

## ANESTHESIOLOGY

# Vital Signs Monitoring with Wearable Sensors in High-risk Surgical Patients

## A Clinical Validation Study

Martine J. M. Breteler, M.Sc., Eline J. KleinJan, M.Sc.,  
Daan A. J. Dohmen, Ph.D., Luke P. H. Leenen, M.D., Ph.D.,  
Richard van Hillegersberg, M.D., Ph.D.,  
Jelle P. Ruurda, M.D., Ph.D., Kim van Loon, M.D., Ph.D.,  
Taco J. Blokhuis, M.D., Ph.D., Cor J. Kalkman, M.D., Ph.D.

ANESTHESIOLOGY 2020; 132:424–39

### EDITOR'S PERSPECTIVE

#### What We Already Know about This Topic

- Changes in vital signs are an important indicator of physiological decline and hence provide opportunities for early recognition and intervention; however, in the hospital ward, vital signs are usually measured intermittently. In between such spot checks, early signs of deterioration may be missed.
- Several “wearable” and wireless sensors have been developed that may capture the patient deterioration earlier.

#### What This Article Tells Us That Is New

- In high-risk surgical patients admitted to a step-down unit, heart rate was accurately measured by the two wearable patch sensors (SensiumVitals [Sensium Healthcare Ltd., United Kingdom] and HealthPatch [VitalConnect, USA]) and by the bed-based contactless mattress sensor (EarlySense [EarlySense Ltd., Israel]) and by the patient-worn monitor (Masimo Radius-7 [Masimo Corporation, USA]). The highest precision for heart rate was seen with the HealthPatch sensor.
- For respiratory rate, the accuracies of the Masimo Radius-7, EarlySense, and SensiumVitals were within a predefined acceptable range, while the HealthPatch overestimated respiratory rate.

Changes in vital signs are an important indicator of physiologic decline and hence provide opportunities for early recognition and intervention.<sup>1–4</sup> However,

### ABSTRACT

**Background:** Vital signs are usually recorded once every 8 h in patients at the hospital ward. Early signs of deterioration may therefore be missed. Wireless sensors have been developed that may capture patient deterioration earlier. The objective of this study was to determine whether two wearable patch sensors (SensiumVitals [Sensium Healthcare Ltd., United Kingdom] and HealthPatch [VitalConnect, USA]), a bed-based system (EarlySense [EarlySense Ltd., Israel]), and a patient-worn monitor (Masimo Radius-7 [Masimo Corporation, USA]) can reliably measure heart rate (HR) and respiratory rate (RR) continuously in patients recovering from major surgery.

**Methods:** In an observational method comparison study, HR and RR of high-risk surgical patients admitted to a step-down unit were simultaneously recorded with the devices under test and compared with an intensive care unit–grade monitoring system (XPRESSON [Spacelabs Healthcare, USA]) until transition to the ward. Outcome measures were 95% limits of agreement and bias. Clarke Error Grid analysis was performed to assess the ability to assist with correct treatment decisions. In addition, data loss and duration of data gaps were analyzed.

**Results:** Twenty-five high-risk surgical patients were included. More than 700 h of data were available for analysis. For HR, bias and limits of agreement were 1.0 (–6.3, 8.4), 1.3 (–0.5, 3.3), –1.4 (–5.1, 2.3), and –0.4 (–4.0, 3.1) for SensiumVitals, HealthPatch, EarlySense, and Masimo, respectively. For RR, these values were –0.8 (–7.4, 5.6), 0.4 (–3.9, 4.7), and 0.2 (–4.7, 4.4) respectively. HealthPatch overestimated RR, with a bias of 4.4 (limits: –4.4 to 13.3) breaths/minute. Data loss from wireless transmission varied from 13% (83 of 633 h) to 34% (122 of 360 h) for RR and 6% (47 of 727 h) to 27% (182 of 664 h) for HR.

**Conclusions:** All sensors were highly accurate for HR. For RR, the EarlySense, SensiumVitals sensor, and Masimo Radius-7 were reasonably accurate for RR. The accuracy for RR of the HealthPatch sensor was outside acceptable limits. Trend monitoring with wearable sensors could be valuable to timely detect patient deterioration.

(ANESTHESIOLOGY 2020; 132:424–39)

in current hospital practice, nurses and physicians rely on intermittent vital signs “spot checks,” typically once every 8-h shift. As a result, early signs of deterioration may be missed, especially at night when deterioration may progress unnoticed until the next morning.<sup>5,6</sup> Furthermore, compliance from nurses to vital signs monitoring protocols is often poor, in particular measurements of respiration rate,<sup>7,8</sup> resulting in incomplete sets of vital signs that may limit detection of physiologic abnormalities.<sup>9,10</sup> As a result, patients can deteriorate unnoticed, potentially leading to

This article is featured in “This Month in Anesthesiology,” page 1A. This article is accompanied by an editorial on p. 407. This article has a related Infographic on p. 19A. This article has an audio podcast. This article has a visual abstract available in the online version. The work presented in this article had an abstract discussion at the Annual Conference of ANESTHESIOLOGY in San Francisco, California, October 15, 2018.

Submitted for publication January 24, 2019. Accepted for publication September 18, 2019. Published online first on November 13, 2019. From the Department of Anesthesiology, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands (M.J.M.B., E.J.K., K.v.L., C.J.K.); FocusCura, Driebergen-Rijsenburg, The Netherlands (M.J.M.B., D.A.J.D.); Department of Technical Medicine, University of Twente, Enschede, The Netherlands (E.J.K.); Department of Trauma Surgery (L.P.H.L.) and Department of Gastrointestinal and Oncologic Surgery (R.v.H., J.P.R.), University Medical Center Utrecht, Utrecht, The Netherlands (L.P.H.L.); and Department of Surgery, Maastricht University Medical Center, Maastricht, The Netherlands (T.J.B.).

Copyright © 2019, the American Society of Anesthesiologists, Inc. All Rights Reserved. Anesthesiology 2020; 132:424–39. DOI: 10.1097/ALN.0000000000003029

preventable adverse events, failure-to-rescue events, and increased hospital costs.<sup>11–13</sup>

Attempts to improve recognition of patients at risk with the implementation of rapid response teams have shown mixed results.<sup>14,15</sup> Successful implementation of a rapid response team critically depends on timely identification of patients at risk (“afferent limb”).<sup>16,17</sup> With the current low monitoring frequency, failure-to-rescue events continue to occur, and improvements in terms of patient outcome may require improved recognition of deterioration.<sup>13,17,18</sup>

Especially in “low care” environments such as surgical wards, continuous remote monitoring could contribute to earlier recognition of the deteriorating patient.<sup>16,19</sup> Several wearable wireless devices intended for vital signs monitoring recently became available.<sup>16</sup> These devices allow the patient to move freely without the inconvenience of physical attachment to a patient monitor. Such technology combined with appropriate remote monitoring facilities could even allow patients to recover at home by allowing safe discharge earlier after surgery.<sup>20,21</sup>

Current wearable sensors are capable of recording heart rate (HR), respiration rate, temperature, and movement. Although some wearable sensors have now obtained approval for medical use, uptake within health care has been minimal. One reason could be that the validity and reliability have not been studied in relevant clinical environments. Several studies demonstrated the feasibility of wireless monitoring in clinical practice, but comparison with intermittent nurse readings cannot validate the continuous performance of wearable sensors.<sup>22–24</sup> Moreover, these sensors are intended to monitor vital sign trends and as such are not designed to deliver “spot” readings to replace nurse observations. Most accuracy studies of continuous vital signs sensors were obtained under controlled laboratory conditions and for short time periods

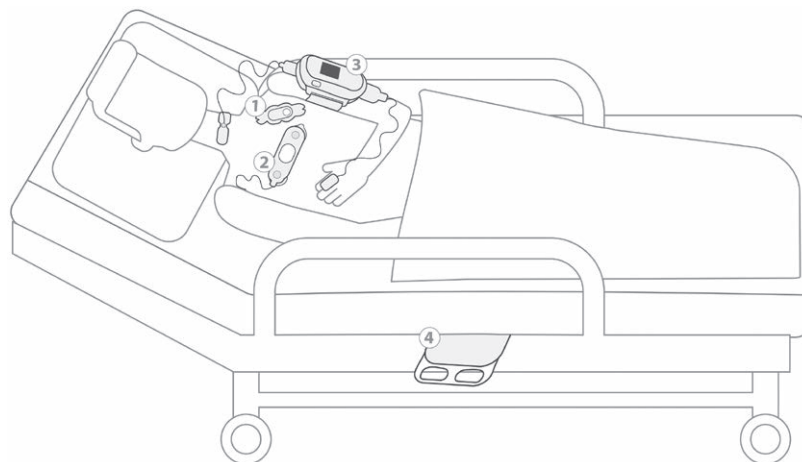
in volunteers or patients at low risk for developing complications.<sup>22,25,26</sup> Therefore, we cannot translate these findings to high-risk patients at risk for developing vital instability.

We recently studied the potential of a wireless patch sensor for monitoring actual high-risk patients and compared the results with an intensive care unit–grade monitoring system.<sup>27</sup> We hypothesize that wireless sensors can monitor HR and respiratory rate (RR) reliably when compared to a traditional “wired” reference standard. However, it is unknown how performance differs and how well they perform against a typical intensive care unit–grade patient monitor. Therefore, the objective of this study is to determine whether four systems with different sensing principles can reliably measure HR and RR continuously in high-risk surgical patients.

## Materials and Methods

### Study Design and Setting

We performed an observational methods comparison study in which high-risk surgical patients were continuously monitored with two wearable patch sensors (SensiumVitals [Sensium Healthcare Ltd., United Kingdom] and HealthPatch [VitalConnect, USA]), a bed-based mattress sensor (EarlySense [EarlySense Ltd., Israel]) and a patient-worn monitor (Masimo Radius-7 [Masimo Corporation, USA]) simultaneously during the initial days of recovery at a surgical step-down unit until transition to the traumatology or surgical oncology ward of the University Medical Center Utrecht, Utrecht, The Netherlands. An overview of the measurement setup can be seen in figure 1. Formal approval for this study was obtained from the Medical Research Ethics Committee of University Medical Center Utrecht (No. 16/062).



