



# Accuracy of remote continuous respiratory rate monitoring technologies intended for low care clinical settings: a prospective observational study

## Précision des technologies de surveillance continue de la fréquence respiratoire à distance destinées aux unités cliniques de basse intensité : étude observationnelle prospective

Kim van Loon, PhD, MD · Linda M. Peelen, PhD · Emmy C. van de Vlasakker, MD · Cor J. Kalkman, PhD, MD · Leo van Wolfswinkel, PhD, MD · Bas van Zaane, PhD, MD

Received: 8 January 2018 / Revised: 18 June 2018 / Accepted: 21 June 2018 / Published online: 7 September 2018  
© The Author(s) 2018

### Abstract

**Purpose** Altered respiratory rate (RR) has been identified as an important predictor of serious adverse events during hospitalization. Introduction of a well-tolerated continuous RR monitor could potentially reduce serious adverse events such as opioid-induced respiratory depression. The purpose of this study was to investigate the ability of different monitor devices to detect RR in low care clinical settings.

**Methods** This was a prospective method-comparison study with a cross-sectional design. Thoracic impedance pneumography (IPG), frequency modulated continuous wave radar, and an acoustic breath sounds monitor were compared with the gold standard of capnography for their ability to detect RR in breaths per minute (breaths·min<sup>-1</sup>) in awake postoperative patients in the postanesthesia care unit. The Bland and Altman method for repeated measurements and mixed effect modelling was used to

obtain bias and limits of agreement (LoA). Furthermore, the ability of the three devices to assist with correct treatment decisions was evaluated in Clarke Error Grids.

**Results** Twenty patients were monitored for 1,203 min, with a median [interquartile range] of 61 [60–63] min per patient. The bias (98.9% LoA) were 0.1 (−7.9 to 7.9) breaths·min<sup>-1</sup> for the acoustic monitor, −1.6 (−10.8 to 7.6) for the radar, and −1.9 (−13.1 to 9.2) for the IPG. The extent to which the monitors guided adequate or led to inadequate treatment decisions (determined by Clarke Error Grid analysis) differed significantly between the monitors ( $P = 0.011$ ). Decisions were correct 96% of the time for acoustic, 95% of the time for radar, and 94% of the time for IPG monitoring devices.

**Conclusions** None of the studied devices (acoustic, IPG, and radar monitor) had LoA that were within our predefined (based on clinical judgement) limits of  $\pm 2$  breaths·min<sup>-1</sup>. The acoustic breath sound monitor predicted the correct treatment more often than the IPG and the radar device.

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s12630-018-1214-z>) contains supplementary material, which is available to authorized users.

K. van Loon, PhD, MD (✉) · L. M. Peelen, PhD · E. C. van de Vlasakker, MD · C. J. Kalkman, PhD, MD · L. van Wolfswinkel, PhD, MD · B. van Zaane, PhD, MD  
Department of Anesthesiology, Intensive Care and Emergency Medicine, University Medical Center Utrecht, Utrecht University, P.O. Box 85500, 3508 GA Utrecht, The Netherlands  
e-mail: K.vanLoon-3@umcutrecht.nl

L. M. Peelen, PhD  
Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

### Résumé

**Objectif** L'altération de la fréquence respiratoire (FR) est un facteur prédictif important d'événements indésirables graves au cours d'une hospitalisation. L'introduction d'un moniteur continu de la FR bien toléré pourrait contribuer à réduire les événements indésirables graves tels que la dépression respiratoire induite par les opioïdes. L'objectif de cette étude était de déterminer la capacité de différents dispositifs de surveillance à détecter la FR dans des cadres de soins cliniques non intensifs.

**Méthodes** Il s'agissait d'une étude prospective transversale de comparaison de méthodes. La pneumographie par impédance thoracique (IPG), le radar à ondes continues de modulation de fréquence, et une surveillance acoustique des bruits respiratoires ont été comparés à la capnographie comme référence pour leur capacité à détecter la FR en respirations par minute (respirations-min<sup>-1</sup>) chez les patients éveillés en postopératoire dans une unité de soins post anesthésiques. La méthode de Bland et Altman pour la création de modèles de répétabilité des mesures et d'effets mixtes a servi à obtenir les biais et limites de concordance (LC). En outre, la capacité des trois dispositifs à contribuer aux bonnes décisions thérapeutiques a été évaluée au moyen des grilles d'erreur de Clarke.

