BMJ Open

Wireless monitoring of high-risk patients using a wearable patch sensor: a clinical validation study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-020162
Article Type:	Research
Date Submitted by the Author:	17-Oct-2017
Complete List of Authors:	Breteler, Martine; Universitair Medisch Centrum Utrecht, Department of Anesthesiology Huizinga, Erik; Universitair Medisch Centrum Utrecht, Anesthesiology van Loon, Kim; Univ Med Ctr Utrecht, Anesthesiology Leenen, Luke; University Medical Centre Utrecht, Trauma Dohmen, Daan; FocusCura Kalkman, Cor; UMC Utrecht, Anesthesiology Blokhuis, Taco; Universitair Medisch Centrum Utrecht, Trauma Surgery
Primary Subject Heading:	Surgery
Secondary Subject Heading:	Nursing
Keywords:	Telemedicine < BIOTECHNOLOGY & BIOINFORMATICS, Vital signs, Wireless technology, Remote monitoring, Continuous monitoring

Wireless monitoring of high-risk patients using a wearable patch sensor: a clinical validation study

M.J.M. Breteler[1,3], E. Huizinga[1], K. van Loon[1], L.P.H. Leenen[2], D.A.J. Dohmen[3], C.J. Kalkman[1], T. J. Blokhuis[2,4]

[1]Department of Anesthesiology, University Medical Center Utrecht, Utrecht University, The Netherlands
[2]Department of Trauma Surgery, University Medical Center Utrecht, Utrecht, The Netherlands
[3]FocusCura, Driebergen-Rijsenburg, The Netherlands
[4]Department of Surgery, Maastricht University Medical Center, Maastricht, The Netherlands

Corresponding author:
Martine Breteler
University Medical Center Utrecht, Mailstop Q.04.2.313
P.O. Box 85500, 3508 GA Utrecht, The Netherlands
Email: m.j.m.breteler@umcutrecht.nl
Phone: +31 6 47306841

Word count: 3733

ABSTRACT

Background and objectives: Intermittent vital signs measurements are the current standard on hospital wards, typically recorded once every 8h. Early signs of deterioration may therefore be missed. Recent innovations have resulted in 'wearable' sensors, which may capture patient deterioration at an earlier stage. The objective of this study was to determine whether a wireless 'patch' sensor is able to reliably measure respiratory and heart rate continuously in high-risk surgical patients. The secondary objective was to explore the potential of the wireless sensor to serve as a safety monitor.

Design: In an observational methods comparisons study, patients were measured with both the wireless sensor and bedside routine standard for at least 24 hours.

Setting: University teaching hospital, single center.

Participants: Twenty-five postoperative surgical patients admitted to a step-down unit.

Outcome measures: Primary outcome were limits of agreement and bias of heart rate and respiratory rate. Secondary outcome measures were sensor reliability, defined as time until first occurrence of data loss.

Results: 1568 hours of vital signs data were analyzed. Bias and 95% limits of agreement for heart rate were -1.1 (-8.8 to 6.5) beats per minute. For respiration rate, bias was -2.3 breaths per minute with wide limits of agreement (-15.8 to 11.2 breaths per minute). Median filtering over a 15 minute period improved limits of agreement of both respiration and heart rate. Sixty-three% of the measurements were performed without data loss greater than two minutes. Overall data loss was limited (6% of time).

Conclusions: The wireless sensor is capable of accurately measuring heart rate, but accuracy for respiratory rate was outside acceptable limits. Remote monitoring has the potential to contribute to early recognition of physiological decline in high-risk patients. Future studies should focus on the ability to detect patient deterioration on low care environments and at home after discharge.

Keywords: telemedicine; vital signs; wireless technology, remote monitoring, continuous monitoring

STRENGTHS AND LIMITATIONS OF THIS STUDY

- This is the first study that investigated the reliability of measuring vital signs continuously for more than 1568 hours with a wearable patch sensor in actual high-risk patients.
- To identify agreement between the two methods for continuous monitoring the Bland and Altman method is not the most appropriate. It measures bias and precision at a certain time point, but cannot inform about the trending ability of the monitoring system. The latter is extremely important since the goal of monitoring is timely recognition of abnormal vital sign patterns.
- The agreement between the two methods improved after (median) filtering the data. Provided one is willing to accept an update rate of once every 15 minutes, such filtering increases the proportion of epochs with reliable measurements and reduces the rate of false positive alarms.
- We used an 'ICU-grade' patient monitoring system as reference. However, our results showed that this method cannot be considered a 'gold' standard for respiratory rate monitoring.

INTRODUCTION

While technological advances have resulted in numerous new diagnostic tools and therapeutic innovations, we are still not able to timely recognize patient deterioration on general hospital wards[1,2]. This contributes to avoidable cardiopulmonary arrest, unplanned admission to the intensive care unit (ICU), an increase in hospitalization costs and detrimental effects on quality of life[3-7]. To timely detect patient deterioration, we may benefit from technical solutions that can track patients' vital signs continuously.

Intermittent vital signs measurements, typically once every nurse shift of eight hours, are the current routine monitoring practice on general hospital wards. As a result, patient deterioration in between measurements can be easily missed. In an attempt to improve the detection of patient deterioration, Early Warning Scoring (EWS) protocols and Medical Emergency Teams (METs) have been implemented in most hospitals around the globe. However, failure-to-rescue events continue to occur even with these systems in place[8,9]. This phenomenon is also known as failure of the 'afferent limb' of the EWS system[10-12].

Alongside the attempts to improve detection of patient deterioration on the ward, there is a trend to reduce the duration of hospitalization by discharging patients home early, for example in 'enhanced recovery after surgery' (ERAS) programs[13-15]. Once a patient is back home, EWS protocols and vital signs monitoring are no longer available. Recovery within the patient's own home environment has many advantages, but unavoidably some surgical complications will now become first manifest in the home setting. This increases the risk that patient deterioration will be recognized too late.

The majority of adverse events are preceded by a significant period of change in vital signs[16-20]. Early recognition of the deteriorating patient might be improved if continuous remote monitoring would become available for at-risk patients in 'low care' environments such as the regular hospital ward or in the first few 'critical' days at home after hospital discharge[21,22]. Recent technological innovations have resulted in lightweight 'wearable' wireless sensors capable of recording and transmitting several vital signs such as heart rate (HR), respiration rate (RR), temperature and patient movement. While the majority of the wearable sensors is strictly 'consumer-grade', some manufacturers have obtained CE and/or FDA approval for use in clinical environments. However, validity and accuracy of these so called 'medical-grade' wearables has not been assessed in real clinical environments, with most of the few validation studies performed in healthy volunteers or limited to short periods of measurements in patients with comorbid conditions[23-25].

The objective of this study was to determine whether a wireless adhesive 'patch' sensor is able to reliably measure RR and HR continuously in patients after high-risk surgery. We aimed to verify whether wireless sensor technology is robust and capable of detecting physiological trend patterns in high-risk patients before introducing wireless vital signs monitoring into clinical practice. A secondary objective was to explore the potential of the wireless sensor to serve as a safety monitor in clinical practice.

Materials & Methods

Study design and setting

We performed a methods comparisons study with an observational design in which patients were continuously monitored after high-risk surgery during the initial days of recovery at the surgical step-down unit (SDU) of the University Medical Center Utrecht, the Netherlands, a large academic hospital. Formal approval for this study was obtained from the local ethical committee (number 15/550).

Study participants

Postoperative patients were asked to participate upon admission to the SDU if their expected stay was at least 24 hours. These patients were considered for enrolment because they represent a high-risk subset of surgical patients that is more prone to experience deterioration compared to patients on the general ward. Exclusion criteria were patients with an implantable cardiac device, an allergy to adhesives, wound or skin lesion near the application site, or inability to give informed consent. After written informed consent was obtained, researchers applied the sensor to the patient's chest to start recording vital signs for one to three days using the wearable sensor and the routine monitoring system described hereafter.

Description of the wireless wearable sensor

The HealthPatch® MD (VitalConnect Inc, San Jose, CA, USA) is a medical-grade lightweight, wireless and wearable adhesive biosensor that measures a number of vital signs continuously: single-lead electrocardiogram (ECG), HR, heart rate variability, RR, skin temperature, body posture and step count. It was designed to facilitate long-term remote monitoring of vital signs and activity metrics within the hospital environment as well as in the post-discharge period at home. The sensor consists of a disposable adhesive patch that houses two ECG electrodes, a thermistor and a zinc-air coin-cell battery. The reusable sensor module contains a triaxial accelerometer and Bluetooth Low Energy (BLE) transceiver (see Appendix 1). The patch can be applied on the patient's chest and measures vital signs continuously up to three days (four days if continuous transmission of its single lead ECG waveform is disabled). The module processes the incoming signals and transmits the data via BLE to a relay device (for this study we used an iPad mini (Apple, Inc. Cupertino, CA, USA) with the 'Healthwatch' (VitalConnect Inc, San Jose, CA, USA) mobile application. This application can display vital signs data in real-time for research purposes, but was not designed to be used as clinical monitoring system. Also, near real-time data can be viewed on the Healthwatch Web cloud based server, to monitor long-term trends. Patient idenficitation information was not entered on the mobile device to ensure privacy protection.

Although the wireless sensor can also measure position and movement (acceleration), in this study we only focused on assessing the accuracy and reliability of RR and HR monitoring. The sensor calculates HR using analysis of the single-lead ECG. The algorithm is based on automated detection of QRS complexes from the ECG waveforms. RR is derived from the combined information from three sources: an embedded algorithm uses a weighted average of two characteristics of the ECG signal: 1.) QRS amplitude modulation and 2.) respiratory sinus arrhythmia (RSA); both ECG-derived signals

change during inspiration and expiration[26], and the algorithm uses 3.) accelerometer data produced by chest movement during respiration [27]. Both HR and RR are updated every 4 s and the manufacturer states an accuracy of ± 3 breaths per minute (breaths/min), in the range of 4 to 42 breaths for RR. The stated accuracy of HR is ± 5 beats per minute (beats/min) or 10% (whichever is greater), in the range of 30-200 beats/min.

Description of the bedside routine standard

HR and RR of patients were continuously monitored with the wearable sensor and simultaneously with a multiparameter bedside monitoring system designed for use in ICU's and Operating Rooms (XPREZZON, Spacelabs Healthcare, Snoqualmie, Washington DC, USA) which served as the reference monitor. This reference uses ECG for HR detection and measures RR by thoracic impedance pneumography. It also measures invasive blood pressures as well as arterial oxygen saturation via pulse oximetry.

Signal analysis

The raw data transmitted by the sensor containing the measurements and their associated time stamps were retrieved in CSV format. Data were stored and processed using Matlab (The MathWorks Inc., Natick, MA, USA). Empty or invalid data ('not-a-number') were removed to obtain continuous 2D vectors of vital sign samples with their corresponding time stamps. Data reports were automatically retrieved from the reference monitor. These contained vital signs data sampled once per minute (i.e., one measurement was saved and transmitted every minute). The sensor data, originally transmitted once every four seconds was therefore resampled to once per minute to produce paired data points with the reference monitor. Furthermore, sensor data and reference data were synchronized to ensure alignment of their respective time stamps. No artifact removal was applied to the data before analysis. Besides the analysis on vital signs data transmitted every minute, a median filter over a 15-minute period was applied to study the effect on HR and RR outliers.

