



OPEN

A randomized controlled trial of artificial intelligence-based analytics for clinical deterioration

Jessica Keim-Malpass^{1,2,6}✉, Sarah J. Ratcliffe³, Matthew T. Clark⁴, Katy N. Krahn^{1,5}, Oliver J. Monfredi⁵, Susan Hamil^{1,5}, Gholamreza Yousefvand^{1,5}, Marieke K. Jones³, Ashley Nelson⁵, Liza P. Moorman⁴, J. Randall Moorman^{1,5} & Jamieson M. Bourque^{1,5}

This pragmatic randomized controlled trial aimed to assess the effect of a passive display of artificial intelligence (AI)-based predictive analytics on hours free of clinical deterioration events among medical and surgical patients in an acute care cardiology medical-surgical ward. 10,422 inpatient visits were randomly assigned by cluster to the intervention group of a display of risk trajectories or to a control group of usual medical care. The trial was undertaken on an 85-bed inpatient cardiology and cardiac surgery ward of an academic hospital with a substantial implementation and education plan. This was a passive display with no specific response mandated. The primary analysis compared events of clinical deterioration (death, emergent ICU transfer, emergent endotracheal intubation, cardiac arrest, or emergent surgery) and compared mortality 21 days after admission. Patients with a large spike in risk score had, on average, twice the length of hospital stay (6.8 compared to 3.4 days). There was no change in the primary outcome between groups. Among those who had a clinical event, there were more event-free hours in the intervention/display-on group compared to the standard-of-care/display-off, but this did not reach statistical significance. Clinicians chose to transfer 11% of patients into or out of display beds, a censoring event removing them from the analysis, thereby undermining aspects of the randomized nature of the study. Predictive analytics monitoring incorporating continuous cardiorespiratory monitoring and displays of risk trajectories coupled with an education plan did not improve patient outcomes. While necessary to conduct the study, the pragmatic design allowed for significant movement towards intervention/displayed beds for sicker patients. Design considerations in the future must focus on understanding clinicians' interpretation, care processes, and communication practices.

Clinical trial registration number: NCT04359641 Registered 4/24/20.

Keywords Predictive monitoring, Artificial intelligence, Clinical deterioration, Predictive analytics monitoring, Implementation of artificial intelligence

Unanticipated clinical deterioration of hospital patients remains a pernicious problem that contributes to in-hospital morbidity and mortality¹⁻⁵. Often, following the clinical deterioration event, clinicians who review these cases can recognize trends in vital signs, laboratory tests, and other electronic health record elements that, in retrospect, might have raised clinical suspicion of illness⁶. For example, the coincidence of rising heart rate and falling blood pressure should suggest incipient shock, but often the changes in physiological parameters are subtle and insidious. Automated detection of such signatures of illness might inform clinicians of patients who could benefit from earlier assessment and proactive clinical action^{1,7}.

Since its inception more than 20 years ago in the neonatal intensive care unit (ICU)^{8,9}, predictive analytics monitoring has been increasingly recognized for its potential to improve patient outcomes by early detection of deterioration. Artificial intelligence (AI) and machine learning methods support synthesizing unprocessed cardiorespiratory monitoring data, laboratory results and clinician-entered vital signs into computational

¹Center for Advanced Medical Analytics, School of Medicine, University of Virginia, Charlottesville, VA, USA.

²Department of Pediatrics, Hematology-Oncology Division, School of Medicine, University of Virginia, Charlottesville, VA, USA. ³Department of Public Health Sciences, Biostatistics Division, University of Virginia, Charlottesville, VA, USA. ⁴Nihon Kohden Digital Health Solutions, Irvine, CA, USA. ⁵Department of Internal Medicine, Cardiovascular Division, School of Medicine, University of Virginia, Charlottesville, VA, USA. ⁶Pediatric Hematology-Oncology, University of Virginia, P.O. Box 800386, Charlottesville, VA 22908, USA. ✉email: jlk2t@virginia.edu

models that can visually represent the estimation of the risk of clinical deterioration¹⁰. Since AI's new and sometimes opaque methodologies lie at its core, predictive analytics monitoring has been held to high standards for transparency. For example, there are numerous guidelines and opinion pieces to inform the developer and the user^{11–14}.

