HHS Public Access

Author manuscript

Resuscitation. Author manuscript; available in PMC 2021 February 20.

Predicting clinical deterioration with Q-ADDS compared to NEWS, Between the Flags, and eCART track and trigger tools

Victoria Campbell[a,*], **Roger Conway**[b], **Kyle Carey**[c], **Khoa Tran**[d], **Adam Visser**[e], **Shaune Gifford**[f], **Mia McLanders**[f,g], **Dana Edelson**[c], **Matthew Churpek**[c]

[a]Intensive Care Unit, Sunshine Coast University Hospital, Queensland, Australia

[b]Deteriorating Patient Response, Sunshine Coast University Hospital, Queensland, Australia

[c]Department of Medicine, University of Chicago, Chicago, IL, United States

[d]Logan Hospital, Queensland, Australia

[e]Intensive Care Unit, Toowoomba Hospital, Queensland, Australia

[f]Patient Safety and Quality Improvement Service, Clinical Excellence Queensland, Australia

[g]School of Psychology, The University of Queensland, Australia

Abstract

Background: Early warning tools have been widely implemented without evidence to guide (a) recognition and (b) response team expertise optimisation. With growing databases from MET-calls and digital hospitals, we now have access to guiding information. The Queensland Adult-Deterioration-Detection-System (Q-ADDS) is widely used and requires validation.

Aim: Compare the accuracy of Q-ADDS to National Early Warning Score (NEWS), Between-the-Flags (BTF) and the electronic Cardiac Arrest Risk Triage Score (eCART)).

Methods: Data from the Chicago University hospital database were used. Clinical deterioration was defined as unplanned admission to ICU or death. Currently used NEWS, BTF and eCART trigger thresholds were compared with a clinically endorsed Q-ADDS variant.

Results: Of 224,912 admissions, 11,706 (5%) experienced clinical deterioration. Q-ADDS (AUC 0.71) and NEWS (AUC 0.72) had similar predictive accuracy, BTF (AUC 0.64) had the lowest, and eCART (AUC 0.76) the highest. Early warning alert (advising ward MO review) had similar

NPV (99.2–99.3%), for all the four tools however sensitivity varied (%: Q-ADDS = 47/NEWS = 49/BTF = 66/eCART = 40), as did alerting rate (% vitals sets: Q-ADDS = 1.4/NEWS = 3.5/BTF = 4.1/eCART = 3.4). MET alert (advising MET/critical-care review) had similar NPV for all the four tools (99.1–99.2%), however sensitivity varied (%: Q-ADDS = 14/NEWS = 24/BTF = 19/eCART = 29), as did MET alerting rate (%: Q-ADDS = 1.4/NEWS = 3.5/BTF = 4.1/eCART = 3.4). High-severity alert (advising advanced ward review, Q-ADDS only): NPV = 99.1%, sensitivity = 26%, alerting rate = 3.5%.

Conclusion: The accuracy of Q-ADDS is comparable to NEWS, and higher than BTF, with eCART being the most accurate. Q-ADDS provides an additional high-severity ward alert, and generated significantly fewer MET alerts. Impacts of increased ward awareness and fewer MET alerts on actual MET call numbers and patient outcomes requires further evaluation.

Keywords

Clinical deterioration; Rapid response system; Track and trigger system; Early warning scores; Q-ADDS; NEWS; Between the flags; eCART; Electronic cardiac arrest triage score; AUC; Area under the receiver operating characteristic curve; Sensitivity and specificity; Predictive value

---

Introduction

Unrecognised or poorly managed patient deterioration is a widely acknowledged problem in hospital-based health care.[1,2] Derangements in physiological parameters are often present in the 24 h leading up to an event,[3] so early recognition and intervention provides an opportunity to reduce preventable adverse outcomes.