**Résultats** Vingt patients ont été suivis pendant 1 203 minutes, avec une durée médiane (plage interquartile) de 61 (60 à 63) minutes par patient. Les biais (LC de 98,9 %) étaient de 0,1 (−7,9 à 7,9) respiration-min<sup>-1</sup> pour le moniteur acoustique, −1,6 (−10,8 à 7,6) pour le radar et −1,9 (−13,1 à 9,2) pour l'IPG. La mesure dans laquelle les moniteurs contribuaient à des décisions thérapeutiques adéquates ou inadéquates (selon l'analyse par la grille d'erreur de Clarke) était significativement différente selon les moniteurs utilisés ( $P = 0,011$ ). Les décisions étaient correctes dans 96 % des cas pour le moniteur acoustique, 95 % des cas pour le radar et 94 % des cas pour la surveillance par IPG.

**Conclusions** Aucun des dispositifs étudiés (moniteurs acoustiques, IPG et radar) n'avait une LC entrant dans nos limites prédéfinies (basées sur le jugement clinique) de  $\pm 2$  respirations-min<sup>-1</sup>. Le moniteur respiratoire acoustique a prédit le traitement correct plus souvent que les dispositifs IPG et radar.

More than 70% of preventable in-hospital deaths are associated with “failure to rescue” and “death in low-mortality risk groups”.<sup>1</sup> Unexpected intensive care unit admissions, cardiopulmonary arrests, and deaths are preceded by abnormal physiologic parameters 6-24 hr prior to these events.<sup>2-5</sup> Altered respiratory rate (RR) is potentially the most important predictor of these serious adverse events.<sup>6,7</sup> Although RR (typically indicated as breaths per minute) is a very informative physiologic parameter, it is not routinely and accurately measured in patients on general hospital wards.<sup>8,9</sup> Caregivers typically assess RR by counting chest wall movements for at most 15-30 sec, while the other physiologic parameters (e.g., systolic blood pressure and temperature) are only

intermittently measured with well-tolerated reliable equipment.

Accurate continuous RR monitors for low care clinical settings can facilitate early recognition of abnormal respiratory patterns and early intervention (e.g., by alerting the rapid response team), which in turn might reduce the incidence of serious adverse events and unfavourable patient outcomes.<sup>10</sup> The functionality of the complete monitoring strategy – which includes the sensing principle, signal analysis, caregiver notification, and protocolized therapeutic intervention – is essential to determine whether introduction of a continuous RR monitor improves patient outcomes.<sup>11</sup>

In this methods-comparison study, we focus on the first and second components (the sensing principle and the signal analysis) of the monitoring strategy. Although our eventual intent is to study the ability of different devices to detect RR in low care clinical settings, such as general hospital wards, these technologies first need evaluating in settings where accurate reference standards can be assured. The current measurement of RR on general hospital wards based on counting chest wall movement is not accurate enough to serve as a reference standard. Furthermore, the incidence of respiratory abnormalities on the general hospital ward is likely too low for such a method-comparison study. Therefore, we studied the ability of frequency modulated continuous wave (FMCW) radar, thoracic impedance pneumography (IPG), and breath sound acoustic monitoring to detect RR in comparison with the reference standard (capnography) in surgical patients during the early postoperative phase on the postanesthesia care unit (PACU). The study was conducted in PACU patients as they are a closely monitored diverse group of patients who share important similarities with our target population and have a variety of respiratory alterations. We hypothesized that remote continuous RR monitors (radar, IPG, and acoustic) can measure RR and recognize abnormal breathing to a similar standard as capnography.

## Methods

### Study population

We studied adult patients who underwent major abdominal surgery and received postoperative care on the PACU at a large tertiary teaching hospital in the Netherlands (University Medical Centre Utrecht). The institutional review board (IRB) reviewed the study protocol and found that it was not subject to the Dutch act on “medical research involving human subjects” (14-278/C, November 2014). Although the IRB waived the need for informed

consent, our study group opted to request written informed consent from all participants. The study was conducted in accordance with the moral, ethical, and scientific principles governing clinical research as set out in the Declaration of Helsinki (2013) and good clinical practice.<sup>12</sup>

### Design and conduct of the study

This method-comparison study had an observational cross-sectional diagnostic design. Hence, all RR monitors were fitted to patients in a similar systematic manner at the same time. Three index devices, namely FMCW radar (RespiRadar; ART Innovation Company, Rotterdam, The Netherlands),<sup>13</sup> thoracic IPG (from the Datex Ohmeda Compact S3, module e-PRESTN, Datex Ohmeda, Inc., Helsinki, Finland),<sup>14</sup> and breath sound acoustic monitor (RRa; RAD-87<sup>TM</sup> Pulse CO-Oximeter, Masimo Corporation, Irvine, CA, USA)<sup>15</sup> were compared with sidestream capnography (the reference standard). Treatment decisions were solely based on vital signs displayed on routinely used bedside monitors (pulse oximetry, electrocardiogram, oscillometric or invasive blood pressure measurement), as per current guidelines for postoperative care.

