CURRENT OPINION

Early warning scores in the perioperative period: applications and clinical operating characteristics

Harm-Jan de Grooth[a,b], Armand R. Girbes[c], and Stephan A. Loer[a]

Purpose of review

Early warnings scores are designed to detect clinical deterioration and promote intervention at the earliest possible moment. Although the ultimate effects on patient outcomes are unclear, early warning scores are now legally mandated in several countries. Here, we review the performance of early warning scores in surgical and perioperative populations.

Recent findings

Early warning scores can be used to screen for postoperative deterioration and surgical complications. We describe a framework to evaluate the balance between missed events and warning signals that are not followed by an adverse event (nonevents). In large surgical cohort studies, the missed event rates ranged between 19 and 69% and the nonevent rates ranged between 72 and 99% for 'optimal' threshold early warning sores. Recent investigations have shown that there may be a substantial discrepancy between the theoretical benefits shown in validation studies and the practical clinical implementation of early warning scores, which may partly explain the absence of measurable benefit from these systems.

Summary

Early warning scores may facilitate protocolized escalation of care for patients at risk of adverse events and can be used in surgical and postoperative patients, but high nonevent rates and practical implementation problems can restrict their usefulness.

Keywords

patient safety, perioperative care, quality improvement

INTRODUCTION

Patients are approximately 1000 times more likely to die in 30 days after surgery than during surgery itself [1]. This justifies the significant efforts that have gone into the development and validation of early warning and intervention systems to identify patients with clinical deterioration in general wards. Ideally, such a system should warn clinicians early, with appropriate risk stratification and with minimal false alarms.

Several track-and-trigger early warning scores are now in use, most of which are modifications of a basic scoring system in which deviations in heart rate, blood pressure, respiratory rate, oxygen saturation, temperature and/or neurological function are weighted and aggregated into a single number [2▪▪]. Depending on local protocols, specific score thresholds lead to escalating care such as repeated scoring after 1 h, warning of the attending physician or the activation of a medical emergency team or a critical care outreach team. In some cases, early intervention can reverse the negative clinical trajectory, and, at the minimum, early warning scores can help identify patients who may benefit from closer monitoring, increased attention from the medical staff, or a higher nurse-to-patient ratio. Together, these warning scores and intervention systems (often termed as the afferent and efferent arms) are designed to facilitate early recognition and treatment of life-threatening problems.

The ultimate effects of these early warning and intervention systems on patient outcomes are unclear. Early before–after studies and a single-center ward-randomized trial showed promising effects

[a]Department of Anesthesiology, Amsterdam University Medical Centers, Location VUmc, Amsterdam, [b]Department of Anesthesiology, Zaans Medisch Centrum, Zaandam and [c]Department of Intensive Care, Amsterdam University Medical Centers, Location VUmc, Amsterdam, the Netherlands

Correspondence to Harm-Jan de Grooth, MD, Department of Anesthesiology, Amsterdam University Medical Centers, Location VUmc, De Boelelaan 1117, 1081 HV Amsterdam, the Netherlands.
Tel: +31 6 46 37 15 07; e-mail: h.degrooth@vumc.nl

Curr Opin Anesthesiol 2018, 31:732–738

DOI:10.1097/ACO.0000000000000657

KEY POINTS

- Early warning scores provide a uniform approach to detect patient at risk of adverse events within 24–48 h.
- Early warning scores were originally developed and validated in medical populations but appear to perform equally well in surgical and perioperative populations.
- The currently used early warning scores have high to very high nonevent rates. In the largest validation cohorts, 89–99% of early warning alarms were not followed by an adverse event.
- There is currently no high-quality evidence that early warning scores improve patient outcomes.
- In the future, patient-tailored algorithms may significantly improve the detection horizon and the balance between missed events and false alarms.

on hospital mortality, cardiac arrests and unplanned ICU admissions [3–6]. But the MERIT study, a landmark cluster-randomized trial, failed to demonstrate any positive effect of the introduction of a comprehensive medical emergency system [7]. Both a 2007 Cochrane review [8] and a 2014 systematic review [9] concluded that there was insufficient high-quality evidence to support the hypothesis that early warning and intervention systems improve patient outcomes. Most recently, a cluster-randomized trial of 21 hospitals in seven countries tested the effect of a pediatric early warning system on all-cause hospital mortality [10▪▪]. Among 144 539 hospitalized children, there was no difference between control vs. early warning score hospitals in mortality rates [adjusted odds ratio 1.01 (95% confidence interval; CI 0.61–1.96)], leading the authors to conclude that their findings do not support the use of this system to reduce mortality [10▪▪].