Incontrovertible proof of the benefit of predictive analytics monitoring in the form of large randomized clinical trials has been lacking. While a meta-analysis of artificial intelligence-based sepsis alerts showed an average mortality reduction of 30%, it included only six randomized trials with 2938 patients¹⁵. To this list, we can add the original trial in this field, that of heart rate characteristics monitoring for neonatal sepsis in 3003 premature infants, which showed a 22% relative mortality reduction when risk estimates were displayed to clinicians¹⁶. Here, we extend the work in neonates to evaluate the impact of predictive analytics monitoring on events of clinical deterioration in a heterogeneous group of adults in acute care cardiac medical-surgical wards. We test the use of an AI-based risk analytic, Continuous Monitoring of Event Trajectories (CoMET, Nihon Kohden Digital Health Solutions, Irvine, CA), that visually displays risk estimates of multiple events of clinical deterioration every 15 min. We hypothesized that the passive display of a visual risk analytic would facilitate early identification of patients at risk for impending adverse outcomes and increase the hours free from clinical deterioration.

Methods

Study design

We conducted a cluster-randomized controlled trial (RCT) among patients admitted to 85-bed acute care cardiology medical-surgical wards. Eleven clusters were defined by room number, and patients were randomized to receive either the intervention (passive display of CoMET) or usual care (standard of care). Randomization began on January 4, 2021, and ended on October 4, 2022. Due to the pragmatic nature and replication of real-world patient contexts, all patients admitted to these floors during the study period were enrolled in the study. Clusters were re-randomized every 2 months following a Latin square design within each stratum to ensure a balanced treatment assignment. Full pre-registered study protocol details and additional randomization details can be found in¹⁷, and the clinical trial was registered (NCT04359641). The study was approved by the University of Virginia Health System Institutional Review Board, and the study was granted a waiver of informed consent (IRB#22196) in accordance with the Declaration of Helsinki. A Data Safety and Monitoring Board reviewed interim results. The results are reported in accordance with the CONSORT-AI guidelines¹². The trial was stopped after meeting the pre-specified sample size enrolled.

Patient care units and populations

The Medical-Cardiology unit predominantly comprised acute cardiology patients admitted with decompensated congestive heart failure, acute coronary syndrome, arrhythmias such as atrial fibrillation, and other cardiac diagnoses. The Medical-Surgical unit was historically the same as Medical-Cardiology. Immediately preceding the study, though, 11 of the 20 beds were repurposed for patients with solid organ transplants under the care of a dedicated transplant team. They differed from the patients and the practices of the cardiology and cardiac surgery beds. The baseline distributions of vital signs and laboratory test results, for example, differed, with a bias toward findings of chronic kidney and liver failure in the transplant patients. Because of the pragmatic nature of this study design in real-world contexts, they remained in the study population. The Cardio-Thoracic Surgery unit predominantly comprised patients awaiting discharge. During the study period, no other early warning systems were used in the study patient care units.

Risk displays

We used the CoMET system to display two risk trajectories, calculated from previously published machine-learning logistic regression analyses relating continuous cardiorespiratory monitoring parameters, nurse-charted vital signs, and laboratory test results to outcome events on these wards from 2013 to 2015¹⁸. These models were developed using 2,194,077 control observations (8,105 admissions) and 23,881 case observations (586 admissions). The predictors included 41 physiological variables that were at least 70% available/non-missing, including data derived from continuous cardiorespiratory monitoring, laboratory assessments, and vital sign assessments¹⁸. The cohort in which the models were originally developed and validated mirrored that in the study population (acute care medical-surgical units) with an event rate of 1.1% (during a pre-COVID-19 time frame) and the imbalanced data was used for training to ensure the predicted risk was well-calibrated to the observed risk¹⁸.

The unit of measurement of the risk estimate is the fold-increase in the probability of the event in the next 24 h compared to the average probability. Thus, a patient with a score of 1 has the average risk, a patient with a score of 2 has twice the average risk, and so on. The plot shows a comet-like icon for each patient with the bed number in the head; the tail is 3 h long. Additional details on the model inputs and model performance, including excellent calibration, are in the following publication¹⁹. An example of what the clinician sees and can view on a single patient is found in Fig. 1.

Intervention implementation details

We followed principles derived from the prior implementation of predictive analytics monitoring in ICUs and non-ICUs at the UVA Hospital and a regional hospital²⁰. Each of the three units identified an interprofessional team of clinical champions to serve as super users, oversee implementation, and ensure staff education. Our team of educators (PIs, CoMET educators, clinical research coordinators (CRCs), and other study investigators) provided sessions with the clinical champions and other clinician groups in each unit before randomization and during the first quarter. We encouraged the integration of the CoMET scores into routine nurse vital sign