Outcomes

The primary outcome was bias and precision (95% Limits of Agreement; LoA) of HR and RR of the wireless sensor compared to the bedside monitor. This reference standard reports an accuracy for HR of ±1% or 3 beats/min (whichever is greater) and an accuracy of ±5% or 1 breaths/min (whichever is greater) for RR[28]. We considered HR and RR to be acceptable for clinical purposes if within ± 10% of the reference monitor or ± 3 breaths/min or ± 5 beats/min[29]. A secondary endpoint was the reliability of detecting true critical clinical conditions such as bradycardia (HR < 50 beats/min), tachycardia (HR > 100 beats/min), bradypnea (RR < 12 breaths/min) and tachypnea (RR > 20 breaths/min)[30]. Another secondary outcome was the reliability defined as time until the first occurrence of data loss (defined as a duration of a gap within the data of 2 minutes, 15 minutes, 1 hour or 4 hours) and the overall amount of data loss from various causes.

Statistical analysis

The series of observation pairs of HR and RR measurements (one data point every minute) derived from the wireless sensor and the reference monitor were compared using the Bland and Altman

Method for repeated measurements[31]. This method was used to account for within-subject variation by correcting for the variance of differences between the average differences across patients and the number of measurements per patient. The mean difference ('bias') between the wireless sensor and reference monitor and the 95% LoA (± 1.96 SD) was determined for both the HR and RR data. In addition, a Clarke Error Grid analysis was conducted to specify the clinical accuracy and the consequences for clinical decision making[32]. The time (hours) to first occurrence of data loss was analysed with Kaplan-Meier survival plots.

A power calculation was not feasible due to the lack of preliminary data with these continuous monitoring systems. Therefore, we aimed to analyze data of at least 25 patients - each with multiple hours of continuous data – which is sufficient to evaluate the validity of the wireless patch sensor.

Results

From September 2015 to September 2016, a total of 33 postoperative patients entered the study. Data from the reference monitor were missing for eight patients due to technical issues with retrieving the data from the monitor's on-board memory, resulting in a total of 1568 hours of monitoring time with the reference monitor and 1702 hours of vital signs monitoring from the wireless sensor. Therefore, measurement pairs of 25 patients were available for agreement analysis. On average, 62 hours remained per patient for further analysis. The range of total monitoring time per patient varied from 12 to 124 hours. Table 1 summarizes patient characteristics and surgical procedures .

Table 1. Patient characteristics (n = 25)

Characteristic	**Value**
Male gender - no (%)	18 (72)
Age (years) - median (IQR) [range]	63.0 (57.8-71.5) [23.0 – 77.0]
Body Mass Index (BMI; kg/m^2)* - median (IQR) [range]	26.2 (24.2-29.4) [17.2 - 40.2]
ASA score	
1-2 (%)	8 (32)
3-4 (%)	17 (68)
Comorbidities	
Hypertension - no (%)	8 (32)
Cardiovascular disease - no (%)	9 (36)
COPD - no (%)	3 (12)
Diabetes - no (%)	3 (12)

IQR: interquartile range. COPD: Chronic Obstructive Pulmonary Disease. * BMI of one patient was missing

Example of a patient measurement

Figure 1 shows the HR and RR trend during the first four postoperative days of a 60-year-old male patient with extensive cardiac and vascular comorbidities, after an open nephrectomy procedure for a

suspected carcinoma in situ. Three events can be recognized: 1.) a sudden HR increase on Thursday evening; later diagnosed as atrial fibrillation, 2.) an episode of bradycardia on Saturday afternoon, and 3.) mild tachycardia and a subtle increase in RR starting Sunday afternoon, caused by bleeding from an aortic branch artery. After coil embolization, the patient was readmitted to the Intensive Care Unit (ICU). This example illustrates agreement between HR and RR measurements recorded with the wireless sensor and the reference standard. Note that RR derived from the reference monitor was highly variable compared with RR from the wireless sensor (Figure 1).

Heart rate

Table 2 shows bias and precison (95% LoA) from comparisons between the wireless sensor and the reference standard. For analysis, 55565 minutes (926 h) of HR measurement pairs were available. The mean difference (bias) in HR was -1.1 beats/min (reference standard minus sensor) with a 95% LoA of -8.8 to 6.5 beats/min. Applying a 15 min median filter resulted in a narrower 95% LoA of -5.7 beats/min to 3.2 beats/min (3986 minutes available for analysis). Bland and Altman plots for the complete and filtered HR datasets are shown in Figure 2a and 2b, respectively.

Table 2 Bland Altman analysis of wireless heart rate and respiratory rate versus reference monitor in postoperative patients

Parameter	Number of measurement pairs	Bias	SD	Lower 95% LoA[b]	Upper 95% LoA[b]
Complete dataset					
Heart rate	55565	-1,1	3,83	-8,8	6,5
Respiratory rate	56674	-2,3	6,8	-15,8	11,2
Filtered dataset[a]					
Heart rate	3986	-1,2	2,2	-5,7	3,2
Respiratory rate	4001	-2,4	4,2	-10,8	5,9

[a] *Dataset after applying a median filter*
[b] *LoA limits of agreement*

Respiratory rate

The mean difference between the reference monitor and the wireless sensor was -2.3 breaths/min with wide levels of agreement (95% LoA: -15.8 to 11.2 breaths/min). The agreement between both methods improved after applying a 15 minute median filter, resulting in a 95% LoA of -10.8 to 5.9 breaths/min. Bland and Altman plots for the complete and filtered RR dataset are displayed in Figures 3a and 3b. Most 'high RR' outliers originated from the reference monitor and were observed in the higher RR range. This is also shown in Figure 1 where RR measurements derived from the reference monitor showed a higher variation compared with RR derived from the wireless sensor.

Diagnostic accuracy of the wireless sensor

Because of the relatively long monitoring time per patient and the high-risk population, we were able to capture several instances of bradycardia, tachycardia, bradypnea and tachypnea. The incidence of bradycardia was rare, a HR below 50 was present in only 2% of all HR measurements in the complete dataset. Tachycardia, bradypnea and tachypnea occurred more frequently in 14%, 15% and 34% of

cases respectively. Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) are shown in Table 3. After applying median filtering, sensitivity and specificity of all episodes with abnormal HR and RR improved.

Table 3 Diagnostic accuracy for bradypnea, tachypnea, bradycardia, tachycardia

	True positives (%)	False positives (%)	True negatives (%)	False negatives (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
Bradycardia								
Complete dataset	824 (71)	363 (1)	55148 (99)	339 (1)	71	99	69	99
Filtered dataset[a]	24 (72)	0 (0)	3953 (100)	9 (0)	73	100	100	100
Tachycardia								
Complete dataset	7111 (90)	1490 (3)	47321 (97)	752 (2)	90	97	83	98
Filtered dataset[a]	496 (98)	65 (2)	3413 (98)	12 (0)	98	98	88	100
Bradypnea								
Complete dataset	2113 (24)	562 (1)	47438 (99)	6561 (12)	24	99	79	88
Filtered dataset[a]	134 (33)	8 (0)	3587 (100)	272 (7)	33	100	94	93
Tachypnea								
Complete dataset	16210 (84)	14602 (39)	22694 (61)	3168 (12)	84	61	53	88
Filtered dataset[a]	1225 (94)	1029 (38)	1668 (62)	79 (5)	94	62	54	95

[a] *Dataset after applying a median filter*

The Clarke error grid analyses of the filtered datasets are shown in Figure 4a and 4b; it shows that 100% of the HR measurements and 99% of RR measurements are within region A or B, respectively within 20% of the reference measurement, or outside 20% of the reference but not leading to unneccesary treatment. None of the HR measurements were in region, B, C or E, which means that none of the measurements would lead to failure to treat, unnecessary treatment or confusion between bradycardia and tachycardia. Very few of the RR measurements (≤ 1 %) were within region C, D or E, indicating a potentially dangerous failure to apply the right treatment.

Technical performance

HR and RR data were recorded for the majority of the time (94%), from 36 sensors in 33 patients (Table 4). Nineteen (53%) wireless patient monitoring series had complete uninterrupted data, but in 17 patients there was data loss (ranging from 8s data loss to 60h). Most found sensor failure to be caused by a sawtooth pattern of the battery level from inadvertent covering of the air opening for the zinc-air battery resulting in measurement gaps.

Table 4 Amount of data loss within all wireless sensor measurements

	Time loss (hours:minutes:seconds)			
	Total loss (%)[a]	Mean loss	Minimum loss	Maximum loss

All wireless sensor measurements (n=36)	101:15:24 (5.9%)	02:48:45	00:00:00	59:59:08
Only wireless sensor measurements with any data loss (n=17)	101:15:24 (15.5%)	05:57:22	00:00:08	59:59:08

[a] *Data loss is defined as time without data as percentage of the total time measured*

Figure 5 shows the survival analysis for 'time to first failure' in data transmission of the sensor. Using a threshold (maximum duration of a gap in the data) of two minutes, this analysis showed that 63% of the wireless sensor measurements in patients were performed without data loss greater than two minutes. A gap duration of one hour resulted in 79% of sensor measurements without data loss greater than an hour.

Discussion

We studied the performance of a wearable wireless sensor to measure HR and RR continuously in high-risk postoperative patients. The results show that this sensor (Healthpatch MD) can accurately measure HR with a deviation within 10% of the reference standard. In contrast, the accuracy for RR was outside the limit range we considered acceptable. However, this finding may be due to frequent outliers and clinically implausible variability of RR values provided by the reference monitor. Median filtering of both signals over a 15 min period resulted in a reduction of the number of RR measurement pairs outside the acceptable LoA and an improvement of LoA. Overall, data loss was limited with HR and RR measurements 94% of the time available.

Strenghts

To the best of our knowledge, this is the first clinical study that investigated the reliability and accuracy of continuous vital signs monitoring using a wearable wireless patch sensor for several days in postoperative surgical patients at a step-down unit. While some of the earlier feasibility studies that used continuous wireless monitoring were performed in hospital environments[24,29], most of the results were actually obtained under controlled laboratory conditions[25,27,33]. Other studies used intermittent nurse observations on the ward as the only reference[34,35]. Although these studies showed satisfactory agreement, they cannot validate the continuous performance of the retrieved data with the wireless devices. Moreover, these wireless monitoring devices are not intended to deliver 'spot' readings for EWS, i.e., their use was evaluated for a purpose outside the intended scope of use. Consequently, a drawback of these study designs was the inability to validate continuous HR and RR measurements of new remote monitoring devices in between nurse observations. In the current study reference HR and RR were measured continuously in a clinical setting.

Limitations

The results of the present study confirm the difficulty of accurate continuous RR monitoring. However, we must consider the limitations of the reference standard[36,37]; accuracy of thoracic impedance RR measurements can be influenced by many factors independent of breathing, such as patient movement, talking and coughing. Impedance artefacts could explain the high number of false negatives (i.e., missed bradypnea), resulting in low sensitivity. In addition, the HealthPatch sensor was

not designed to indicate respiration rates < 5 breaths/min. Nonetheless, a progressive slowing of breathing rate may still be identified and used as indicator of vital instability, for example to recognize life-threatening opioid-induced respiratory depression.

Impedance technique which is the current bedside routine standard for continuous measurement of HR and RR in most hospitals today and therefore clinically relevant. However, it cannot be considered a gold standard for RR measurement. While capnography is widely regarded as a 'true' gold standard for RR, it has several drawbacks for continuous unsupervised respiratory monitoring in spontaneously breathing patients, since its nasal cannula can be easily dislodged, leading to incomplete data and a high number of false positive 'apnea' alarms[38].

Filtering

Applying a median filter over 15 minute data epochs improved reliability of HR and RR by removing outliers, for example, a transient very high RR caused by a movement artefact. This is appropriate since these transient artefacts (i.e., RR >45 breaths/min) are extremely unlikely from a physiological perspective. Although filtering effectively eliminates such artifacts, this comes at the expense of reducing the number of available measurements and the reduced ability to detect sudden changes in vital signs (e.g., cardiac arrest). An alternative might be to use a 'moving' median filter to keep the update rate at once every minute or once every 2 min. On the other hand, improved eliminaton of outliers could result in a higher proportion of epochs with reliable HR and RR resulting in lower false positive alarms. The latter is extremely important if this remote monitoring system is to be clinically deployed on the ward or at home. Furthermore, continuous remote monitoring on the ward with a reduced frequency (i.e., once every 15 minutes) still provides much more information regarding the patient's vital signs than the current intermittent monitoring practice, where vital signs are usually only observed once every 8h nurse shift.

