European Journal of Surgical Oncology 49 (2023) 278–284

Contents lists available at ScienceDirect

European Journal of Surgical Oncology

journal homepage: www.ejso.com

A wearable patch based remote early warning score (REWS) in major abdominal cancer surgery patients

Jonna A. van der Stam [a, b, c, 1, *], Eveline H.J. Mestrom [d, 1], Simon W. Nienhuijs [e], Ignace H.J.T. de Hingh [e, h], Arjen-Kars Boer [b, c], Natal A.W. van Riel [a, c, g], Koen T.J. de Groot [i], Wim Verhaegh [i], Volkher Scharnhorst [a, b, c], R. Arthur Bouwman [d, f]

[a] Department of Biomedical Engineering, Eindhoven University of Technology, Eindhoven, the Netherlands
[b] Clinical Laboratory, Catharina Hospital, Eindhoven, the Netherlands
[c] Expert Center Clinical Chemistry Eindhoven, Eindhoven, the Netherlands
[d] Department of Anesthesiology, Intensive Care & Pain Medicine, Catharina Hospital, Eindhoven, the Netherlands
[e] Department of Surgery, Catharina Hospital, Eindhoven, the Netherlands
[f] Department of Electrical Engineering, Eindhoven University of Technology, Eindhoven, the Netherlands
[g] Department of Vascular Medicine, Amsterdam University Medical Centers, Amsterdam, the Netherlands
[h] GROW – School for Oncology and Developmental Biology, Maastricht University, Maastricht, the Netherlands
[i] Department of AI, Data Science & Digital Twin, Philips Research, Eindhoven, the Netherlands

ARTICLE INFO

Article history:
Received 12 May 2022
Received in revised form
30 July 2022
Accepted 26 August 2022
Available online 2 September 2022

Keywords:
Wearable monitoring
Remote early warning score
Early warning score
Deterioration detection
patch

ABSTRACT

Introduction: The shift toward remote patient monitoring methods to detect clinical deterioration requires testing of wearable devices in real-life clinical settings. This study aimed to develop a remote early warning scoring (REWS) system based on continuous measurements using a wearable device, and compare its diagnostic performance for the detection of deterioration to the diagnostic performance of the conventional modified early warning score (MEWS).

Materials and methods: The study population of this prospective, single center trial consisted of patients who underwent major abdominal cancer surgery and were monitored using routine in-hospital spotcheck measurements of the vital parameters. Heart and respiratory rates were measured continuously using a wireless accelerometer patch (HealthDot). The prediction by MEWS of deterioration toward a complication graded Clavien-Dindo of 2 or higher was compared to the REWS derived from continuous measurements by the wearable patch.

Main results: A total of 103 patients and 1909 spot-check measurements were included in the analysis. Postoperative deterioration was observed in 29 patients. For both EWS systems, the sensitivity (MEWS: 0.20 95% CI: [0.13–0.29], REWS: 0.20 95% CI: [0.13–0.29]) and specificity (MEWS: 0.96 95% CI: [0.95–0.97], REWS: 0.96 95% CI: [0.95–0.97]) were assessed.

Conclusions: The diagnostic value of the REWS method, based on continuous measurements of the heart and respiratory rates, is comparable to that of the MEWS in patients following major abdominal cancer surgery. The wearable patch could detect the same amount of deteriorations, without requiring manual spot check measurements.

1. Introduction

Despite continuous efforts and improvements in surgical techniques, postoperative patient deterioration is frequently encountered in the general surgical ward [1,2]. Many studies have shown that deterioration is often preceded by abnormalities in vital signs [3–6]. To anticipate on these abnormalities, regular measurements of vital parameters combined with early warning scoring systems have been adopted in hospitals. However, the best strategy is a much-debated subject, and limitations such as lack of validation, generalizability, or improvement of clinical outcomes are often discussed [7–9]. A drawback of early warning scoring systems is that they are manually measured and calculated, and

* Corresponding author. Computational biology, Building 15 Gemini South , Groene Loper, 5612, AZ, Eindhoven, the Netherlands.
E-mail address: j.a.v.d.stam@tue.nl (J.A. van der Stam).
[1] These authors contributed equally.

https://doi.org/10.1016/j.ejso.2022.08.034

therefore, labor intensive. Recent advances in wearable sensors that measure vital parameters provide an opportunity to replace the current manual procedures with remote monitoring systems and potentially improve the recognition of deterioration [10]. Observational studies have shown that wearables can measure abnormalities in vital signs preceding deterioration [11]. Despite the clear theoretical advantages of automated early warning systems, few wearables have been tested in prospective clinical settings [11,12], even fewer have already been implemented in clinical practice [13].