**Fig. 1.** Overview of the measurement setup in patients with two wearable patch sensors placed on the chest (1: HealthPatch, 2: SensiumVitals), a patient-worn monitor connected to a pulse oximeter probe attached to the finger and an acoustic sensor applied in the neck (3: Masimo Radius-7), and a bed-based sensor placed under the patient's mattress (4: EarlySense).

## Study Population

Patients were recruited at the preoperative screening clinic before surgery or upon admission to the step-down unit. The inclusion criterion was patients from the surgical traumatology or surgical oncology specialty. These patients were considered for enrollment since they represent a high-risk subset of surgical patients that is more prone to deterioration compared to patients on the general ward. Exclusion criteria were patients with bacterial infections requiring barrier nursing, an allergy to skin adhesives, wound or skin lesion near the application site, or patients with a pacemaker or implantable cardioverter defibrillator. After written informed consent was obtained from the patient, the four sensor systems (described in detail below) were applied to the patient and vital signs recording started.

**SensiumVitals System (No. 1, Wearable “Patch”-Type Sensor).** The SensiumVitals sensor is a single-use, lightweight (50 g), wireless, and wearable adhesive biosensor that measures HR, RR, and axillary temperature every 2 min.<sup>22</sup> It is designed to facilitate ambulatory wireless monitoring of patients throughout their hospital ward stay, without the need to change batteries. Measurements are wirelessly transmitted *via* a low-power radio protocol to hotspots (“bridges” connected to the hospital network) installed throughout the step-down unit, and sent to a central monitoring server. Data were subsequently extracted from the server and saved locally. To guarantee care as usual, health-care professionals had no access to the data. Study personnel had access to the system to check its functionality.

The patch sensor is applied to the patient’s chest by means of two conventional electrocardiogram electrodes and measures vital signs continuously (transmitted once every 2 min) up to 5 days. Although the patch also has a sensor to measure axillary temperature, in this study we only assessed the accuracy and reliability of HR and RR. The sensor internally records HR and RR in a sequential fashion. Every 2 min, a 30-s segment of electrocardiogram, followed by a 60-s segment of the respiratory signal, is recorded and analyzed. The 30-s period of electrocardiogram is recorded and processed to calculate the average HR by analyzing a single-lead electrocardiogram, which is preprocessed first to filter the raw electrocardiogram signal to minimize noise due to, *e.g.*, motion artifacts. The sensor’s embedded algorithm rejects the HR signal if it is invalid due to excessive contamination by noise. RR is recorded by impedance pneumography that measures the small impedance changes across the chest as the lungs contract and expand during breathing. The embedded algorithm excludes segments that are corrupted by motion artifacts or other irregular patterns from the calculation of RR values. The manufacturer states an accuracy of  $\pm 2$  beats/min, in the range of 30 to 210 beats for HR. The stated accuracy of RR is  $\pm 2$  breaths/min, in the range of 5 to 62 breaths/min.

**HealthPatch MD (No. 2, Wearable “Patch”-Type Sensor).** The HealthPatch MD is a lightweight (10 g), wireless, and

wearable adhesive biosensor that measures HR, HR variability, single-lead electrocardiogram, RR, skin temperature, body posture, and step count. It is designed to facilitate long-term remote monitoring of patients within the hospital setting as well as in the home setting after hospital discharge. Details of the wearable sensor and its system are described in a previous study where the sensor has been tested in a clinical environment.<sup>27</sup> The sensor consists of a disposable patch and a reusable module that needs to be applied on the patient’s chest and measures vital signs continuously up to 3 days (4 days if transmission of its single-lead electrocardiogram is disabled). The sensor calculates HR using an algorithm which is based on automated detection of QRS complexes from the electrocardiogram waveforms. RR is derived from the combined information from three sources. An embedded algorithm uses a weighted average of two characteristics of the electrocardiogram signal: (1) QRS amplitude modulation and (2) respiratory sinus arrhythmia. Both electrocardiogram-derived signals change during inspiration and expiration, and the algorithm uses (3) accelerometer data produced by chest movement during respiration.<sup>28</sup> The sensor calculates HR during a period of 10 beats, and RR is calculated during a 45-s segment. Both HR and RR are updated every 4 s, and the manufacturer states an accuracy of  $\pm 3$  breaths/min, in the range of 4 to 42 breaths for RR. The stated accuracy of HR is  $\pm 5$  beats/min or 10% (whichever is greater), in the range of 30 to 200 beats/min.

**EarlySense System (No. 3, Contactless Mattress Sensor).** The EarlySense system is a contactless piezoelectric sensor that is placed under the patient’s mattress. It connects to an EarlySense bedside monitor that displays HR, RR, and body motion with an update time of 60 s as long as the patient remains in bed. The averaging time to calculate HR is 10 beats, and RR is calculated during a segment of 5 breaths. In this study, the EarlySense monitor display was hidden from care professionals. The piezoelectric sensor detects ballistic vibrations of body movements, chest wall movement from respiration, and cardio ballistic movements, which are associated with ejection of blood with each heart cycle. The EarlySense system was previously shown to measure HR and RR accurately in patients in intensive care unit settings.<sup>29,30</sup> The manufacturer states an accuracy of  $\pm 5$  beats/min or 4% (whichever is greater) in the range of 30 to 170 beats for HR. The stated accuracy of RR is  $\pm 1.5$  breaths/min, or 4% (whichever is greater), in the range of 6 to 45 breaths/min.

**Masimo Radius-7 (No. 4, Patient-Worn Pulse Oximeter and RR Monitor).** Masimo Radius-7 is a patient-worn monitor connected to a pulse oximeter probe attached to the finger for pulse rate and oxygen saturation monitoring and a novel acoustic adhesive sensor applied in the neck (RRA; Masimo Corporation) for RR monitoring. This acoustic sensor detects upper airway sounds during inhalation and exhalation. Signal processing algorithms convert these acoustic

patterns into breathing cycles to calculate RR, whereby it distinguishes breathing patterns from other background noise such as talking, coughing, or snoring. The device (a rechargeable module with display) is worn on the upper arm and wirelessly connects *via* Bluetooth to the base monitor (Masimo “Root” platform) at the patient’s bedside. The sensor records HR during a 30-s epoch and RR during a 60-s segment. These vital signs are updated every second, but stored once per 2 s. With a 12-h battery life, the battery modules must be “swapped” with a freshly charged device once every nurse shift. For this study, alarms were deactivated and the monitor’s display was hidden from care professionals. Ramsay *et al.* reported high accuracy for RR values of the acoustic monitor when compared with capnography in a study of 33 surgical patients at the postanesthesia care unit.<sup>31</sup> Another study reported that acoustic monitoring of RR was more accurate when compared with impedance pneumography and frequency-modulated continuous wave radar.<sup>32</sup> The manufacturer states an accuracy of  $\pm 3$  beats/min (in patients at rest) or  $\pm 5$  beats/min (during motion) in the range of 25 to 240 beats for pulse rate. The stated accuracy of RR is  $\pm 1$  breath/minute in the range of 4 to 70 breaths/min.

**Description of the Bedside Routine Standard.** HR and RR of patients were continuously monitored with all four systems as described and simultaneously with a multiparameter bedside monitoring system designed for use in intensive care units and operating rooms (XPRESSON [Spacelabs Healthcare, USA]), which served as reference monitor. This reference uses electrocardiogram for HR detection and measures RR by thoracic impedance pneumography. An electrocardiogram epoch of 8 beats is used to calculate HR and it measures RR during a period of 4 breaths. Vital signs are updated every second, but stored once per minute. This reference standard reports an accuracy of  $\pm 3$  beats/min or  $\pm 1\%$  (whichever is greater) in the range of 15 to 300 for HR. The reported accuracy of RR is  $\pm 1$  breath/min or  $\pm 5\%$  (whichever is greater), in the range of 0 to 200.

## Signal Analysis

Data of the four wireless sensors and the reference system were retrieved in comma-separated values and JavaScript object notation formats and stored in a dedicated local research database. Data were processed using MATLAB (MathWorks, USA). Data reports from the reference monitor contained vital signs data sampled once per minute (*i.e.*, one measurement was saved and transmitted every minute). The wireless sensors use different averaging times for HR and RR and send out their data at different update rates. To produce data pairs of the reference standard with each of the wireless sensors for comparison, the update frequency was resampled. Consequently, data of the HealthPatch and data of the Masimo Radius-7 were downsampled to once per minute, which means that one sample per minute of each sensor was retained corresponding to the nearest time point of the reference monitor. To produce paired data points

with SensiumVitals (transmitted once every 2 min), data of the reference standard needed to be downsampled from once every minute to once per 2 min. The update frequency of EarlySense was unchanged at 1 min.

To ensure alignment between time series, a synchronization method was used based on cross-correlation to estimate the time shift of each index sensor with the reference sensor in order to calculate the number of samples each device needed to be delayed or forwarded in time. This was based on the assumption both signals (*e.g.*, HR from the index and reference sensor) were similar in shape, but with different time stamps. After synchronization, a “moving” median filter with a window of 15 min was applied to eliminate movement artifacts.

## Outcomes

The primary outcome measure was HR and RR of all wireless sensors as compared with the reference standard. We considered HR and RR to be acceptable for clinical purposes if within  $\pm 10\%$  of the reference standard or  $\pm 5$  beats/min or  $\pm 3$  breaths/min. A secondary outcome measure was the Clarke Error Grid analysis to quantify the clinical accuracy of HR and RR as compared with the reference standard. The Clarke Error Grid breaks down a scatterplot of the reference standard and the devices under test in regions A (values within 20% of the reference standard) up to and including region E (values that would lead to reverse treatment decisions of, *e.g.*, bradycardia and tachycardia). The regions were defined based on the cutoff boundaries of the Modified Early Warning Score.<sup>33</sup> Another secondary outcome measure was the technical performance of each wireless sensor, which was evaluated by the proportion of total amount of data loss and maximum duration of data loss, defined as gap durations with a maximum length of 5 min, 15 min, 1 h, or 4 h or longer.