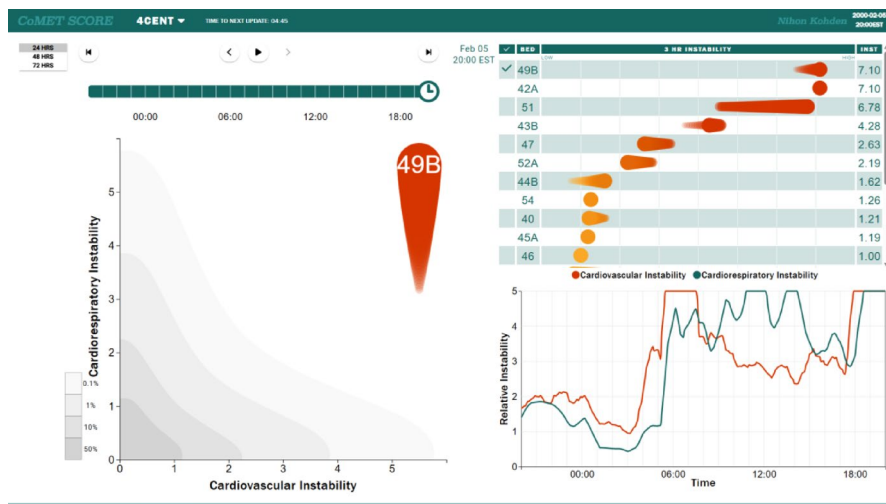


Fig. 1. CoMET display of a hypothetical patient in Bed 49B. The head-to-tail of the CoMET represents a 3-hour trajectory. This patient has greater than a 5 fold risk of both a cardiorespiratory and cardiovascular clinical deterioration event in the next 24 h. On the right upper quadrant, the user will find the leaderboard of where this patient's risk compares to the other patients. In the right lower quadrant, the user can look at the trajectory of the 2 individual risk scores.

assessments, rounding discussions, hand-off reports, and escalation protocols. Our recommended protocol for reporting and responding to changes in the CoMET scores included nurse notification of physicians or advanced practice providers of an increase in score by 2 units over a three-hour period and physician evaluation of the patient as appropriate. A CRC attended morning rounds on each unit on a rotating basis to show the display and answer questions about interpreting score trajectories. The CoMET score was available on the unit at all times to those randomized to the intervention group on large scale monitors and could be accessed via a web-based platform if not on site. The CoMET scores flowed through to the electronic medical record every hour and were recorded within the vital sign flowsheet¹⁷.

Study outcomes

The primary outcome was the number of hours free of clinical deterioration (adverse clinical events including: an emergent ICU transfer, emergent intubation, cardiac arrest, emergent surgery, and death) within 21 days of admission. With this outcome, patients were awarded hours that they remained stable without a clinical deterioration event. A maximum score was 21 event-free days (504 h), and patients who were successfully discharged from the hospital prior to the 21 days without an event were awarded all 504 h. Patients who died during admission were counted as having 0 event-free hours. Patients with non-emergent ICU transfers, non-emergent/planned surgeries, or changes in bed assignments that caused a change in arm assignment were censored. Dedicated clinician CRCs adjudicated each case of clinical deterioration/study endpoint event. For any events that were ambiguous, the PIs of the study (JKM and JMB) blindly adjudicated outcomes through a weekly meeting. This report only presents the primary outcome and we will report additional analyses separately focused on the time to clinical action, post-ICU trajectories, and costs to the healthcare system. A Data Safety Monitoring Board met at pre-specified intervals to assess for any unintended harms.

Statistical analysis

We used a generalized estimating equation with a zero-inflated negative binomial to analyze event-free hours and a Cox proportional hazards model to analyze time-untill-first-critical events. The model handled the correlation of multiple hospital admissions from the same patient, within-unit clusters within the study design, and differences in the amount of time at risk for censored patients. An intent-to-treat analytic strategy was utilized where all randomized patients were included for the primary endpoint comparison between the intervention and usual care. As a post-hoc analysis, we also analyzed the primary outcome in an 'at-risk' cohort, or those patients who experienced a rise in 2 on either axis of the CoMET display during their hospitalization. The analyses in the at-risk cohort were exploratory, as they were not predefined. The post-hoc analysis was conducted to uncover more details surrounding censoring events and patient movement.

Results

Over the 22 months of randomization, which coincided with the COVID-19 pandemic, there were 10,422 patient visits within our study cohort (Fig. 2). They were majority male, White, non-Hispanic, and on average 65 years of age with relatively low hospital length of stays (Table 1). Model performance is reported separately¹⁹, but overall the predictive models performed well and outperformed a common early warning score, qSOFA.