Physiological track and trigger systems support early recognition and guide appropriate escalation of care.[4,5] The first of these were bedside charts introduced to promote the timely presence of clinical help when vital signs became abnormal.[6–8] Since then, national and international accreditation and safety organisations have recommended that track and trigger systems be routinely implemented in clinical practice.[2,9–11] As a result, these systems have evolved rapidly and have not been implemented with the same methodological rigour of other clinical interventions. Moreover, track and trigger systems vary significantly. Most have tiered escalation protocols, with early/mild derangement thresholds for ward-level escalation and more severe derangement thresholds for medical emergency team (MET) or critical care escalation. However, they differ in design, algorithm thresholds, and levels of escalation., Single track and trigger systems, such as the Australian Between the Flags (BTF), rely on a single vital sign parameter threshold breach to trigger escalation of care,[12] whereas aggregate scoring systems, such as the National Early Warning Score (NEWS), escalate based on a cumulative score generated when multiple parameter derangement occurs.[13] Evidence suggests that systems with aggregate scoring have better predictive validity,[14,15] which has implications for patient safety and clinical resourcing.

The Queensland Adult Deterioration Detection System (Q-ADDS) is a track and trigger system used widely throughout Australia and some developing countries. It has a number of important features. It is a "combination" system, with MET alert thresholds for both an aggregate score across multiple parameters (like NEWS), and for severe vital sign

derangement on a single parameter (like BTF), which provides a safety net for incomplete vital sign sets and severe derangement on a single vital sign. It adds a quantified oxygen component to the vital sign thresholds, which aligns with recent Australian Thoracic Society of Australia and New Zealand guidelines[16] that recommend escalation based on oxygen requirement rather than oxygen saturation levels. Q-ADDS has been designed according to human factor principles and heuristically evaluated throughout each stage of its development.[17–23] Additionally, the escalation guidance provides 2 tiers of ward level visibility and escalation (mild derangement (score 4–5), and high-severity (score 6–7)), rather than just mild or MET thresholds. The underlying algorithm of Q-ADDS, however, has not been validated for predictive accuracy.

Our aim was to compare the accuracy and alerting efficiency of Q-ADDS to other commonly used early warning track and trigger tools and to eCART, an advanced digital algorithm tool that includes patient demographics, vital signs, and laboratory values. Additionally, because of the lack of evidence to guide optimal MET trigger thresholds, we performed a direct comparison of two low systolic blood pressure (SBP) and two high respiratory rate (RR) thresholds. Low SBP is of particular interest, as it is a frequent cause of MET calls where critical care input is not required, and is thus a potential opportunity for improved MET system efficiency.[24]

Methods

Study cohort and data collection

Vital sign data from all adult patients (age 18 and older) hospitalized on the medical-surgical wards at the University of Chicago from November 2008 to November 2018 were included in the study. Patients were excluded if they did not have at least one full set of numerical vital signs and a measurement of the amount of oxygen delivered at least once during their ward stay (e.g., "room air"). The study was approved by the University of Chicago's Institutional Review Board (IRB #17–1342), and a waiver of consent was granted. Patient demographics, vital signs, laboratory values, location data, and discharge disposition were collected from the electronic health record (Epic; Verona, WI) through the University's Clinical Research Data Warehouse.

Outcomes

The primary outcome of the study was clinical deterioration on the wards, which was defined as a direct ward to intensive care unit (ICU) transfer or death on the ward, in line with prior work.[25] ICU transfer and death were analysed separately as secondary outcomes. Patients who were transferred from the wards to another location (e.g., the operating room) prior to their ICU transfer were not counted as having an event.

Early warning tools

Four early warning tools were compared: Q-ADDS, BTF, NEWS and eCART. BTF, NEWS and eCART provide bilevel notification/escalation alerting (early/mild derangement thresholds and MET thresholds), whereas Q-ADDS has an additional alert threshold (high-severity) advising advanced ward level review. Additionally, four different variations of the

Q-ADDS (comparing two thresholds of low SBP and of high RR) were investigated (Supplementary Fig. 1). Q-ADDS Variant 1 has been used for the primary analysis in this study since this is the version planned for prospective clinical implementation. Clinical alert/ escalation thresholds for each tool are presented in Table 1.