In case an alarm is generated by the remote monitoring system, a nurse can personally check on the patient and correct if the cause of the alarm was not related to a change in the patient's medical condition. However, it must be realized that a large number of false alarms is very disruptive on the general ward, especially when there is a low nurse to patient ratio (e.g., at night). This may even decrease patient safety by taking away valuable nursing time from patients who are in real need of attention. We suggest that eliminating outliers to improve reliability and eliminate false positive alerts far outweighs the limited benefits of having 'continuous' vital signs data streams.

'Methods comparison' methodology and continuous monitoring

The goal of this study was to determine whether the wireless sensor is able to reliably measure RR and HR over time in postoperative patients. Although Bland and Altman analyses can reliably indicate bias and precision of 'spot measurements', it does not inform about the 'trending ability' of the monitoring system over time, while this is of ultimate importance to timely recognize abnormal vital sign patterns. This was also confirmed in the study of Churpek et al. that showed the added value of using trends of vital signs for detecting clinical deterioration on the wards[21]. In our study, the example case in Figure 1 clearly demonstrates the ability of the wireless sensor to detect important physiological

changes of a deteriorating patient over time, even while the limits of agreement for detecting bradypnea as determined by Bland Altman analyses were deemed not acceptable. Therefore, we wish to emphasize that accurate trend measurements (e.g., the ability to detect deterioration over time) is more important than just one accurate single measurement at one specific point in time.

Continuous monitoring on the general ward is still unknown territory. Future studies should therefore focus on the performance of wireless monitoring in patients at the general ward, including validation during periods of mobilization. Particular emphasis should be on the early detection of critical adverse events. However, the usability and patient perspective on remote monitoring are important to determine when creating the infrastructure of a complete remote monitoring solution.

Conclusion

Wireless continuous monitoring of HR and RR have the potential to contribute to early recognition of physiological decline in high-risk patients. However, the wireless sensor studied here needs further improvement to allow early recognition of progressive slowing of breathing patterns. Optimizing the signal processing algorithms to reduce artefacts and to prevent disruptive false alarms are necessary to be able to use wireless vital signs monitoring as a patient safety net.

Contributors

MB, CK, DD and TB were involved in the clinicial trial design and ethics submission. EH was involved in the planning and conduct of the study and drafting of the results. MB assisted in the conduct of the study, performed analysis of clinical data and drafted the paper. Throughout the study, LL, KL, CK and TB assisted in the conduct of the study. All authors reviewed and approved the final version of the manuscript.

Funding

This research was funded solely from institutional funds.

Competing interests

At the time of the study, MB was part-time employee of FocusCura (Health IT company, Driebergen-Zeist, The Netherlands) and DD is founder and CEO of FocusCura. No funding from FocusCura was obtained. VitalConnect Inc, the manufacturer of the HealthPatch MD sensor, neither had insight or influence on the analysis of the data or drafting of the manuscript.

Data sharing statement

Authors are planning upon producing further publications using this dataset. Afterwards, anonymized patient recordings will be available from the corresponding author upon request.

References

1 Fuhrmann L, Lippert A, Perner A, *et al.* Incidence, staff awareness and mortality of patients at risk on general wards.

Resuscitation 2008;**77**:325–30. doi:10.1016/j.resuscitation.2008.01.009

2 Kause J, Smith G, Prytherch D, *et al.* A comparison of Antecedents to Cardiac Arrests, Deaths and EMergency Intensive care Admissions in Australia and New Zealand, and the United Kingdom—the ACADEMIA study. *Resuscitation* 2004;**62**:275–82. doi:10.1016/j.resuscitation.2004.05.016

3 Sax FL, Charlson ME. Medical patients at high risk for catastrophic deterioration. *Crit Care Med* 1987;**15**:510–5.

4 Jonsson T, Jonsdottir H, Möller AD, *et al.* Nursing documentation prior to emergency admissions to the intensive care unit. *Nurs Crit Care* 2011;**16**:164–9. doi:10.1111/j.1478-5153.2011.00427.x

5 McQuillan P, Pilkington S, Allan A, *et al.* Confidential inquiry into quality of care before admission to intensive care. *BMJ* 1998;**316**:1853–8.

6 Calzavacca P, Licari E, Tee A, *et al.* The impact of Rapid Response System on delayed emergency team activation patient characteristics and outcomes—A follow-up study. *Resuscitation* 2010;**81**:31–5. doi:10.1016/j.resuscitation.2009.09.026

7 Ghaferi AA, Birkmeyer JD, Dimick JB. Variation in hospital mortality associated with inpatient surgery. *N Engl J Med* 2009;**361**:1368–75. doi:10.1056/NEJMsa0903048

8 Taenzer AH, Pyke JB, McGrath SP. A Review of Current and Emerging Approaches to Address Failure-to-Rescue. *Anesthesiology* 2011;**115**:421–31. doi:10.1097/ALN.0b013e318219d633

9 Ferraris VA, Bolanos M, Martin JT, *et al.* Identification of Patients With Postoperative Complications Who Are at Risk for Failure to Rescue. *JAMA Surgery* 2014;**149**:1103. doi:10.1001/jamasurg.2014.1338

10 Trinkle RM, Flabouris A. Documenting Rapid Response System afferent limb failure and associated patient outcomes. *Resuscitation* 2011;**82**:810–4. doi:10.1016/j.resuscitation.2011.03.019

11 Hammer JA, Jones TL, Brown SA. Rapid response teams and failure to rescue: one community's experience. *J Nurs Care Qual* 2012;**27**:352–8. doi:10.1097/NCQ.0b013e31825a8e2f

12 Hillman K, Chen J, Cretikos M, *et al.* Introduction of the medical emergency team (MET) system: a cluster-randomised controlled trial. *Lancet* 2005;**365**:2091–7. doi:10.1016/S0140-6736(05)66733-5

13 Pajarón-Guerrero M, Fernández-Miera MF, Dueñas-Puebla JC, *et al.* Early Discharge Programme on Hospital-at-Home Evaluation for Patients with Immediate Postoperative Course after Laparoscopic Colorectal Surgery. *Eur Surg Res* 2017;**58**:263–73. doi:10.1159/000479004

14 Roulin D, Donadini A, Gander S, *et al.* Cost-effectiveness of the implementation of an enhanced recovery protocol for colorectal surgery. *Br J Surg* 2013;**100**:1108–14. doi:10.1002/bjs.9184

15 Gustafsson UO, Scott MJ, Schwenk W, *et al.* Guidelines for Perioperative Care in Elective Colonic Surgery: Enhanced Recovery After Surgery (ERAS®) Society Recommendations. *World J Surg* 2012;**37**:259–84. doi:10.1007/s00268-012-1772-0

16 Buist M, Bernard S, Nguyen TV, *et al.* Association between clinically abnormal observations and subsequent in-hospital mortality: a prospective study. *Resuscitation* 2004;**62**:137–41. doi:10.1016/j.resuscitation.2004.03.005

17 Churpek MM, Yuen TC, Huber MT, *et al.* Predicting cardiac arrest on the wards: a nested case-control study. *Chest* 2012;**141**:1170–6. doi:10.1378/chest.11-1301

18 Cuthbertson BH, Boroujerdi M, McKie L, *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. doi:10.1097/01.CCM.0000254826.10520.87

19 Hollis RH, Graham LA, Lazenby JP, *et al.* A Role for the Early Warning Score in Early Identification of Critical Postoperative Complications. *Ann Surg* 2016;**263**:918–23. doi:10.1097/SLA.0000000000001514

20 Goldhill DR, White SA, Sumner A. Physiological values and procedures in the 24 h before ICU admission from the ward. *Anaesthesia* 1999;**54**:529–34.

21 Churpek MM, Adhikari R, Edelson DP. The value of vital sign trends for detecting clinical deterioration on the wards. *Resuscitation* 2016;**102**:1–5. doi:10.1016/j.resuscitation.2016.02.005

22 Watkins T, Whisman L, Booker P. Nursing assessment of continuous vital sign surveillance to improve patient safety on the medical/surgical unit. *Journal of Clinical Nursing* 2016;**25**:278–81. doi:10.1111/jocn.13102

23 Selvaraj N. Long-term Remote Monitoring of Vital Signs using a Wireless Patch Sensor. 2014;:1–4.

24 Hernandez-Silveira M, Ahmed K, Ang S-S, *et al.* Assessment of the feasibility of an ultra-low power, wireless digital patch for the continuous ambulatory monitoring of vital signs. *BMJ Open* 2015;**5**:e006606–6. doi:10.1136/bmjopen-2014-006606

25 Chan AM, Selvaraj N, Ferdosi N, *et al.* Wireless patch sensor for remote monitoring of heart rate, respiration, activity, and falls. *Conf Proc IEEE Eng Med Biol Soc* 2013;**2013**:6115–8. doi:10.1109/EMBC.2013.6610948

26 Yasuma F, Hayano J-I. Respiratory sinus arrhythmia: why does the heartbeat synchronize with respiratory rhythm? *Chest* 2004;**125**:683–90.

27 Chan AM, Ferdosi N, Narasimhan R. Ambulatory respiratory rate detection using ECG and a triaxial accelerometer. *Conf Proc IEEE Eng Med Biol Soc* 2013;**2013**:4058–61. doi:10.1109/EMBC.2013.6610436

28 Healthcare S. Data Sheet Ultraview SL. 2016.

29 Ben-Ari J, Zimlichman E, Adi N, *et al.* Contactless respiratory and heart rate monitoring: validation of an innovative tool. *J Med Eng Technol* 2010;**34**:393–8. doi:10.3109/03091902.2010.503308

30 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. doi:10.1016/j.resuscitation.2012.12.016

31 Bland JM, Altman DG. Agreement between methods of measurement with multiple observations per individual. *J Biopharm Stat* 2007;**17**:571–82. doi:10.1080/10543400701329422

32 Clarke WL, Cox D, Gonder-Frederick LA, *et al.* Evaluating clinical accuracy of systems for self-monitoring of blood glucose. *Diabetes Care* 1987;**10**:622–8.

33 Selvaraj N. Long-term Remote Monitoring of Vital Signs using a Wireless Patch Sensor. 2014;:1–4.

34 Weenk M, van Goor H, Frietman B, *et al.* Continuous Monitoring of Vital Signs Using Wearable Devices on the General Ward: Pilot Study. *JMIR Mhealth Uhealth* 2017;**5**:e91. doi:10.2196/mhealth.7208

35 Granholm A, Pedersen NE, Lippert A, *et al.* Respiratory rates measured by a standardised clinical approach, ward staff, and a wireless device. *Acta Anaesthesiol Scand* 2016;**60**:1444–52. doi:10.1111/aas.12784

36 Freundlich JJ, Erickson JC. Electrical impedance pneumography for simple nonrestrictive continuous monitoring of respiratory rate, rhythm and tidal volume for surgical patients. *Chest* 1974;**65**:181–4.

37 Wiklund L, Hök B, Ståhl K, *et al.* Postanesthesia monitoring revisited: frequency of true and false alarms from different monitoring devices. *J Clin Anesth* 1994;**6**:182–8.