The current early warning scores are based on intermittent spotcheck measurements and a response protocol based on manually calculated scores. A commonly used tool is the modified early warning score (MEWS), which assigns points to impaired vital parameters and triggers follow-up care when the score exceeds a threshold [14]. The implementation of bedside devices that automatically calculate early warning scores based on manually recorded measurements results in an increase in the number of registered MEWSs [15]. Developing automated early warning systems by applying machine learning techniques to large-scale clinical datasets provides alternative early warning scores [16,17]. Patient data can be derived from electronic medical records (EMRs), laboratory data, and wearable sensors.

One recently developed wearable sensor is Healthdot (Philips Electronic Nederland B.V.), a small patch that is attached to the patient's left lower rib using an adhesive patch. The device has the ability to continuously, remotely and unobtrusively measure vital parameters for a period of up to two weeks, which makes it suitable for remote monitoring. In addition, this wearable device has previously been shown to be clinically acceptable in terms of accuracy [18,19]. However, the clinical performance of wearable data compared to current practice in the surgical ward is unknown.

The aim of this study was to develop a remote early warning score (REWS) using vital parameters collected by a wearable sensor, and compare its diagnostic performance for the detection of deterioration to the diagnostic performance of the conventional modified early warning score (MEWS).

2. Materials and Methods

2.1. Study design

The TRICA study, NCT03923127, is a prospective single-center study of wearable sensors in postoperative patients in a tertiary hospital (Catharina Hospital, Eindhoven, the Netherlands). The trial was approved by the Medical Ethics Committee (Maxima Medical Center, Veldhoven, the Netherlands [W19.001]).

2.2. Patient population

All adult patients scheduled to undergo major abdominal cancer surgery between April 2019 and August 2020 were eligible to participate in the study. The exclusion criteria were: pregnancy or breastfeeding, allergy to tissue adhesives, antibiotic-resistant skin infection, active implantable devices, and any skin condition in the area of device application. With an expected complication rate of up to 25% for the different types of major abdominal cancer surgery a sample size of 143 inclusions was selected in order to collect a reasonable amount relevant complications. Patients that met the inclusion criteria were identified by their surgeon during a preoperative routine visit, patients interested in participation met with a member of the clinical trial team to be further informed about the study. Informed consent was obtained prior to the start of the study procedures.

2.3. Data collection

After the surgical procedure the investigational device Healthdot (Philips Electronic Nederland B.V.), a wearable patch of $5 \times 3 \times 1$ cm that weighs 13.6 g, was applied to the patient's lower left rib on the midclavicular line. Heart rate (HR) and respiratory rate (RR) measurements were stored in the internal memory of the patch at time intervals of 8 s and 1 s, respectively. The accuracy of the device in postoperative patients has previously been established [18,19]. The HR data were resampled to a 1 s interval using linear interpolation. Every 5 min, the averages (mean) of the data collected over the past 5 min were transferred to a cloud server using built-in wireless communication technology (LoRa). Immediately after surgery, the wearable patch was applied in the post-anesthesia or intensive care unit, whichever was applicable to the patient. The wearable patch sensor then collected vital parameters (HR and RR) over a two-week period. Although the device can also collect data at home, only the vital parameters collected during the patients' hospital stay were considered in the present analysis. During the data collection both the regular care team and research team did not have access to the data collected by the wearable patch.

Once a patient was admitted to the general ward, MEWS scores were collected and registered in the electronic medical records (EMR) by the nursing staff as part of routine care. In general, MEWSs were measured every shift of 8 h. The local early warning score table is included in Appendix A, and the score can have a value between 0 and 19. The MEWS scores were retrospectively extracted.

All postoperative complications were retrospectively documented and graded by physicians according to the Clavien-Dindo classification [20]. Complications of grade 2 or higher were considered clinically relevant deteriorations, as those patients needed at least pharmacological, surgical, endoscopic, or radiological intervention.

2.4. Pre-processing of data

HR and RR measurements extracted from the internal memory of the wearable patch were used in the present analysis. A quality index (range 0–100) was reported together with the HR and RR measurements, low-quality measurements (quality index = 0) were excluded. Thereafter, the HR and RR data were down sampled to 5-min averaged data by calculating the mean of the data within the 5-min windows to match the sampling frequency of the cloud data storage transmissions.