## Statistical Analysis

No formal rules for power calculations can be found in literature regarding methods comparison studies where multiple sequential measurements are performed per patient. Therefore, we performed no formal power calculations before the study. The sample size was pragmatically based on our desire to at least include 15 adverse events, which would allow us to extend validation of the measured sensor values well into the abnormal physiology range.

The data pairs of HR and RR measurements derived from the wireless sensors and the reference sensor were analyzed using the Bland and Altman method for repeated measurements<sup>34</sup> and using mixed effects models as suggested by Myles and Cui.<sup>35</sup> The Bland and Altman method was used to account for within-subject variation by correcting for the differences across patients and the variance of differences between the average differences. The bias (mean difference) between the index sensors and reference monitor and the



95% limits of agreement ( $\pm 1.96$  SD) were determined for both HR and RR data. Furthermore, we calculated the limits of agreement by using a mixed effects model that involves using time as random effect and a random intercept per subject while adjusting for the mean of each subject over time and the mean measurement between each wireless sensor and reference monitor for each measurement occasion. To apply this mixed effects model, we checked that the variance of the repeated measurements for each patient was independent of the mean of the repeated measurements. This random effect model of Myles was suggested as modification for handling repeated measurements.<sup>35</sup>

In addition, a Clarke Error Grid analysis was conducted to specify clinical accuracy of the wireless sensors against the reference standard and to study the potential consequences for treatment decisions.<sup>36</sup> This was expressed as the percentage of data representing adequate and inadequate treatment decisions. Technical performance was analyzed by the duration of data loss and the total amount of data loss. Duration of data loss with a maximum length of 4 min, 15 min, 60 min or longer than 4 h was identified to evaluate the potential of wireless monitoring. The analyses were conducted using MATLAB version 2017b.

To evaluate the *trending ability* of HR and RR of the wireless sensors, we created four-quadrant plots and calculated the concordance rate for each of the wireless sensors using an exclusion zone of 1 beat/min for HR and 1 breath/min for RR.<sup>37</sup>

## Results

From February to September 2017, a total of 31 high-risk surgical patients entered the study, of whom 25 were

monitored on the step-down unit. The other six patients were only monitored at the surgical ward, because they were immediately transferred from the intensive care unit to the ward and bypassed the step-down unit, or the time to conduct measurements at the step-down unit was too short. In total, 720 h of vital signs monitoring on the step-down unit with all wireless sensors attached were available, with a median duration of 19 h per patient (minimum, 21 min; maximum, 111 h). Three patients were not monitored with the HealthPatch sensor, due to shortage of these sensors. Table 1 summarizes patient characteristics.

## HR

Table 2 shows bias and precision (95% limits of agreement) from comparisons between the four wireless sensors and the reference standard after applying a “moving” median filter of 15 min. Bias and the 95% limits of agreement derived from the mixed effects models for HR of the SensiumVitals, HealthPatch, EarlySense, and Masimo Radius-7 were all within the predefined acceptable range. The 95% limits of agreement as calculated with the Bland-Altman method showed wider limits of agreement for all sensors. The HealthPatch showed the narrowest limits of agreement, as can be seen in the Bland and Altman plots of figure 2A–D, with limits of agreement from the Bland-Altman method in black, and from the mixed effects models in red.

Although HR derived with the Masimo Radius-7 sensor was accurate for sinus rhythm, substantial variability of HR of the Masimo Radius-7 sensor was observed during episodes of atrial fibrillation in five patients.

**Table 1.** Patient Characteristics (N = 25)

Gender, No. (%)	Female	9 (36)
	Male	16 (64)
Age, yr, median [IQR]		62 [51–71]
BMI, kg/m <sup>2</sup> , median [IQR]		27 [25–30]
Specialty, No. (%)	Surgical gastrointestinal oncology	14 (56)
	Traumatology	11 (44)
Admission diagnosis, No. (%)	Esophagectomy	10 (32)
	Herniation colon	1 (4)
	Hemihpatectomy	1 (4)
	Enucleation of neuroendocrine tumor	1 (4)
	Surgery for liposarcoma	1 (4)
	Trauma (pneumothorax, hemothorax, multiple rib fractures, cruris fracture)	11 (44)
ASA Physical Status, No. (%)	ASA I or II	19 (76)
	ASA III or IV	6 (24)
Comorbidities, No. (%)	Hypertension	4 (16)
	Cardiovascular disease	2 (8)
	COPD	3 (12)
	Diabetes	2 (8)
Length of hospital stay, days, median [IQR]		13 [11–20]
Readmission to hospital within 30 days, No. (%)		2 (8)

\*BMI of one patient was missing.

ASA, American Society of Anesthesiologists; BMI, body mass index; COPD, chronic obstructive pulmonary disease; IQR, interquartile range.

**Table 2.** Bland–Altman Analysis for Heart Rate and Respiratory Rate from the Wireless Sensors *versus* the Reference Monitor

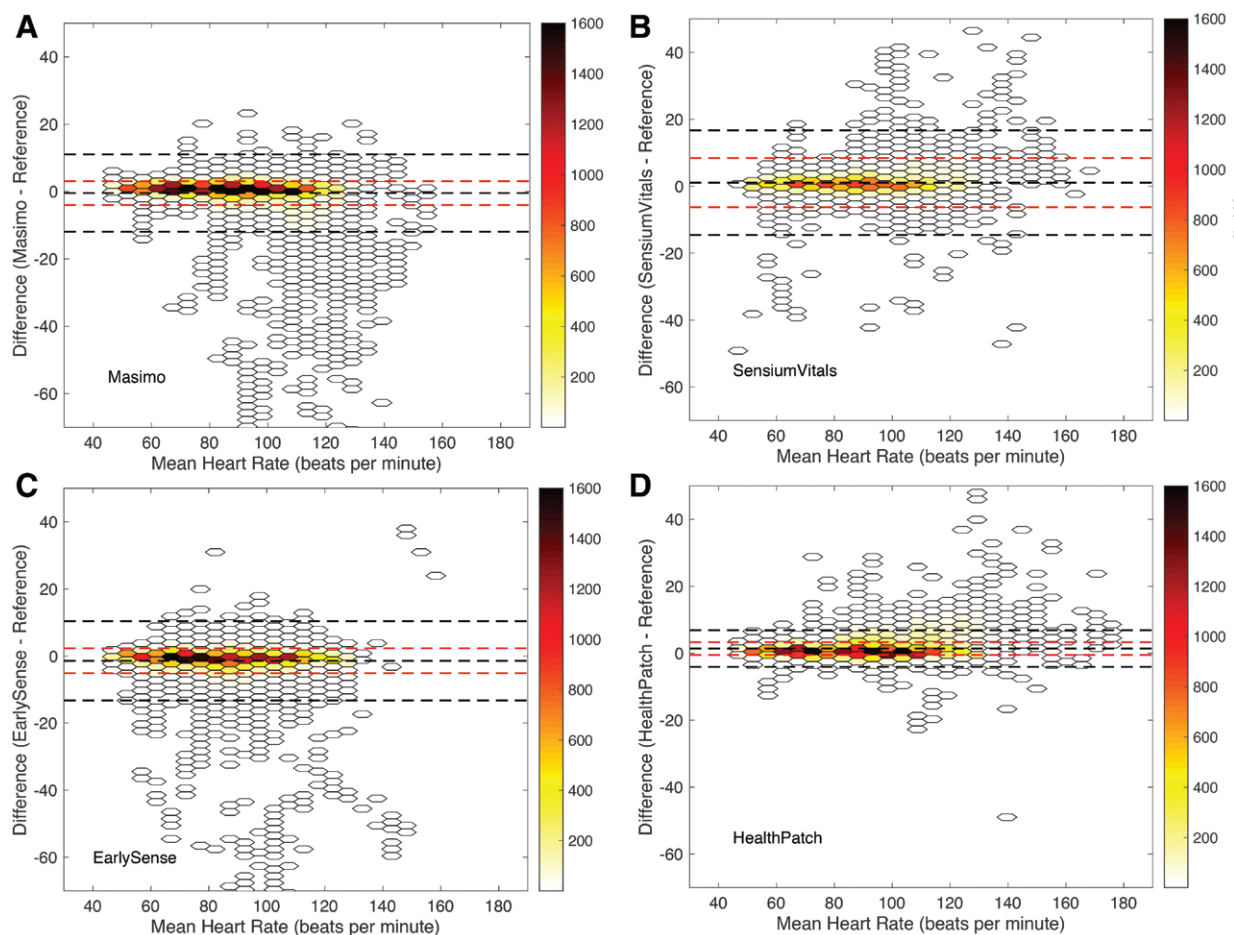
	No. of Measurement Pairs	No. of Patients	Bias	Lower 95% LoA BA	Upper 95% LoA BA	Lower 95% LoA MEM	Upper 95% LoA MEM
Heart rate							
SensiumVitals	16,917	25	1.0	−14.6	16.7	−6.3	8.4
HealthPatch	29,619	22	1.3	−4.1	6.9	−0.5	3.3
EarlySense	29,470	25	−1.4	−13.2	10.4	−5.1	2.3
Masimo	34,992	25	−0.4	−11.9	11.0	−4.0	3.1
Respiratory rate							
SensiumVitals	17,595	25	−0.8	−8.5	6.9	−7.4	5.6
HealthPatch	29,135	22	4.4	−5.8	14.7	−4.4	13.3
EarlySense	27,921	25	0.4	−5.6	6.4	−3.9	4.7
Masimo	33,032	25	0.2	−6.6	6.3	−4.7	4.4

BA, Bland–Altman method; LoA, limits of agreement; MEM, mixed effects model.