In prior studies,[15] NEWS was calculated on cumulative score only, based on prior publications and clinical use in the United States.[14] However, the UK National Institute for Health and Care Excellence (NICE) guidelines[2] mandate that for NEWS (and the more recent NEWS2), an early/ward level escalation should occur for a cumulative score of 5, and also for any single parameter score of 3 alone, so this has been included as an additional analysis (labelled NEWS+). The version of eCART used in this project is a previously published random forest algorithm that includes patient characteristics, vital signs, and laboratory values as variables. Scores were calculated at each time point that a patient had a full set of numerical vital signs needed for calculation, with the exception of oxygen delivery, mental status, and laboratory values (for eCART), which were carried forward from prior values if missing.

Statistical analysis

Patient characteristics and score distributions were presented as counts, percentages, and median with inter-quartile range (IQR) as appropriate. For each tool sensitivity, specificity, and the area under the receiver operating characteristic curve (AUC), negative predictive values (NPV) and positive predictive values (PPV) were calculated by defining all observations within 24 h leading up to an outcome as having the event and all other observations as not having the event, as per prior work.[15] AUCs were compared between tools using the method of DeLong.[26] The time-to-event was the time difference between when a score first reached a particular threshold and the outcome event was also calculated to determine the lead time of detection. Finally, an early warning score efficiency curve was created graphing the percentage of positive alerts (vital sign sets that breached an alert/ escalation threshold) on the *y*-axis against the sensitivity on the *x*-axis (see Figs. 1 and 2). All analyses were performed using Stata version 15.1 (Stata Corps; College Station, Texas), and two-sided $P < 0.05$ was considered statistically significant.

Results

Study cohort

A total of 226,981 adult admissions on the wards occurred during the study period, with 2069 excluded due to missing vital signs and/or oxygen delivery values, which left 224,912 admissions (5,349,995 vital sign sets) in the final cohort. Of these admissions, 11,706 (5.2%) experienced at least one clinical deterioration outcome. The study cohort was 58% female, 56% African-American, and the median age was 57 (IQR 41 –69) years (see Table 2). Fifteen percent of the study cohort had an ICU admission at least once during their hospital stay, which could have occurred before or after their ward stay, and the full cohort had an in-hospital mortality of 1%.

Early warning tool comparisons

Accuracy—The AUC distributions of the different early warning tools are shown in Table 3. For the combined outcome, eCART had the highest AUC for predicting an event in the next 24 h (0.79), followed by NEWS (0.73), Q-ADDS (0.72 for all four variants (Supplementary Table 1)), and BTF (0.64) (all pairwise comparisons using only Q-ADDS Variant 1 with $P < 0.05$). Findings for predicting ICU transfer (eCART 0.78, NEWS 0.72, Q-ADDS 0.71, BTF 0.63) and for predicting death (eCART 0.92, NEWS 0.88, Q-ADDS 0.87, BTF 0.78), had a similar order.

The PPV, NPV, sensitivity and specificity for predicting an outcome in the next 24 h for the different alerting thresholds for each tool are shown in Table 4. The early alerting threshold (advising ward MO review) had similar NPV for all 4 tools (eCART/NEWS 99.2% vs Q-ADDS/BTF 99.3%), however PPV varied (Q-ADDS = 3.9%, NEWS + = 4.4%, BTF = 1.8%, eCART = 7.2%). Sensitivity also varied: NEWS (44%), NEWS+ (49%), Q-ADDS (47%), and eCART (40%), with BTF being the most sensitive (66%).

MET alert thresholds (advising MET/critical-care review) had a similar NPV for all four tools (eCART/NEWS 99.1% vs Q-ADDS/BTF 99.2%), however PPV varied (Q-ADDS = 12.4%, NEWS = 8.4%, BTF = 5.7%, eCART = 10.2%), as did sensitivity (Q-ADDS = 14%, NEWS = 24%, BTF = 19%, eCART = 29%). The high-severity alert (advising advanced ward review, Q-ADDS only) had an NPV of 99.1%, PPV of 8.1%, and a sensitivity of 26%.

Time-to-event—The early alerting threshold time-to-event was similar for NEWS, NEWS +, Q-ADDS and eCART (46/48/45/44 h respectively), where BTF was earlier (53 h). The MET alerting time-to-event for NEWS, Q-ADDS, BTF and eCART was 36/20/36/35 h respectively.