2.5. Remote early warning score

As an alternative to the current manually collected early warning scores, the introduction of wearable sensors that automatically collect vital parameters allows the introduction of a remote early warning score (REWS). The REWS was built based solely on automatically collected HR and RR measurements from the wearable. Using the point scoring table of the local MEWS score for the HR and RR measured by the wearable patch, EWS points were assigned at every 5-min interval. (HR < 40 bpm: 2 points, HR 40–50 bpm: 1 point, HR 51–100: 0 points, HR 101–110: 1 point, HR 111–130 bpm: 2 points, HR > 130 bpm 3 points. RR < 9 rpm: 2 points, RR 9–14 rpm: 0 points, RR 15–20 rpm: 1 point, RR 21–30 rpm: 2 points, RR > 30 rpm: 3 points [19].) These scores were subsequently averaged over a 1-h period to avoid major influence of short deviations of vital parameters, resulting in the REWS, a numeric score ranging from 0 to 6.

2.6. Data analysis

The REWS was calculated retrospectively for the same time points when a MEWS score was registered in the EMR whenever wearable sensor data were available. If no MEWS was extracted from the EMR, the patient was excluded, as no comparison to the REWS could be made.

Baseline characteristics are expressed as totals and percentages in the case of binominal data or as medians and interquartile ranges (IQRs) in the case of numerical data. The significance of the differences between the deterioration and no-deterioration group was assessed using Fisher's exact test or Wilcoxon test for binominal or numerical data, respectively.

The performance of both early warning scores was then assessed by evaluating whether the patient experienced deterioration (complication CD $\geq$ 2) within 24 h after the spot-check measurement. Spotchecks within 24 h after a deterioration are excluded as deviating vital parameters shortly after deterioration are expected to be associated with the recent deterioration rather than an upcoming one. A visual representation on the in- and excluded spotchecks can be found in appendix B. The early warning scores were plotted against the occurrence of deterioration using a receiver operating characteristic (ROC) curve, which describes the discriminative power of the EWS scores for the range of all possible cutoff values [21]. The 95% confidence intervals (CIs) around the area under the ROC (AUROC) were computed using 2000 stratified bootstrap replicates. Statistical comparison of the two ROC curves was performed using the Delongs test for correlated ROC curves [22].

The distributions of the early warning scores were plotted as normalized histograms, and the data were split into the deterioration (CD $\geq$ 2 complications within the next 24 h) and no-deterioration groups. For the MEWS, the cutoff value was set at $\geq$ 3 to match routine practice, as action by the medical team is undertaken when the MEWS rises above this cutoff value. The cutoff value for the REWS was selected to match the number of true-positive cases from the MEWS score. The sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and 95% CI were then calculated based on these cutoffs [23,24]. The analysis was performed using R (V4.0.5) and R Studio (V1.4.1717) [25,26].

3. Results

A total of 143 patients were enrolled in this study. After excluding 23 patients to whom no wearable sensor was applied, 2 patients without spot-check data, and 15 patients without wearable sensor data, 103 patients were included in the analysis (Fig. 1). The baseline characteristics and outcomes of the study population are summarized in Table 1.

29 out of the 103 patients(28%) experienced clinically relevant deterioration during the study period. Patients who experienced deterioration were older (median 68 vs. 63 years, p = 0.01) and had a significantly longer length of hospital stay (median 13 vs. 8 days, p < 0.01) than those who did not experience deterioration. Interventions in case of deterioration included unanticipated ICU admission (n = 8), pain medication (n = 1), radiological (n = 5), surgical (n = 4) and antimicrobial (n = 18) intervention. No significant differences in comorbidities were found between the groups (Appendix C).

1909 spotchecks were used in the analysis, of which 665 (35%) originated from patients who experienced deterioration during their hospital stay. A median of 16 spotchecks were obtained for each patient. For patients who experienced deterioration, a significantly higher number of spotchecks was available (median 22 vs. 15, p < 0.01), which can be explained by the significantly longer length of stay. In total, 105 spotchecks (5.5%) were collected within the 24 h prior to deterioration (complication with CD $\geq$ 2) and were thus labelled as positive cases.