## RR

The mean difference and limits of agreement derived from the mixed effects models for RR of the SensiumVitals, EarlySense, and Masimo Radius-7 were all within the predefined

accepted range as shown in table 2. The HealthPatch overestimated RR, with a mean difference of 4.4 breaths/min and with wide levels of agreement of −4.4 to 13.3 breaths/min. The 95% limits of agreement calculated from the Bland and



**Fig. 2.** Bland and Altman plots of heart rate for Masimo Radius-7 (A), SensiumVitals (B), EarlySense (C), and HealthPatch (D). Limits of agreement from the Bland–Altman method in black, and from mixed effects models in red.

Altman method showed wider limits of agreement for all sensors. EarlySense showed the narrowest limits of agreement for RR. Figure 3A–D illustrates the Bland and Altman plots.

### Clinical Accuracy of the Wireless Sensors

The Clarke Error Grid analyses of HR and RR of each wireless sensor are plotted in figure 4A–D and figure 5A–D. The percentage of data pairs in the regions A to E are shown in table 3. Region A or B, respectively, indicates within 20% of the reference standard or outside 20% of the reference standard but not leading to unnecessary or wrong treatment. These results show that adequate treatment decisions regarding changes in HR (zone A or B) would have been taken in 99.4%, 100%, 99.5%, and 99.5% with the SensiumVitals, HealthPatch, EarlySense, and Masimo Radius-7, respectively. None of the HealthPatch HR measurements and only a very few (0.5% or fewer) of the HR measurements of Masimo Radius-7, EarlySense, and SensiumVitals were

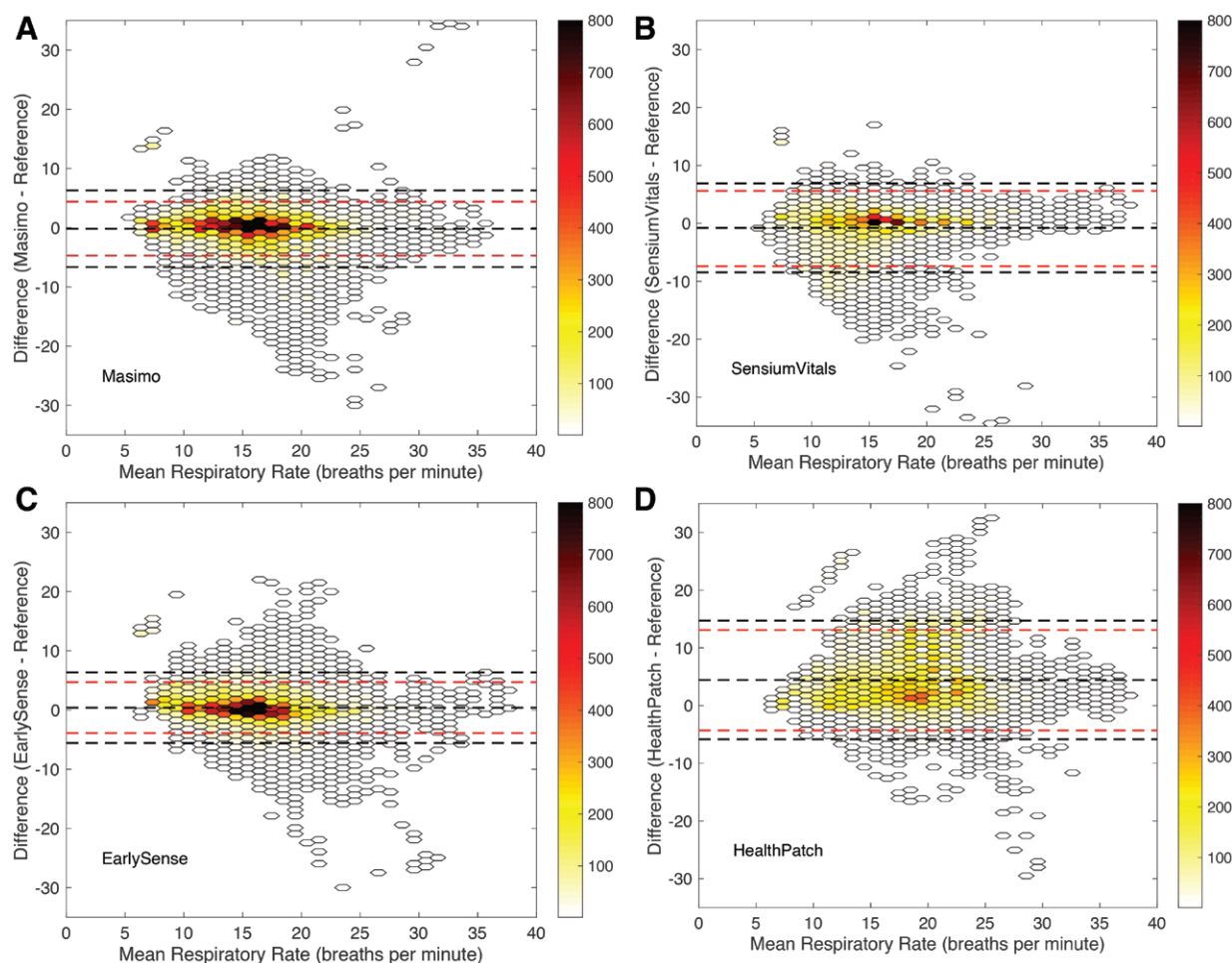
within regions C, D, or E, suggesting that very few measurements would lead to failure to treat, unnecessary treatment, or confusion between bradycardia and tachycardia.

For RR, adequate treatment decisions would have been 92.7%, 77.3%, 96.6%, and 96.3% with the SensiumVitals, HealthPatch, EarlySense, and Masimo Radius-7, respectively. A number of RR measurements (greater than 10%) of the HealthPatch sensor were within regions C, D, or E, indicating a potentially dangerous failure to apply the right treatment.

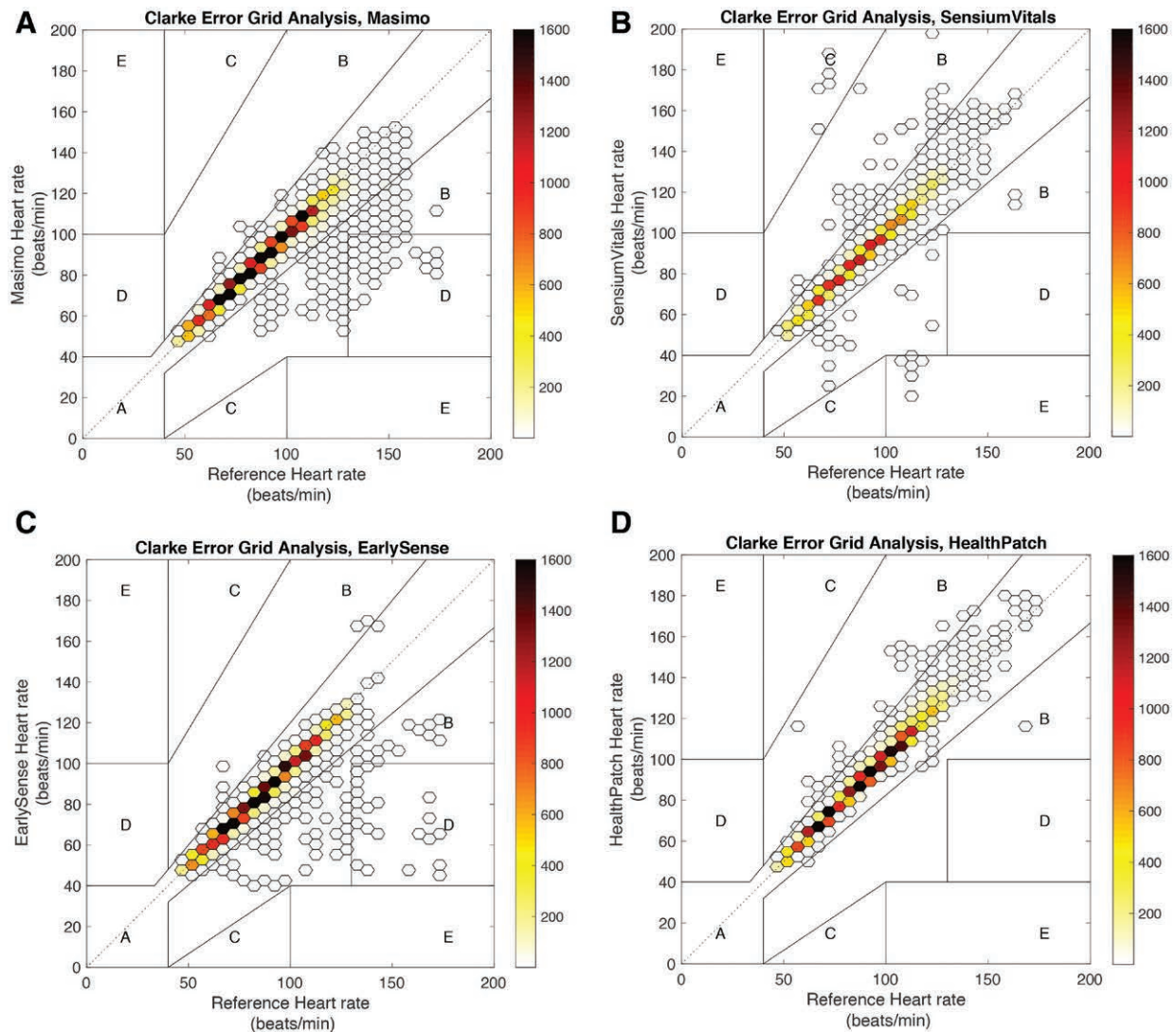
Four-quadrant plots showing the trending ability of  $\Delta$  HR and  $\Delta$  RR (*i.e.*, difference between consecutively obtained RR and HR values) for both the wearable sensors and reference standard are shown in a density plot that can be found in appendix 1 and appendix 2.

### Technical Performance

Data loss of HR measurements was 12.9% (83 of 633 h), 12.3% (79 of 640 h), 27.5% (182 of 664 h), and 6.5% (47



**Fig. 3.** Bland and Altman plots of respiratory rate for Masimo Radius-7 (A), SensiumVitals (B), EarlySense (C), and HealthPatch (D). Limits of agreement from the Bland–Altman method in black, and from mixed effects models in red.