Many extraneous factors may influence the relationship between early warning scores and measurable improvements in outcomes. The patient-oriented effects of the early warning systems cannot be evaluated separately from the intervention system, making every system only as effective as the treatments that follow the alarms. Therefore, we believe that early warning scores merit a review that evaluates these scores on the terms that they were designed for: the ability to detect patients at high risk for future adverse events.

In this review, we provide an overview of the current state of early warning score research. We focus on the clinical operating characteristics of the scores rather than on the effect of early-warning and intervention systems on patient outcomes, which has been reviewed before [8,9]. We review the performance of several widely used early warning score iterations with special attention to perioperative patients and we present a framework to evaluate the performance of early warning scores in a clinically interpretable manner.

AVAILABLE EARLY WARNING SCORES FOR PERIOPERATIVE PATIENTS

Many different early warning scores have been developed, each suiting a specific hospital situation or tailored to the availability of data in a specific cohort. Single-parameter scoring systems have been shown to uniformly underperform multiparameter scores [11,12▪▪]. Of the multiparameter scores, almost all include heart rate, respiratory rate and blood pressure. Most scores include temperature and oxygen saturation, and some scores include supplemental oxygen, urine output, impaired cognition and a 'nurse concerned' parameter [13]. In Table 1, we present an overview of score validation studies that have been performed in surgical or postoperative patients.

The nomenclature of the different scores is not completely uniform, and several validation studies have left out specific parameters from scores because of the unavailability of data (such as mental status) in the study database. Nevertheless, multiple studies have found that the National Early Warning Score (NEWS) and the Vitalpac Vitalpack Early Warning Score (ViEWS) appear to predict cardiac arrest within 24 h better than other scores with a similar scoring structure [12▪▪,13,19]. In-depth analyses of the exact differences between various scores can be found elsewhere [13].

Most early warning score validation studies have been performed in medical or mixed populations [13], but two large studies have shown that the ViEWS and NEWS perform similarly in medical and surgical patients in terms of discrimination [equal area under the receiver operating characteristics curves (AUROCs) in both subpopulations] [14,15]. However, as the adverse event rate is generally lower in surgical patients, the false alarm rate will tend to be higher in this population [14,15].

Only one large study [12▪▪] has compared different scores in postoperative patients and found that the seven-item NEWS outperforms the six-item MEWS, and that both scores are outperformed by the 16-item electronic cardiac arrest triage (eCART) score (see Bartkowiak *et al.* in Table 2), a logistic-regression derived function that incorporates both vital signs and laboratory results [20].

More complex tools to identify at-risk patients are currently being developed. A recent study has shown that machine-learning algorithms such as random forests or gradient-boosting machines are

Table 1. The performance of different early warning scores in general surgical and perioperative patients

	Population	Sample size	Adverse events detected using score	Incidence of adverse events	Score, items	AUROC	Threshold[a]	Missed event rate	Nonevent rate
Bartkowiak *et al.* [12■■]	Postoperative patients	32 537	ICU transfer, cardiac arrest or death	3.82%	NEWS, 7	0.76	7	25%	95%
					MEWS, 6	0.75	3	19%	94%
					eCART, 16	0.79	2.1	25%	89%
Kellett and Kim [14]	Surgical patients scored on admission	30 485	Death within 48 h	0.04%	ViEWS, 6	0.89	7	69%	99%
			Death within 5 days	0.21%	ViEWS, 6	0.87	7	74%	97%
Kovacs *et al.* [15]	Nonelective surgical patients	20 626	Death within 24 h	1.37%	NEWS, 7	0.91	n.r.		
			ICU admission within 24 h	1.77%	NEWS, 7	0.76	n.r.		
			Cardiac arrest within 24 h	0.30%	NEWS, 7	0.86	n.r.		
			Death, cardiac arrest or ICU admission within 24 h	3.44%	NEWS, 7	0.87	n.r.		
Smith *et al.* [16]	Surgical patients	572	Death, cardiac arrest, ICU admision, emergency surgery or severe complications	8.04%	EEWS, 8	0.87	3	26%	74%
Hollis *et al.* [17]	Postoperative patients	552	Major postoperative complications with end-organ failure	6.70%	ViEWS, 6	0.90	8	19%	72%
Gardner-Thorpe *et al.* [18]	Surgical patients	334	ICU admission	4.79%	MEWS, 6	n.r.	4	25%	82%