Only 5.3% of patients had a clinical deterioration event, and the vast majority of participants had the maximum amount of event-free hours (Table 2). There was no statistically significant difference in the primary

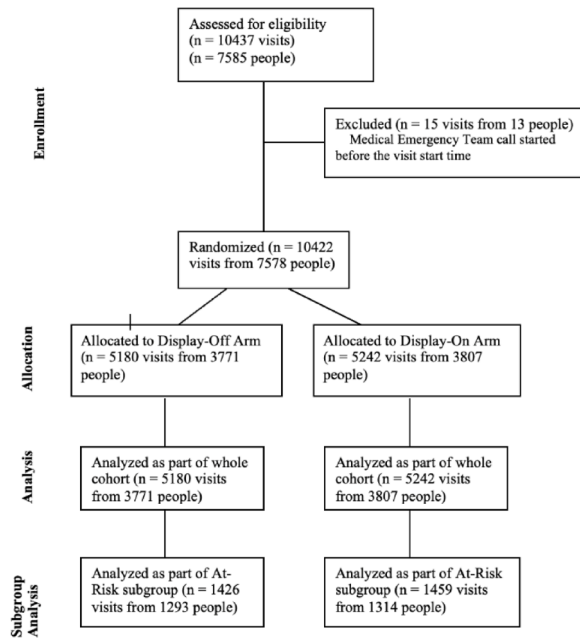


Fig. 2. CONSORT randomization information.

	Cohort		At risk subgroup	
	Display on n = 5242	Display off n = 5180	Display on n = 1459	Display off n = 1426
Sex				
Female	2105 (40.2)	2181 (42.1)	602 (41.3)	587 (41.2)
Male	3136 (59.8)	2999 (57.9)	857 (58.7)	839 (58.8)
Race				
White	3942 (75.2)	3910 (75.5)	1053 (72.2)	1041 (73.0)
Black	1009 (19.2)	1013 (19.6)	316 (21.7)	322 (22.6)
Asian	58 (1.1)	47 (0.9)	18 (1.2)	11 (0.8)
Other	215 (4.1)	198 (3.8)	69 (4.7)	47 (3.3)
Unknown	18 (0.3)	12 (0.2)	3 (0.2)	5 (0.4)
Ethnicity				
Hispanic	174 (3.3)	162 (3.1)	66 (4.5)	44 (3.1)
Non-Hispanic	5056 (96.5)	5001 (96.5)	1390 (95.3)	1378 (96.6)
Refused / Unknown	12 (0.2)	17 (0.3)	3 (0.2)	4 (0.3)
Age at Admission	65.4 (54.7, 74.4)	65.3 (54.4, 74.9)	65.2 (54.0, 73.9)	64.9 (54.6, 73.8)
Hospital Unit				
Medical-Cardiology	1888 (36.0)	1851 (35.7)	424 (29.1)	427 (29.9)
Cardio-Thoracic Surgery	1752 (33.4)	1708 (33.0)	561 (38.5)	506 (35.5)
Medical/Surgical	1602 (30.6)	1621 (31.3)	474 (32.5)	493 (34.4)
Hospital Length of Stay (days)	2.7 (1.1, 5.0)	2.6 (1.1, 4.9)	4.7 (2.8, 7.8)	4.5 (2.6, 7.9)

Table 1. Admission characteristics of the cohort and at-risk subgroup by arm. * n(%) or median (IQR) are shown.

outcome by study arm in both the full cohort and those at risk for clinical deterioration. Among those who had a clinical event, there were more event-free hours in the intervention/display-on group compared to the standard-of-care/display-off, but this did not reach statistical significance. Censoring events and components meeting the endpoint of clinical deterioration are found in Table 2. The estimated marginal means from the negative binomial generalized estimating equation are found in Fig. 3.

n(%) or median (inter-quartile range). IRR/ORs for display-on versus display-off using negative binomial or logistic regression adjusted for study design.

Bed movement was common within this RCT (n=382 went from usual care to CoMET display, and n=419 went from CoMET display to usual care). In a *post-hoc* analysis, we found higher relative risk of