Alerting efficiency—The tools varied in alerting efficiency (Table 4). For early derangement alerts (when corrected by excluding MET alerts from the total) eCART had the highest efficiency (5.2% of vital sign sets) and BTF the lowest efficiency (41% of vital sign sets), with NEWS+ and Q-ADDS being similar (13% of vital sign sets). For MET alerts, however, Q-ADDS variant 1 was the most efficient (1.4% of vital sign sets), compared to NEWS+, eCART and BTF (3.5%, 3.4%, 4.1% respectively). Q-ADDS high-severity ward alert occurred on 7% of vital sign sets.

The four Q-ADDS variants compared two different high RR and low SBP MET thresholds. The overall accuracy of the different Q-ADDS variants were identical, with an AUC of 0.72 for the combined outcome, 0.71 for ICU transfer, and 0.87 for death. Changing the high RR MET thresholds (variants 1,3) did not alter the NPV, PPV or sensitivity, however lowering the SBP MET threshold from < 90 mmHg to <80 mmHg alone reduced MET alerts by 45% (18–10% of encounters) with a reduction in sensitivity from 17% to 14%, but no impact on NPV (99.0%/99.0%).

Discussion

The results demonstrate that the Q-ADDS algorithm has a similar discriminative accuracy for either intensive care unit transfer or death to NEWS and higher accuracy than BTF, with the advanced digital algorithm eCART proving more accurate than the vital-sign based tools. This relationship was similar to the previous comparison study.[15] It is likely that additional parameters improve accuracy rather than the manipulation of parameter thresholds. This is supported by the superior accuracy from the addition of patient demographics and laboratory results in eCART, versus the similarity between AUCs of the four variants of Q-ADDS, where vital sign thresholds were manipulated but the number of parameters remained constant.

Early detection provides an opportunity for early intervention, and therefore prevention, of severe physiological derangement and subsequent adverse outcomes, but only if that detection is appropriately acted upon. Good clinical acumen and reliable recognition/alerting tools are key. These tools have to balance under-alerting, which may miss at-risk patients, and over-alerting, which can lead to alarm fatigue, unnecessary responder workload or "wrong place, wrong time" clinical situations. Alert/escalation fatigue can lead to failure to escalate or respond and can also lead to medical orders for patient specific modifications to alerting thresholds which reduces the sensitivity of the tools. To date there is no published evidence to define the optimal balance.

All four early warning tools included in this study have early escalation thresholds, where a ward-level clinical review is required. At the applied escalation alerting thresholds, this Q-ADDS algorithm has a similar early ward level alerting accuracy and efficiency to NEWS, but generates almost 70% fewer alerts than BTF. eCART, again, was superior. Additionally, Q-ADDS provides a high-severity ward level alert (high-severity score 6), with higher sensitivity than NEWS or BTF MET thresholds. Escalation for single vital sign derangements may have serious implications for patient safety and resource allocation. Early alerts for ward-level review for any isolated score of 3 are used with NEWS in the UK, and although this improves sensitivity, it is at the cost of a 30% increase in ward medical officer escalation alerts–with no improvement in accuracy.

With regard to MET alerting, at the applied thresholds, Q-ADDS was the most efficient, generating >50% fewer MET alerts than NEWS or eCART, and >60% fewer than BTF. This was achieved with little difference in the NPV (an indicator of missed alerting opportunities in studies such as this where the outcome prevalence is low). Q-ADDS does have the lowest sensitivity, which raises concerns for missed opportunities to intervene. However, it is important to distinguish between validation of performance of the early warning tool algorithm (recognition), and performance of an early warning "system" (multi-level recognition, escalation and response) which is much more complex, and requires large-scale outcome-based research methodology.

Q-ADDS has a number of features aimed at redirecting services rather than reducing sensitivity. Stratified early warning thresholds facilitate prevention of deterioration, rather than response to it. A redirection of response expertise to advanced members of the ward

team (high-severity ward level alert) allows further opportunity for ward level awareness and escalation prior to MET/external escalation. This redirection is supported both by these results (sensitivity 26%) and previously published MET call outcome data.[27] Low SBP triggers approximately 30% of all MET calls,[27] most being isolated hypotension, and less than 4% of those calls result in more than ward level intervention or ICU transfer.[24] These data show that reducing the low SBP MET threshold from 90 mmHg to 80 mmHg will result in a 40% reduction in MET alerts, so improving specificity for this threshold alone will reduce the strain on critical care resources. A MET attendance will still be recommended for any SBP <80 mmHg, hypotension associated with additional derangements, or as triggered by the "clinician worried" criterion.