**Fig. 4.** Clarke Error Grid analysis to quantify clinical accuracy of the heart rate measurements with the Masimo Radius-7 (A), SensiumVitals (B), EarlySense (C), and HealthPatch (D) as compared with the reference standard. The *colored dots* are measurement pairs each superimposed on the Error Grid boundaries, where the color intensity is proportional to the number of observations. *Region A* encloses points within 20% of the reference monitor; *region B* contains points outside 20% of the reference, but not leading to unnecessary treatment. *Region C* contains points leading to unnecessary treatment, *region D* indicates a potentially dangerous failure to detect bradycardia or tachycardia, and *region E* represents points where events are confused (e.g., bradycardia with tachycardia).

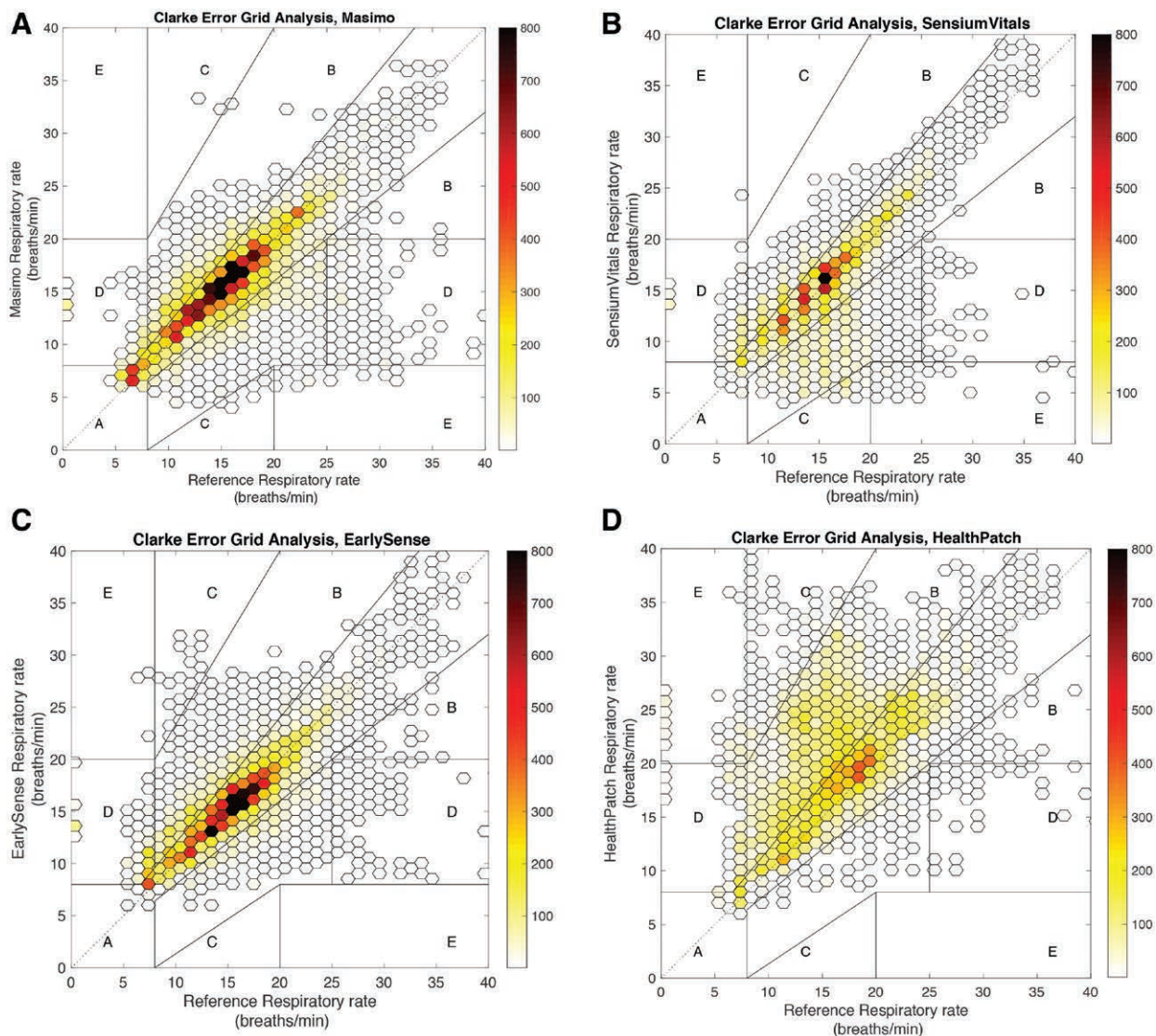
of 727h) for SensiumVitals, HealthPatch, EarlySense, and Masimo Radius-7, respectively. In addition, available HR data from the reference standard were not continuously available, either (data loss of 154 of 727h; 21.2%). For RR, data loss was 34.0% (122 of 359h), 13.1% (83 of 633h), 37.9% (250 of 659h), and 12.8% (92 of 720h) for SensiumVitals, HealthPatch, EarlySense, and Masimo Radius-7, respectively. From the reference standard, data loss of RR measurements was 20.6% (148 of 720h). An overview of overall data loss is shown in Table 4. Figure 6 shows the percentage of epochs without data divided over

gaps of data loss with a maximum length of 5 min, 15 min, 1 h, 4 h, or longer than 4 h. More than 90% of the gap durations were not longer than 15 min for both HR and RR, with the majority of gaps less than 5 min. However, for HealthPatch 84% of the gaps of HR data were below 15 min, and for Masimo Radius-7 89% of the gaps for RR were under 15 min.

## Discussion

We studied the performance of two wearable patch sensors, a patient-worn monitor, and a contactless mattress





**Fig. 5.** Clarke Error Grid analysis to quantify clinical accuracy of the respiratory rate measurements with the Masimo Radius-7 (A), SensiumVitals (B), EarlySense (C), and HealthPatch (D) as compared with the reference standard. The colored dots are measurement pairs each superimposed on the Error Grid boundaries, where the color intensity is proportional to the number of observations. Region A encloses points within 20% of the reference monitor; region B contains points outside 20% of the reference, but not leading to unnecessary treatment. Region C contains points leading to unnecessary treatment, region D indicates a potentially dangerous failure to detect bradypnoea or tachypnoea, and region E represents points where events are confused (e.g., bradypnoea with tachypnoea).

sensor to measure HR and RR continuously in high-risk surgical patients. The results show that these sensors can accurately measure HR, with the highest precision for the HealthPatch sensor. For RR, the accuracies of Masimo Radius-7, EarlySense, and SensiumVitals were within our predefined acceptable range. In contrast, the HealthPatch tended to overestimate RR. Furthermore, clinical accuracy of all sensors for HR was nearly 100%. EarlySense had the lowest percentage of RR measurements in regions D (failure to detect bradypnoea/tachypnoea) and E (opposite treatment). The wireless systems provided data more than

62% of time for RR and more than 73% of time for HR. Surprisingly, the wired reference system provided data only 79% of time for both HR and RR, which is lower than most wireless sensor systems. One possible explanation is pull on electrocardiogram electrodes from patient movement. Available HR and RR data were lowest for the bed-based EarlySense, since there are no data when the patient is not in bed, for instance during mobilization.

We recently reported on the accuracy of HealthPatch,<sup>27</sup> but including a wider range of wearable wireless sensors in this study provided a fuller understanding of the

**Table 3.** Clarke Error Grid Analysis to Quantify Clinical Accuracy of Heart Rate and Respiratory Rate of All Wireless Sensors

	Zone A No. (%)	Zone B No. (%)	Zone C No. (%)	Zone D No. (%)	Zone E No. (%)	Zone A + B No. (%)
Heart rate						
Masimo	34,645 (99.0)	166 (0.5)	0 (0)	181 (0.5)	0 (0)	34,811 (99.5)
EarlySense	29,123 (98.8)	206 (0.7)	19 (0.1)	122 (0.4)	0 (0)	29,329 (99.5)
HealthPatch	29,576 (99.9)	39 (0.1)	4 (0)	0 (0)	0 (0)	29,615 (100)
SensiumVitals	16,682 (98.6)	139 (0.8)	54 (0.3)	5 (0)	37 (0.2)	16,821 (99.4)
Respiratory Rate						
Masimo	27,198 (82.3)	4,610 (14.0)	235 (0.7)	907 (2.7)	82 (0.3)	31,808 (96.3)
EarlySense	22,417 (80.3)	4,597 (16.5)	302 (1.1)	597 (2.1)	8 (0)	27,014 (96.6)
HealthPatch	14,179 (48.7)	8,344 (28.6)	5,631 (19.3)	453 (1.5)	528 (1.8)	22,523 (77.3)
SensiumVitals	13,003 (74.0)	3,314 (18.7)	758 (4.3)	384 (2.2)	136 (0.7)	16,317 (92.7)