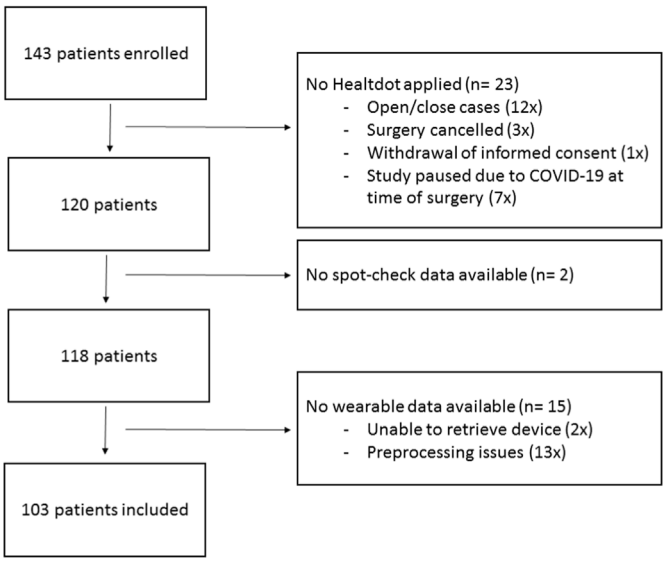

Fig. 1. Flow chart.

Fig. 2 shows the ROC curves for both MEWS and REWS to predict clinical deterioration within 24 h after the assessment. The AUROCs were comparable, the AUC for REWS (0.71, 95% CI: 0.66–0.77) was numerically higher than that of MEWS (0.62, 95% CI: 0.57–0.68). Despite overlapping confidence intervals the p-value of a delongs test was p < 0.01.

To match the number of true positives of the MEWS spotchecks, the cutoff for the REWS score was placed at $\geq$ 2.4. As shown in Fig. 3 and Table 2, 1816 (95.1%) of the REWS spot-checks did not exceed this threshold. A total of 93 (4.9%) REWS spotchecks did exceed the threshold, of which 21 (22.6%) were followed by deterioration within the next 24 h. The $\geq$2.4 cutoff for the REWS results in a sensitivity of 0.20 and a specificity of 0.96. An overview of the outcomes for different REWS and MEWS cutoff values is provided in Appendix D and E.

The MEWS and REWS both exceeded their respective thresholds in 9 patients that experienced a deterioration, 1 patient only exceeded the MEWS threshold and 1 only exceeded the REWS threshold. The e 18 patients that did experience a clinical deterioration did not exceed the threshold in the 24 h prior to a deterioration for both systems. Both MEWS and REWS spotchecks had the same number of true positives by design (21 spotchecks); however, the number of false positives was higher for REWS than for MEWS (72 vs. 65 spotchecks).

4. Discussion

In the present study, which enrolled patients after major abdominal cancer surgery, the REWS, based on vital parameters measured by a wireless patch sensor (Healthdot) through accelerometry, was compared with routine MEWS spotchecks performed by nursing staff. The cutoff point for the REWS was chosen to match the number of true positives obtained in current practice, which resulted in comparable sensitivity and specificity for the REWS. These findings are important because early warning scores for predicting clinical deterioration are of great interest in current clinical practice. However, the collection of information required

Table 1
Baseline characteristics, data availability and outcomes table. Continuous data is represented as mean (sd) or median [IQR].

Variable	Total	No deterioration	Deterioration (CD ≥ 2)	P value
Number of patients	103	74	29	
Female	46 (45%)	35 (45%)	11 (38%)	0.51
Age (years)	64 [57–70]	63 [55–68]	67.5 [63–74]	**0.01**
BMI (kg m^{-2})	26 [23–29]	25 [23–29]	26 [23–30]	0.80
ASA physical status				
II	78 (76%)	58 (78%)	20 (69%)	0.32
III	24 (23%)	16 (22%)	8 (28%)	0.61
IV	1 (1%)	0 (0%)	1 (3%)	0.28
Type of cancer surgery				
Esophageal resection	19 (18%)	15 (20%)	4 (14%)	0.58
HIPEC	22 (21%)	15 (20%)	7 (24%)	0.79
Pancreas surgery	19 (18%)	14 (19%)	5 (17%)	1.00
LAR/APR with IORT	20 (19%)	14 (19%)	6 (21%)	1.00
LAR/APR without IORT	6 (6%)	3 (4%)	3 (10%)	0.35
Debulking	11 (11%)	9 (12%)	2 (7%)	0.72
Gastric resection	4 (4%)	2 (3%)	2 (7%)	0.31
Other	2 (2%)	2 (3%)	0 (0%)	–
Surgery duration (minutes)	317 [256.5–386.5]	321 [266–395]	288 [247–375]	0.49
LOS hospital (days)	8 [7–12]	8 [7–10]	12 [10–22]	**<0.01**
Postoperative ICU admission	80 (78%)	56 (76%)	24 (83%)	0.60
LOS postoperative ICU admission (days)	1 [1–2]	1 [1–2]	1 [1–2]	0.23
Antiarrhythmic medication in ward	22 (21%)	14 (19%)	8 (28%)	0.42
Type of deterioration				
Bleeding			1	
Wound infection			1	
Abscess			2	
Pneumonia			10	
Anastomotic leakage			3	
Perforation			1	
Fever e.c.i.			2	
Thromboembolic event			3	
Urinary tract infection			3	
Stoma necrosis			2	
Atrial fibrillation			1	
Chyle leakage			1	
Other			2	

CD: Clavien-Dindo score, LOS: Length Of Stay.