[a]'Optimal' threshold alarm score as reported by study authors or, when no optimal threshold was reported, the threshold with a missed event rate as close as possible to 25%. The missed event rate is the proportion of adverse events not detected by an early warning alarm. The nonevent rate is the proportion of early warning alarm signals that are not followed by an adverse event. AUROC, area under the receiver operating characteristics curve; eCART, electronic Cardiac Arrest Triage; EEWS, expanded early warning score; MEWS, modified early warning score; n.r., not reported; NEWS, National Early Warning Score; ViEWS, Vitalpac Early Warning Score.

significantly better at discriminating patients at risk of adverse events [21]. These systems are not yet readily available at the bedside as their complexity requires computerized implementation.

Automated early warning devices may provide yet another tool to identify at-risk patients. Although an overview of the current developments in the field of remote monitoring is beyond the scope of this review, we highlight two studies with specific relevance to early warning scores. In a multinational before–after study including almost 20 000 patients, an 'automated advisory vital sign

Table 2. Performance metrics of early warning scores

Performance metrics	Calculation	Meaning in the context of early warning scores
Missed event rate	1 – sensitivity	Proportion of adverse events not detected by an early warning alarm
Nonevent rate	1 – positive predictive value	Proportion of early warning alarm signals that are not followed by an adverse event
Sensitivity	True positive rate	Proportion of adverse events detected by an early warning alarm
Specificity	True negative rate	Proportion of patients without an adverse event classified as 'safe' by the early warning score
Positive predictive value	$\frac{\text{sensitivity}\times\text{prevalence}}{\text{sensitvity}\times\text{prevalence}+(1-\text{specificity})\times(1-\text{prevalence})}$	Proportion of early warning alarm signals followed by an adverse event
Negative predictive value	$\frac{\text{specificity}\times(1-\text{prevalence})}{(1-\text{sensitivity})\times\text{prevalance}+\text{specificity}\times(1-\text{prevalence})}$	Proportion of 'safe' signals not followed by an adverse event
Area under the receiver operating characteristics curve (AUROC)	Area under the curve that describes sensitivity as a function of the false positive rate (1 – specificity).	Summary measure of discrimination: the probability that a patient with a future adverse event has a higher score than a patient without a future adverse event

monitor' based on the ViEWS was implemented in 10 hospitals [22]. Deployment of the device was not associated with an increase in the calling rate to the emergency team, but was associated with improved survival in the patients receiving an emergency team intervention. The clinical significance of these results has yet to be determined. In a feasibility study, an automated MEWS-scoring device incorporating end-tidal capnography and wireless nurse alerts was deployed in postoperative patients at risk of respiratory depression [23▪]. This resulted in 3.3 alarms per hour of monitoring, the majority of which were technical errors or measurement errors. This high rate of false alarms is compounded by the intrinsic nonevent rate of the MEWS described below.

RESOURCE UTILIZATION AND SCORE PROTOCOL ADHERENCE

A recent large Australian multicenter observational study with more than three million patients showed that there was a substantial increase in emergency calling rates from 2008 to 2016 and a concurrent decrease in hospital mortality during the same period [24▪]. However, the association between calling rates and mortality disappeared after correction for patient factors, illness and comorbidities, leading the authors to conclude that increased medical emergency calls do not reduce mortality.