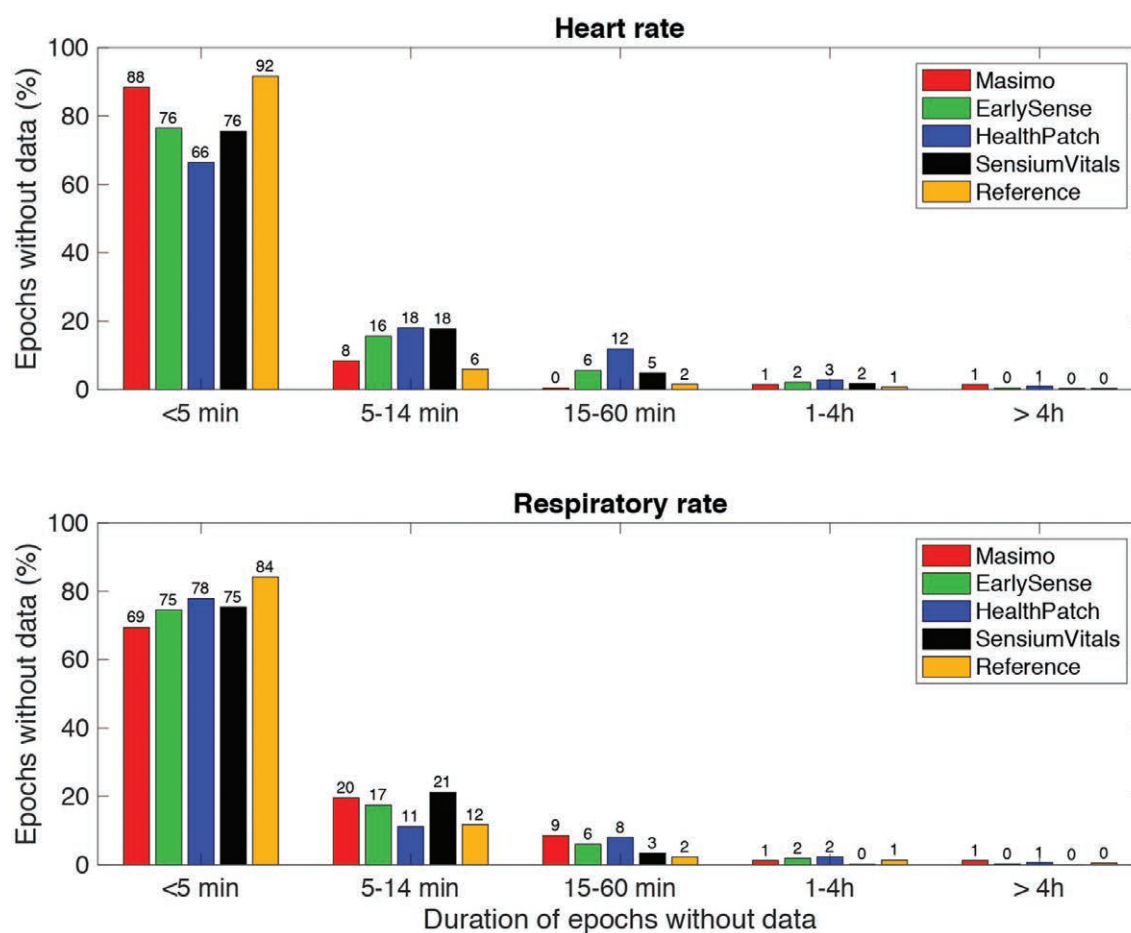
differences in reliability. Until now, most accuracy studies were obtained under laboratory conditions, or with low-risk patients for a limited time.<sup>22,25,26,28</sup> Comparison of RR from Masimo Radius-7 and capnography showed high accuracy at the postanesthesia care unit. Two recent studies reported high accuracy for Masimo Radius-7 RR when compared with capnography and impedance pneumography.<sup>31,32</sup> The EarlySense system can track HR and RR accurately in intensive care unit patients,<sup>29</sup> although data were obtained during supervised conditions. Other studies used nurse-recorded vital signs observations as reference method. Granholm *et al.*<sup>24</sup> mentioned an unacceptable lack of agreement between SensiumVitals and nurse readings, while Weenk *et al.*<sup>23</sup> reported that HealthPatch recordings were in agreement with nurse measurements. Although such studies hint at the potential of wireless monitoring in clinical practice, comparison with nurse observations cannot reliably indicate (continuous) performance of wearable sensors. First, nurses have digit preferences for RR readings, and their approximations deviate from the actual RR.<sup>38,39</sup> Second, nurse readings were evaluated against “spot” measurements, which is outside the intended scope of continuous monitoring. Finally, continuous vital signs cannot be validated in between nurse observations. In the current study, HR and RR were measured continuously with both the index sensors and a reference monitor in a high-risk clinical setting.

The results of the current study confirm that the wireless devices provide similar monitoring accuracy to the wired reference standard. However, each system has specific strengths and drawbacks. Both “patch-type” sensors use electrocardiogram to derive HR, which was highly accurate. The Masimo Radius-7 uses photoplethysmography from the finger probe to derive HR, which was accurate, except during atrial fibrillation with rapid ventricular rate, where it underestimated actual ventricular rate. The bed-based EarlySense monitor estimates HR with ballistocardiography and, as a result, it also underestimated HR during atrial fibrillation. For all devices, RR was clearly the more difficult vital sign to measure. The HealthPatch

sensor overestimated RR, whereas RR estimates from the SensiumVitals patch were more robust.

Although the reference standard used in the current study is part of our hospital-wide intensive care unit monitoring system, and thus clinically relevant, its thoracic impedance measurement cannot be considered a true “gold” standard for RR monitoring, and observed RR shows wide variations in patients who are moving and talking.<sup>40</sup> An unknown part of the measurement error is therefore potentially related to deficiencies of the reference standard, rather than the wireless sensors. To account for this, we also derived the limits of agreement with mixed effects models, which adds the mean of each patient over time and the mean measurement of each measurement pair as an explanatory variable. These results showed that the limits of agreement were reduced for both HR and RR of all four sensors under test, suggesting that RR derived with the reference differs from the actual RR. Although we considered using capnography as a reference standard instead of thoracic impedance for RR, in pilot tests we found that unsupervised capnography in spontaneously breathing patients was prone to frequent sensor malposition, high amounts of data loss, and poor patient acceptance.<sup>41</sup> Capnography as a reference standard for RR is only feasible for shorter periods of time when a research assistant is continuously present to observe and maintain correct sensor position.

We considered HR and RR to be acceptable for clinical purposes if within  $\pm 10\%$  of the reference standard or  $\pm 5$  beats/min or  $\pm 3$  breaths/min. These limits may be considered “wide” during controlled conditions, but not during unsupervised monitoring of patients who were at times moving and talking. However, guidelines for acceptable limits of agreement with continuous vital signs monitoring do not yet exist. It may be clinically desirable to redefine acceptable accuracy limits depending on the value of the vital sign measured; for example, for very high RRs it is less relevant if the RR is 32 or 35 breaths/min. However, for low RRs, it is critically important to know whether a patient’s RR is 8 or 5 breaths/min.



**Fig. 6.** Overview of epochs without data (%) divided over gaps of data loss with a maximum length of 5 min, 15 min, 1 h, 4 h, or longer than 4 h for each sensor. The EarlySense data contain epochs without data during bed exits and mobilization.

When interpreting the findings of this study, some limitations should be taken into account. First, the observation time available for agreement analysis varied per patient. This was due to variability in length of stay and whether patient participation started before surgery or upon admission. However, given the large amount of monitored time available, we believe we can draw valid inferences regarding the reliability of the different sensors. A second limitation is our inability to assess the exact amount of data loss in the EarlySense system, because no data were available on the number and duration of bed exits, which overestimated data loss. Although the accuracy of available EarlySense HR and RR data was good, it is important to realize that during periods of mobilization, there is no indication of a patient's vital status.

Data loss was highest for RR from the SensiumVitals sensor; however, a 66% available data rate still provides much more information regarding the patient's vital status than the current frequency of nurse readings on the ward. This data loss may be a result of the “conservative”

algorithm used, which strictly rejects potentially invalid RR readings in an attempt to reduce the incidence of false alarms, *e.g.*, from motion artifacts.<sup>22</sup> Moreover, more than 90% of episodes with data loss did not exceed 15 min, which means that with a 15-min moving median filter used in the current study, still two thirds of values are transmitted and brief transmission loss would not result in false alarms. Surprisingly, data loss for HR and RR data from the reference monitor was greater than 20%, possibly due to pull on electrocardiogram electrodes. In such cases, wireless HR monitoring might even outperform conventional wired systems.

We opted to use a “moving” median filter during 15-min data epochs with an update rate once every minute to eliminate HR and RR outliers resulting from patient motion. Applying such filtering, however, reduces the ability to detect sudden changes in vital signs, for example apnea or cardiac arrest. The current first generation of wearable devices is not designed to substitute for intensive



**Table 4.** Overall Amount of Data Loss of Each Sensor, Measured as Percentage of Total Monitoring Time in Minutes

	Percentage of Total Data Loss, No. (%)†
Heart Rate	
SensiumVitals	2,776 (12.9)
HealthPatch	4,726 (12.3)
EarlySense*	10,939 (27.5)
Masimo	2,830 (6.5)
Reference standard	9,232 (21.2)
Respiratory Rate	
SensiumVitals	7,338 (34.0)
HealthPatch	4,986 (13.1)
EarlySense*	14,999 (37.9)
Masimo	5,524 (12.8)
Reference standard	8,891 (20.6)

\*No data available during bed exits and mobilization. †Expected data loss was not excluded (e.g., sensors that were temporarily disconnected during certain diagnostic procedures).

care unit–grade monitoring systems for continuous monitoring. Their intended use is to identify abnormal trends in physiology in order to get the right care to the right patient at the right time. Consequently, eliminating spurious vital signs outliers at the expense of real-time monitoring might be a necessary trade-off. This approach minimizes false-positive alerts, which is paramount to avoid alarm fatigue and direct nurses to only go see patients who are in need of extra observation. This is relevant when remote patient monitoring solutions are deployed on a general surgical ward, where nurse-to-patient ratios are much lower, and in the home setting with only limited (remote) supervision.

The potential benefits of continuous remote wireless patient monitoring are increasingly recognized.<sup>16,42–44</sup> A recent review by Downey *et al.*<sup>45</sup> concluded that continuous monitoring outside critical care settings is feasible and may show patient benefits in terms of improved outcomes and cost efficiency. Another review suggested that implementation of remote monitoring with automated notifications increased involvement of ward physicians, rather than increased rapid response teams activation.<sup>43</sup> Nonetheless, large well-controlled studies in high-risk populations are needed to obtain evidence of the impact of remote monitoring on postoperative outcomes. Particular emphasis in future studies should be on the unintended consequences of remote monitoring, such as the risk of reduced patient contact and inappropriate consultation of end-users during the introduction of new technology.<sup>46</sup>

In conclusion, the tested wearable devices accurately represented HR. RR was clearly harder to measure, but the devices were accurate enough to identify abnormal patterns in RR. None of the tested devices is designed to substitute for continuous intensive care unit–grade monitoring systems, and our data suggest they cannot be used as such. However, wireless wearable vital signs sensors could

become valuable tools to reduce failure-to-rescue events within patients outside of high-care facilities.