Table 2
Distribution of the patients over the different areas of the Early Warning Scores. The cutoff values match the cutoff values shown in Fig. 3. A Clavien Dindo (CD) score of 0 is assigned when no deterioration is present.

variable	MEWS		REWS	
Cutoff	<3	≥3	<2.4	≥2.4
Spot-checks				
Total number	1823	86	1816	93
No deterioration	1739	65	1732	72
Deterioration (CD ≥ 2)	84	21	84	21
% deterioration	4.6%	24.4%	4.6%	22.6%
Performance				
Sensitivity	0.20 [0.13–0.29]		0.20 [0.13–0.29]	
Specificity	0.96 [0.95–0.97]		0.96 [0.95–0.97]	
PPV	0.24 [0.16–0.35]		0.23 [0.15–0.32]	
NPV	0.95 [0.94–0.96]		0.95 [0.94–0.96]	
Complication Severity				
CD 0	1710	63 (3.6%)	1702	71 (4.0%)
CD 1	29	2 (6.5%)	30	1 (3.2%)
CD 2	45	11 (19.6%)	45	11 (19.6%)
CD 3	17	2 (10.5%)	17	2 (10.5%)
CD 4	22	8 (26.7%)	22	8 (26.7%)

CD: Clavien-Dindo score, MEWS: Modified Early Warning Score, REWS: Remote Early Warning Score, PPV: positive predictive value, NPV: negative predictive value.

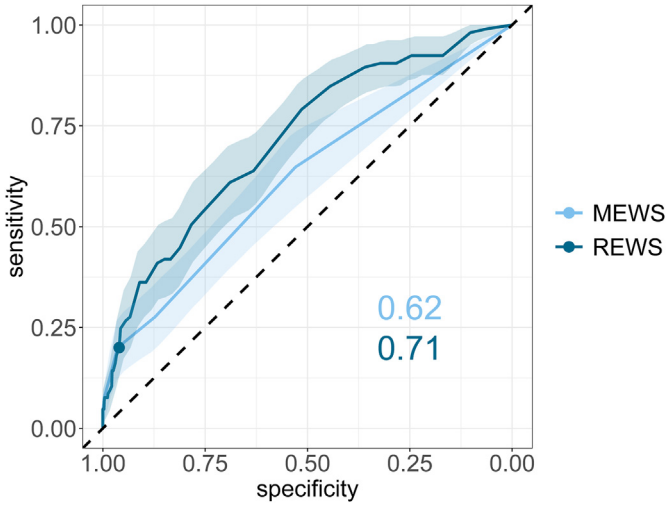

Fig. 2. Receiver operating characteristic (ROC) curve for the Modified Early Warning Score (MEWS) and Remote Early Warning Score (REWS) to describe the development of a deterioration in the next 24 h. Displaced numbers are the area under the ROC-curves. Dashed line represents the random-guessing line.

for these systems is labor intensive. Therefore, intermittent assessments with relatively large time intervals are used, although this weakens the concept of early warning of deterioration because deterioration may occur between assessments.

In the present cohort, 1816 spot-check measurements did not exceed the REWS cutoff value. These spotchecks would automatically be classified as low-risk and thus would not require any action from the nursing staff. Simultaneously, the number of spotchecks exceeding the alarm threshold increased (MEWS: n = 86, REWS: n = 93), leading to a larger number of patients requiring follow-up care. However, translating these results into an

