delay in minutes was 20 minutes for non-RNs (median = 0, interquartile range [IQR] = 0-26) and 36.5 minutes for RNs (median = 3, IQR = 0.50). The difference in delay time between non-RNs and RNs was significant (P < .0001).

The final number of early warning score instances was 2556 in the first data collection period and 2078 in the second data collection period (recall that only RNs document early warning scores). The mean delay in time for our first set of medical record audits was 77 minutes (median = 50, IQR = 9-109). In our second data collection period, the delay increased to 83 minutes (median = 54, IQR = 16-120). The increase between the 2 time periods was significant (P = .0009).

As for consistency of the early warning scores with the most recently charted VS, the Table shows the findings for the consistency of RN charting of the score with the charting of elevated VS for both periods of review. At both time periods, consistency was low for both HR and RR (range, 0%-41%) but was significantly improved for HR posteducational campaign (P = .001).

![Figure](image_url) Average time delay of vital sign documentation: nurses versus PCTs. PCT indicates patient care technician; RN, registered nurse.
Table. Consistency of PEWS Documentation With Elevated VS Before and After Educational Campaign

<table>
<thead>
<tr>
<th>Elevated VS</th>
<th>Preeducational Campaign, n/Total (%)</th>
<th>Posteducational Campaign, n/Total (%)</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>HR &gt; 20 above normal</td>
<td>6/71 (8)</td>
<td>22/62 (35)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>HR &gt; 30 above normal</td>
<td>0/24 (0)</td>
<td>25/66 (29)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>RR &gt; 10 above normal</td>
<td>21/70 (30)</td>
<td>44/106 (41)</td>
<td>.12</td>
</tr>
<tr>
<td>RR &gt; 20 above normal</td>
<td>2/18 (11)</td>
<td>7/28 (25)</td>
<td>.25</td>
</tr>
</tbody>
</table>

Abbreviations: HR, heart rate; PEWS, Pediatric Early Warning Score; RR, respiratory rate; VS, vital signs.

Behavioral observations

A total of 15 RNs were observed for ninety 1-minute time blocks. Of the observations, the median number of time blocks assigned an activity was 114, demonstrating that there were blocks spent in more than 1 activity (ie, multitasking). Over the 90-minute observation session, the median number of time blocks in which the nurse was engaged in each activity was 18 (16%, IQR = 10.5%-19.5%) for physical assessment, 47 (41%, IQR = 40%-55.5%) for documentation, 30 (26%, IQR = 18.5%-53.5%) for communication, and 14 (12%, IQR = 6%-21.5%) for “other,” findings that are similar to other research on RN workflow. All but one nurse used a paper system to take notes during physical assessments and then referenced this paper while charting at the computer. Vital sign monitors rather than physical VS assessments were used at least once per observation for 13 of the 15 observations. Finally, bedside documentation occurred for 91 of 611 (14%) observed 1-minute time blocks of EMR charting time.

Nurse feedback

RNs stated that concurrent assessment and charting were encouraged for the PCT staff and that they had different strategies for prioritizing tasks in a busy environment. They also mentioned that barriers to concurrent documentation involved lack of computer availability and/or functionality, excessive log-on times, and preferences for not charting in front of patients and families. Interestingly, RNs stated that they were likely to postpone documentation of both elevated and nonelevated early warning scores. In the case of an elevated score, other tasks, such as communication with other care providers, collection of additional information about patient condition, and initial patient interventions, were given more priority. In the case of nonelevated scores, they again referred to the work-around of batching many documentation tasks. With respect to consistency with PCT-entered VS, RNs revealed that they did not regularly reference these when making their early warning score calculations, instead, using their own VS assessment that they had not documented. Finally, some RNs reported a lack of confidence in the ability of the tool to detect deterioration above and beyond their own assessment skills.

