Rapid response systems. Recognition and management of the deteriorating patient

Ludikhuize, J.

Citation for published version (APA):
Chapter 3

How nurses and physicians judge their own quality of care for deteriorating patients on medical wards: Self-assessment of quality of care is suboptimal

Jeroen Ludikhuize
Dave A. Dongelmans
Susanne M. Smorenburg
Miranda Gans – Langelaar
Evert de Jonge
Sophia E. de Rooij

Critical Care Medicine 2012
Abstract

Objective: To describe how nurses and physicians judge their own quality of care for deteriorating patients on medical wards compared with the judgment of independent experts.

Design: Cross-sectional study using interviews of care-providers regarding their perceived quality of care for clinically deteriorating patients compared with retrospective judgment by independent experts.

Setting: Academic Medical Center of Amsterdam, the Netherlands.

Patients: Between April and July 2009, all patients with cardiopulmonary arrests and unplanned intensive care unit admissions from six medical nursing wards were included. The care-providers (nurses and physicians) taking care of these patients in the previous 12 hrs were included.

Measurements and Main Results: Forty-seven events and 198 interviews were analyzed. Skill and knowledge level regarding the recognition of a deteriorating patient were rated on a scale of 1–10 with means (sd) of 7.9 (0.8) and 7.7 (0.9), respectively. Nurses and residents attributed coordination of care largely to themselves (74% and 76%, respectively). Communication, cooperation, and coordination were graded in a positive manner (medians between 7.3 and 8), whereas the medical staff graded these factors higher compared to the grading by nurses and residents. Negative predictive values regarding the presence of a delay compared with an expert panel was 37% for nurses and 38% for residents and specialists.

Conclusions: Care-providers mostly rate their care provided to patients in the hours preceding a life-threatening adverse event as good. In contrast, independent experts had a more critical appraisal of the provided care in regards to timely recognition. These findings may partly explain the reluctance of care-providers to implement patient safety initiatives.
Introduction

Adverse events (AEs), such as cardiopulmonary arrests or unplanned intensive care unit (ICU) admissions, are generally preceded by clear and detectable signs of deterioration in the 24 hrs prior to these events.1–3 Quality of care before an AE has been deemed substandard. Possible causes include failure to recognize clinical urgency, lack of knowledge, and failure to call for assistance.4

To improve the care of patients on nursing wards who are at risk of an AE, rapid response systems (RRSs) have been initiated to facilitate the process of recognition and early treatment of deteriorating patients.5 One important component of these systems is a track and trigger system, which enables care-providers to easily assess patients.6 The Modified Early Warning Score (MEWS) based on Subbe et al is widely used.7–9

To date, unequivocal clinical evidence regarding the effectiveness of RRS is not available.10–12 Current evidence suggests that many factors influence the process of identification and response to a clinically deteriorating patient.13 These include communication and teamwork14, knowledge and skill15, and leadership and coordination of care at moments of clinical instability.16,17 Currently, little is known about the perception of care-providers regarding these factors. Therefore, in a setting in which a RRS was not implemented, we analyzed the perceptions of care-providers regarding the presence of a delay in recognition of deterioration and compared this with the judgment of an independent expert panel. Secondly, we also studied which factors were considered by the care-providers to be important or influential for the process of recognition.

Materials and Methods

Hospital setting

The study was conducted in the Academic Medical Center in Amsterdam. Six medical wards with 150 beds were included, and the wards admitted patients from internal medicine, infectious diseases, pulmonology, nephrology, and gastroenterology and included four high dependency beds.

At time of this study, no components of a RRS (including MEWS) had been implemented. This meant that no specific protocol was present to measure vital signs, and there were no specific calling criteria described to call the ICU for bed-side consultation.