38 Macknet M, Kimball-Jones P, Applegate R II, *et al.* Accuracy of a novel bioacoustic sensor for monitoring respiratory

rate. *European Journal of Anaesthesiology* 2007;**24**:34. doi:10.1097/00003643-200706001-00124

Figure titles and legends

Figure 1 Example of a patient that is being measured for four days continuously with the wireless sensor (red) and reference standard (blue). The upper panel shows heart rate in beats per minute, the lower shows respiratory rate in breaths per minute.

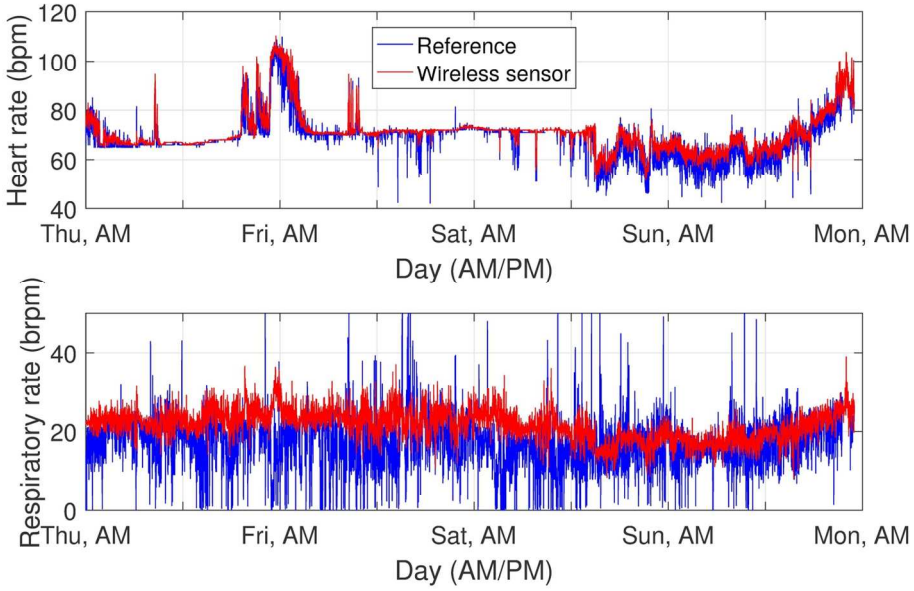
Figure 2 Bland and Altman plots for complete (a) and filtered (b) datasets for heart rate during admission at the surgical step-down unit with few (white) to many (dark red) measurement pairs

Figure 3 Bland and Altman plots for complete (a) and filtered (b) datasets for respiratory rate during admission at the surgical step-down unit with few (white) to many (dark red) measurement pairs

Figure 4 Clark error grid analysis to quantify clinical accuracy of the HR and RR measurements with the HealthPatch MD as compared to the reference monitor of the filtered dataset. Region (A) are points within 20% of the reference monitor, region (B) contains points outside 20% of the reference, but not leading to unnecessary treatment, region (C) are points leading to unnecessary treatment, region (D) indicats a potentially dangerous failure to detect bradycardia or tachycardia (a) or bradypnea or tachypnea (b), region (E) represents points where events are confused (e.g., bradycardia with bradypnea).

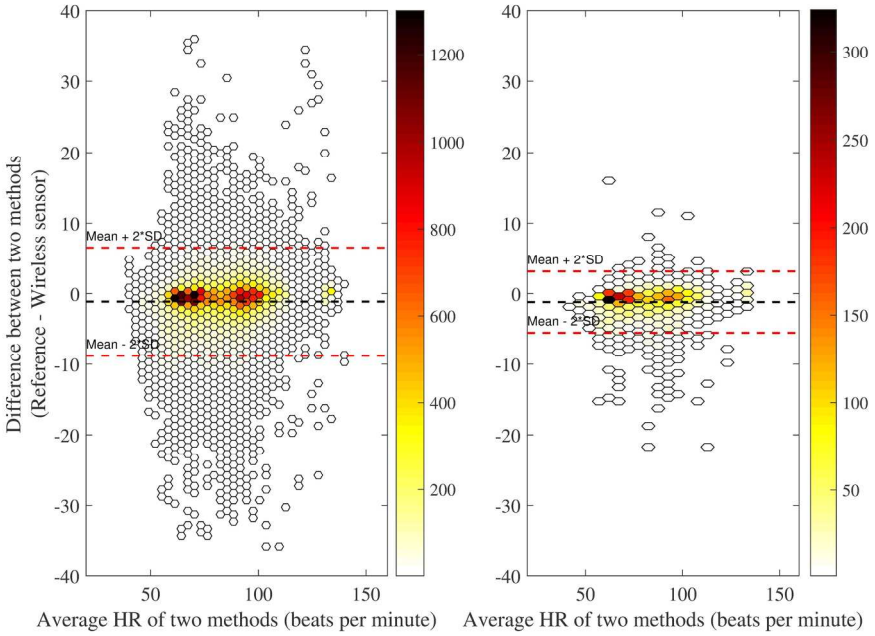
Figure 5 Survival analysis of 36 wireless sensor measurements in 33 patients versus time with various threshold times (maximum duration of a gap in the data). Data loss longer than the threshold counts as failure. The vertical marks indicate the end of a measurement other than failure (i.e., when the patient is being discharged to the ward and data transmission stopped).

Appendix 1 The HealthPatch MD consists of an adhesive patch (with 1-lead ECG and a Zinc-air battery) and a sensor module. BLE: Bluetooth Low Energy; bpm: beats per minute; brpm: breaths per minute



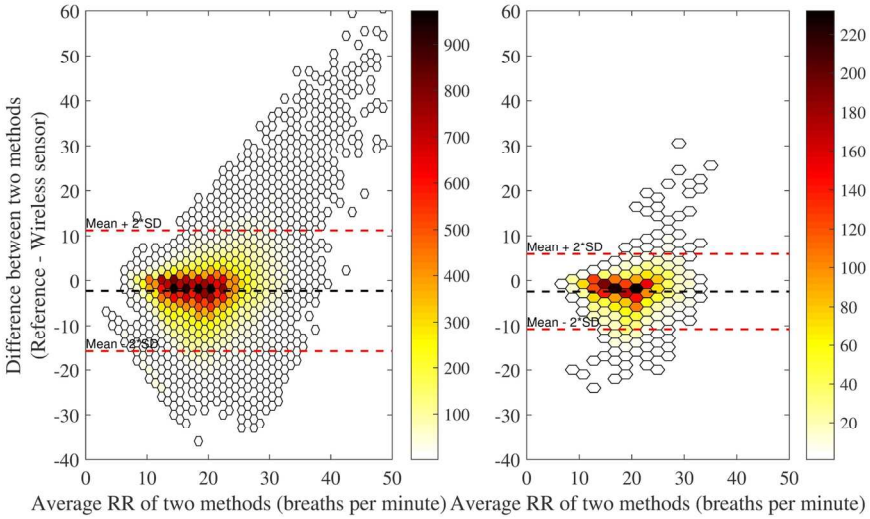
Example of a patient that is being measured for four days continuously with the wireless sensor (red) and reference standard (blue). The upper panel shows heart rate in beats per minute, the lower shows respiratory rate in breaths per minute.

140x89mm (300 x 300 DPI)



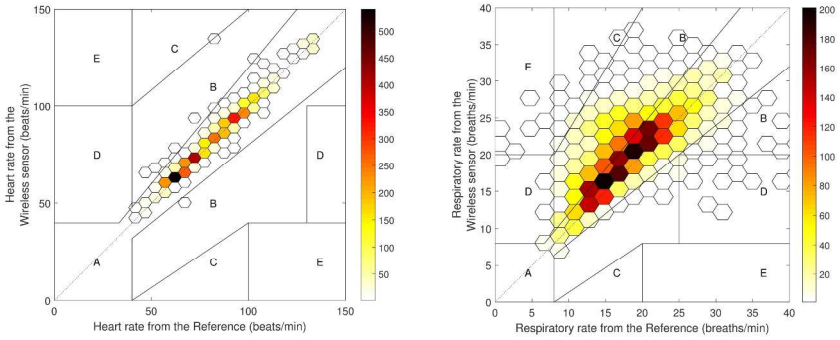
Bland and Altman plots for complete (a) and filtered (b) datasets for heart rate during admission at the surgical step-down unit with few (white) to many (dark red) measurement pairs

192x133mm (300 x 300 DPI)



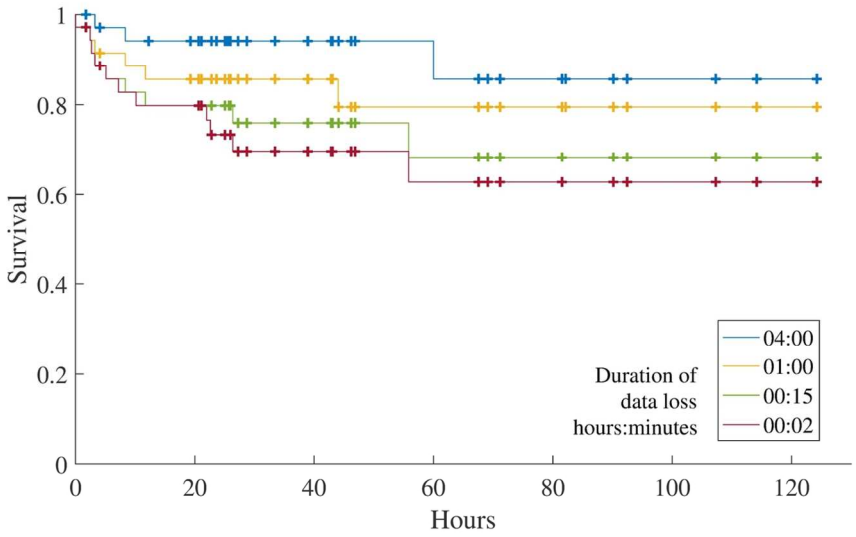
Bland and Altman plots for complete (a) and filtered (b) datasets for respiratory rate during admission at the surgical step-down unit with few (white) to many (dark red) measurement pairs

152x84mm (300 x 300 DPI)



Clark error grid analysis to quantify clinical accuracy of the HR and RR measurements with the HealthPatch MD as compared to the reference monitor of the filtered dataset. Region (A) are points within 20% of the reference monitor, region (B) contains points outside 20% of the reference, but not leading to unnecessary treatment, region (C) are points leading to unnecessary treatment, region (D) indicats a potentially dangerous failure to detect bradycardia or tachycardia (a) or bradypnea or tachypnea (b), region (E) represents points where events are confused (e.g., bradycardia with bradypnea).

286x161mm (300 x 300 DPI)



Survival analysis of 36 wireless sensor measurements in 33 patients versus time with various threshold times (maximum duration of a gap in the data). Data loss longer than the threshold counts as failure. The vertical marks indicate the end of a measurement other than failure (i.e., when the patient is being discharged to the ward and data transmission stopped).