### Reference standard

Capnography was performed using a nasal cannula (Smart Capnoline; Oridion Medical 1987 Ltd., Israel) connected to the capnograph (Compact S3 module E-miniC, Datex Ohmeda, Inc., Helsinki, Finland). The RR derived from the capnograph was calculated as a 1 min average and was updated breath-by-breath. The manufacturer reported an accuracy of  $\pm 1$  breath·min<sup>-1</sup>. Supplemental oxygen was delivered over the cannula at a maximum of 4 L·min<sup>-1</sup>. Patients needing more oxygen received oxygen through a mask placed over the cannula.

### Index devices

We studied three different index monitors with different sensing principles. The first, the FMCW radar was a prototype that determines RR by frequency and amplitude differences between electromagnetic radio waves emitted to and reflected by the chest wall.<sup>13</sup> It is a non-invasive, wireless, contactless RR monitor that is mounted to the ceiling above the bed. The radar determines the RR over 20 sec epochs with an update rate of 1.6 Hz. The accuracy of the FMCW radar has not been previously reported. The second index device was thoracic IPG, which measured impedance changes (0.2-7 Ohm) over electrocardiographic electrodes.<sup>14</sup> Impedance changes originate from an altered gas to fluid volume ratio together with an altered length of

conductance induced by the inhalation and exhalation of air. The IPG determined the RR over 30 sec epochs with an update rate of 0.1 Hz. The reported accuracy in the device manual is  $\pm 5$  breaths·min<sup>-1</sup>. The third index device, the breath sound acoustic monitor, includes an adhesive transducer applied to the patient's neck that detects the acoustic signals induced by air displacement during breathing.<sup>15</sup> The manufacturer reported a bias of 0.8 breaths·min<sup>-1</sup> (standard deviation [SD] 3.4) in postsurgical patients compared with sidestream capnography.<sup>15</sup> The update rate was 0.5 Hz, and the averaging mode was 20 sec with an alarm delay for apnea of 30 sec. Other alarms were switched off.

### Data collection and processing

Signals were synchronously sampled and recorded with corresponding time stamps in the Anesthesia Information Management System (AIMS) (AnStat; Carepoint, Ede, the Netherlands). The reference and index devices used different averaging times for RR calculation. We were not able to adjust these averaging times to allow for comparison of the moving averages. Therefore, we stored the median values of RR once every minute to enable comparison. The raw FMCW radar waveform information (describing the amplitude and phase of the reflected radio waves) was stored separately on secure digital cards. Signal-processing, based on a zero-crossing algorithm, was performed offline and RRs derived from the radar device were subsequently stored in the AIMS. Offline processing of the raw FMCW radar data was performed by an engineer blinded to the RRs derived from the other devices. Two observers (K.L. and E.V.) observed the patients and their caregivers continuously for voluntary movement, talking, and specific activities (without interfering with the regular care process) to identify potential artifacts. Observations were categorized and listed in advance. In addition, free text observations were made. The appendix in the electronic supplement information describes the activities that were observed during the study measurements.

The ability of a monitor to support correct treatment decisions is clinically very important and graphically depicted in a Clarke Error Grid (CEG). The reference lines in a CEG for RR that mark the boundaries between correct and incorrect treatment decisions were determined by consensus between the authors.

### Outcome measures

The primary outcome was RR (breaths·min<sup>-1</sup>) compared with the reference standard. We considered limits of agreement (LoA) within  $\pm 2$  breaths·min<sup>-1</sup> as clinically relevant, considering the allowance for false negatives. A

secondary outcome was the ability of the monitors to guide treatment decisions using CEG analysis, expressed as the percentage of data pairs representing adequate and inadequate treatment decisions.

### Statistical analysis

The published statistical literature does not provide formal rules to calculate statistical power for method-comparison studies when multiple measurements are recorded per patient. Therefore, the sample size by convenience was set at 20 patients, each of who had at least 60 consecutive data points (median RR every min) per device. The primary outcome (RR) was analyzed using both the Bland and Altman (BA) method for repeated measurements and the mixed effect models as suggested by Myles.<sup>16,17</sup> In the BA method, the bias is determined by the mean of the subject means. The LoA are calculated with the SD, which is corrected for the number of observations per subject and multiple testing (LoA  $\pm 2.63 \cdot \text{SD}$ ). We also report the variance ratio (between-subject variance/total variance) as the repeated measurements correction is only applied to the subject variance. Furthermore, we calculated the LoA that were derived by adding the residual random effects from two mixed effect models for the index and reference device to the total variance. Myles previously suggested using this method as alternative for handling repeated measurements.<sup>17</sup> For mixed effect modelling, we used the “nlme” package version 3.1-131 of R (<https://cran.r-project.org/web/packages/nlme/index.html>) with the restricted maximum likelihood method.