Study design

In this cross-sectional study, a convenience sample of all patients enduring a cardiopulmonary arrest or unplanned ICU admission between April 1 and July 1, 2009, were included. Cardiopulmonary arrest was defined as an event in which respiratory and/or cardiopulmonary activity was absent and for which the cardiac arrest team was called.
and initiated cardiopulmonary resuscitation, including pharmacological, fluid, and/or mechanical resuscitation. An unplanned ICU admission was defined according to the Dutch national ICU registry as an admission that could have been deferred without risk for at least 12 hrs.

All care-providers who were responsible for the patient in the 12 hrs prior to the AE were requested to participate. This included nurses, residents, and medical specialists. According to their work schedules, the required care-providers were contacted through telephone, e-mail, and/or direct personal contact.

**Development and administration of the questionnaire**

Employing a semi-structured literature review incorporating cross-referencing of (key) publications and employing peer review, five domains were identified that are involved in providing care in acute hospital settings. These domains were as follows: 1) individual skills and knowledge regarding the recognition of deteriorating patients, 2) communication and transfer of information, 3) teamwork, 4) leadership and coordination, and 5) supervision. Subsequently, a questionnaire was formulated employing a panel that included two nurses and two physicians from the ICU, one physician from the field of internal medicine, and two staff members from the patient safety department of our institution. For questions that required grading, a scale from 0 to 10 was employed. Face validity was further tested on three nurses and two residents for clarity, feasibility, and ambiguity (see supplementary materials).

The questionnaire was administered in a face-to-face manner by two trained investigators. No relevant patient material, including patient charts, was available to the participants. The investigator initiated the questionnaire by providing a short patient description to ensure that the participant was directly involved in the care of the particular patient.

All data were anonymously entered into a Microsoft Access database that was built for the purposes of this study. Given the observational nature of the study, and according to the local institutional review board, a waiver was obtained to conduct the research without informed consent. The study conforms to the provision of the Declaration of Helsinki in 1975 (revised in 2008 Seoul, South Korea).

**Expert panel and Modified Early Warning Score**

The expert panel (consisting of an internist [S.E.d.R.], an intensivist [D.A.D.], and an experienced ICU nurse [M.G.L.]) was asked to independently rate each case for the possible presence of a delay in the recognition of the deteriorating state of the patient and whether the event could have been avoided. The experts were provided with all patient materials including photocopies of the patient chart for preparation purposes. Although the MEWS was not operational in the hospital, all registered vital parameters in the 48 hrs prior the event were also noted, and according to the MEWS (system
provided in Supplementary Figure 1.) values were calculated for each measurement and provided to the experts for incorporation into their review. The experts were blinded to the results of the questionnaires. After careful individual preparation, decisions regarding the avoidability and the delay of recognition were made by consensus. In cases where no consensus could be reached, an independent experienced intensivist (E.d.J.) made a final decision.

Data analysis and statistics
Continuous variables are expressed as the mean and SD if they are normally distributed or as medians and interquartile ranges if they are not normally distributed. Categorical variables are expressed as counts (%).
To test more than two independent groups of continuous normally distributed variables, the analysis of variance test and post hoc analysis were used. Likewise, if continuous data were not normally distributed, the Kruskal–Wallis test to compare three or more groups was used. If data appeared to be non-independent, the Friedman test and the Wilcoxon signed rank test were used to test for differences. As incidental clustering of data occurred within some of the care-providers (repeatedly assessing the same patient); we reduced these multiple measures by replacing them with the median value. The assessment data from each care-provider per event from the shift in which the event took place were taken as the primary data point. Interobserver correlation (Cohen’s $\kappa$) was calculated between judgments of event qualifiers by the different care-providers. We deemed expert opinion to be the reference standard against which to calculate the assessments of other care-providers in terms of positive and negative predictive values.
Statistical uncertainty was expressed by 95% confidence limits as appropriate, and statistical significance was defined at 0.05. The analyses were performed using SPSS version 19.0 (Chicago, Illinois, USA) and confidence interval analysis (CIA) software version 2.2.0 (University of Southampton, UK).