122x68mm (300 x 300 DPI)

211x123mm (72 x 72 DPI)

BMJ Open

Reliability of wireless monitoring using a wearable patch sensor in high-risk surgical patients at a step-down unit in the Netherlands: a clinical validation study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-020162.R1
Article Type:	Research
Date Submitted by the Author:	27-Dec-2017
Complete List of Authors:	Breteler, Martine; Universitair Medisch Centrum Utrecht, Department of Anesthesiology Huizinga, Erik; Universitair Medisch Centrum Utrecht, Anesthesiology van Loon, Kim; Univ Med Ctr Utrecht, Anesthesiology Leenen, Luke; University Medical Centre Utrecht, Trauma Dohmen, Daan; FocusCura Kalkman, Cor; UMC Utrecht, Anesthesiology Blokhuis, Taco; Universitair Medisch Centrum Utrecht, Trauma Surgery
Primary Subject Heading:	Surgery
Secondary Subject Heading:	Nursing
Keywords:	Telemedicine < BIOTECHNOLOGY & BIOINFORMATICS, Vital signs, Wireless technology, Remote monitoring, Continuous monitoring

Reliability of wireless monitoring using a wearable patch sensor in high-risk surgical patients at a step-down unit in the Netherlands: a clinical validation study

M.J.M. Breteler[1,3], E. Huizinga[1], K. van Loon[1], L.P.H. Leenen[2], D.A.J. Dohmen[3], C.J. Kalkman[1], T. J. Blokhuis[2,4]

[1]Department of Anesthesiology, University Medical Center Utrecht, Utrecht University, The Netherlands
[2]Department of Trauma Surgery, University Medical Center Utrecht, Utrecht, The Netherlands
[3]FocusCura, Driebergen-Rijsenburg, The Netherlands
[4]Department of Surgery, Maastricht University Medical Center, Maastricht, The Netherlands

Corresponding author:
Martine Breteler
University Medical Center Utrecht, Mailstop Q.04.2.313
P.O. Box 85500, 3508 GA Utrecht, The Netherlands
Email: m.j.m.breteler@umcutrecht.nl
Phone: +31 6 47306841

Word count: 3911

ABSTRACT

Background and objectives: Intermittent vital signs measurements are the current standard on hospital wards, typically recorded once every 8h. Early signs of deterioration may therefore be missed. Recent innovations have resulted in 'wearable' sensors, which may capture patient deterioration at an earlier stage. The objective of this study was to determine whether a wireless 'patch' sensor is able to reliably measure respiratory and heart rate continuously in high-risk surgical patients. The secondary objective was to explore the potential of the wireless sensor to serve as a safety monitor.

Design: In an observational methods comparisons study, patients were measured with both the wireless sensor and bedside routine standard for at least 24 hours.

Setting: University teaching hospital, single center.

Participants: Twenty-five postoperative surgical patients admitted to a step-down unit.

Outcome measures: Primary outcome were limits of agreement and bias of heart rate and respiratory rate. Secondary outcome measures were sensor reliability, defined as time until first occurrence of data loss.

Results: 1568 hours of vital signs data were analyzed. Bias and 95% limits of agreement for heart rate were -1.1 (-8.8 to 6.5) beats per minute. For respiration rate, bias was -2.3 breaths per minute with wide limits of agreement (-15.8 to 11.2 breaths per minute). Median filtering over a 15 minute period improved limits of agreement of both respiration and heart rate. Sixty-three% of the measurements were performed without data loss greater than two minutes. Overall data loss was limited (6% of time).

Conclusions: The wireless sensor is capable of accurately measuring heart rate, but accuracy for respiratory rate was outside acceptable limits. Remote monitoring has the potential to contribute to early recognition of physiological decline in high-risk patients. Future studies should focus on the ability to detect patient deterioration on low care environments and at home after discharge.

Keywords: telemedicine; vital signs; wireless technology, remote monitoring, continuous monitoring

STRENGTHS AND LIMITATIONS OF THIS STUDY

- We validated the accuracy of a wireless patch sensor to measure heart and respiratory rate in the intended target population for remote wireless monitoring: post-surgical patients at high risk for complications.
- Monitoring was continued for several days.
- The reference standard was an 'ICU-grade' patient monitoring system.
- Although used in clinical practice, the accuracy of the respiratory rate from the reference standard (thoracic impedance pneumography) has known limitations.
- The Bland and Altman methods comparison approach is not ideal for time series data; a test for 'trending ability' is desirable.

INTRODUCTION

While technological advances have resulted in numerous new diagnostic tools and therapeutic innovations, we are still not able to timely recognize patient deterioration on general hospital wards[1,2]. This contributes to avoidable cardiopulmonary arrest, unplanned admission to the intensive care unit (ICU), an increase in hospitalization costs and detrimental effects on quality of life[3-7]. To timely detect patient deterioration, we may benefit from technical solutions that can track patients' vital signs continuously.

Intermittent vital signs measurements, typically once every nurse shift of eight hours, are the current routine monitoring practice on general hospital wards. As a result, patient deterioration in between measurements can be easily missed. In an attempt to improve the detection of patient deterioration, Early Warning Scoring (EWS) protocols and Medical Emergency Teams (METs) have been implemented in most hospitals around the globe. However, failure-to-rescue events continue to occur even with these systems in place[8,9]. This phenomenon is also known as failure of the 'afferent limb' of the EWS system[10-12]. Alongside the attempts to improve detection of patient deterioration on the ward, there is a trend to reduce the duration of hospitalization by discharging patients home early, for example in 'enhanced recovery after surgery' (ERAS) programs[13-15]. Once a patient is back home, EWS protocols and vital signs monitoring are no longer available. Recovery within the patient's own home environment has many advantages, but unavoidably some surgical complications will now become first manifest in the home setting. This increases the risk that patient deterioration will be recognized too late.

The majority of adverse events are preceded by a significant period of change in vital signs[16-20]. Early recognition of the deteriorating patient might be improved if continuous remote monitoring would become available for at-risk patients in 'low care' environments such as the regular hospital ward or in the first few 'critical' days at home after hospital discharge[21,22]. Recent technological innovations have resulted in lightweight 'wearable' wireless sensors capable of recording and transmitting several vital signs such as heart rate (HR), respiration rate (RR), temperature and patient movement. While the majority of the wearable sensors is strictly 'consumer-grade', some manufacturers have obtained CE and/or FDA approval for use in clinical environments. However, validity and accuracy of these so called 'medical-grade' wearables has not been extensive assessed in real clinical environments. Two studies reported satisfactory agreement between heart rate, respiration rate of a wearable patch sensor and their respective reference devices. However, these measurements were obtained from healthy participants in controlled conditions. Another study showed reliable heart rates and respiration rates with a wearable patch sensor in the majority of patients, but these data was limited to short periods of measurements in patients with comorbid conditions[23-25]. As such, we cannot translate these findings accordingly to patients in a clinical environment at risk for complications.

The objective of this study was to determine whether a wireless adhesive 'patch' sensor is able to reliably measure RR and HR continuously in patients after high-risk surgery. We aimed to verify whether wireless sensor technology is robust and capable of detecting physiological trend patterns in

high-risk patients before introducing wireless vital signs monitoring into clinical practice. A secondary objective was to explore the potential of the wireless sensor to serve as a safety monitor in clinical practice.

Materials & Methods

Study design and setting

We performed a methods comparisons study with an observational design in which patients were continuously monitored after high-risk surgery during the initial days of recovery at the surgical step-down unit (SDU) of the University Medical Center Utrecht, the Netherlands, a large academic hospital. Formal approval for this study was obtained from the local ethical committee (number 15/550).

Study participants

Postoperative patients were asked to participate upon admission to the SDU if their expected stay was at least 24 hours. These patients were considered for enrolment because they represent a high-risk subset of surgical patients that is more prone to experience deterioration compared to patients on the general ward. Exclusion criteria were patients with an implantable cardiac device, an allergy to adhesives, wound or skin lesion near the application site, or inability to give informed consent. After written informed consent was obtained, researchers applied the sensor to the patient's chest to start recording vital signs for one to three days using the wearable sensor and the routine monitoring system described hereafter.

Description of the wireless wearable sensor

The HealthPatch® MD (VitalConnect Inc, San Jose, CA, USA) is a medical-grade lightweight, wireless and wearable adhesive biosensor that measures a number of vital signs continuously: single-lead electrocardiogram (ECG), HR, heart rate variability, RR, skin temperature, body posture and step count. It was designed to facilitate long-term remote monitoring of vital signs and activity metrics within the hospital environment as well as in the post-discharge period at home. The sensor consists of a disposable adhesive patch that houses two ECG electrodes, a thermistor and a zinc-air coin-cell battery. The reusable sensor module contains a triaxial accelerometer and Bluetooth Low Energy (BLE) transceiver (see Appendix 1). The patch can be applied on the patient's chest and measures vital signs continuously up to three days (four days if continuous transmission of its single lead ECG waveform is disabled). The module processes the incoming signals and transmits the data via BLE to a relay device (for this study we used an iPad mini (Apple, Inc. Cupertino, CA, USA) with the 'Healthwatch' (VitalConnect Inc, San Jose, CA, USA) mobile application. This application can display vital signs data in real-time for research purposes, but was not designed to be used as clinical monitoring system. Also, near real-time data can be viewed on the Healthwatch Web cloud based server, to monitor long-term trends. Quality of the sensor data was verified several times by the researchers during data collection. Patient identification information was not entered on the mobile device to ensure privacy protection.

Although the wireless sensor can also measure position, in this study we only focused on assessing the accuracy and reliability of RR and HR monitoring. The sensor calculates HR using analysis of the single-lead ECG. The algorithm is based on automated detection of QRS complexes from the ECG waveforms. RR is derived from the combined information from three sources: an embedded algorithm uses a weighted average of two characteristics of the ECG signal: 1.) QRS amplitude modulation and 2.) respiratory sinus arrhythmia (RSA); both ECG-derived signals change during inspiration and expiration[26], and the algorithm uses 3.) accelerometer data produced by chest movement during respiration [27]. Both HR and RR are updated every 4 s and the manufacturer states an accuracy of ± 3 breaths per minute (breaths/min), in the range of 4 to 42 breaths for RR. The stated accuracy of HR is ± 5 beats per minute (beats/min) or 10% (whichever is greater), in the range of 30-200 beats/min.

Description of the bedside routine standard

HR and RR of patients were continuously monitored with the wearable sensor and simultaneously with a multiparameter bedside monitoring system designed for use in ICU's and Operating Rooms (XPREZZON, Spacelabs Healthcare, Snoqualmie, Washington DC, USA) which served as the reference monitor. This reference uses ECG for HR detection and measures RR by thoracic impedance pneumography.

Signal analysis

The raw data transmitted by the sensor containing the measurements and their associated time stamps were retrieved in CSV format. Data were stored and processed using Matlab (The MathWorks Inc., Natick, MA, USA). Empty or invalid data ('not-a-number') were removed to obtain continuous 2D vectors of vital sign samples with their corresponding time stamps. Data reports were automatically retrieved from the reference monitor. These contained vital signs data sampled once per minute (i.e., one measurement was saved and transmitted every minute). The sensor data, originally transmitted once every four seconds was therefore resampled to once per minute (i.e., one sample per minute of the sensor was retained corresponding to the neareset time point of the reference monitor) to produce paired data points with the reference monitor. Furthermore, sensor data and reference data were synchronized to ensure alignment of their respective time stamps. No artifact removal was applied to the data before analysis.

Besides the analysis on vital signs data transmitted every minute, a median filter over a 15-minute period was applied to study the effect on HR and RR outliers and to further explore the potential of the wireless sensor in clinical practice. This filtering was calculated as a median over subsequent epochs of 15 minutes.

Outcomes

The primary outcome was bias and precision (95% Limits of Agreement; LoA) of HR and RR of the wireless sensor compared to the bedside monitor. This reference standard reports an accuracy for HR of ±1% or 3 beats/min (whichever is greater) and an accuracy of ±5% or 1 breaths/min (whichever is greater) for RR[28]. We considered HR and RR to be acceptable for clinical purposes if within ± 10% of the reference monitor or ± 3 breaths/min or ± 5 beats/min[29]. A secondary endpoint was the

reliability of detecting true critical clinical conditions such as bradycardia (HR < 50 beats/min), tachycardia (HR > 100 beats/min), bradypnea (RR < 12 breaths/min) and tachypnea (RR > 20 breaths/min)[30]. Another secondary outcome was the reliability defined as time until the first occurrence of data loss (defined as a duration of a gap within the data of 2 minutes, 15 minutes, 1 hour or 4 hours) and the overall amount of data loss from various causes.