## Research Support

The work described in this study was funded by the e-Health “Citrien” fund from the Netherlands Federation of University Medical Centres, Utrecht, The Netherlands. This work was also supported by the Horizon 2020 program of the European Commission (“Nightingale,” grant No. 727534).

## Competing Interests

At the time of the study, Ms. Breteler was a part-time employee of health IT company FocusCura (Driebergen-Zeist, The Netherlands), and Dr. Dohmen is founder and CEO of FocusCura. No funding from FocusCura was obtained, and FocusCura has no financial relationship with any of the device manufacturers. The other authors declare no competing interests.

## Correspondence

Address correspondence to Ms. Breteler: University Medical Center Utrecht, Mail Stop Q.04.2.313, P.O. Box 85500, 3508 GA Utrecht, The Netherlands. [m.j.m.breteler@umcutrecht.nl](mailto:m.j.m.breteler@umcutrecht.nl). This article may be accessed for personal use at no charge through the Journal Web site, [www.anesthesiology.org](http://www.anesthesiology.org).

## References

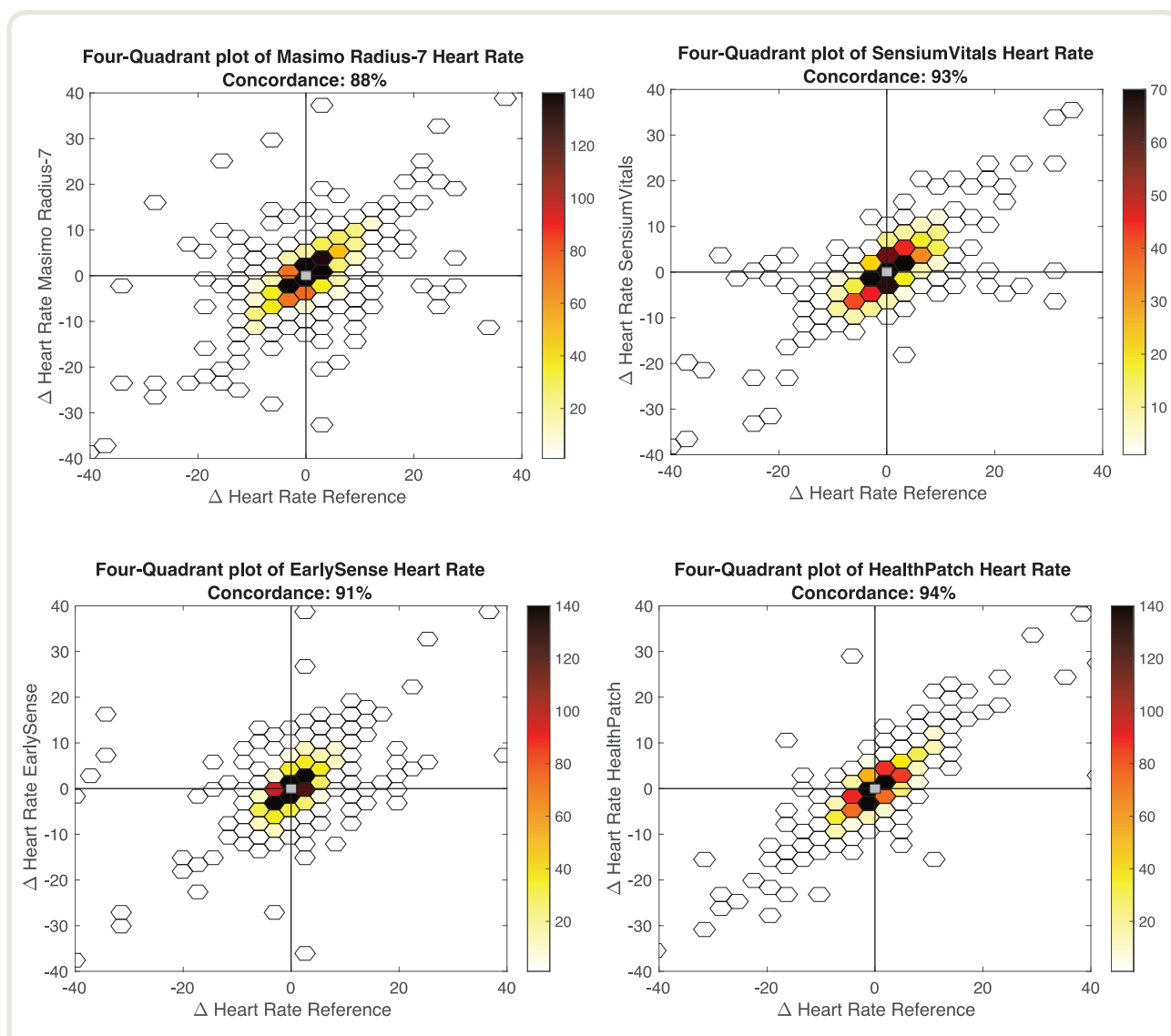
1. Buist M, Bernard S, Nguyen TV, Moore G, Anderson J: Association between clinically abnormal observations and subsequent in-hospital mortality: A prospective study. *Resuscitation* 2004; 62:137–41
2. Goldhill DR, White SA, Sumner A: Physiological values and procedures in the 24h before ICU admission from the ward. *Anaesthesia* 1999; 54:529–34
3. Cuthbertson BH, Boroujerdi M, McKie L, Aucott L, Prescott G: Can physiological variables and early warning scoring systems allow early recognition of the deteriorating surgical patient? *Crit Care Med* 2007; 35:402–9
4. Hillman KM, Bristow PJ, Chey T, Daffurn K, Jacques T, Norman SL, Bishop GF, Simmons G: Antecedents to hospital deaths. *Intern Med J* 2001; 31:343–8
5. Beckett Dj, Gordon C, Paterson R, Chalkley S, Macleod D, Bell D: Assessment of clinical risk in the out of hours hospital prior to the introduction of Hospital at Night. *Acute Med* 2009; 8:33–8
6. Fuhrmann L, Lippert A, Perner A, Østergaard D: Incidence, staff awareness and mortality of patients at risk on general wards. *Resuscitation* 2008; 77:325–30
7. Hogan J: Why don't nurses monitor the respiratory rates of patients? *Br J Nurs* 2006; 15:489–92



8. Cretikos MA, Bellomo R, Hillman K, Chen J, Finfer S, Flabouris A: Respiratory rate: The neglected vital sign. *Med J Aust* 2008; 188:657–9
9. Ludikhuize J, Smorenburg SM, de Rooij SE, de Jonge E: Identification of deteriorating patients on general wards; measurement of vital parameters and potential effectiveness of the Modified Early Warning Score. *J Crit Care* 2012; 27:424.e7–13
10. Cardona-Morrell M, Prgomet M, Lake R, Nicholson M, Harrison R, Long J, Westbrook J, Braithwaite J, Hillman K: Vital signs monitoring and nurse-patient interaction: A qualitative observational study of hospital practice. *Int J Nurs Stud* 2016; 56:9–16
11. Calzavacca P, Licari E, Tee A, Egi M, Downey A, Quach J, Haase-Fielitz A, Haase M, Bellomo R: The impact of Rapid Response System on delayed emergency team activation patient characteristics and outcomes—A follow-up study. *Resuscitation* 2010; 81:31–5
12. Kause J, Smith G, Prytherch D, Parr M, Flabouris A, Hillman K; Intensive Care Society (UK); Australian and New Zealand Intensive Care Society Clinical Trials Group: 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
13. Taenzer AH, Pyke JB, McGrath SP: A review of current and emerging approaches to address failure-to-rescue. *ANESTHESIOLOGY* 2011; 115:421–31
14. Hillman K, Chen J, Cretikos M, Bellomo R, Brown D, Doig G, Finfer S, Flabouris A; MERIT study investigators: Introduction of the medical emergency team (MET) system: A cluster-randomised controlled trial. *Lancet* 2005; 365:2091–7
15. Chan PS, Jain R, Nallmothu BK, Berg RA, Sasson C: Rapid response teams: A systematic review and meta-analysis. *Arch Intern Med* 2010; 170:18–26
16. Taenzer AH, Spence BC: The afferent limb of rapid response systems: Continuous monitoring on general care units. *Crit Care Clin* 2018; 34:189–98
17. DeVita MA, Smith GB, Adam SK, Adams-Pizarro I, Buist M, Bellomo R, Bonello R, Cerchiari E, Farlow B, Goldsmith D, Haskell H, Hillman K, Howell M, Hravnak M, Hunt EA, Hvarfner A, Kellett J, Lighthall GK, Lippert A, Lippert FK, Mahroof R, Myers JS, Rosen M, Reynolds S, Rotondi A, Rubulotta F, Winters B: “Identifying the hospitalised patient in crisis”—A consensus conference on the afferent limb of rapid response systems. *Resuscitation* 2010; 81:375–82
18. 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
19. Watkins T, Whisman L, Booker P: Nursing assessment of continuous vital sign surveillance to improve patient safety on the medical/surgical unit. *J Clin Nurs* 2016; 25:278–81
20. Pajarón-Guerrero M, Fernández-Miera MF, Dueñas-Puebla JC, Cagigas-Fernández C, Allende-Mancisidor I, Cristóbal-Poch L, Gómez-Fleitas M, Manzano-Peral MA, Gonzalez-Fernandez CR, Aguilera-Zubizarreta A, Sanroma-Mendizábal P; Hospital Valdecilla ERAS Study Group: 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
21. Roulin D, Donadini A, Gander S, Griesser AC, Blanc C, Hübner M, Schäfer M, Demartines N: Cost-effectiveness of the implementation of an enhanced recovery protocol for colorectal surgery. *Br J Surg* 2013; 100:1108–14
22. Hernandez-Silveira M, Ahmed K, Ang SS, Zandari F, Mehta T, Weir R, Burdett A, Toumazou C, Brett SJ: 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
23. Weenk M, van Goor H, Frietman B, Engelen IJ, van Laarhoven CJ, Smit J, Bredie SJ, van de Belt TH: Continuous monitoring of vital signs using wearable devices on the general ward: Pilot study. *JMIR Mhealth Uhealth* 2017; 5:e91
24. Granholm A, Pedersen NE, Lippert A, Petersen LF, Rasmussen LS: Respiratory rates measured by a standardised clinical approach, ward staff, and a wireless device. *Acta Anaesthesiol Scand* 2016; 60:1444–52
25. Chan AM, Selvaraj N, Ferdosi N, Narasimhan R: Wireless patch sensor for remote monitoring of heart rate, respiration, activity, and falls. *Conf Proc IEEE Eng Med Biol Soc* 2013; 2013:6115–8
26. Selvaraj N: Long-term remote monitoring of vital signs using a wireless patch sensor. *IEEE Healthcare Innovation Conference* 2014:1–4. DOI: 10.1109/HIC.2014.7038880
27. Breteler MJM MSc, Huizinga E, van Loon K, Leenen LPH, Dohmen DAJ, Kalkman CJ, Blokhuis TJ: 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. *BMJ Open* 2018; 8:e020162
28. 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
29. Ben-Ari J, Zimlichman E, Adi N, Sorkine P: Contactless respiratory and heart rate monitoring: Validation of an innovative tool. *J Med Eng Technol* 2010; 34:393–8
30. Zimlichman E, Szyper-Kravitz M, Shinar Z, Klap T, Levkovich S, Unterman A, Rozenblum R, Rothschild JM, Amital H, Shoenfeld Y: Early recognition of acutely deteriorating patients in non-intensive care units: Assessment of an innovative monitoring technology. *J Hosp Med* 2012; 7:628–33
31. Ramsay MA, Usman M, Lagow E, Mendoza M, Untalan E, De Vol E: The accuracy, precision and