Results
Demographics
Forty-seven consecutive patients were included (see Table 1). Nine of these patients underwent a cardiopulmonary arrest, and 38 (81%) underwent an unplanned ICU admission. Because all events originating from the nursing wards were included, six patients experienced two separate events, and two patients experienced a total of three separate events. In the following, the term “patient” indicates a “unique” event and is used indiscriminately. Thirty-three patients (70%) were male, and the median (interquartile range) age was 61 yrs (49–73 yrs). The median length of hospital stay was 61 hrs (14–314 hrs). Among patients undergoing an unplanned ICU admission, the median Acute
Physiology And Chronic Health Evaluation II score was 19 (15–23), and the median Simplified Acute Physiology Score II score was 47 (38–56). The median length of stay in the ICU was 66 hrs (22–124).

In the 12 hrs prior to the event, 233 care-providers were involved in the care of these patients and were asked to participate in the study. One resident declined to participate. Fourteen care-providers (6%) were not traceable, and contact could not be established. Twenty care-providers (9%) stated that during their shift they did not have any interactions with the patient. Therefore, 198 questionnaires (85%) were available for analysis with between two and six questionnaires being administered per patient. A total of 126 (64%) of the questionnaires were related to care-providers working during the shift in which the event took place. Respondent demographics are shown in Table 2. A total of 117 “unique” care-providers participated in the study. In general, senior medical staff members were more often male (68%), older (median 45), and had more experience (median 20) compared to the other professionals.

**MEWS values and moment of concern**

For the 47 events, a total of 639 measurements of vital signs were taken in the 48 hrs prior to the event. Retrospectively, a median MEWS score of 3 (1–4) was found for all measurements. Thirty-eight (81%) patients reached the MEWS cutoff score of three or more during this period, indicating that the patient should be considered at risk for deterioration. The median MEWS value for the measurements taken from these “positive” measurements was 4 (3–6).

Among care-providers working during the “event-shift,” two nurses, five residents, and 13 (32%) specialists indicated that they did not experience any concerns regarding the
patients’ clinical condition, which was similar to reports from the care-providers working during the shift preceding the event.

Perception of care-providers

Irrespective of the event they were involved in, all care-providers were asked to rate their skill and knowledge level on a scale from 0 (low) to 10 (high) regarding their ability to recognize and manage a deteriorating patient. Overall, knowledge level was rated to be a mean (sd) of 7.9 (0.8), and skill level was rated to be a mean of 7.7 (0.9). Skill level showed no significant differences between type of care-provider (one-way analysis of variance, \( p = .089 \)). However, knowledge level differed significantly between nurses (7.7 [0.8]) and specialists (8.5 [0.7]) (unpaired t test \( p < .0001 \)) and between residents (7.8 [0.7]) and specialists (\( p = .001 \)). Table 3 shows the perception of cooperation, communication, and coordination of care. To account for clustering, the median grade per type of care-provider per event was calculated and used for further comparison. In general, specialists graded these factors higher than both nurses and residents. Between-group significant differences were present for all factors. Using the Wilcoxon signed ranked test to test for significant differences between two groups, the only comparisons that were not significant were those between residents and specialists for cooperation and communication (\( p = .24 \) and .187, respectively).

When care-providers working during the event were asked whether a coordinator of care who orchestrated the care was present prior to the event, 38 (91%) residents, 37 (90%) specialists, and 38 (88%) nurses responded that a coordinator was present. Interestingly,
Table 3. Perception of cooperation, communication and coordination of care.