To date there is no guiding evidence for the optimal lead time for early or late (MET) thresholds. Alerting/escalation must occur early enough for clinicians to respond and prevent adverse outcomes, but not so early that the alerts lead to unnecessary interventions or alarm fatigue. The early threshold lead times for Q-ADDS, NEWS and BTF are similar (40–50 h). For MET escalation, Q-ADDS had a lead time of 20 h, compared to 36 h for NEWS and BTF, but provides the additional high-severity ward alert at 31 h. It could be argued that earlier alerting should translate into earlier intervention and therefore better patient outcomes, however, who is best to provide that early response (i.e., medical treating team or critical care team) and when the ideal time is for critical care input, remains unclear. Critical care input greater than 12–24 h prior to an adverse event has questionable validity and significant resource implications. Optimal early escalation and MET escalation threshold lead times requires further exploration.

As described above for low SBP, a direct comparison was made for the high RR MET threshold. High respiratory rate is the most accurate of all the parameters in predicting deterioration but lacks evidence as to the ideal threshold in a combination single/multi-track and trigger system. Therefore, we compared two high RR thresholds (36 vs 30 breaths per minute) but there was little difference between the two.

It is expected that advanced algorithms will provide safer healthcare in the future. As technology advances, electronic health platforms will support the use of statistically derived algorithms like eCART, which combines patient demographics and laboratory values in addition to vital sign monitoring to provide a real-time scoring tool. eCART, using the alerting thresholds described, was significantly more accurate than the paper-based tools, with the advantage being particularly for early derangement detection. The Australian health system is still in the process of transitioning to electronic health records, and hospitals – particularly those in regional and remote areas – are unlikely to have this technology imminently and are reliant on paper-based tools.

This study has several strengths. This is the first published study to (1) evaluate the underlying algorithm of Q-ADDS compared to other commonly used tools, (2) use the inclusion of quantified oxygen requirement as a parameter, and (3) compare different thresholds for the key variables of high RR and SBP. Patient data were analysed on both complete vital sign sets and on individual vital signs with similar accuracy. The analysis

included data from over 220,000 patients over a 10-year period. The data reflects what is typically found in similar comparison studies.

Our findings are limited by this being a single-site study. This was unavoidable, as although a similar prior study included multiple sites, only one of these sites reliably recorded oxygen delivery.[15] Nevertheless, the single-site data had similar characteristics when directly compared to the multi-site dataset. These data do not include calls based on other escalation criteria, such as clinician worry, patient confusion, or patients with modified alerting thresholds (available for BTF, Q-ADDS), or limitations of medical therapy orders, and therefore cannot be directly translated into impact on actual call numbers. The validation is statistical, and is based on the generation of alerts only. What action follows these alerts (the early warning system) is a critical component of the impact on patient outcomes. Implementation data would be required to evaluate associated improvement in patient outcomes. Additionally, the applicability of the ICU transfer outcome data to the Australian population must be interpreted in the context of the differences in patient populations, ICU admission and discharge policy between the USA and Australia. These limitations must be taken into consideration when interpreting the results in relation to Australian hospitalised patients.

Conclusion

The accuracy of the Q-ADDS algorithm was found to be superior to BTF and comparable to NEWS for predicting clinical deterioration, with eCART, the advanced digital monitoring tool being the most accurate. The Q-ADDS algorithm generated significantly fewer MET alerts than BTF, NEWS or eCART. The effect of increased ward awareness and intervention responsibility, and fewer MET alerts on actual MET call numbers and patient outcomes requires further evaluation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

REFERENCES

1. Goldhill D, McNarry A, Mandersloot G, et al. A physiologically-based early warning score for ward patients: the association between score and outcome. J Anaesth 2005;60:547–53.

2. Centre for Clinical Practice at NICE. Acutely ill patients in hospital: recognition of and response to acute illness in adults in hospital. 2007.