	Cohort			At-risk subgroup		
	Display on n = 5242	Display Off n = 5180	IRR/OR (95% CI)	Display on n = 1459	Display off n = 1426	IRR/OR (95% CI)
Primary Outcome:						
Event-Free Hours	504 (0-504)	504 (0-504)	0.995 (0.88–1.12)	504 (0-504)	504 (0-504)	1.07 (0.91–1.26)
Outcome censored due to:						
No censoring	4461 (85.1)	4411 (85.1)		1204 (82.5)	1168 (81.9)	
Bed change	617 (11.8)	595 (11.5)		195 (13.4)	197 (13.8)	
Re-randomization	164 (3.1)	174 (3.4)		60 (4.1)	61 (4.3)	
Components of Primary Outcome:						
Had any Event	282 (5.4)	271 (5.2)	1.01 (0.85–1.20)	179 (12.3)	171 (12.0)	1.05 (0.84–1.32)
Event-Free Hours in Patients with an Event ¹	100.6 (0-500)	95.0 (0-499)	0.99 (0.88–1.12)	142.8 (0-500)	90.4 (0-498)	1.07 (0.91–1.26)
Death	57 (1.1)	42 (0.8)	1.42 (0.93–2.16)	37 (2.5)	28 (2.0)	1.45 (0.86–2.46)
Emergent ICU Admission	224 (4.3)	233 (4.5)	0.94 (0.78–1.13)	151 (10.3)	150 (10.6)	0.99 (0.90–1.27)
Emergent Intubation	65 (1.2)	48 (0.9)	1.31 (0.89–1.92)	38 (2.6)	30 (2.1)	1.29 (0.78–2.13)
Cardiac arrest followed by ICU/Death	22 (0.4)	25 (0.5)	0.84 (0.45–1.54)	12 (0.8)	14 (1.0)	0.96 (0.42–2.19)
Emergent transfer for Surgery	22 (0.4)	16 (0.3)	1.34 (0.68–2.61)	11 (0.8)	8 (0.6)	1.67 (0.66–4.21)
First Event that Occurred						
Death	21 (0.4)	16 (0.3)		14 (1.0)	10 (0.7)	
Emergent ICU	198 (3.8)	201 (3.9)		133 (9.1)	131 (9.3)	
Emergent Intubation	24 (0.5)	16 (0.3)		11 (0.8)	9 (0.6)	
Cardiac arrest followed by ICU/Death	18 (0.3)	22 (0.4)		10 (0.7)	11 (0.8)	
Surgery	21 (0.4)	16 (0.3)		11 (0.8)	8 (0.6)	

Table 2. Study outcomes for the cohort and at-risk subjects. ¹ N = those who had an event.

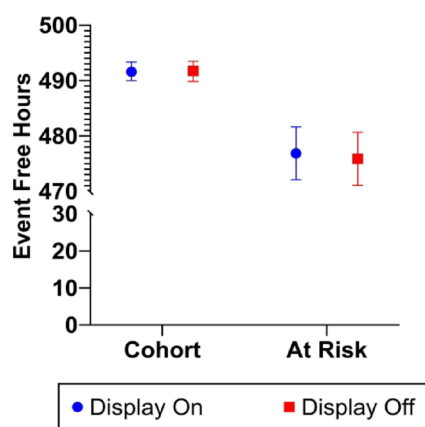


Fig. 3. Estimated marginal mean (and 95% CI) event-free hours by arm and population.

clinical deterioration among those transferred from a non-display to a display bed. This finding indicates that clinicians made decisions that undermined the cluster RCT design by preferentially placing sicker patients into the CoMET display beds, presumably to obtain additional information about their underlying physiological instability. (Fig. 4).

Discussion

We conducted a large randomized controlled trial to test the impact of predictive analytics monitoring on cardiology and cardiac surgery patients in in-patient acute care wards. Our major findings are that a display of the trajectories of the risk of imminent deterioration without alerts was not associated with an increase in hours free of clinical deterioration. We found that there was an increase in relocating sicker patients from usual care to CoMET display-on, thereby distorting the RCT design. We note this study was undertaken during the early stages of the COVID-19 pandemic in our area, which disrupted staffing and resource utilization¹⁰ and was associated with a lower event rate than our pre-COVID baseline. Clinician-driven transfers of sicker patients into display beds may suggest that the intervention was perceived as useful at the bedside, even if it did not improve the prespecified primary outcome and these behavior patterns should be studied in greater detail alongside other

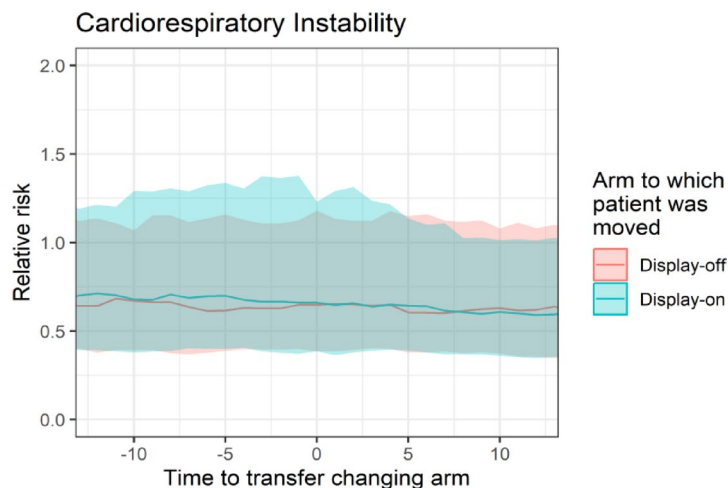


Fig. 4. Time zero is transfer from one arm to another and the unit of analysis is days. The line is the median and the shaded region is the IQR. Blue are those patients transferred from a display-off to display-on bed and their overall relative risk of clinical deterioration is higher, representing a sicker population.