We studied the potential influence on treatment decisions using CEG.<sup>18</sup> Adequate (areas a, b), unnecessary (area c), inadequate (area d), and opposite (area e) treatment decisions are reflected by different regions in CEGs with cut-off points for bradypnea ( $< 9$  breaths·min<sup>-1</sup>) and tachypnea (individually assessed using definitions of  $> 20$  and  $> 25$  breaths·min<sup>-1</sup>) with 98.9% confidence intervals (CI). Furthermore, CEG percentages for areas a-e were compared between monitors with a Fisher exact test. We also studied the positive predictive value and negative predictive value for abnormal respiration defined as bradypnea ( $< 9$  breaths·min<sup>-1</sup>) and tachypnea ( $> 25$  breaths·min<sup>-1</sup>) combined. Analyses were performed with R version 3.1.1 (R Foundation for Statistical Software, Vienna, Austria).

### Results

From December 15 2014 until February 20 2015, 20 patients were enrolled. Study measurements were performed for 1,203 min, with a median [interquartile

range] of 61 min [60-63] per patient. For the acoustic breath sound monitor we stored fewer ( $m = 952$  min) complete data pairs because of data acquisition problems in four patients. The RR measurements derived from the capnograph ranged between 4 and 38 breaths·min<sup>-1</sup>. Table 1 shows other patient, surgical, and respiratory pattern characteristics.

The primary outcomes are presented in Table 2. Compared with the RR measured by reference capnography, the acoustic breath sound monitor showed a bias of 0.1 breaths·min<sup>-1</sup> with a 98.9% LoA of  $-7.9$  to  $7.9$  breaths·min<sup>-1</sup>. The 98.9% LoA derived from the mixed effect model was  $-6.2$  to  $6.2$  breaths·min<sup>-1</sup>. The bias and 98.9% LoA for IPG and FMCW radar showed a larger deviation from the reference. The BA plots for IPG, FMCW radar, and the acoustic monitor are depicted in Fig. 1, with LoAs from the BA method in red, and from the mixed effect models in black. The intensity of the dots displays the number of superimposed measurement pairs. The BA plot for the acoustic monitor shows a horizontal pattern around the bias near zero for the average RR range of 5-28 breaths·min<sup>-1</sup>. The dispersion of measurement pairs around the bias was larger for the radar and IPG measurements.

The CEG results on measurement error and its influence on treatment decisions at the same time is presented in Fig. 2, where the CEGs are plotted. The percentages of data pairs in the regions a-e are provided in Table 3. The extent to which the monitors guided adequate or led to inadequate treatment decisions (determined by CEG analysis) differed significantly between the three monitors ( $P = 0.011$ ). Adequate treatment decisions (areas a and b) with 98.9% CI would have been arrived at 96 (95 to 97)%, 95 (93 to 96)% and 94 (91 to 95)% of the time for acoustic, radar, and IPG monitoring respectively. Area e of the CEG contains those points that would treat patients with bradypnea as if they had tachypnea, and vice-versa. None of the measurements with the acoustic monitor are in this region, while the radar and IPG have respectively 0.1% and 0.7% measurement pairs in this region.

### Discussion

Diagnostic studies evaluating non-invasive RR monitors are scarce and even fewer studies describe wireless RR monitors for use on the general hospital ward.<sup>11</sup> In this study, we focused on continuous RR monitors, their accuracy, and to which extent they cause erroneous treatment decisions in a clinical setting that partly mimics postoperative admission to a surgical hospital ward. For the three monitors that were studied, the LoAs were outside our predefined acceptable LoA of  $\pm 2$

**Table 1** Baseline characteristics of 20 participants monitored during their postoperative stay on the PACU

	<i>n</i> = 20
Age (yr)	58.5 (14.7)
Sex (male)	9 (45%)
BMI (kg.m <sup>-2</sup> )	27.9 (4.0)
<i>Surgical procedure</i>	
Duration (min)	159 [141-179]
Laparoscopic procedure	14 (70%)
Specialty (GE-urology-gynecology)	13-5-2 (65-25-10%)
Analgesia: epidural	9 (45%)
Analgesia: intravenous opioids	8 (40%)
<i>Comorbidities</i>	
Cardiovascular disease	9 (45%)
Pulmonary disease	4 (20%)
Obstructive sleep apnea syndrome (OSAS)	0 (0 %)
<i>Respiratory parameters during study measures</i>	
Respiratory rates (capnograph) [range]	4-38 breaths·min <sup>-1</sup>
No. patients with apneas	4 (20%)
End-tidal carbon dioxide level (kPa) [mean (SD)]	4.8 (0.5)
Peak end-tidal carbon dioxide level (kPa) [range]	4-8.4
Missing capnography RR measurements (min)	13 (1%)
Oxygen desaturation (SpO <sub>2</sub> < 90 %)	6 (16%)
High supplemental oxygen flow (> 4 L·min <sup>-1</sup> )	6 (30%)