Similarly, in the cluster-randomized MERIT study, the calling rate to the medical emergency team increased 2.8-fold in the intervention vs. control hospitals, but this was almost completely because of an increase in calls that were *not* subsequently associated with an adverse event [7]. Only 9% of the additional calls were followed by an adverse event.

In combination with the uncertain patient-oriented benefit, these data give rise to the concern that early warning scores lead to deskilling of ward staff and to reduced attention to the patients for whom the emergency response teams are primarily responsible. A relatively small single-center study reported that for every medical emergency call, there occurred on average 1.1 incidents as a consequence of staff leaving normal duties, although all of these incidents were minor and none led to patient-oriented adverse events [25].

In addition, several investigations suggest that there may be an important discordance between score validation studies and clinical practice. An analysis of more than 2.5 million NEWS records from multiple hospitals revealed statistically unlikely accumulations of heart rate records just below the threshold of 91 beats per minute and temperature records just above the threshold of 36.0°C [26▪]. An in-depth investigation into 67 patients who suffered an in-hospital cardiac arrest or unexpected death in a Danish hospital uncovered several problems with application of the protocolized (NEWS-based) early warning score [27]. The monitoring frequency in the 24 h prior to the events was correct in only 27% of cases and an elevated score was followed by an appropriate clinical response (alerting the correct staff member) in only 29–58% of cases.

Together, these data may explain to some extent the absence of measurable patient-oriented benefit from the implementation of early warning scores.

EARLY WARNING SCORES FROM THE CLINICIAN'S PERSPECTIVE: THE BALANCE BETWEEN MISSED EVENTS AND NONEVENTS

The performance of an EWS is generally described in terms of the AUROC, a value between 0.50 (no discrimination) and 1.00 (perfect discrimination). The AUROC can be interpreted as the probability that a patient with a future adverse event has a higher score than a patient without a future adverse event. As a summary measure of discrimination, the AUROC can be used to broadly compare the performance of different scoring systems, but it has no practical clinical value. Many studies, therefore, describe the sensitivity and specificity of one or more score thresholds. Sensitivity has direct patient-level relevance as it can be interpreted as the detection probability of a future adverse event. But the specificity of a score threshold is only meaningful in relation to the incidence of adverse events, as we will demonstrate below.

We propose that the missed event rate and the nonevent rate are more clinically relevant and directly interpretable performance measures of early warning systems. The missed event rate (the complement of sensitivity) is the proportion of all adverse events that are not prospectively detected by an early warning score threshold. The nonevent rate is the proportion of early warning score alarms that are not followed by an adverse event and is calculated as the complement of the positive predictive value (Table 2). An ideal early warning system has a missed event rate and nonevent rate both close to 0%.

The relevance of the missed event rate and the nonevent rate can be illustrated by evaluating the NEWS used in 70% of UK National Health Service institutions [2▪▪]. In a landmark analysis that led to the implementation of the NEWS, the performances of 34 scoring systems were compared using 198 755 observation sets from 35 585 acute medical admissions [19,28]. The NEWS was found to have the highest AUROC (0.873) to predict death within 24 h. A NEWS of 5 or more, which had an approximate sensitivity of 82% and specificity of 80% [19,28], was chosen as the threshold to activate a 'response by a clinician or team with competence in the assessment and treatment of acutely ill patients' [2▪▪]. This threshold had a missed event rate of 18% and a nonevent rate of 96% (at a 24-h mortality rate of 1%). A higher NEWS of 7 was chosen as the threshold necessitating an escalated emergency response by 'staff with critical care skills including airway management' [2▪▪]. Compared with the lower activation threshold of 5, the threshold at 7 had a missed event rate that increased from 18 to 33%, but despite substantially higher specificity (92 vs. 80%), the nonevent rate improved only marginally from 96 to 92%.

Similarly, in a cohort of 59 301 mixed medical and surgical admissions, the ViEWS was shown to have an AUROC of 0.78 to predict cardiac arrest within 24 h [20]. At a threshold score of 8, the missed event rate was 28% and the nonevent rate was 99% (at a cardiac arrest incidence of only 0.21%).