implementation outcomes. Further, treating bed movement as a censoring event may have biased the overall study results to the null.

We justify using continuous cardiorespiratory monitoring in the predictive models because of the causality leakage of clinician-ordered tests. For example, the presence of a serum lactate measurement in the chart bespeaks clinical suspicion of sepsis. It excels as a predictor, but not because it detects changes before clinicians. Likewise, the use of pressors, fluid administration, and blood cultures as predictors rather than endpoints allows causality leakage, artificially inflating model performance. Kamran and coworkers recently showed that the Epic Sepsis Model had no predictive capability at all when assessed prior to any clinician suspicion or action, that is, before fluids, antibiotics, and lactate tests²¹.

We justify the randomized clinical trial approach to the study of AI-based risk analytics. We find irreducible uncertainty in other trial designs and results, even our own demonstration of a 50% reduction in septic shock in a surgical ICU after the introduction of CoMET trained for urgent unplanned intubation for acute respiratory failure and hemorrhage requiring large transfusion²². While a control medical ICU without CoMET monitoring had no consistent benefit over the same time period, the absence of randomization has allowed skepticism of the before-and-after study results. We especially endorse the simple and pragmatic nature of the current trial - we enrolled everyone to mimic complex real-world acute care environments and used important endpoints to clinicians and patients.

Other studies have used non-randomized designs. For example, Winslow, Edelson, Churpek, and coworkers found reduced mortality throughout the hospital after initiating their early warning system²³. The magnitude was substantial but was the same even for patients with low risks who were never alerted. Adams, Henry, Saria, and coworkers used a machine learning-based sepsis early warning system that included mechanical ventilation, lactate, and vasopressors as sepsis predictors and placed passive alerts in the EHR, but only after “verifiable symptoms” were present²⁴. They found reduced mortality in the 1.2% of patients who met retrospective sepsis criteria and whose alerts were acknowledged within three hours but did not present outcomes of the other patients^{24,25}. Groups led by Shimabukuro²⁶ and Watkinson²⁷ performed small randomized trials with divergent results. Escobar, Liu and coworkers used a non-randomized stepped wedge approach to show a convincing mortality reduction with an EHR-based early warning score in 19 hospitals²⁸. They employed an informed intermediary to receive alerts and to review the clinical record before alerting bedside clinicians.

Cluster randomization was chosen because clinicians needed to easily recognize who was in the intervention group. We did not want to add additional burden to the complexity of the acute care environment and did not want to falsely reassure a clinician who did not see a rising score to assume the patient was in the intervention group. Although we have previously tested the positive predictive value of creating a threshold-based alert from CoMET¹, we did not want to add to clinician burden or create additional markers of alarm fatigue and decided to offer suggestions for response without an active alert strategy. The lack of mandated protocolized responses, however, makes it challenging to attribute the neutral results only to the accuracy of the analytic.

We approached education and implementation of novel analytics in real-world environments within the context of the COVID-19 pandemic and wanted to minimize any contribution to clinician burnout²⁹. Implementation considerations focused on promoting trust, transparency, integration with workflow, prolonged stakeholder engagement, tailored education embedded within hospital norms, supporting actionability through integration within clinical care, and sustainability through ongoing supportive interactions²⁹. However, variability in clinician patterns may limit generalizability if there are underlying practices that differ substantially between hospitals.

AI clinical trials call for an understanding of new and separate design elements not previously contemplated. There are several elements of this type of intervention, such as the use of passive displays, that warrant further

study and have RCT design implications when studying AI-based predictive analytics: the proactive nature of the clinical action response, the rarity of the event of interest, and the complexity of implementation within evolving acute care environments. While the pragmatic design allowed for implementation in real-world environments, this also meant there was heterogeneity in patient populations, unit work environments, and clinician responses. The real-world context of clinical trials such as this implies that the implementation of AI-based risk analytics must also be adaptive to ensure long-term integration and practice flexibility across different work environments.