Data are presented as mean (standard deviation), median [interquartile range], or percentage (%), as indicated. BMI = body mass index; GE = gastroenterology; PACU = postanesthesia care unit; RR = respiratory rate; SD = standard deviation

**Table 2** Primary outcome, respiratory rate as measured by different monitors (IPG, FMCW radar, and acoustic breath sound monitoring) compared with capnography

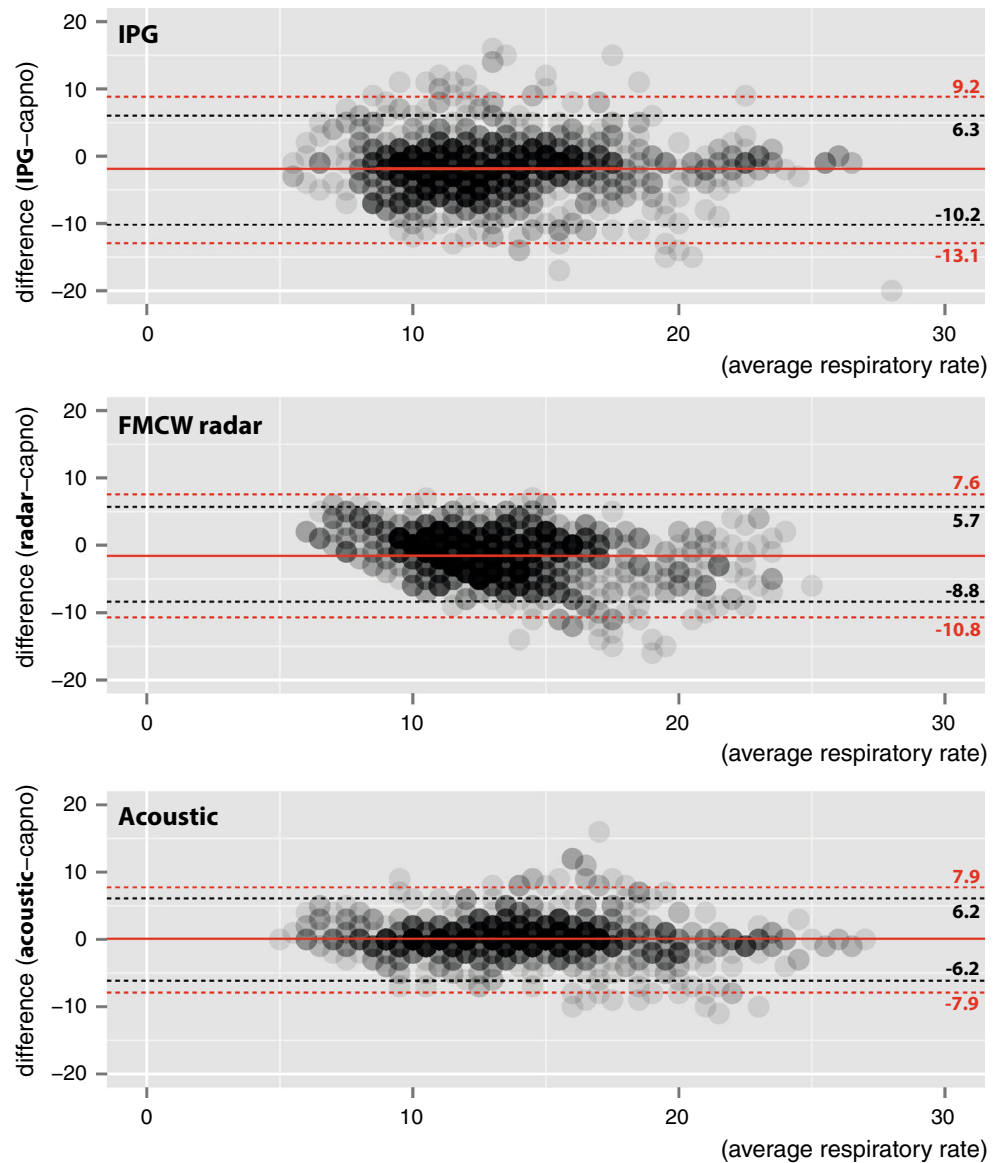
Device	IPG	Radar	Acoustic
Bias (breaths·min <sup>-1</sup> )	-1.9	-1.6	0.1
SD [BA] (breaths·min <sup>-1</sup> )	4.1	3.5	2.9
Lower LoA [BA]	-13.1	-10.8	-7.9
Upper LoA [BA]	9.2	7.6	7.9
SD [M] (breaths·min <sup>-1</sup> )	3.1	2.8	2.3
Lower LoA [M]	-10.2	-8.8	-6.2
Upper LoA [M]	6.3	5.7	6.2
<i>m</i> pairs	1101	1188	952*
Missing data pairs (%)	2.5	0.2	1.0
VAR ratio	0.31	0.29	0.26