<table>
<thead>
<tr>
<th></th>
<th>Grade</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooperation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nurse</td>
<td>7,5 (7,3 - 8,0)</td>
<td>0,002</td>
</tr>
<tr>
<td>Resident</td>
<td>8,0 (7,5 - 8,3)</td>
<td></td>
</tr>
<tr>
<td>Specialist</td>
<td>8,0 (7,7 - 9,0)</td>
<td></td>
</tr>
<tr>
<td>Coordination</td>
<td></td>
<td>&lt; 0,001</td>
</tr>
<tr>
<td>Nurse</td>
<td>7,3 (6,8 - 8,0)</td>
<td></td>
</tr>
<tr>
<td>Resident</td>
<td>8,0 (7,5 - 8,0)</td>
<td></td>
</tr>
<tr>
<td>Specialist</td>
<td>8,0 (8,0 - 9,0)</td>
<td></td>
</tr>
<tr>
<td>Communication</td>
<td></td>
<td>0,003</td>
</tr>
<tr>
<td>Nurse</td>
<td>7,5 (7,0 - 8,0)</td>
<td></td>
</tr>
<tr>
<td>Resident</td>
<td>8,0 (7,0 - 8,5)</td>
<td></td>
</tr>
<tr>
<td>Specialist</td>
<td>8,0 (7,7 - 9,0)</td>
<td></td>
</tr>
</tbody>
</table>

Care-providers were asked to rate the degree of cooperation, coordination, and communication regarding the care provided prior to each event taking place. To account for cluster effects, a single-grade per event, per type of care-provider was generated by averaging the grades given by the type of care-provider. Statistical differences between groups were tested employing the Friedman test. Grades are presented as medians with interquartile ranges.

Table 4. Presence of delay and degree of avoidability compared with expert opinion: Perception of care-providers working during the event regarding presence of delay in recognition.

<table>
<thead>
<tr>
<th></th>
<th>Nurse</th>
<th>Resident</th>
<th>Specialist</th>
<th>Expert</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delay present</td>
<td>13 (31)</td>
<td>8 (19)</td>
<td>6 (15)</td>
<td>28 (60)</td>
</tr>
<tr>
<td>No delay present</td>
<td>28 (65)</td>
<td>30 (71)</td>
<td>33 (80)</td>
<td>18 (38)</td>
</tr>
<tr>
<td>Could not be determined</td>
<td>2 (5)</td>
<td>4 (10)</td>
<td>2 (5)</td>
<td>1 (2)</td>
</tr>
</tbody>
</table>

Perception of delay in recognition of the deteriorating states of care-providers working during the shift in which the event occurred were compared with the expert panel. Percentages of delay for each type of care-provider is presented and also for the expert panel. Percentages were rounded to integer numbers.

76% of the residents and 74% of the nurses indicated that they themselves were the coordinators of care compared to only 17% of the specialists.

**Delay and avoidability**

Tables 4–6 show the presence of a delay according to both the care-providers working during the event and the expert panel. Nurses indicated that a delay was present in 13 (31%) cases. Residents indicated that there was a delay in eight cases, and specialists reported a delay in six cases. The expert panel concluded that a delay was present in 28 (60%) cases.

Concordance with the reports of the care-providers who worked during the shift prior to the event was high (between 89% and 96% agreement).
Positive and negative predictive values were calculated using the expert decision as the gold standard (Table 5). Positive predictive values for all types of care-providers (i.e., the percentage in which the experts and care-providers agreed regarding the presence of a delay) were between 62% (95% confidence interval 36–82) among nurses and 67% (95% confidence interval 30–90) among specialists. Negative predictive values (percentage in which the experts and care-providers concurred in the absence of delay regarding the recognition of the patients. Negative predictive values represent the “true negative ratio” in which the experts and care-providers concurred in the absence of delayed recognition, and thus timely recognition of the deteriorating clinical state, Cohen κ’s were used to assess the degree of inter-rater compatibility between the type of care-providers in the shift of the event compared with the shift preceding the event. Percentages were rounded to integer numbers. CI; Confidence Interval.