One aspect of the study design we would like to improve in the future is the integration of in-depth qualitative methodology and implementation outcomes alongside patient outcomes to better understand how clinicians process, communicate, and integrate risk scores into their care practices. Understanding the cognitive interpretation of scores in real-time, the response through diverse care processes, and communication patterns involving AI scores preceding clinical events is an area in particular need of further study. In hindsight, hours free of events of clinical deterioration may not have been the most sensitive primary outcome. Outcomes specific to the at-risk population and subsequent intensive care unit trajectories may have been more appropriate, given the overall rarity of our event.

Data availability

The datasets generated during the current study are not publicly available because we are still analyzing secondary outcomes, but are available for replication from the corresponding author at reasonable request.

Received: 1 June 2025; Accepted: 2 February 2026

Published online: 05 February 2026

References

- Keim-Malpass, J., Clark, M. T., Lake, D. E. & Moorman, J. R. Towards development of alert thresholds for clinical deterioration using continuous predictive analytics monitoring. *J. Clin. Monit. Comput.* **34** (4), 797–804. <https://doi.org/10.1007/s10877-019-00361-5> (2020).
- Smith, M. E. B. et al. Early warning system scores for clinical deterioration in hospitalized patients: a systematic review. *Ann. Am. Thorac. Soc.* **11** (9), 1454–1465. <https://doi.org/10.1513/AnnalsATS.201403-102OC> (2014).
- Escobar, G. J. et al. Early detection of impending physiologic deterioration among patients who are not in intensive care: development of predictive models using data from an automated electronic medical record. *J. Hosp. Med.* **7** (5), 388–395. <https://doi.org/10.1002/jhm.1929> (2012).
- Escobar, G. J. et al. Intra-hospital transfers to a higher level of care: contribution to total hospital and intensive care unit (ICU) mortality and length of stay (LOS). *J. Hosp. Med.* **6** (2), 74–80. <https://doi.org/10.1002/jhm.817> (2011).
- Curtis, K. et al. Treatments costs associated with inpatient clinical deterioration. *Resuscitation* **166**, 49–54. <https://doi.org/10.1016/j.resuscitation.2021.07.022> (2021).
- Blackwell, J. N. et al. Early detection of In-Patient deterioration: one prediction model does not fit all. *Crit. Care Explor.* **2** (5), e0116. <https://doi.org/10.1097/CCE.0000000000000116> (2020).
- Moss, T. J. et al. Signatures of subacute potentially catastrophic illness in the ICU: model development and validation. *Crit. Care Med.* **44** (9), 1639–1648. <https://doi.org/10.1097/CCM.0000000000001738> (2016).
- Randall Moorman, J. The principles of whole-hospital predictive analytics monitoring for clinical medicine originated in the neonatal ICU. *Npj Digit. Med.* **5** (1), 41. <https://doi.org/10.1038/s41746-022-00584-y> (2022).
- Fairchild, K. D. et al. Septicemia mortality reduction in neonates in a heart rate characteristics monitoring trial. *Pediatr. Res.* **74** (5), 570–575. <https://doi.org/10.1038/pr.2013.136> (2013).
- Keim-Malpass, J. & Moorman, L. P. Nursing and precision predictive analytics monitoring in the acute and intensive care setting: an emerging role for responding to COVID-19 and beyond. *Int. J. Nurs. Stud. Adv.* **3**, 100019. <https://doi.org/10.1016/j.ijnsa.2021.100019> (2021).
- Collins, G. S. et al. Protocol for development of a reporting guideline (TRIPOD-AI) and risk of bias tool (PROBAST-AI) for diagnostic and prognostic prediction model studies based on artificial intelligence. *BMJ Open.* **11** (7), e048008. <https://doi.org/10.1136/bmjopen-2020-048008> (2021).
- Liu, X. et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. *Nat. Med.* **26** (9), 1364–1374. <https://doi.org/10.1038/s41591-020-1034-x> (2020).
- Rivera, S. C. et al. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. *BMJ* **370**, m3210. <https://doi.org/10.1136/bmj.m3210> (2020).
- Vasey, B. et al. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. *Nat. Med.* **28** (5), 924–933. <https://doi.org/10.1038/s41591-022-01772-9> (2022).
- Zhang, Z. et al. Effectiveness of automated alerting system compared to usual care for the management of sepsis. *Npj Digit. Med.* **5** (1), 101. <https://doi.org/10.1038/s41746-022-00650-5> (2022).
- Moorman, J. R. et al. Mortality reduction by heart rate characteristic monitoring in very low birth weight neonates: a randomized trial. *J. Pediatr.* **159** (6), 900–906. <https://doi.org/10.1016/j.jpeds.2011.06.044> (2011).
- Keim-Malpass, J. et al. Predictive Monitoring-Impact in acute care cardiology trial (PM-IMPACT): protocol for a randomized controlled trial. *JMIR Res. Protoc.* **10** (7), e29631. <https://doi.org/10.2196/29631> (2021).
- Moss, T. J. et al. Cardiorespiratory dynamics measured from continuous ECG monitoring improves detection of deterioration in acute care patients: A retrospective cohort study. *PLoS ONE.* **12** (8), e0181448. <https://doi.org/10.1371/journal.pone.0181448> (2017).
- Keim-Malpass, J. et al. Prospective validation of clinical deterioration predictive models prior to intensive care unit transfer among patients admitted to acute care cardiology wards. *Physiol. Meas.* **45** (6). <https://doi.org/10.1088/1361-6579/ad4e90> (2024).
- Prudente, M. L. Principles for Real-World Implementation of Bedside Predictive Analytics Monitoring. *Appl. Clin. Inform.* Published online 2021. Accessed September 20, <https://doi.org/10.1055/s-0041-1735183> (2021).
- Kamran, F. et al. Evaluation of sepsis prediction models before onset of treatment. *NEJM AI.* **1** (3). <https://doi.org/10.1056/AIoa2300032> (2024).
- Ruminski, C. M. et al. Impact of predictive analytics based on continuous cardiorespiratory monitoring in a surgical and trauma intensive care unit. *J. Clin. Monit. Comput.* **33** (4), 703–711. <https://doi.org/10.1007/s10877-018-0194-4> (2019).
- Winslow, C. J. et al. The impact of a machine learning early warning score on hospital mortality: A multicenter clinical intervention trial. *Crit. Care Med.* **50** (9), 1339–1347. <https://doi.org/10.1097/CCM.00000000000005492> (2022).
- Henry, K. E. et al. Human-machine teaming is key to AI adoption: clinicians' experiences with a deployed machine learning system. *Npj Digit. Med.* **5** (1), 97. <https://doi.org/10.1038/s41746-022-00597-7> (2022).