BA = Bland and Altman method; Bias = mean difference; FMCW= frequency modulated continuous wave; IPG = thoracic impedance pneumography; LoA = 98.9% limit of agreement; M = (Myles) mixed effects model method; *m* pairs = number of complete measurement pairs used in the analysis; SD = standard deviation; VAR<sub>ratio</sub> = between-subject variance/total variance. \*The study included 20 patients; however, we experienced problems acquiring data with the acoustic breath sound monitor in three patients

breaths·min<sup>-1</sup>. The acoustic monitor had the highest percentage of RR measurements in CEG area a (i.e., correct treatment decisions and within 20% of the reference

measurement) and the lowest percentage of RR measurements in CEG area d (i.e., dangerous failure to detect bradypnea/tachypnea) and area e (i.e., opposite

**Fig. 1** Bland and Altman plots for impedance pneumography (IPG), radar, and acoustic monitors. The intensity of the dots indicates the number of superimposed measurement pairs (1-10). 98.8% limits of agreement from the Bland and Altman method in red, and from mixed effect models in black. FMCW = frequency modulated continuous wave



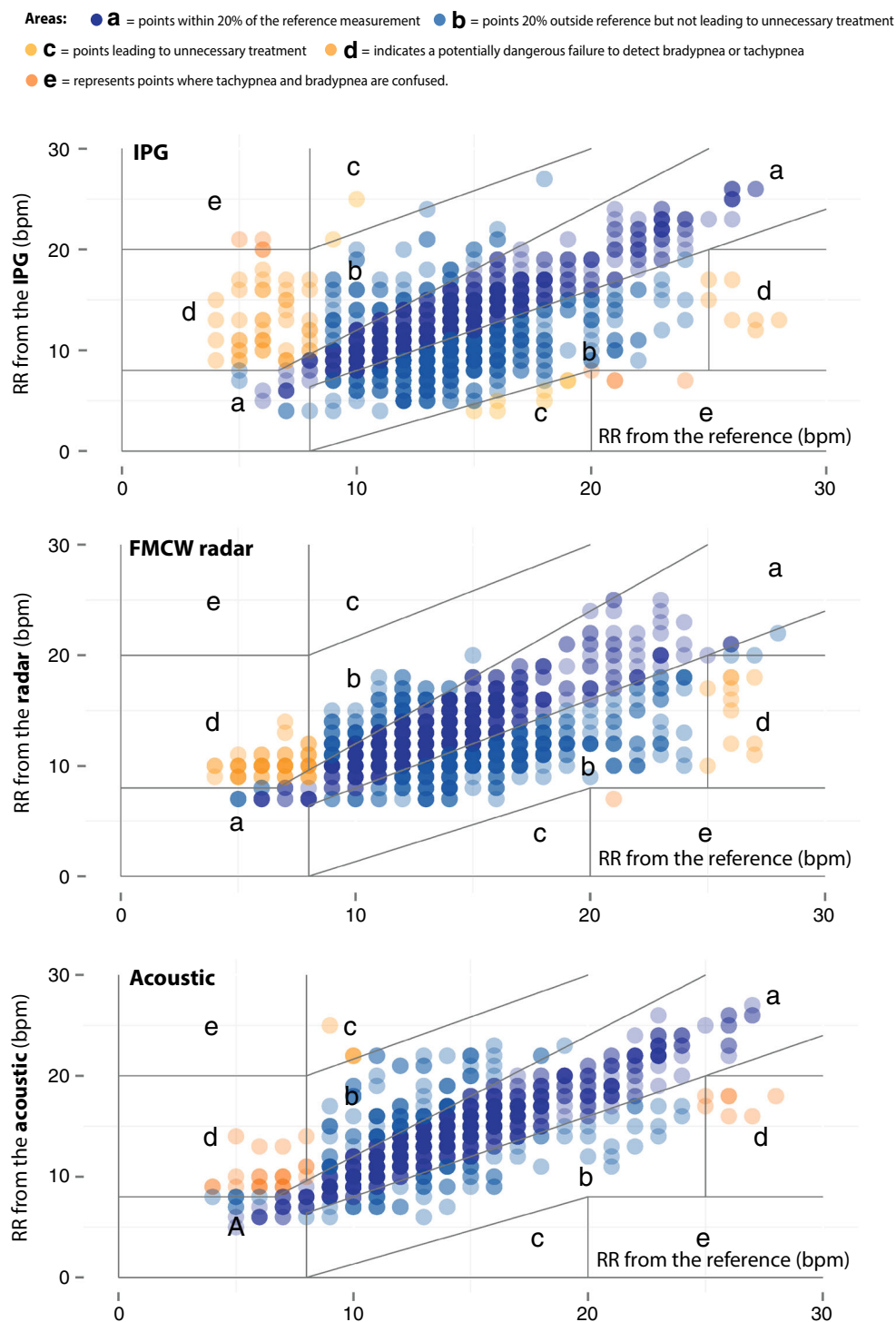
treatment decisions). These percentages differed significantly from the other devices (IPG and FMCW radar).