3. Goldhill D, White S, Sumner A. Physiological values and procedures in the 24 h before ICU admission from the ward. J Anaesth 1999;54:529–34.

4. Downey CL, Tahir W, Randell R, et al. Strengths and limitations of early warning scores: a systematic review and narrative synthesis. Int J Nurs Stud 2017;76:106–19. [PubMed: 28950188]

5. Gerry S, Birks J, Bonnici T, et al. Early warning scores for detecting deterioration in adult hospital patients: a systematic review protocol. BMJ Open 20177: e019268-e68.

6. Lee A, Bishop G, Hillman K, et al. The medical emergency team. J Anaesth Crit Care 1995;23:183–6.

7. Morgan R, Williams F, Wright M. An early warning scoring system for detecting developing critical illness. Clin Intensive Care 1997;8:100.

8. Morgan RJM, Wright MM. In defence of early warning scores. Br J Anaesth 2007;99:747–8. [PubMed: 17933804]

9. ANZICS-CORE MET dose Investigators. Rapid Response Team composition, resourcing and calling criteria in Australia. Resuscitation 2012;83:563–7. [PubMed: 22067975]

10. National consensus statement: essential elements for recognising and responding to acute physiological deterioration second edition. Australian Commission on Safety and Quality in Health Care, 2017.

11. Winters BD, Weaver SJ, Pfoh ER, et al. Rapid-response systems as a patient safety strategy: a systematic review. Ann Internal Med 2013;158:417–25. [PubMed: 23460099]

12. Hughes C, Pain C, Braithwaite J, et al. 'Between the flags': implementing a rapid response system at scale. BMJ Quality Safety 2014;23:714–7.

13. Jarvis S, Kovacs C, Briggs J, et al. Aggregate National Early Warning Score (NEWS) values are more important than high scores for a single vital signs parameter for discriminating the risk of adverse outcomes. Resuscitation 2015;87:75–80. [PubMed: 25433295]

14. 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;84:465–70. [PubMed: 23295778]

15. Green M, Lander H, Snyder A, et al. Comparison of the Between the Flags calling criteria to the MEWS, NEWS and the electronic Cardiac Arrest Risk Triage (eCART) score for the identification of deteriorating ward patients. Resuscitation 2018;123:86–91. [PubMed: 29169912]

16. Beasley R, Chien J, Douglas J, et al. Thoracic Society of Australia and New Zealand oxygen guidelines for acute oxygen use in adults: 'Swimming between the flags'. Respirology 2015;20:1182–91. [PubMed: 26486092]

17. Christofidis MJ, Hill A, Horswill MS, et al. A human factors approach to observation chart design can trump health professionals' prior chart experience. Resuscitation 2013;84:657–65. [PubMed: 23041534]

18. Cornish L, Hill A, Horswill MS, et al. Eye-tracking reveals how observation chart design features affect the detection of patient deterioration: an experimental study. Appl Ergonom 2019;75:230–42.

19. Horswill MS, Preece MH, Hill A, et al. Human factors research regarding observation charts: research project overview. Australian Commission on Safety Quality in Healthcare; 2010.

20. Preece MH, Hill A, Horswill MS, et al. Supporting the detection of patient deterioration: observation chart design affects the recognition of abnormal vital signs. Resuscitation 2012;83:1111–8. [PubMed: 22353643]

21. Christofidis M, Horswill M, Hill A, et al. Task analysis and heuristic analysis of insulin charts: Final report prepared for the Australian Commission on Safety and Quality in Health Care: 2 2 2012.

22. Christofidis MJ, Hill A, Horswill MS, et al. Less is more: the design of early-warning scoring systems affects the speed and accuracy of scoring. J Adv Nurs 2015;71:1573–86. [PubMed: 25645881]

23. Christofidis MJ, Hill A, Horswill MS, et al. Observation chart design features affect the detection of patient deterioration: a systematic experimental evaluation. J Adv Nurs 2016;72:158–72. [PubMed: 26556775]

24. Bergmeir C, Bilgrami I, Bain C, Webb GI, Orosz J, Pilcher D. Designing a more efficient, effective and safe Medical Emergency Team (MET) service using data analysis. PLoS One 2017;12(12): e0188688.