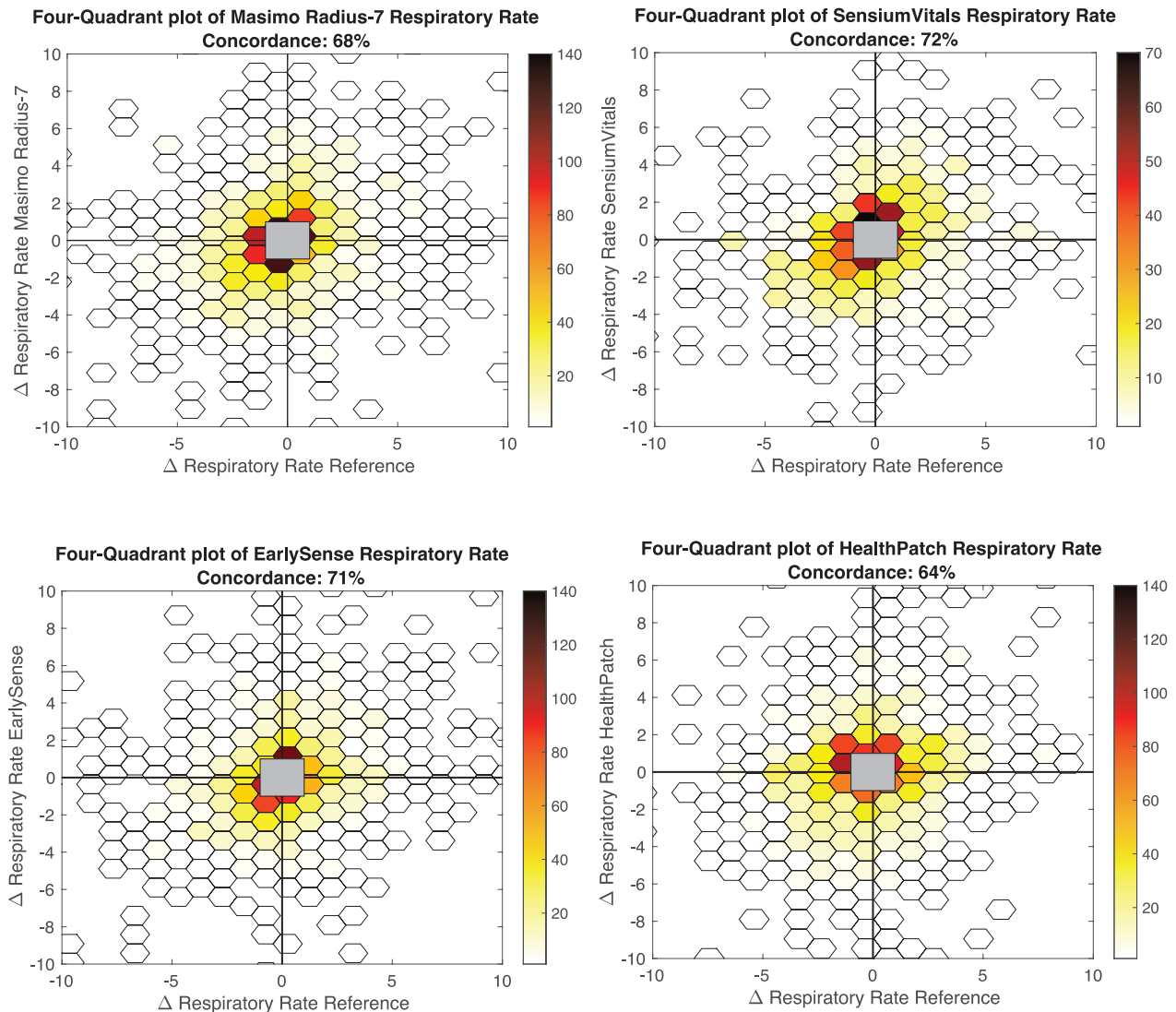
- reliability of measuring ventilatory rate and detecting ventilatory pause by rainbow acoustic monitoring and capnometry. *Anesth Analg* 2013; 117:69–75
32. van Loon K, Peelen LM, van de Vlasakker EC, Kalkman CJ, van Wolfswinkel L, van Zaane B: Accuracy of remote continuous respiratory rate monitoring technologies intended for low care clinical settings: A prospective observational study. *Can J Anaesth* 2018; 62:137–9
  33. Subbe CP, Kruger M, Rutherford P, Gemmel L: Validation of a modified Early Warning Score in medical admissions. *QJM* 2001; 94:521–6
  34. Bland JM, Altman DG: Agreement between methods of measurement with multiple observations per individual. *J Biopharm Stat* 2007; 17:571–82
  35. Myles PS, Cui J: Using the Bland-Altman method to measure agreement with repeated measures. *Br J Anaesth* 2007; 99:309–11
  36. Clarke WL, Cox D, Gonder-Frederick LA, Carter W, Pohl SL: Evaluating clinical accuracy of systems for self-monitoring of blood glucose. *Diabetes Care* 1987; 10:622–8
  37. Saugel B, Grothe O, Wagner JY: Tracking changes in cardiac output: Statistical considerations on the 4-quadrant plot and the polar plot methodology. *Anesth Analg* 2015; 121:514–24
  38. Badawy J, Nguyen OK, Clark C, Halm EA, Makam AN: Is everyone really breathing 20 times a minute? Assessing epidemiology and variation in recorded respiratory rate in hospitalised adults. *BMJ Qual Saf* 2017; 26:832–6
  39. Flenady T, Dwyer T, Applegarth J: Accurate respiratory rates count: So should you! *Australas Emerg Nurs J* 2017; 20:45–7
  40. Wiklund L, Hök B, Ståhl K, Jordeby-Jönsson A: Postanesthesia monitoring revisited: Frequency of true and false alarms from different monitoring devices. *J Clin Anesth* 1994; 6:182–8
  41. Macknet M, Kimball-Jones P, Applegate R II, Martin R, Allard M: Accuracy of a novel bioacoustic sensor for monitoring respiratory rate. *Eur J Anaesthesiol* 2007; 24:34
  42. McGillion MH, Duceppe E, Allan K, Marcucci M, Yang S, Johnson AP, Ross-Howe S, Peter E, Scott T, Ouellette C, Henry S, Le Manach Y, Paré G, Downey B, Carroll SL, Mills J, Turner A, Clyne W, Dvirnik N, Mierdel S, Poole L, Nelson M, Harvey V, Good A, Pettit S, Sanchez K, Harsha P, Mohajer D, Ponnambalam S, Bhavnani S, Lamy A, Whitlock R, Devereaux PJ; PROTECT Network Investigators: Postoperative remote automated monitoring: Need for and state of the science. *Can J Cardiol* 2018; 34:850–62
  43. Boer C, Touw HR, Loer SA: Postanesthesia care by remote monitoring of vital signs in surgical wards. *Curr Opin Anaesthesiol* 2018; 31:716–22
  44. Subbe CP, Duller B, Bellomo R: Effect of an automated notification system for deteriorating ward patients on clinical outcomes. *Crit Care* 2017; 21:52
  45. Downey C, Randell R, Brown J, Jayne DG: Continuous versus intermittent vital signs monitoring using a wearable, wireless patch in patients admitted to surgical wards: Pilot cluster randomized controlled trial. *J Med Internet Res* 2018; 20:e10802
  46. Prgomet M, Cardona-Morrell M, Nicholson M, Lake R, Long J, Westbrook J, Braithwaite J, Hillman K: Vital signs monitoring on general wards: Clinical staff perceptions of current practices and the planned introduction of continuous monitoring technology. *Int J Qual Health Care* 2016; 28:515–21

## Appendix 1



**Fig. A1.** Four-quadrant plot showing the  $\Delta$  heart rate values for both the reference standard and the wearable sensors. The values on the horizontal axis refer to  $\Delta$  heart rate values of the reference; the vertical axis refers to the  $\Delta$  heart rate values of the studied sensors.

## Appendix 2



**Fig. A2.** Four-quadrant plot showing the  $\Delta$  respiratory rate values for both the reference standard and the wearable sensors. The values on the horizontal axis refer to  $\Delta$  respiratory rate values of the reference; the vertical axis refers to the  $\Delta$  respiratory rate values of the studied sensors.