Statistical analysis

The series of observation pairs of HR and RR measurements (one data point every minute) derived from the wireless sensor and the reference monitor were compared using the Bland and Altman Method for repeated measurements[31]. This method was used to account for within-subject variation by correcting for the variance of differences between the average differences across patients and the number of measurements per patient. The mean difference ('bias') between the wireless sensor and reference monitor and the 95% LoA (± 1.96 SD) was determined for both the HR and RR data. In addition, a Clarke Error Grid analysis was conducted to specify the clinical accuracy and the consequences for clinical decision making[32]. The time (hours) to first occurrence of data loss was analysed with Kaplan-Meier survival plots.

A power calculation was not feasible due to the lack of preliminary data with these continuous monitoring systems. Therefore, we aimed to analyze data of at least 25 patients - each with multiple hours of continuous data – which is sufficient to evaluate the validity of the wireless patch sensor.

Results

From September 2015 to September 2016, a total of 33 postoperative patients entered the study. Data from the reference monitor were missing for eight patients due to technical issues with retrieving the data from the monitor's on-board memory, resulting in a total of 1568 hours of monitoring time with the reference monitor and 1702 hours of vital signs monitoring from the wireless sensor. Therefore, measurement pairs of 25 patients were available for agreement analysis. On average, 62 hours remained per patient for further analysis. The range of total monitoring time per patient varied from 12 to 124 hours. Table 1 summarizes patient characteristics and surgical procedures .

Table 1. Patient characteristics (n = 25)

Characteristic	**Value**
Male gender - number (%)	18 (72)
Age (years) - median (IQR) [range]	63.0 (57.8-71.5) [23.0 – 77.0]
Body Mass Index (BMI; kg/m^2)* - median (IQR) [range]	26.2 (24.2-29.4) [17.2 - 40.2]
ASA score	
1-2 (%)	8 (32)
3-4 (%)	17 (68)
Comorbidities	
Hypertension - no (%)	8 (32)

Cardiovascular disease - no (%)	9 (36)
COPD - no (%)	3 (12)
Diabetes - no (%)	3 (12)

IQR: interquartile range. COPD: Chronic Obstructive Pulmonary Disease. * BMI of one patient was missing

Example of a patient measurement

Figure 1 shows the HR and RR trend during the first four postoperative days of a 60-year-old male patient with extensive cardiac and vascular comorbidities, after an open nephrectomy procedure for a suspected carcinoma in situ. Three events can be recognized: 1.) a sudden HR increase on Thursday evening; later diagnosed as atrial fibrillation, 2.) an episode of bradycardia on Saturday afternoon, and 3.) mild tachycardia and a subtle increase in RR starting Sunday afternoon, caused by bleeding from an aortic branch artery. After coil embolization, the patient was readmitted to the Intensive Care Unit (ICU). This example illustrates agreement between HR and RR measurements recorded with the wireless sensor and the reference standard. Note that RR derived from the reference monitor was highly variable compared with RR from the wireless sensor (Figure 1).

Heart rate

Table 2 shows bias and precison (95% LoA) from comparisons between the wireless sensor and the reference standard. For analysis, 55565 minutes (926 h) of HR measurement pairs were available. The mean difference (bias) in HR was -1.1 beats/min (reference standard minus sensor) with a 95% LoA of -8.8 to 6.5 beats/min. Applying a 15 min median filter resulted in a narrower 95% LoA of -5.7 beats/min to 3.2 beats/min (3986 minutes available for analysis). Bland and Altman plots for the complete and filtered HR datasets are shown in Figure 2a and 2b, respectively.

Table 2 Bland Altman analysis of wireless heart rate and respiratory rate versus reference monitor in postoperative patients

Parameter	Number of measurement pairs	Bias	SD	Lower 95% LoA[b]	Upper 95% LoA[b]
Complete dataset					
Heart rate	55565	-1,1	3,83	-8,8	6,5
Respiratory rate	56674	-2,3	6,8	-15,8	11,2
Filtered dataset[a]					
Heart rate	3986	-1,2	2,2	-5,7	3,2
Respiratory rate	4001	-2,4	4,2	-10,8	5,9

[a] *Dataset after applying a median filter*

[b] *LoA limits of agreement*

Respiratory rate

The mean difference between the reference monitor and the wireless sensor was -2.3 breaths/min with wide levels of agreement (95% LoA: -15.8 to 11.2 breaths/min). The agreement between both methods improved after applying a 15 minute median filter, resulting in a 95% LoA of -10.8 to 5.9 breaths/min. Bland and Altman plots for the complete and filtered RR dataset are displayed in Figures 3a and 3b.

Most 'high RR' outliers originated from the reference monitor and were observed in the higher RR range. This is also shown in Figure 1 where RR measurements derived from the reference monitor showed a higher variation compared with RR derived from the wireless sensor. This high variation reduced after applying a median filter over 15-minutes.

Diagnostic accuracy of the wireless sensor

Because of the relatively long monitoring time per patient and the high-risk population, we were able to capture several instances of bradycardia, tachycardia, bradypnea and tachypnea. The incidence of bradycardia was rare, a HR below 50 was present in only 2% of all HR measurements in the complete dataset. Tachycardia, bradypnea and tachypnea occurred more frequently in 14%, 15% and 34% of cases respectively. Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) are shown in Table 3. After applying median filtering, sensitivity and specificity of all episodes with abnormal HR and RR improved.

Table 3 Diagnostic accuracy for bradypnea, tachypnea, bradycardia, tachycardia

	True positives (%)	False positives (%)	True negatives (%)	False negatives (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
Bradycardia								
Complete dataset	824 (71)	363 (1)	55148 (99)	339 (1)	71	99	69	99
Filtered dataset[a]	24 (72)	0 (0)	3953 (100)	9 (0)	73	100	100	100
Tachycardia								
Complete dataset	7111 (90)	1490 (3)	47321 (97)	752 (2)	90	97	83	98
Filtered dataset[a]	496 (98)	65 (2)	3413 (98)	12 (0)	98	98	88	100
Bradypnea								
Complete dataset	2113 (24)	562 (1)	47438 (99)	6561 (12)	24	99	79	88
Filtered dataset[a]	134 (33)	8 (0)	3587 (100)	272 (7)	33	100	94	93
Tachypnea								
Complete dataset	16210 (84)	14602 (39)	22694 (61)	3168 (12)	84	61	53	88
Filtered dataset[a]	1225 (94)	1029 (38)	1668 (62)	79 (5)	94	62	54	95

[a] *Dataset after applying a median filter*

The Clarke error grid analyses of the filtered datasets are shown in Figure 4a and 4b; it shows that 100% of the HR measurements and 99% of RR measurements are within region A or B, respectively within 20% of the reference measurement, or outside 20% of the reference but not leading to unneccesary treatment. None of the HR measurements were in region, B, C or E, which means that none of the measurements would lead to failure to treat, unnecessary treatment or confusion between bradycardia and tachycardia. Very few of the RR measurements (≤ 1 %) were within region C, D or E, indicating a potentially dangerous failure to apply the right treatment.

Technical performance

HR and RR data were recorded for the majority of the time (94%), from 36 sensors in 33 patients (Table 4). Nineteen (53%) wireless patient monitoring series had complete uninterrupted data, but in 17 patients there was data loss (ranging from 8s data loss to 60h). Most found sensor failure to be caused by a sawtooth pattern of the battery level from inadvertent covering of the air opening for the zinc-air battery resulting in measurement gaps.

Table 4 Amount of data loss within all wireless sensor measurements

	Time loss (hours:minutes:seconds)			
	Total loss (%)[a]	Mean loss	Minimum loss	Maximum loss
All wireless sensor measurements (n=36)	101:15:24 (5.9%)	02:48:45	00:00:00	59:59:08
Only wireless sensor measurements with any data loss (n=17)	101:15:24 (15.5%)	05:57:22	00:00:08	59:59:08

[a] *Data loss is defined as time without data as percentage of the total time measured*

Figure 5 shows the survival analysis for 'time to first failure' in data transmission of the sensor. Using a threshold (maximum duration of a gap in the data) of two minutes, this analysis showed that 63% of the wireless sensor measurements in patients were performed without data loss greater than two minutes. A gap duration of one hour resulted in 79% of sensor measurements without data loss greater than an hour.

Discussion

We studied the performance of a wearable wireless sensor to measure HR and RR continuously in high-risk postoperative patients. The results show that this sensor (Healthpatch MD) can accurately measure HR with a deviation within 10% of the reference standard. In contrast, the accuracy for RR was outside the limit range we considered acceptable. However, this finding may be due to frequent outliers and clinically implausible variability of RR values provided by the reference monitor. Median filtering of both signals over a 15 min period resulted in a reduction of the number of RR measurement pairs outside the acceptable LoA and an improvement of LoA. Overall, data loss was limited with HR and RR measurements 94% of the time available.

Strengths

To the best of our knowledge, this is the first clinical study that investigated the reliability and accuracy of continuous vital signs monitoring using a wearable wireless patch sensor for several days in postoperative surgical patients at a step-down unit. Most studies were actually obtained under controlled laboratory conditions[25,27,33]. These studies demonstrated the ability of the HealthPatch sensor to accurately measure HR and RR in adult participants. Hernandez et al. [24] reported a higher accuracy for HR and RR measurements with the SensiumVitals digital patch in stable patients with comorbid conditions for a limited time period (2h) compared to our study. Other studies used intermittent nurse observations on the ward as the only reference. Weenk et al. [34] reported that both HR and RR of the HealthPatch were in agreement with nurse measurements, although wide limits of

agreement were found. Another study compared RR measurements of nurses with readings from the SensiumVitals digital patch and found inadequate agreement [35]. Although these studies showed the feasibility of wireless technology in clinical practice, comparison with nurse readings cannot validate the continuous performance of the wireless devices. Moreover, these wireless monitoring devices are not intended to deliver 'spot' readings for EWS, i.e., their use was evaluated for a purpose outside the intended scope of use. Consequently, a drawback of these study designs was the inability to validate continuous HR and RR measurements of new remote monitoring devices in between nurse observations. In the current study reference HR and RR were measured continuously in a clinical setting.

Limitations

The results of the present study confirm the difficulty of accurate continuous RR monitoring. However, we should also consider the limitations of the reference standard[36,37]; accuracy of thoracic impedance RR measurements can be influenced by many factors independent of breathing, such as patient movement, talking and coughing. Impedance artefacts could explain the high number of false negatives (i.e., missed bradypnea), resulting in low sensitivity. Impedance technique which is the current bedside routine standard for continuous measurement of HR and RR in most hospitals today and therefore clinically relevant. However, it cannot be considered a gold standard for RR measurement. While capnography is widely regarded as a 'true' gold standard for RR, it has several drawbacks for continuous unsupervised respiratory monitoring in spontaneously breathing patients, since its nasal cannula can be easily dislodged, leading to incomplete data and a high number of false positive 'apnea' alarms[38]. Furthermore, the HealthPatch sensor was not designed to indicate respiration rates < 5 breaths/min. Nonetheless, a progressive slowing of breathing rate may still be identified and used as indicator of vital instability, for example to recognize life-threatening opioid-induced respiratory depression.

Filtering

Applying a median filter over 15 minute data epochs improved reliability of HR and RR by removing outliers, for example, a transient very high RR caused by a movement artefact. This is appropriate since these transient artefacts (i.e., RR >45 breaths/min) are extremely unlikely from a physiological perspective. Although filtering effectively eliminates such artifacts, this comes at the expense of reducing the number of available measurements and the reduced ability to detect sudden changes in vital signs (e.g., cardiac arrest). An alternative might be to use a 'moving' median filter to keep the update rate at once every minute or once every 2 min. On the other hand, improved eliminaton of outliers could result in a higher proportion of epochs with reliable HR and RR resulting in lower false positive alarms. The latter is extremely important if this remote monitoring system is to be clinically deployed on the ward or at home. Furthermore, continuous remote monitoring on the ward with a reduced frequency (i.e., once every 15 minutes) still provides much more information regarding the patient's vital signs than the current intermittent monitoring practice, where vital signs are usually only observed once every 8h nurse shift.