25. Churpek MM, Yuen TC, Winslow C, et al. Multicenter development and validation of a risk stratification tool for ward patients. Am J Resp Crit Care Med 2014;190:649–55. [PubMed: 25089847]

26. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44:837–45. [PubMed: 3203132]

27. Missen K, Porter JE, Raymond A, et al. Adult Deterioration Detection System (ADDS): an evaluation of the impact on MET and Code blue activations in a regional healthcare service. Collegian 2018;25:157–61.
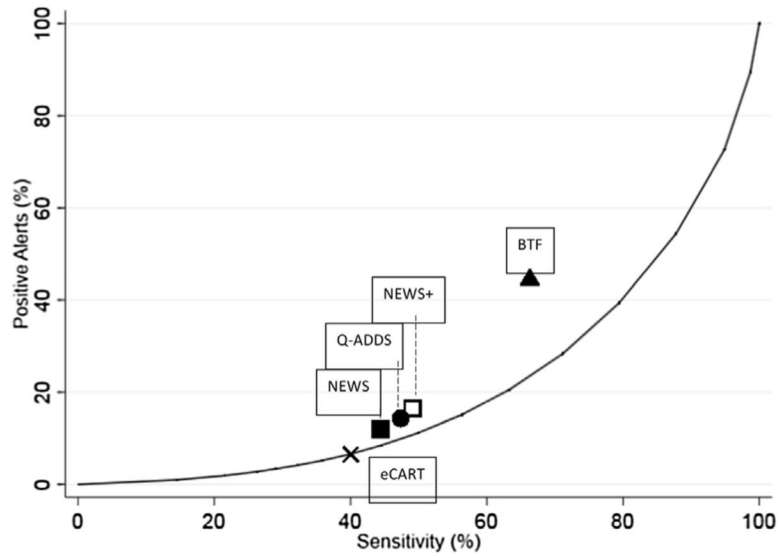
Fig. 1 –.
Early/ward alert efficiency curve. This curve illustrates the sensitivity of different alerting thresholds for whether an event occurs in the next 24 h (*x*-axis) vs. the percentage of alerts that are above the different thresholds (*y*-axis).
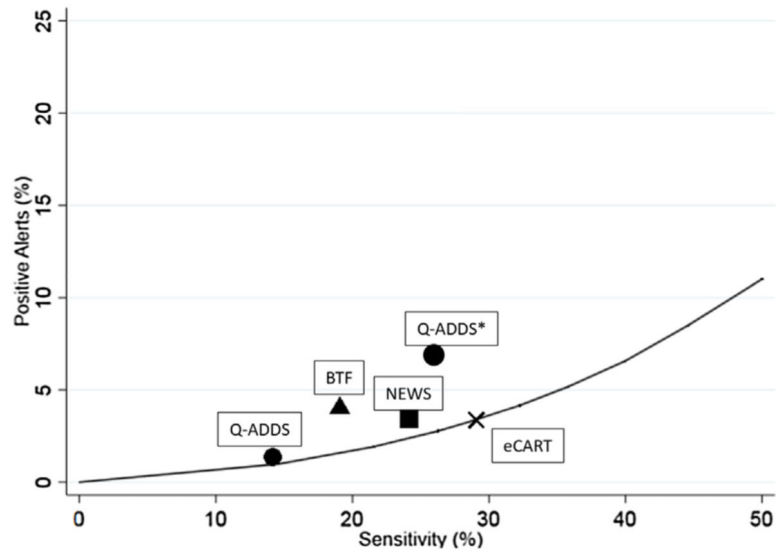
Fig. 2 –.

MET alert efficiency curve. This curve illustrates the sensitivity of different MET alerting thresholds for whether an event occurs in the next 24 h (*x*-axis) vs. the percentage of alerts that are above the different thresholds (*y*-axis). Note the *y* axis is reduced to 25% for improved visibility. *Q-ADDS high-severity ward escalation alert (Q-ADDS score 6).

Table 1 –

Clinically used alert/escalation thresholds.