This shows that the discriminatory characteristics of early warning scores (quantified with the AUROC, sensitivity and specificity) do not necessarily reflect practical usefulness. The low incidence of adverse events causes a disparity between apparently good discrimination (high AUROC) and poor nonevent rates.

For the surgical population, this is illustrated by a study that investigated the performance of the admission ViEWS to predict subsequent in-hospital mortality among 30 485 consecutive surgical admissions to a Canadian regional hospital [14]. The AUROC for mortality within 48 h after admission was 0.89, but the 48-h mortality rate was only 0.04%. At a threshold score of 7, the missed event rate was 69% and the nonevent rate was 99%. In other words, despite a good AUROC, almost none of the alarms raised by this system were followed by an event within 48 h, whereas the large majority of actual adverse events were not detected.

Similarly, a recent study investigated the accuracy of the ViEWS to predict complications associated with end-organ damage or critical care admission among 552 postoperative patients [17]. The score had good to excellent ability to discriminate patients who developed a severe complication within 24 h from those who did not (AUCROC 0.90). But as the incidence of severe complications was relatively low (6.4%), only a high threshold of eight points was deemed by the authors to 'yield alerts at an acceptable rate' [17]. Practically, this means that a patient with a heart rate of 140/min, a respiratory rate of 35/min and a SBP of 105 mmHg – but with normal temperature, oxygen saturation and alertness – would not reach the alarm threshold (7 EWS points).

Figure 1 shows that between-study and between-score differences in nonevent rates depend less on discriminatory power than on the incidence of adverse events.

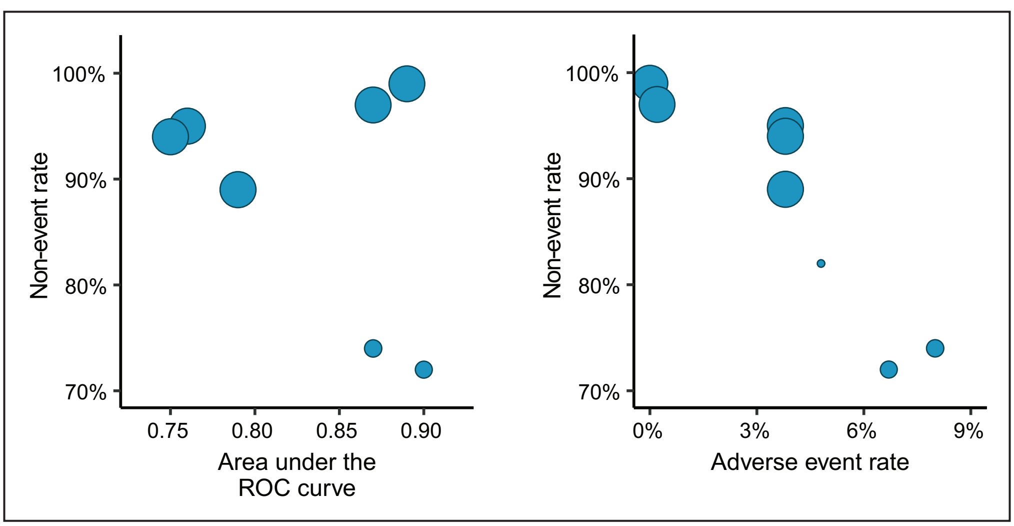

FIGURE 1. The relationship between the proportion of alarms not followed by an adverse event (the nonevent rate) and the area under the receiver operating characteristics curve (left) and the adverse event rate (right) for different early warning scores in different cohorts of surgical and perioperative patients. Data from Table 2. Circle sizes are proportional to the log of study sample sizes. It is shown that differences in nonevent rates are more related to the incidence of adverse events in the population (right) than to the discriminatory power of the early warning scores as summarized by the AUROC (left). AUROC, area under the receiver operating characteristics curve; eCART, electronic cardiac arrest triage; EEWS, expanded early warning score; MEWS, modified early warning score; NEWS, National Early Warning Score; ViEWS, Vitalpac Early Warning Score.