DISCUSSION

We confirmed our concerns that the early warning score used in our hospital was not working as intended. There was an assumption that VS and early warning scoring were both consistent and timely and therefore reliably serving as a foundation for making our escalation algorithm “work” as described in the literature. We suspect that many of the workflow issues we identified are not unique either to our hospital or to our nursing staff.

The medical record audit data demonstrated significant delays in documentation of
VS and early warning scores by RNs. The observations and RN feedback suggest that the lack of in-time documentation and batching of work are driven by multiple forces. Second, our findings of very low consistency between elevated RR and HR scores and their associated early warning subscores might be alarming, in particular because these patients are at a higher risk of deterioration. RNs provided feedback that they did not regularly reference VS that they had not collected when calculating the early warning score.

RNs were ambivalent about the effectiveness of an early warning score in relation to their own assessment skills; were frustrated by the time spent scoring and documenting the early warning score, a large percentage of which are normal or low; and variably implemented the early warning score on the basis of competing pressures on their time. In a systematic literature review on the RN’s role in detecting patient deterioration, Odell et al25 demonstrated that an early warning score assessment or its documentation often followed rather than preceded detection of a sick child. Detecting deterioration through routine assessments of patient condition was the least used process across all of the studies reviewed. Instead, RNs relied on intuitions or “gut feelings” that something was wrong. These feelings are tied to recognition of deviations from normal patterns, which then lead to additional assessments of patient status such as VS and calculation of early warning scores.26

Finally, we found significant differences pre- and posteducational campaign: an increase in the delay in the charting of the early warning score and more consistent documentation between the cardiovascular subscore and HR. We doubt the clinical significance of the first finding, as the delay was more than an hour at both time periods. While there was improvement in consistency, it was still poor. Both results require further investigation concerning the impact of education on RN documentation practices and task prioritization.

Additional system-based improvement initiatives to improve the functionality of early warning scores could include changes to the physical environment or improved technology interfaces that are more supportive of real-time data entry. For example, optimized bedside computer access, or tablet or smartphone devices with applications for documentation and relay of information to the EMR, might address the issues with timeliness. Some EMRs have already been adapted so that VS are auto-populated into the early warning score tool to ensure consistency. In addition, the unintended consequences of non-RN and RN task assignments should be examined, and perhaps a return to VS assessment by RNs to reintegrate VS collection and documentation with RN physical assessments should be considered.

Furthermore, there is a need for more effective translation of nurse intuition into the kinds of empirical evidence that other providers desire; this may allow for a more balanced and functional integration of subjective and objective patient assessments such as an early warning score.27,28 Awareness of a culture of ambivalence about early warning scores must be acknowledged, and more direct attention to how communication between providers occurs and is acted on is important. Expanding the criteria for assessment of risk of deterioration might include provider intuitions, family concerns, and concerns about potential communication problems within the team. Communication about deterioration should occur at the bedside to achieve a “shared mental model” about the physical assessment findings and resulting explicit actions to be taken by team members, including making predictions about outcomes from proposed interventions and documenting an expected time frame in an escalation plan.27,29,30

Limitations

While we believe that the issues we have identified might be similar to other pediatric hospitals, we are a single site and our findings may not be generalizable to other institutions. Other limitations to our study are the small sample of behavioral observations
of RN workflow around patient assessment. Also, we do not know whether the delay or inconsistencies of early warning scores have any clinical significance on patient care—deterioration detection may occur by other means, and documentation may be delayed for more important patient care that serves to improve patient status.

CONCLUSIONS

To our knowledge, findings on the lack of timeliness of early warning scores and VS inconsistency with early warning scores have not been presented before in the literature. This calls into question the effectiveness of the early warning scores as a screening tool in actual practice, and fundamental questions about the role of clinical acumen versus objective tools still remain. Cumbersome technology interfaces, workflow under time pressure, human cognitive biases, caregiver-task mismatches, and increased time away from the bedside all are barriers to early warning score utilization. Additional work, with input from experts in human factors and workflow analysis, is needed.

REFERENCES