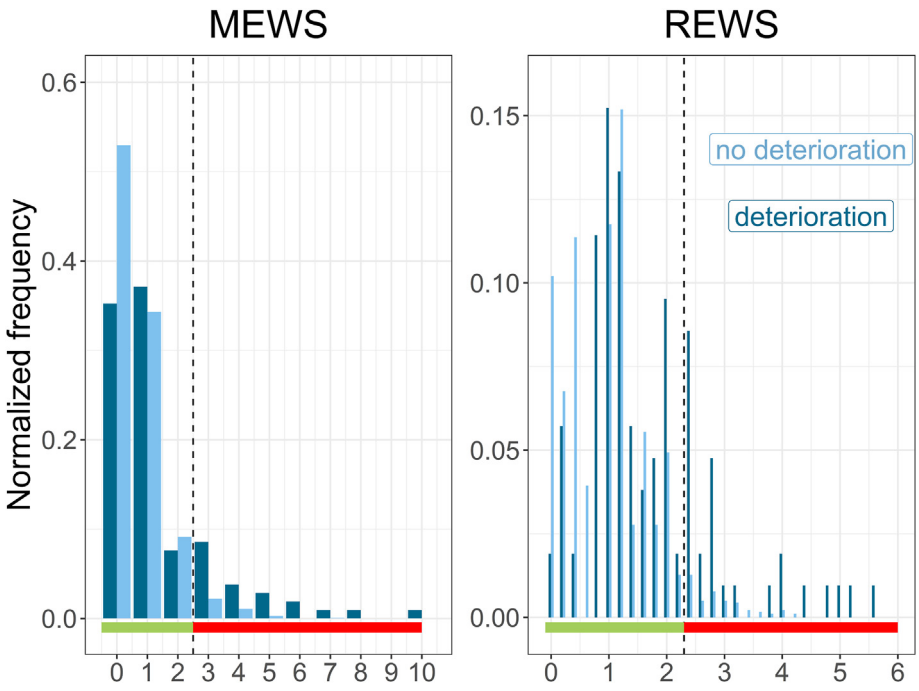

Fig. 3. Normalized histograms for the Modified Early Warning Score and Remote Early Warning Score. The colored bars show the distribution of the Early Warning Scores, and the dashed lines indicate the cutoff values. For the MEWS, the cutoff is placed at ≥3 based on local protocols. The cutoff value for the REWS is placed at ≥2.4 to match the number of true positives classified by the MEWS.

impact on the net workload, the 1816 automatically and remotely performed spotchecks that did not require follow-up checks by the nurse would amply compensate for the 7 extra checks that exceeded the EWS threshold. While the exact amount of time saved is highly dependent on the future monitoring protocols that are implemented in the general wards, replacing part of or all monitoring in the general ward with an automated early warning scoring system could result in reduction of the manual EWS measurement originating workload of nursing staff. With an estimate of at least 10 min of time needed to manually measure and report early warning score variables (partly) replacing this system with a remote monitoring system could free up time and allow nursing staff to focus on other important tasks rather than on the routine measurement of EWS scores.

Another notable finding is that in the present cohort, the REWS based on two continuously measured vital parameters had a comparable AUC as the standard MEWS assessments consisting of seven vital parameters. A previous study evaluating early warning scores supported the finding that HR and RR were the most important predictors for deterioration among the vital parameters in the MEWS and that simplification of the number of parameters could be justified [27].

In the literature, only a few studies have evaluated wearable sensors in a clinical setting [11,12,16]. Previous clinical studies investigated comparable wearables in similar populations. Although different outcomes were analyzed, all studies reported promising results of the wearable device in terms of the detection of adverse events, decreased hospital length of stay, and 30-day readmission rate [11,12]. Alternative types of continuous sensors have also been studied, and a study using continuous SpO2 sensor measurements demonstrated improved detection of deteriorating patients [16]. Continuous monitoring using a mattress sensor in the general ward decreases the length of hospital stay and reduces medical emergency rates [9]. Although promising, the authors encountered significant limitations in the use of such a system when the patients gained mobility.

The present study aimed to assess the performance of REWS compared to MEWS, and other trials comparing continuous measurements with intermittent spotchecks also found promising results. Weenk and colleagues. compared continuous measurements of vital parameters with MEWS and found good agreement between both types of data collection, suggesting that the MEWS system might be replaced [28]. However, they did not report the outcomes or predictive values, complicating comparison to our findings. A different and more personalized approach was analyzed by Keim and colleagues, who investigated the deviations of vital parameters from a patient's baseline using continuous measurements compared to intermittent measurements [29]. They concluded that continuous monitoring could provide alert thresholds as the degree of change from a patient's own baseline.

The strengths of the present study include its prospective design and the real-world setting in a clinically relevant patient population. The wearable sensor was compared with current clinical practice, while allowing unobtrusive automated EWS assessments. Another potential advantage of wearable monitoring is that patients can move freely, which is important for early post-operative recovery. Furthermore, sleep disturbance for the measurement of vital parameters by the nursing staff is reduced.