CONCLUSION

Early warning scores can provide a structured evaluation of ward patients who are at risk of clinical deterioration or imminent adverse events and who may benefit from specific interventions or closer monitoring. These scores have been developed and validated in large cohorts of medical patients, but recent studies show that the discriminating ability of the most widely used scores (ViEWS and NEWS) is similar in both medical and surgical populations. However, the advantage of broad usability in different (sub)populations necessarily leads to limited specificity. Patients admitted with acute leukemia and patients admitted after hip replacement surgery have very different sequences of events that lead to ICU admission, cardiac arrest or death. Yet in a given hospital, both patients are monitored using the same early warning score, which may explain why these scores do not perform well over time horizons longer than 24–48 h: These patients only share the final common pathway to cardiac arrest or death. When the incidence of adverse events is low, this lack of specificity causes missed event rates that are marginally acceptable and high or very high nonevent rates. In the future, machine-learning algorithms using many situation-tailored variables may significantly lengthen the screening horizon and improve the balance between missed events and alarms that are not followed by an event. Until such advanced screening systems become widely available, clinicians and administrators should take care not to overinterpret validations studies, as the implementation of these scores appears to be so complex that patient-oriented benefits are difficult to measure at best, or nonexistent at worst.

Acknowledgements

None.

Financial support and sponsorship

None.

Conflicts of interest

There are no conflicts of interest.

REFERENCES AND RECOMMENDED READING

Papers of particular interest, published within the annual period of review, have been highlighted as:
■ of special interest
■■ of outstanding interest