Comparative analysis is performed in which the expert opinion was set as “reference standard.” Positive predictive values represent the “true-positive ratio” in which the experts and care-providers concurred in the presence of delay regarding the recognition of the patients. Negative predictive values represent the “true negative ratio” in which the experts and care-providers concurred in the absence of delayed recognition, and thus timely recognition of the deteriorating clinical state, Cohen κ’s were used to assess the degree of inter-rater compatibility between the type of care-providers in the shift of the event compared with the shift preceding the event. Percentages were rounded to integer numbers. CI; Confidence Interval.

Table 5. Presence of delay and degree of avoidability compared with expert opinion: Predictive capabilities and inter-rater comparability between care-providers (event-shift) and expert panel.

<table>
<thead>
<tr>
<th>Expert versus</th>
<th>Positive Predictive Value (95% CI)</th>
<th>Positive Predictive Value (95% CI)</th>
<th>κ (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nurse</td>
<td>62 (36 - 82)</td>
<td>37 (22 - 56)</td>
<td>-0.01 (-0.30 - 0.27)</td>
</tr>
<tr>
<td>Resident</td>
<td>63 (31 - 86)</td>
<td>38 (23 - 56)</td>
<td>0.00 (-0.28 - 0.28)</td>
</tr>
<tr>
<td>Specialist</td>
<td>67 (30 - 90)</td>
<td>38 (23 - 55)</td>
<td>0.02 (-0.25 - 0.29)</td>
</tr>
</tbody>
</table>

Comparative analysis is performed in which the expert opinion was set as “reference standard.” Positive predictive values represent the “true-positive ratio” in which the experts and care-providers concurred in the presence of delay regarding the recognition of the patients. Negative predictive values represent the “true negative ratio” in which the experts and care-providers concurred in the absence of delayed recognition, and thus timely recognition of the deteriorating clinical state, Cohen κ’s were used to assess the degree of inter-rater compatibility between the type of care-providers in the shift of the event compared with the shift preceding the event. Percentages were rounded to integer numbers. CI; Confidence Interval.

Table 6. Presence of delay and degree of avoidability compared with expert opinion: Perception of avoidability per type of care-providers (both shifts) and expert panel.

<table>
<thead>
<tr>
<th>Avoidability</th>
<th>Grade</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nurse</td>
<td>2.5 (1.0 - 5.0)</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Resident</td>
<td>1.0 (0 - 3.0)</td>
<td></td>
</tr>
<tr>
<td>Specialist</td>
<td>2.0 (0 - 4.8)</td>
<td></td>
</tr>
<tr>
<td>Expert</td>
<td>5.0 (4.0 - 8.0)</td>
<td></td>
</tr>
</tbody>
</table>

The degree of avoidability was rated on a scale from 0 (low avoidability) to 10 (high avoidability) and compared with expert opinion in which the median score per event, per type of care-provider was generated to account for clustering. Grades are presented as medians with interquartile range. Testing for significant differences between groups was performed using the Friedman test.

Positive and negative predictive values were calculated using the expert decision as the gold standard (Table 5). Positive predictive values for all types of care-providers (i.e., the percentage in which the experts and care-providers agreed regarding the presence of a delay) were between 62% (95% confidence interval 36–82) among nurses and 67% (95% confidence interval 30–90) among specialists. Negative predictive values (percentage in which the experts and care-providers agreed regarding the absence of a delay, which we refer to as timely recognition) were low; the negative predictive value was between 37% (95% confidence interval 22–56) among nurses and 38% (95% confidence interval 23–56/55) among residents and specialists, respectively. An interobserver correlation using Cohen’s κ for all comparisons between the care-providers who worked during the event-shift and the expert panel was poor, which was approximately 0.