A limitation of the present study is that it was performed at a single center and investigated the local performance of the EWS, which might impair external validity. Second, the primary outcome consisted of heterogeneous deteriorations. This heterogeneity challenged the definition of the forecast window and generalizability of patterns in vital parameters before deterioration occurred. A 24 h forecast window length was chosen because changes in vital parameters can be found up to 24 h prior to severe deterioration, and a forecast window of 24 h is commonly used in the literature [30]. However, mild deterioration is assumed to impair vital parameters less than severe complications, and their detection could benefit from a shorter forecast window. Additionally, a shorter

forecast window could align with the length of working shifts, which are typically 8 h in most hospitals. Third, the population size did not allow splitting into training and validation datasets. Therefore, the performance of the proposed REWS cutoff value should be evaluated using an independent dataset. Fourth, the analysis was challenged by unavailable wearable data in 15 patients. The unavailability of wearable data is caused by preprocessing issues, such as corruption of the internal flash drive of the device and problems with data synchronization. This cause of missing data will be mitigated in future introduction of the wearable device into routine clinical practice and/or its use in follow up trials, as the wirelessly sent data will be used instead of data extraction from internal memory.

The present trial is an important first step toward remote monitoring and wearable-based decision support for patients in general wards. This trial provides a ground for larger scale follow up trials that are needed to optimize and validate protocols for the use of wearable based EWS in clinical practice. In the future, wearable based remote monitoring could be further explored for home monitoring as well. The present analysis focused on the comparison of wearable data collected at the same time as the routine MEWS assessment. However, as wearable patches continuously measure vital parameters, follow-up trials can aim to develop novel, continuous wearable-based early warning algorithms. Considering the current sensitivity of both EWS systems for the detection of deteriorations it will be interesting to assess whether the use of continuous vital parameters can lead to better and/or earlier identification of deteriorating patients.

5. Conclusions

A remote early warning scoring system based on vital parameters measured by a wearable patch offers a predictive performance comparable to that of the MEWS in patients following major abdominal cancer surgery. The wearable patch detected the same number of true deteriorations without requiring manual spotchecks. Overall, the present study strengthens the idea that remote continuous monitoring with a wearable accelerometer sensor patch has potential as a clinical decision support tool while improving patient mobility and comfort as well as workload of nursing staff.

Funding

The trial was partly financed by a grant from the Rijksdienst voor Ondernemend Nederland (Grant No: RVO ITEA 161006).

CRediT authorship contribution statement

Jonna A. van der Stam: Conceptualization, Data curation, Funding acquisition, Formal analysis, Writing – review & editing, Quality control of data and algorithms, Data analysis and interpretation, statistical analysis, Manuscript preparation, Manuscript editing. **Eveline H.J. Mestrom:** Conceptualization, Data curation, Funding acquisition, Formal analysis, Writing – review & editing, study design, data acquisition, Quality control of data and algorithms, Data analysis and interpretation, statistical analysis, Manuscript preparation, Manuscript editing. **Simon W. Nienhuijs:** Conceptualization, Data curation, Funding acquisition, Formal analysis, Writing – review & editing, study design, data acquisition, Data analysis and interpretation, Manuscript Review & Editing. **Ignace H.J.T. de Hingh:** Conceptualization, Data curation, Funding acquisition, Formal analysis, Writing – review & editing, study design, data acquisition, Data analysis and interpretation, Manuscript Review & Editing. **Arjen-Kars Boer:** Conceptualization, Data curation, Formal analysis, Writing – review & editing, Data analysis and interpretation, Manuscript Review & Editing. **Natal A.W. van Riel:** Data curation, Formal analysis, Writing – review & editing, Data analysis and interpretation, Manuscript Review & Editing. **Wim Verhaegh:** Data curation, Formal analysis, Writing – review & editing, Quality control of data and algorithms, Data analysis and interpretation, statistical analysis, Manuscript Review & Editing, Data curation, Formal analysis, Writing – review & editing, Quality control of data and algorithms, Data analysis and interpretation, statistical analysis, Manuscript Review & Editing. **Volkher Scharnhorst:** Conceptualization, Data curation, Formal analysis, Writing – review & editing, study design, Data analysis and interpretation, Manuscript Review & Editing. **R. Arthur Bouwman:** Conceptualization, Data curation, Formal analysis, Writing – review & editing, study design, Data analysis and interpretation, Manuscript Review & Editing.

Declaration of competing interest

Arthur Bouwman acts as a clinical consultant for Philips Research and Koen de Groot and Wim Verhaegh are employees of Philips Research. The other authors declare no potential conflicts of interest.

Acknowledgments

The authors would like to thank Alberto Bonomi for contributing to the data pre-processing and providing feedback on the manuscript and Linda Eerikainen for contributing to the data preprocessing.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejso.2022.08.034.