In case an alarm is generated by the remote monitoring system, a nurse can personally check on the patient and correct if the cause of the alarm was not related to a change in the patient's medical condition. However, it must be realized that a large number of false alarms is very disruptive on the general ward, especially when there is a low nurse to patient ratio (e.g., at night). This may even decrease patient safety by taking away valuable nursing time from patients who are in real need of attention. We suggest that eliminating outliers to improve reliability and eliminate false positive alerts far outweighs the limited benefits of having 'continuous' vital signs data streams.

'Methods comparison' methodology and continuous monitoring

The goal of this study was to determine whether the wireless sensor is able to reliably measure RR and HR over time in postoperative patients. Although Bland and Altman analyses can reliably indicate bias and precision of 'spot measurements', it does not inform about the 'trending ability' of the monitoring system over time, while this is of ultimate importance to timely recognize abnormal vital sign patterns. This was also confirmed in the study of Churpek et al. that showed the added value of using trends of vital signs for detecting clinical deterioration on the wards[21]. In our study, the example case in Figure 1 clearly demonstrates the ability of the wireless sensor to detect important physiological changes of a deteriorating patient over time, even while the limits of agreement for detecting bradypnea as determined by Bland Altman analyses were deemed not acceptable. Therefore, we wish to emphasize that accurate trend measurements (e.g., the ability to detect deterioration over time) is more important than just one accurate single measurement at one specific point in time.

Continuous monitoring on the general ward is still unknown territory. Future studies should therefore focus on the performance of wireless monitoring in patients at the general ward, including validation during periods of mobilization. Particular emphasis should be on the early detection of critical adverse events. However, the usability and patient perspective on remote monitoring are important to determine when creating the infrastructure of a complete remote monitoring solution.

Conclusion

Wireless continuous monitoring may have the potential to contribute to early recognition of physiological decline in high-risk patients. The tested wireless sensor was able to accurately record heart rate, but the accuracy of respiratory rate needs further optimization to reduce the incidence of false alarms and allow timely recognition of altered breathing patterns.

Contributors

MB, CK, DD and TB were involved in the clinicial trial design and ethics submission. EH was involved in the planning and conduct of the study and drafting of the results. MB assisted in the conduct of the study, performed analysis of clinical data and drafted the paper. Throughout the study, LL, KL, CK and TB assisted in the conduct of the study. All authors reviewed and approved the final version of the manuscript.

Funding

This research was funded solely from institutional funds.

Competing interests

At the time of the study, MB was part-time employee of FocusCura (Health IT company, Driebergen-Zeist, The Netherlands) and DD is founder and CEO of FocusCura. No funding from FocusCura was obtained. To ensure independent analysis, the wearable sensor (HealthPatch MD) was purchased from VitalConnect Inc, the manufacturer of the HealthPatch MD. VitalConnect Inc, neither had insight or influence on the analysis of the data or drafting of the manuscript.

Data sharing statement

Authors are planning upon producing further publications using this dataset. Afterwards, anonymized patient recordings will be available from the corresponding author upon request.

References

1 Fuhrmann L, Lippert A, Perner A, *et al.* Incidence, staff awareness and mortality of patients at risk on general wards. *Resuscitation* 2008;**77**:325–30. doi:10.1016/j.resuscitation.2008.01.009

2 Kause J, Smith G, Prytherch D, *et al.* A comparison of Antecedents to Cardiac Arrests, Deaths and EMergency Intensive care Admissions in Australia and New Zealand, and the United Kingdom—the ACADEMIA study. *Resuscitation* 2004;**62**:275–82. doi:10.1016/j.resuscitation.2004.05.016

3 Sax FL, Charlson ME. Medical patients at high risk for catastrophic deterioration. *Crit Care Med* 1987;**15**:510–5.

4 Jonsson T, Jonsdottir H, Möller AD, *et al.* Nursing documentation prior to emergency admissions to the intensive care unit. *Nurs Crit Care* 2011;**16**:164–9. doi:10.1111/j.1478-5153.2011.00427.x

5 McQuillan P, Pilkington S, Allan A, *et al.* Confidential inquiry into quality of care before admission to intensive care. *BMJ* 1998;**316**:1853–8.

6 Calzavacca P, Licari E, Tee A, *et al.* The impact of Rapid Response System on delayed emergency team activation patient characteristics and outcomes—A follow-up study. *Resuscitation* 2010;**81**:31–5. doi:10.1016/j.resuscitation.2009.09.026

7 Ghaferi AA, Birkmeyer JD, Dimick JB. Variation in hospital mortality associated with inpatient surgery. *N Engl J Med* 2009;**361**:1368–75. doi:10.1056/NEJMsa0903048

8 Taenzer AH, Pyke JB, McGrath SP. A Review of Current and Emerging Approaches to Address Failure-to-Rescue. *Anesthesiology* 2011;**115**:421–31. doi:10.1097/ALN.0b013e318219d633

9 Ferraris VA, Bolanos M, Martin JT, *et al.* Identification of Patients With Postoperative Complications Who Are at Risk for Failure to Rescue. *JAMA Surgery* 2014;**149**:1103. doi:10.1001/jamasurg.2014.1338

10 Trinkle RM, Flabouris A. Documenting Rapid Response System afferent limb failure and associated patient outcomes. *Resuscitation* 2011;**82**:810–4. doi:10.1016/j.resuscitation.2011.03.019

11 Hammer JA, Jones TL, Brown SA. Rapid response teams and failure to rescue: one community's experience. *J Nurs Care Qual* 2012;**27**:352–8. doi:10.1097/NCQ.0b013e31825a8e2f

12 Hillman K, Chen J, Cretikos M, *et al.* Introduction of the medical emergency team (MET) system: a cluster-randomised controlled trial. *Lancet* 2005;**365**:2091–7. doi:10.1016/S0140-6736(05)66733-5

13 Pajarón-Guerrero M, Fernández-Miera MF, Dueñas-Puebla JC, *et al.* Early Discharge Programme on Hospital-at-Home Evaluation for Patients with Immediate Postoperative Course after Laparoscopic Colorectal Surgery. *Eur Surg Res* 2017;**58**:263–73. doi:10.1159/000479004

14 Roulin D, Donadini A, Gander S, *et al.* Cost-effectiveness of the implementation of an enhanced recovery protocol for colorectal surgery. *Br J Surg* 2013;**100**:1108–14. doi:10.1002/bjs.9184

15 Gustafsson UO, Scott MJ, Schwenk W, *et al.* Guidelines for Perioperative Care in Elective Colonic Surgery: Enhanced Recovery After Surgery (ERAS®) Society Recommendations. *World J Surg* 2012;**37**:259–84. doi:10.1007/s00268-012-1772-0

16 Buist M, Bernard S, Nguyen TV, *et al.* Association between clinically abnormal observations and subsequent in-hospital mortality: a prospective study. *Resuscitation* 2004;**62**:137–41. doi:10.1016/j.resuscitation.2004.03.005

17 Churpek MM, Yuen TC, Huber MT, *et al.* Predicting cardiac arrest on the wards: a nested case-control study. *Chest* 2012;**141**:1170–6. doi:10.1378/chest.11-1301

18 Cuthbertson BH, Boroujerdi M, McKie L, *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. doi:10.1097/01.CCM.0000254826.10520.87

19 Hollis RH, Graham LA, Lazenby JP, *et al.* A Role for the Early Warning Score in Early Identification of Critical Postoperative Complications. *Ann Surg* 2016;**263**:918–23. doi:10.1097/SLA.0000000000001514

20 Goldhill DR, White SA, Sumner A. Physiological values and procedures in the 24 h before ICU admission from the ward. *Anaesthesia* 1999;**54**:529–34.

21 Churpek MM, Adhikari R, Edelson DP. The value of vital sign trends for detecting clinical deterioration on the wards. *Resuscitation* 2016;**102**:1–5. doi:10.1016/j.resuscitation.2016.02.005

22 Watkins T, Whisman L, Booker P. Nursing assessment of continuous vital sign surveillance to improve patient safety on the medical/surgical unit. *Journal of Clinical Nursing* 2016;**25**:278–81. doi:10.1111/jocn.13102

23 Selvaraj N. Long-term Remote Monitoring of Vital Signs using a Wireless Patch Sensor. Health Innovation Conference (HIC) 2014;IEEE:83–86.

24 Hernandez-Silveira M, Ahmed K, Ang S-S, *et al.* Assessment of the feasibility of an ultra-low power, wireless digital patch for the continuous ambulatory monitoring of vital signs. *BMJ Open* 2015;**5**:e006606–6. doi:10.1136/bmjopen-2014-006606

25 Chan AM, Selvaraj N, Ferdosi N, *et al.* Wireless patch sensor for remote monitoring of heart rate, respiration, activity, and falls. *Conf Proc IEEE Eng Med Biol Soc* 2013;**2013**:6115–8. doi:10.1109/EMBC.2013.6610948

26 Yasuma F, Hayano J-I. Respiratory sinus arrhythmia: why does the heartbeat synchronize with respiratory rhythm? *Chest* 2004;**125**:683–90.

27 Chan AM, Ferdosi N, Narasimhan R. Ambulatory respiratory rate detection using ECG and a triaxial accelerometer. *Conf Proc IEEE Eng Med Biol Soc* 2013;**2013**:4058–61. doi:10.1109/EMBC.2013.6610436

28 Healthcare S. Data Sheet Ultraview SL. 2016.

29 Ben-Ari J, Zimlichman E, Adi N, *et al.* Contactless respiratory and heart rate monitoring: validation of an innovative tool. *J Med Eng Technol* 2010;**34**:393–8. doi:10.3109/03091902.2010.503308

30 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. doi:10.1016/j.resuscitation.2012.12.016

31 Bland JM, Altman DG. Agreement between methods of measurement with multiple observations per individual. *J Biopharm Stat* 2007;**17**:571–82. doi:10.1080/10543400701329422

32 Clarke WL, Cox D, Gonder-Frederick LA, *et al.* Evaluating clinical accuracy of systems for self-monitoring of blood glucose. *Diabetes Care* 1987;**10**:622–8.

33 Selvaraj N. Long-term Remote Monitoring of Vital Signs using a Wireless Patch Sensor. 2014;:1–4.

34 Weenk M, van Goor H, Frietman B, *et al.* Continuous Monitoring of Vital Signs Using Wearable Devices on the General Ward: Pilot Study. *JMIR Mhealth Uhealth* 2017;**5**:e91. doi:10.2196/mhealth.7208

35 Granholm A, Pedersen NE, Lippert A, *et al.* Respiratory rates measured by a standardised clinical approach, ward staff, and a wireless device. *Acta Anaesthesiol Scand* 2016;**60**:1444–52. doi:10.1111/aas.12784

36 Freundlich JJ, Erickson JC. Electrical impedance pneumography for simple nonrestrictive continuous monitoring of respiratory rate, rhythm and tidal volume for surgical patients. *Chest* 1974;**65**:181–4.

37 Wiklund L, Hök B, Ståhl K, *et al.* Postanesthesia monitoring revisited: frequency of true and false alarms from different monitoring devices. *J Clin Anesth* 1994;**6**:182–8.

38 Macknet M, Kimball-Jones P, Applegate R II, *et al.* Accuracy of a novel bioacoustic sensor for monitoring respiratory rate. *European Journal of Anaesthesiology* 2007;**24**:34. doi:10.1097/00003643-200706001-00124

Figure titles and legends

Figure 1 Example of a patient that is being measured for four days continuously with the wireless sensor (red) and reference standard (blue). The upper panel shows heart rate in beats per minute, the lower shows respiratory rate in breaths per minute.