When interpreting these findings, the presumed superiority of the acoustic monitor may, at least partly, be due to the choice of capnography as a potentially imperfect reference standard. In addition, the acoustic monitor may have an advantage compared with the two other index devices, as the capnograph and the acoustic monitor both measure displacement of air by tracing carbon dioxide concentrations or sound production during the respiratory cycle respectively. Patient activities such as talking and partially obstructive breath (e.g., snoring) can interrupt the displacement of air and may cause measurement artifacts in a similar direction. Unlike

capnography and acoustic monitoring, IPG and radar are likely to suffer from movement artifacts as these modalities rely on the measurement of chest wall movements. We advocate selecting and combining different sensor technologies to obtain reliable information on RR in all clinical situations. Future studies should evaluate the signal analytic techniques necessary to select the correct RR measurement for a specific clinical situation.

The finding that the accuracy for all three index monitors was outside the predefined limits needs our consideration. Capnography has proved its additive value in the high care setting during mechanical ventilation,<sup>19</sup> but this has been translated into clinical use in lower or intermediate care settings. The accuracy of capnography may also be reduced during spontaneous breathing in

**Fig. 2** Clarke Error Grids for respiratory rate measured with impedance pneumography (IPG), radar, and acoustic sensor. Areas labelled a-e provide information on measurement error and its influence on treatment. Areas: dark blue a = points with 20% of the reference measurement; light blue b = points 20% outside reference but not leading to unnecessary treatment; light orange c = points leading to unnecessary treatment; light orange d = indicates a potentially dangerous failure to detect bradypnea or tachypnea; dark orange e = represents where tachypnea and bradypnea are confused; bpm = breaths·min<sup>-1</sup>; FMCW = frequency modulated continuous wave; RR = respiratory rate



patients who receive supplemental oxygen, which dilutes expired carbon dioxide, resulting in erroneously low end-tidal carbon dioxide.<sup>12,20</sup> It is thus important to realize that an undefined part of the measurement error is potentially attributed to the capnograph, rather than to the index devices. For this reason, we also calculated the bias and 95% LoA with mixed effect models that add the mean of

each measurement pair as an explanatory variable. This corrects, at least partly, the measurement error that is attributable to the capnograph. The SD obtained from the mixed effect models are reduced for all three methods, suggesting that the RR obtained with capnography did indeed differ from the true RR. Despite this improvement

**Table 3** Diagnostic accuracy for abnormal breathing (defined as RR below 10 or above 25 breaths·min<sup>-1</sup>)

	Clarke Error Grid					Diagnostic Accuracy					
	a (%)	b (%)	c (%)	d (%)	e (%)	TP	FP	FN	TN	PPV % [95% CI]	NPV % [95% CI]
IPG	55	38	1	5	0.7	21	164	63	853	11 [7 to 17]	93 [91 to 95]
FMCW radar	59	36	0	5	0.1	22	48	64	1054	31 [22 to 43]	94 [93 to 95]
Acoustic	78	18	0.4	3	0	41	29	34	848	59 [47 to 69]	96 [95 to 97]

Clarke Error Grid areas a-e reflect the potential influence on treatment decisions that are classified as 'a' adequate, 'b' adequate but 20% off the reference, 'c' unnecessary, 'd' inadequate, and 'e' opposite. CI = confidence interval; FMCW = frequency modulated continuous wave; FN = false negatives; FP = false positives; NPV = negative predictive value; PPV = positive predictive value; TN = true negatives; TP = true positives

using the Myles method,<sup>17</sup> the monitors frequently exceeded our predefined LoA.