References

[1] Pearse RM, et al. Mortality after surgery in Europe: a 7 day cohort study. Lancet 2012;380:1059–65.
[2] Noordzij PG, et al. Postoperative mortality in The Netherlands: a population-based analysis of surgery-specific risk in adults. Anesthesiology 2010;112:1105–15.
[3] Buist MD, et al. Recognising clinical instability in hospital patients before cardiac arrest or unplanned admission to intensive care. A pilot study in a tertiary-care hospital. Med J Aust 1999;171:22–5.
[4] Hillman KM, et al. Duration of life-threatening antecedents prior to intensive care admission. Intensive Care Med 2002;28:1629–34.
[5] Churpek MM, et al. Using electronic health record data to develop and validate a prediction model for adverse outcomes in the wards. Crit Care Med 2014;42:841–8.
[6] Cuthbertson BH, et al. Can physiological variables and early warning scoring systems allow early recognition of the deteriorating surgical patient? Crit Care Med 2007;35:402–9.
[7] Gerry S, et al. Early warning scores for detecting deterioration in adult hospital patients: systematic review and critical appraisal of methodology. BMJ 2020;369:m1501.
[8] Smith ME, et al. Early warning system scores for clinical deterioration in hospitalized patients: a systematic review. Ann Am Thorac Soc 2014;11:1454–65.
[9] Brown H, et al. Continuous monitoring in an inpatient medical-surgical unit: a controlled clinical trial. Am J Med 2014;127:226–32.
[10] Leenen JPL, et al. Current evidence for continuous vital signs monitoring by wearable wireless devices in hospitalized adults: systematic review. J Med Internet Res 2020;22:e18636.
[11] Breteler MJM, et al. Are current wireless monitoring systems capable of detecting adverse events in high-risk surgical patients? A descriptive study. Injury 2020;51(Suppl 2):S97–s105.
[12] Downey C, et al. Continuous versus intermittent vital signs monitoring using a wearable, wireless patch in patients admitted to surgical wards: pilot cluster randomized controlled trial. J Med Internet Res 2018;20:e10802.
[13] Escobar GJ, et al. Automated identification of adults at risk for in-hospital clinical deterioration. N Engl J Med 2020;383:1951–60.

[14] Subbe CP, et al. Validation of a modified early warning score in medical admissions. QJM 2001;94:521–6.
[15] Mestrom E, et al. Implementation of an automated early warning scoring system in a surgical ward: practical use and effects on patient outcomes. PLoS One 2019;14:e0213402.
[16] Clifton L, et al. Predictive monitoring of mobile patients by combining clinical observations with data from wearable sensors. IEEE J Biomed Health Inform 2014;18:722–30.
[17] Cho KJ, et al. Detecting patient deterioration using artificial intelligence in a rapid response system. Crit Care Med 2020;48:e285–9.
[18] Jacobs F, et al. Reliability of heart rate and respiration rate measurements with a wireless accelerometer in postbariatric recovery. PLoS One 2021;16:e0247903.
[19] van der Stam JA, et al. Accuracy of vital parameters measured by a wearable patch following major abdominal cancer surgery. Eur J Surg Oncol 2021;48(4):917–23.
[20] Dindo D, et al. Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. Ann Surg 2004;240:205–13.
[21] Robin X, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinf 2011;12:77.
[22] DeLong ER, et al. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44:837–45.
[23] Simel DL, et al. Likelihood ratios with confidence: sample size estimation for diagnostic test studies. J Clin Epidemiol 1991;44:763–70.
[24] Collett D. Modelling binary data. Boca raton Florida. Chapman & Hall/CRC; 1999.
[25] Team R. RStudio: integrated development environment for R. In: P. RStudio; 2021. City.
[26] Team RC. A language and environment for statistical computing. In: R. F. f. S. Computing; 2021. City.
[27] Churpek MM, et al. Multicenter comparison of machine learning methods and conventional regression for predicting clinical deterioration on the wards. Crit Care Med 2016;44:368–74.
[28] Weenk M, et al. Wireless and continuous monitoring of vital signs in patients at the general ward. Resuscitation 2019;136:47–53.
[29] Keim-Malpass J, et al. Towards development of alert thresholds for clinical deterioration using continuous predictive analytics monitoring. J Clin Monit Comput 2020;34:797–804.
[30] Kollef MH, et al. A randomized trial of real-time automated clinical deterioration alerts sent to a rapid response team. J Hosp Med 2014;9:424–9.