Avoidability of the event (Table 6) was rated on a ten-point scale (0 indicating low avoidability and 10 indicating high avoidability). In general, avoidability was deemed low by the care-providers compared with the experts. Significant differences between all types of care-providers and the experts were present (p < .0001, Wilcoxon signed rank test), but there were no differences among the care-providers.
Discussion

Using a cross-sectional study design, 47 consecutive cardiopulmonary arrests and unplanned ICU admissions were included during a 3-month time period. Using a cut-off value of three MEWS points, 81% of patients could be identified as being at risk for deterioration at a median of 14 hrs prior to the AE. Self-perceived quality of care indicated that care-providers were satisfied with the care they provided to these patient prior to the event in terms of communication, cooperation, and coordination of care. Judgment by the care-providers regarding the presence of a delay in recognizing clinical deterioration did not correspond to the judgment of the expert panel. Although seemingly unimportant, approximately 70% of nurses and residents viewed themselves as being the coordinator of care prior the event.

An earlier study reported that care on general wards prior to ICU admission was suboptimal. Failure of the prompt detection of deteriorating patients has been associated with an increase in mortality. The vast majority of care-providers considered their role in the care of the patients as being adequate to excellent. Given the fact that these patients suffered serious AEs, it seems that there is a lack of critical self-assessment of the involved care-providers. There was a significant discrepancy between the opinions of the care-providers and those of the expert panel regarding the avoidability of the event; this difference was even larger regarding the presence of a delay in recognition of the deteriorating state. This likely represents a lack of critical appraisal by care-providers of their professional performance. One study reported that the accuracy of self-assessment of physicians regarding their competence may be limited. It is difficult to explain why these differences are present. Research from associated fields may indicate that upon additional education utilizing simulation training, participants may become more aware of all the factors that are involved in “good care,” which can result in increased survival. Additionally, the experience levels of nurses and physicians, of which our group was relatively inexperienced, may also be directly related to an increase in failure rates, possibly due to a lack of insight. A positive attitude toward change, as well as self-confidence, has been shown to directly increase the quality of care and to improve self-judgment.

Another crucial factor in the care of patients at risk is coordination, which results from clear leadership. Interestingly, 74% of nurses and 76% of residents considered themselves to be the coordinator of care, whereas 10% of care-providers indicated that no coordinator was present. It is apparent that both professional groups feel that they are dedicated and responsible regarding their patients. However, the fact that both nurses and residents consider themselves as coordinators of care may indicate that they feel dedicated and responsible for that patient. This may, however, have great implications for the patient because in the absence of a clear care coordinator, confusion might arise quickly, and mistakes may accumulate in this sometimes hectic and stressful period.
A strength of this study is the fact that all physicians and nurses who cared for the patient in the 12 hrs prior to the event were asked to participate, and the response rate was very high. This study has also some weaknesses. The use of an expert panel may introduce some bias and may over- or underestimate the actual presence of a delay in recognition. Secondly, as this questionnaire was set up for a quantitative analysis, we had to restrict the number and type of questions. As a consequence, this may in some occasions limit the in-depth analysis of our observations. Finally, selective recall by the care-providers may be present because the time after the event to the administration of the questionnaire was a median of 13 days. However, all care-providers immediately recognized the patient based upon a simple case description, which is likely due to the relative rarity of these events and their impact on people. However, the differences between how experts and care-providers view delays in recognition of patients at risk are so large that we consider it to be very unlikely that this could be the mere result of overestimation by experts, and we propose that they are likely to be related to an inaccuracy of self-assessment.

In conclusion, care-providers’ lack of critical self-assessment regarding the quality of care they provide to patients undergoing life-threatening AEs is apparent. Improving critical self-assessment should be part of RRS or other initiatives to improve outcomes for deteriorating patients on nursing wards.

**Conflict of interest statement**

No conflict of interest declared.

**Acknowledgements**

We would like to thank Kimm Henricks and Jacqueline Lankelma for their significant contributions in the logistics and acquisition of data for this study. For statistical guidance, we would like to thank Dr. Jan Binnekade for his time and effort invested in this project.
Reference List


Supplementary file containing the questionnaire with manuscript “How nurses and physicians judge their own quality of care in deteriorating patients on medical wards. Self-assessment of quality of care is suboptimal.”