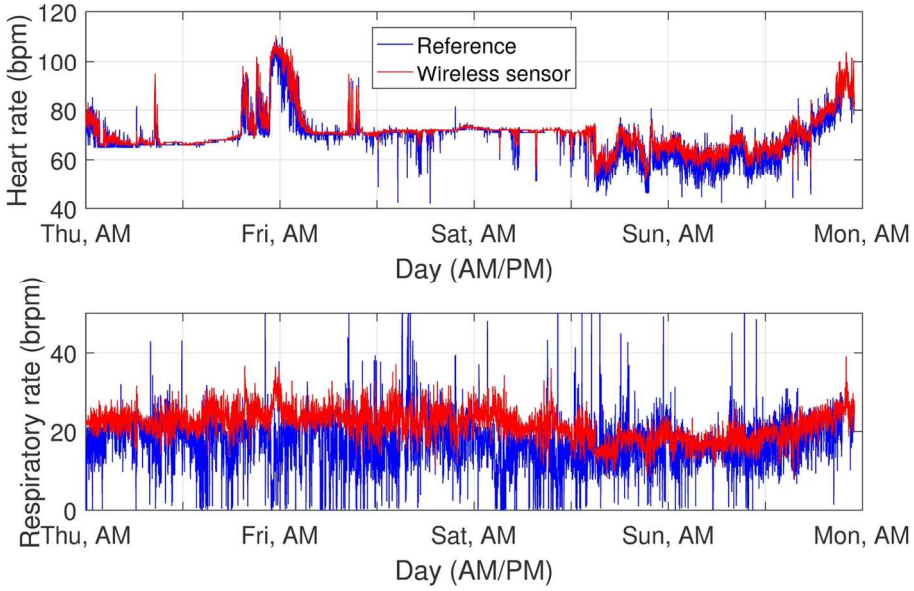
Figure 2 Bland and Altman plots for complete (a) and filtered (b) datasets for heart rate during admission at the surgical step-down unit with few (white) to many (dark red) measurement pairs

Figure 3 Bland and Altman plots for complete (a) and filtered (b) datasets for respiratory rate during admission at the surgical step-down unit with few (white) to many (dark red) measurement pairs

Figure 4 Clark error grid analysis to quantify clinical accuracy of the HR (a) and RR (b) measurements with the HealthPatch MD as compared to the reference monitor of the filtered dataset. Region (A) are points within 20% of the reference monitor, region (B) contains points outside 20% of the reference, but not leading to unnecessary treatment, region (C) are points leading to unnecessary treatment, region (D) indicats a potentially dangerous failure to detect bradycardia or tachycardia (a) or bradypnea or tachypnea (b), region (E) represents points where events are confused (e.g., bradycardia with bradypnea).

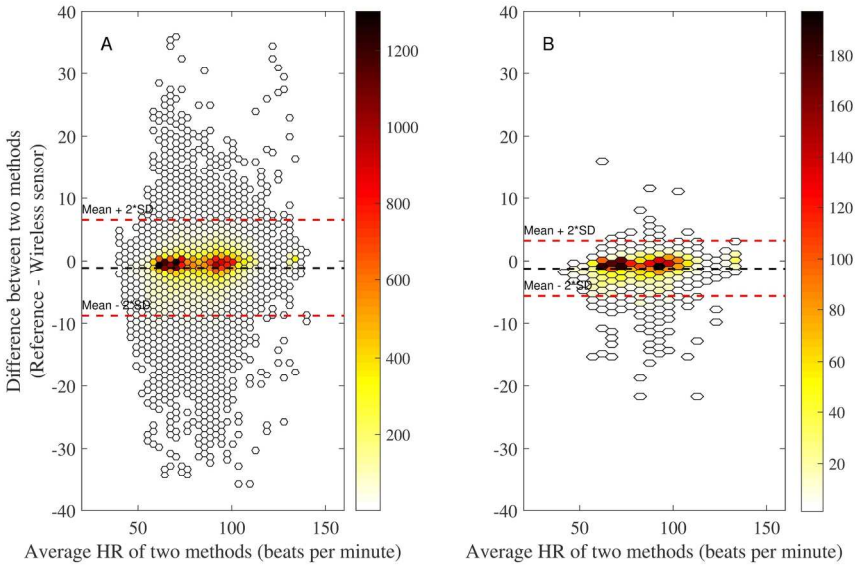
Figure 5 Survival analysis of 36 wireless sensor measurements in 33 patients versus time with various threshold times (maximum duration of a gap in the data). Data loss longer than the threshold counts as failure. The vertical marks indicate the end of a measurement other than failure (i.e., when the patient is being discharged to the ward and data transmission stopped).

Appendix 1 The HealthPatch MD consists of an adhesive patch (with 1-lead ECG and a Zinc-air battery) and a sensor module. BLE: Bluetooth Low Energy; bpm: beats per minute; brpm: breaths per minute



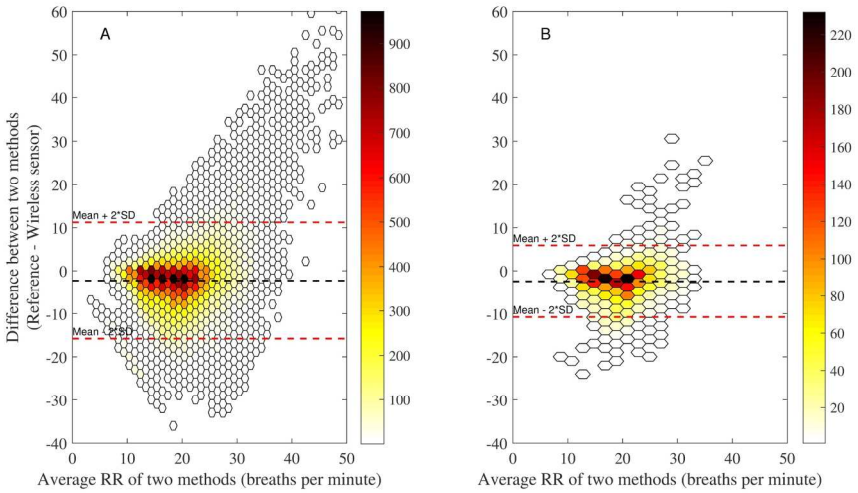
Example of a patient that is being measured for four days continuously with the wireless sensor (red) and reference standard (blue). The upper panel shows heart rate in beats per minute, the lower shows respiratory rate in breaths per minute.

140x89mm (300 x 300 DPI)



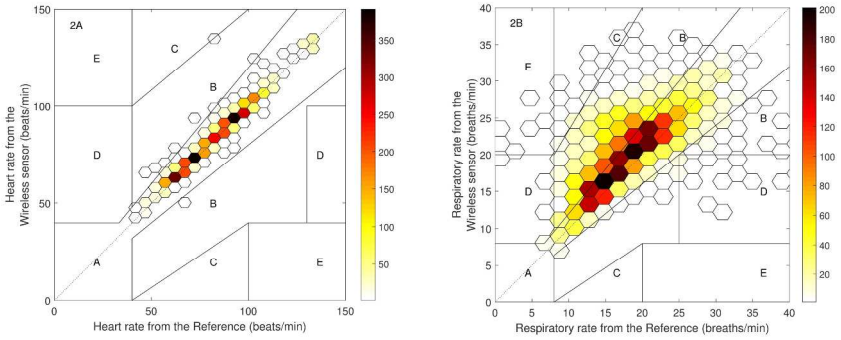
Bland and Altman plots for complete (a) and filtered (b) datasets for heart rate during admission at the surgical step-down unit with few (white) to many (dark red) measurement pairs

179x110mm (300 x 300 DPI)



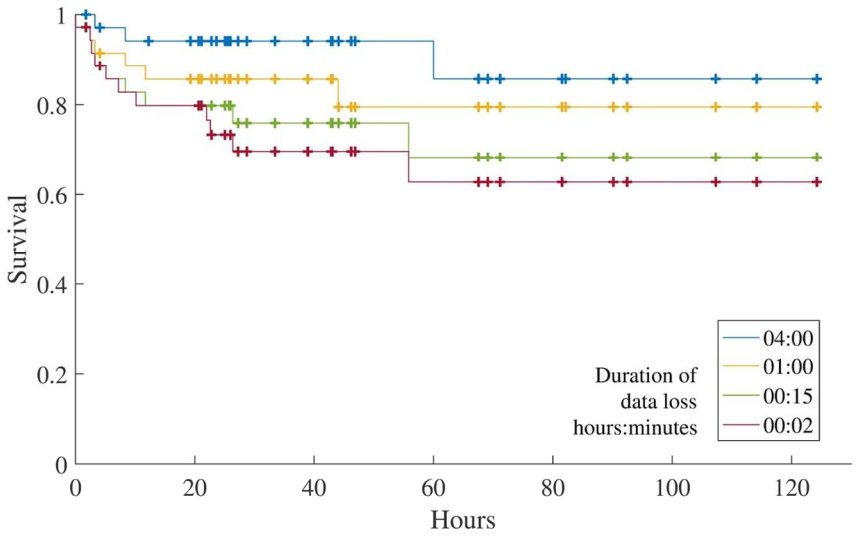
Bland and Altman plots for complete (a) and filtered (b) datasets for respiratory rate during admission at the surgical step-down unit with few (white) to many (dark red) measurement pairs

177x93mm (300 x 300 DPI)



Clark error grid analysis to quantify clinical accuracy of the HR (a) and RR (b) measurements with the HealthPatch MD as compared to the reference monitor of the filtered dataset. Region (A) are points within 20% of the reference monitor, region (B) contains points outside 20% of the reference, but not leading to unnecessary treatment, region (C) are points leading to unnecessary treatment, region (D) indicats a potentially dangerous failure to detect bradycardia or tachycardia (a) or bradypnea or tachypnea (b), region (E) represents points where events are confused (e.g., bradycardia with bradypnea).

286x161mm (300 x 300 DPI)



Survival analysis of 36 wireless sensor measurements in 33 patients versus time with various threshold times (maximum duration of a gap in the data). Data loss longer than the threshold counts as failure. The vertical marks indicate the end of a measurement other than failure (i.e., when the patient is being discharged to the ward and data transmission stopped).

122x68mm (300 x 300 DPI)

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No.	Recommendation	Page No.	Relevant text from manuscript
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1 & 2	
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2	
Introduction				
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	3	
Objectives	3	State specific objectives, including any prespecified hypotheses	3	
Methods				
Study design	4	Present key elements of study design early in the paper	4	
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	4	
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	4	
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case		
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	5	
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	*5*	
Bias	9	Describe any efforts to address potential sources of bias	NA	
Study size	10	Explain how the study size was arrived at	6	

Continued on next page

Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	5
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	5 & 6
		(*b*) Describe any methods used to examine subgroups and interactions	NA
		(*c*) Explain how missing data were addressed	5
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	NA
		(*e*) Describe any sensitivity analyses	
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	6
		(b) Give reasons for non-participation at each stage	6
		(c) Consider use of a flow diagram	NA
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	6
		(b) Indicate number of participants with missing data for each variable of interest	6
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	6
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	*7 & 8*
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	
		Cross-sectional study—Report numbers of outcome events or summary measures	
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	7
		(*b*) Report category boundaries when continuous variables were categorized	7
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	NA

Continued on next page

Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	8 & 9
Discussion			
Key results	18	Summarise key results with reference to study objectives	9
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	9 & 10
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	9 & 10
Generalisability	21	Discuss the generalisability (external validity) of the study results	11
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	11

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.