	Early/mild derangement alert (escalation to medical officer)	Higher severity alert (escalation to advanced medical officer)	Severe derangement alert (escalation to Medical Emergency team)
Q-ADDS	4–5	6–7	8 or any single parameter in the purple ("E") zone
NEWS	5–6		7
NEWS+ (+ single parameter)	5–6 or any single parameter = 3 but score <5		7
BTF	Not applicable		Not applicable
eCART	93–96		97

Table 2 –

Patient characteristics.

Demographic	Value	All encounters ($n = 224,912$)
Age	Median (IQR)	57 (41, 69)
Sex	Female	129,427 (58%)
	Male	94,787 (42%)
	Unknown	13 (0.01%)
Race	White	83,444 (37%)
	Black/African-American	124,938 (56%)
	Asian/Mideast Indian	4944 (2.2%)
	American Indian or Alaska Native	472 (0.2%)
	Native Hawaiian or other Pacific Islander	290 (0.1%)
	More than one Race	4179 (1.9%)
	Unknown	6645 (3.0%)
Ethnicity	Hispanic	10,910 (4.9%)
	Not Hispanic	206,613 (92%)
	Unknown	7389 (3.3%)
Encounter LOS (hours)	Median (IQR)	77 (44, 143)
ICU ever	N (%)	34,687 (15%)
Died ever	N (%)	2183 (1.0%)

Columns left to right: Score | Median (IQR) | ICU transfer AUC | Death AUC | Combined outcome AUC

Table 3 –

Distributions and AUCs for included early warning tool algorithms.

Score	Median (IQR)	ICU transfer AUC	Death AUC	Combined outcome AUC
NEWS	2 (1, 3)	0.72 (0.72, 0.73)	0.88 (0.88, 0.89)	0.73 (0.73, 0.73)
NEWS+	2 (1, 3)	0.72 (0.72, 0.72)	0.88 (0.87, 0.89)	0.73 (0.73, 0.73)
BTF	0 (0, 1)	0.63 (0.63, 0.63)	0.78 (0.77, 0.78)	0.64 (0.64, 0.64)
eCART	46 (11, 72)	0.78 (0.78, 0.78)	0.92 (0.91, 0.92)	0.79 (0.79, 0.79)
QADDS variation #1	2 (1, 3)	0.71 (0.71, 0.71)	0.87 (0.86, 0.88)	0.72 (0.72, 0.72)
QADDS variation #2	2 (1, 3)	0.71 (0.71, 0.71)	0.87 (0.86, 0.87)	0.72 (0.72, 0.72)
QADDS variation #3	2 (1, 3)	0.71 (0.71, 0.72)	0.87 (0.86, 0.88)	0.72 (0.72, 0.72)
QADDS variation #4	2 (1, 3)	0.71 (0.71, 0.71)	0.87 (0.86, 0.87)	0.72 (0.72, 0.72)

Table 4 –

Characteristics of different thresholds of the studied scoring systems using whether an event occurred within the next 24 h of a vital sign set.

Tool	Score	Encounters (%)[a]	Scores above threshold (%)	Median (IQR) hours to outcome	Sensitivity	Specificity
Early derangement threshold						
Q-ADDS	>=4	106,313 (47%)	778,952 (14%)	45 (14, 119)	47%	86%
NEWS	>=5	99,711 (44%)	650,041 (12%)	46 (14, 119)	44%	88%
NEWS (+ single parameter)	>=5	129,644 (58%)	897,913 (16%)	48 (15, 124)	49%	84%
BTF	Yellow	204,790 (91%)	2,418,036 (44%)	53 (19, 131)	66%	56%
eCART	>=92	71,688 (32%)	463,403 (9%)	44 (13, 119)	45%	92%
Higher severity threshold						
Q-ADDS	>=6	44,240 (20%)		31 (7, 97)	26%	96%
MET threshold						
Q-ADDS	>=8	21,535 (10%)	74,142 (1%)	20 (4, 79)	14%	99%
NEWS	>=7	43,542 (19%)	186799 (3%)	36 (9, 102)	24%	97%
BTF	Red	61,216 (27%)	217,210 (4%)	36 (8, 107)	19%	96%
eCART	>=97	34,744 (15%)	184,195 (3%)	35 (9, 104)	29%	97%

[a]Encounters denotes the number (%) of encounters that reached each score threshold at least once on the wards.