**Part A: Patient demographics**
1. Date of birth __ _ / _ _ / 19 _ _ (dd / mm / yyyy)
2. Gender  Male / female
3. Date of admission __ _ / _ _ / 20 _ _ (dd / mm / yyyy)
4. Specialty of admission Internal Medicine / other
5. Type of event Cardiopulmonary arrest / unplanned ICU admission
6. Date of event __ _ / _ _ / 20 _ _ (dd / mm / yyyy)
7. Time of event __ _ / _ _ (hh / mm)
8. Department at time of event Nursing ward / Medium Care

**Part B: Care provider demographics**
1. Date of birth __ _ / _ _ / 19 _ _ (dd / mm / yyyy)
2. Gender  Male / female
3. Profession Nurse / Resident / Specialist
4. Years post-graduation __ _ (yy)
5. Employment at AMC __ _ (mm)
6. Employment at current ward __ _ (mm)

**Part C: Questionnaire**
At the beginning of the questionnaire, a check was performed according to the patient material, whether the care provider was indeed responsible for the patient within the 12 hour timeframe prior the event. The following questions were used to structure the interview and guide documentation of the answers provided.

**Introduction and general questions**
1. At what moment prior to the event did you feel worried concerning the patient's clinical condition?
   - I did not feel worried at any time regarding this patient
   - 0 – 1 hour prior the event
   - 2 – 4 hours prior the event
   - 4 – 8 hours prior the event
   - 8 – 12 hours prior the event
2. If yes, did you discuss your worries at anytime with a colleague? And if so, with whom?
   - Yes, I discussed my worries with:
     i. Other nurse
     ii. Resident
     iii. Medical specialist
     iv. ICU
     v. Other
   - No, I did not discuss my worries at anytime (Proceed to question 5)
3. On a scale from 0 (insufficiently discussed) to 10 (excellently discussed), how would you grade to your opinion the situation whether your worries were sufficiently discussed in that conversation?
4. In your view, were you the first person that felt worried regarding the clinical conditions of the patient?
   - Yes
   - No
   - Don’t know
   - No opinion

Domain 1: Knowledge and skills
5. How would you grade your knowledge regarding recognizing a deteriorating patient on a scale from 0 (no knowledge) to 10 (optimal knowledge)?
6. How would you grade your skill level regarding coordinating the optimal care for a deteriorating patient on a scale from 0 (no skills) to 10 (optimal skills)?

Domain 2: Communication and cooperation
7. Was the patient transferred to your care at the start of your shift?
   - Yes
   - No (Proceed to question 9)
8. How would you grade the information of this transferral on quality and usefulness on a scale from 0 (poor, unusable) to 10 (excellent, usable)?
9. How would you grade the general level of communication between all involved care providers in the hours prior the event on a scale from 0 (poor, unusable) to 10 (excellent, usable)?
10. How would you grade the general level of communication between all involved care providers in the hours prior the event on a scale from 0 (poor, unusable) to 10 (excellent, usable)?
Domain 3: Teamwork

11. How would you grade the quality of cooperation between all involved care providers in the hours prior the event on a scale from 0 (poor) to 10 (excellent)?

Domain 4: Coordination and leadership

12. In the hours preceding the event, according to your opinion, was there one clear coordinator or director present who coordinated the care for this patient?
   o Yes
   o No (Proceed to question 15)
   o Don’t know

13. In case one coordinator/director of care was present, who was that person according to your opinion?
   o Myself
   o The nurse
   o The resident
   o The specialist
   o Don’t know

14. Regarding the person who was the coordinator/director of care, could you rate his/her accessibility on a scale from 0 (no threshold) to 10 (high threshold)?

15. How would you grade the quality of the coordination of care in the hours preceding the event on a scale form 0 (poor) to 10 (excellent)?

Domain 5: Supervision

16. In regards to the moments of supervision. Did you feel that you were taken seriously and also understood on a scale from 0 (poor) to 10 (excellent)?