25. Adams, R. et al. Prospective, multi-site study of patient outcomes after implementation of the TREWS machine learning-based early warning system for sepsis. *Nat. Med.* **28** (7), 1455–1460. <https://doi.org/10.1038/s41591-022-01894-0> (2022).
26. Shimabukuro, D. W., Barton, C. W., Feldman, M. D., Mataraso, S. J. & Das, R. Effect of a machine learning-based severe sepsis prediction algorithm on patient survival and hospital length of stay: a randomised clinical trial. *BMJ Open. Respir. Res.* **4** (1), e000234. <https://doi.org/10.1136/bmjresp-2017-000234> (2017).
27. Watkinson, P. J. et al. A randomised controlled trial of the effect of continuous electronic physiological monitoring on the adverse event rate in high risk medical and surgical patients. *Anaesthesia* **61** (11), 1031–1039. <https://doi.org/10.1111/j.1365-2044.2006.04818.x> (2006).
28. Escobar, G. J. et al. Automated identification of adults at risk for In-Hospital clinical deterioration. *N. Engl. J. Med.* **383** (20), 1951–1960. <https://doi.org/10.1056/NEJMsa2001090> (2020).
29. Moorman, L. P. Principles for real-world implementation of bedside predictive analytics monitoring. *Appl. Clin. Inf.* **12** (4), 888–896. <https://doi.org/10.1055/s-0041-1735183> (2021).

Author contributions

JKM and JMB secured funding and maintained oversight of the study. JKM, SJR, MTC, KNK, LPM, JRM, and JMB conceptualized the study. KNK, SH, GY, and AN collected primary data for the study. SJR and MKJ conducted analyses for the study. JKM drafted the initial manuscript. JKM, SJR, MTC, KNK, OJM, SH, GY, MKJ, AN, LPM, JRM, JMB provided edits and approved the final manuscript.

Funding

This study was funded by the Frederick Thomas Advanced Medical Analytics Fund, University of Virginia and AHRQ R01HS028803 (Keim-Malpass/Bourque MPI). The investigators received in-kind support from Nihon Kohden Digital Health Solutions for use and support of the CoMET system. The funders had no direct role/oversight of this published work.

Declarations

Competing interests

MC is an employee of Nihon Kohden Digital Health Solutions. LPM and JRM are consultants for Nihon Kohden Digital Health Solutions. JKM owns equity shares of ArteraAI and JRM owns equity shares of Medical Predictive Science Corporation, whose products are not discussed in this unrelated work. The other authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to J.K.-M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2026