1. Sessler DI, Khanna AK. Perioperative myocardial injury and the contribution of hypotension. Intensive Care Med 2018; 44:811–822.
2. ■■ Royal College of Physicians. National Early Warning Score (NEWS) 2: standardising the assessment of acute-illness severity in the NHS. Updated report of a working party. 2017.
Comprehensive description of the second iteration of the National Early Warning Score (NEWS), the most widely and uniformly implemented early warning system.
3. Priestley G, Watson W, Rashidian A, *et al.* Introducing Critical Care Outreach: a ward-randomised trial of phased introduction in a general hospital. Intensive Care Med 2004; 30:1398–1404.
4. DeVita MA, Braithwaite RS, Mahidhara R, *et al.*, Medical Emergency Response Improvement Team (MERIT). Use of medical emergency team responses to reduce hospital cardiopulmonary arrests. Qual Saf Healthcare 2004; 13:251–254.
5. Bellomo R, Goldsmith D, Uchino S, *et al.* Prospective controlled trial of effect of medical emergency team on postoperative morbidity and mortality rates. Crit Care Med 2004; 32:916–921.
6. Buist MD, Moore GE, Bernard SA, *et al.* Effects of a medical emergency team on reduction of incidence of and mortality from unexpected cardiac arrests in hospital: preliminary study. BMJ 2002; 324:387–390.
7. Hillman K, Chen J, Cretikos M, *et al.* Introduction of the medical emergency team (MET) system: a cluster-randomised controlled trial. Lancet (London, England) 2005; 365:2091–2097.
8. McGaughey J, Alderdice F, Fowler R, *et al.* Outreach and Early Warning Systems (EWS) for the prevention of intensive care admission and death of critically ill adult patients on general hospital wards. Cochrane database Syst Rev 2007; 3:CD005529.
9. Alam N, Hobbelink ELL, van Tienhoven AJJ, *et al.* The impact of the use of the Early Warning Score (EWS) on patient outcomes: a systematic review. Resuscitation 2014; 85:587–594.
10. ■■ Parshuram CS, Dryden-Palmer K, Farrell C, *et al.* Effect of a pediatric early warning system on all-cause mortality in hospitalized pediatric patients: the EPOCH Randomized Clinical Trial. JAMA 2018; 319:1002–1012.
Recent large multinational cluster RCT does not show a beneficial effect of a pediatric early warning system on all-cause hospital mortality.
11. Rothschild JM, Gandara E, Woolf S, *et al.* Single-parameter early warning criteria to predict life-threatening adverse events. J Patient Saf 2010; 6:97–101.
12. ■■ Bartkowiak B, Snyder AM, Benjamin A, *et al.* Validating the Electronic Cardiac Arrest Risk Triage (eCART) Score for Risk Stratification of Surgical Inpatients in the Postoperative Setting: retrospective cohort study. Ann Surg 2018. [Epub ahead of print]
The predictive value of the MEWS, NEWS and eCART scores are tested in the largest postoperative cohort to date (32 537 patients).
13. Smith MEB, Chiovaro JC, O'Neil M, *et al.* Early warning system scores for clinical deterioration in hospitalized patients: a systematic review. Ann Am Thorac Soc 2014; 11:1454–1465.
14. Kellett J, Kim A. Validation of an abbreviated Vitalpac™ Early Warning Score (ViEWS) in 75,419 consecutive admissions to a Canadian regional hospital. Resuscitation 2012; 83:297–302.
15. Kovacs C, Jarvis SW, Prytherch DR, *et al.* Comparison of the National Early Warning Score in nonelective medical and surgical patients. Br J Surg 2016; 103:1385–1393.
16. Smith T, Den Hartog D, Moerman T, *et al.* Accuracy of an expanded early warning score for patients in general and trauma surgery wards. Br J Surg 2012; 99:192–197.
17. Hollis RH, Graham LA, Lazenby JP, *et al.* A role for the early warning score in early identification of critical postoperative complications. Ann Surg 2016; 263:918–923.
18. Gardner-Thorpe J, Love N, Wrightson J, *et al.* The value of Modified Early Warning Score (MEWS) in surgical in-patients: a prospective observational study. Ann R Coll Surg Engl 2006; 88:571–575.
19. Smith GB, Prytherch DR, Meredith P, *et al.* The ability of the National Early Warning Score (NEWS) to discriminate patients at risk of early cardiac arrest, unanticipated intensive care unit admission, and death. Resuscitation 2013; 84:465–470.
20. Churpek MM, Yuen TC, Park SY, *et al.* Using electronic health record data to develop and validate a prediction model for adverse outcomes in the wards. Crit Care Med 2014; 42:841–848.
21. Churpek MM, Yuen TC, Winslow C, *et al.* Multicenter comparison of machine learning methods and conventional regression for predicting clinical deterioration on the wards. Crit Care Med 2016; 44:368–374.
22. Bellomo R, Ackerman M, Bailey M, *et al.*, Vital Signs to Identify, Target, and Assess Level of Care Study (VITAL Care Study) Investigators. A controlled trial of electronic automated advisory vital signs monitoring in general hospital wards. Crit Care Med 2012; 40:2349–2361.
23. ■ Blankush JM, Freeman R, McIlvaine J, *et al.* Implementation of a novel postoperative monitoring system using automated Modified Early Warning Scores (MEWS) incorporating end-tidal capnography. J Clin Monit Comput 2017; 31:1081–1092.
Feasibility study shows that the implementation of an automated early score device including end-tidal capnography leads to very high alarm rates.

24. Santamaria J, Moran J, Reid D. Increasing the number of medical emergency
■ calls does not improve hospital mortality. Crit Care Med 2018; 46:1063–1069.

This observational cohort study including more than three million patients showed that, after adjustment for patient-related factors, increased calling rates to emergency teams were not associated with reduced mortality rates.

25. Concord Medical Emergency Team (MET) Incidents Study Investigators. Cheung W, Sahai V, *et al.* Incidents resulting from staff leaving normal duties to attend medical emergency team calls. Med J Aust 2014; 201:528–531.

26. Pedersen NE, Rasmussen LS, Petersen JA, *et al.* A critical assessment of early
■ warning score records in 168,000 patients. J Clin Monit Comput 2018; 32:109–116.

In-depth analysis of recorded early warning score parameters shows that care-givers are prone to record values just below warning thresholds.

27. Petersen JA, Mackel R, Antonsen K, Rasmussen LS. Serious adverse events in a hospital using early warning score - what went wrong? Resuscitation 2014; 85:1699–1703.

28. Prytherch DR, Smith GB, Schmidt PE, Featherstone PI. ViEWS–towards a national early warning score for detecting adult inpatient deterioration. Resuscitation 2010; 81:932–937.