17. How would you grade the general culture in terms of hierarchy on your ward i.e. Specialty on a scale from 0 (no apparent hierarchical structure) to 10 (stringent hierarchy)?

Domain 6: Delay and avoidability

18. If you take into account this case including all possible associated factors, how would you perceive the possibility that delay in recognition of the deteriorating was present?
   o No delay present, on time
   o Delay present
   o Could not be determined

19. If you take into account this including all possible associated factors, how would you grade the avoidability of this event on a scale from 0 (absolutely unavoidable) tot 10 (absolutely avoidable)?
20. At this moment, would you like to comment on any of the particular domains regarding certain aspects which, to your opinion, have been underexposed regarding the recognition of the deteriorating condition in this patient?

Part D: Expert panel

21. If you take into account this case including all possible associated factors, how would you perceive the possibility that delay in recognition of the deteriorating was present?
   o No delay present, on time
   o Delay present
   o Could not be determined

22. If you take into account this including all possible associated factors, how would you grade the avoidability of this event on a scale from 0 (absolutely unavoidable) to 10 (absolutely avoidable)?

Definitions of the domains and specific topics (if applicable):
Wikipedia and a dictionary were used as source for the definitions and discussed by the expert panel.

Knowledge:
Knowledge is a familiarity with someone or something, which can include information, facts, descriptions, and/or skills acquired through experience or education.

Skill:
A skill is the learned capacity to carry out pre-determined results often with the minimum outlay of time, energy, or both. Skills can often be divided into domain-general and domain-specific skills.

Communication:
Communication is the activity of conveying meaningful information. Communication requires a sender, a message, and an intended recipient, although the receiver need not be present or aware of the sender’s intent to communicate at the time of communication.

Cooperation:
Cooperation is the process of working or acting together. In its simplest form it involves things working in harmony, side by side.

Teamwork:
Teamwork is work performed by a team towards a common goal. In health care teamwork has been defined as: a dynamic process involving two or more healthcare professionals with complementary backgrounds and skills, sharing common health goals and exercising concerted physical and mental effort in assessing, planning, or evaluating patient care.
Coordination:
Coordination is the act of coordinating a number of process and/or people, making different people or things work together for a goal or effect to fulfill a common shared goal.

Leadership:
Leadership has been described as the “process of social influence in which one person can enlist the aid and support of others in the accomplishment of a common task”.

Supervision:
A supervisor is first and foremost an overseer whose main responsibility is to ensure that a group of subordinates get out the assigned amount of production, when they are supposed to do it and within acceptable levels of quality, costs and safety.

**Supplementary figure 1.** The Modified Early Warning Score (MEWS).

**MEWS system:**

<table>
<thead>
<tr>
<th>MEWS score</th>
<th>3</th>
<th>2</th>
<th>1</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heart rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;40</td>
<td>40-50</td>
<td>51-100</td>
<td>101-110</td>
<td>111-130</td>
<td>&gt;130</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Systolic blood pressure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;70</td>
<td>70-80</td>
<td>81-100</td>
<td>101-200</td>
<td>&gt;200</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Respiration rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;9</td>
<td>9-14</td>
<td>15-20</td>
<td>21-30</td>
<td>&gt;30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Temperature</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;35,1</td>
<td>35,1-36,5</td>
<td>36,6-37,5</td>
<td>&gt;37,5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AVPU score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A (A)</td>
<td>V (response to voice)</td>
<td>P (reacting to pain)</td>
<td>U (Unresponsive)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Worried about patient’s condition: 1 point</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urine production below 75 milliliter during previous 4 hours: 1 point</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saturation below 90% despite adequate oxygen therapy: 3 points</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upon reaching 3 or more points → call resident in charge</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The MEWS instrument was implemented as a tool that ward staff can use to identify the patient at risk of deterioration. The described method was adapted from Subbe et al. 7