Although our investigation focused on the accuracy of the various non-invasive RR monitors in the low care setting, other factors also play a role in the usability of continuous RR monitoring in this setting. First, it is important that the sensing principle detects most respiratory problems. Monitoring displacement of air (as with the capnograph and acoustic monitor) can detect upper airway obstruction, while monitoring chest wall movements (as with IPG and radar) and may falsely suggest adequate breathing even in the presence of severe upper airway obstruction. Second, patient comfort is also an important factor in the choice of RR monitors as it can hinder both accuracy and patient acceptance. Patient contact and wires are needed for the capnograph and acoustic monitors. This would, in many cases, reduce patient comfort, limit patient mobility, and thereby threaten the sustainability of the monitoring strategy. We encountered this issue during a previous observational pilot study, in which we intended to study respiratory patterns with capnography on hospital wards (IRB decision 10-032/C, February 2010). That study involved additional nasal cannula (Smart Capnoline; Oridion Medical 1987 Ltd. Israel), which annoyed patients and it was terminated early as all six patients refrained from further participation within a few hours of starting monitoring (unpublished results). Similar issues with patient discomfort could conceivably occur with long-term use of the acoustic and IPG adhesive sensors. The radar was the only wireless and contactless RR monitor evaluated in this study. The radar was a prototype monitor that provided us with detailed information on all patient movements, whereas for the current study, we were only interested in chest wall movements to determine RR. An update of this prototype with smart algorithms that are able to detect voluntary movement will likely improve the accuracy and sustainability in low care clinical settings. Such

algorithms should be able to reliably distinguish between breaths and voluntary movements.

There were some limitations to our study. First, we were unable to obtain a complete acoustic RR data set in all the patients because of data storage problems. This, however, is unlikely to have influenced the study conclusions, as the LoA in our analyses only included patients with similar complete data sets. Whether our sample size (of approximately 1,200 data pairs in 20 patients) is sufficient remains unclear as formal rules for sample size calculation in method-comparison studies with repeated measurements have not been established. Another potential limitation is the study setting itself – while our overarching aim is to better understand the characteristics of these monitors in low care settings such as general hospital wards, due to the need to use a continuous reference standard, we studied patients in the PACU during the immediate postoperative period with capnography as our reference. In addition, we measured a wide range of RRs that were also outside normal values. We presume that these results can be generalized to patients admitted to the general hospital ward with other pathologies causing abnormal RRs. Confirming the functionality on the general hospital ward in a method-comparison study is difficult without a suitable reference standard. Therefore, we studied the monitors' accuracy in a PACU setting with a pragmatic study design. Such a study should ultimately be followed by an intervention study assessing the additive value on general hospital wards.<sup>11</sup>

## Conclusion

This study directly compared acoustic, IPG, and FMCW radar monitoring with capnography monitoring for RR during spontaneous breathing in postsurgical patients in the immediate postoperative period. RR measurements with the acoustic monitor were most accurate but the LoA were wide. Accordingly, none of the studied continuous RR



monitors were accurate enough to be introduced into clinical practice.

**Editorial responsibility** This submission was handled by Dr. Hilary P. Grocott, Editor-in-Chief, *Canadian Journal of Anesthesia*.

**Author contributions** This study was designed by *Kim van Loon, Linda M. Peelen, Cor J. Kalkman, and Bas van Zaane*. *Kim van Loon and Emmy C. van de Vlasakker* approached the participants, managed study inclusion, and performed study measurements and observations. *Leo van Wolfswinkel and Kim van Loon* designed the data collection. *Kim van Loon, Emmy C. van de Vlasakker, Leo van Wolfswinkel, and Linda M. Peelen* analyzed the data. *Kim van Loon, Linda M. Peelen, Cor J. Kalkman, and Bas van Zaane* wrote the manuscript. All authors read and approved the manuscript.

**Disclosure/funding** This study was funded by the Department of Anesthesiology of the University Medical Centre Utrecht. No external funding was obtained. Applied Radar Technology (ART) Innovation Company contributed to the study by providing a prototype of the FMCW radar on loan. Masimo International SARL contributed to the study by providing 20 adhesive RRA sensors for the RAD-87 monitor. All authors disclose no commercial or non-commercial affiliations that may be perceived as a conflict of interest related to this work.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

1. *Institute of Medicine (US) Committee on Quality of Health Care in America*. To Err Is Human: Building a Safer Health System, 1st edition. Kohn LT, Corrigan JM, Donaldson MS (Eds). Washington, DC: National Academies Press, 2000.
2. *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.
3. *Hillman KM, Bristow PJ, Chey T, et al*. Antecedents to hospital deaths. *Intern Med J* 2001; 31: 343-8.
4. *Hodgetts TJ, Kenward G, Vlachonikolis IG, Payne S, Castle N*. The identification of risk factors for cardiac arrest and formulation of activation criteria to alert a medical emergency team. *Resuscitation* 2002; 54: 125-31.
5. *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.
6. *Fieselmann JF, Hendryx MS, Helms CM, Wakefield DS*. Respiratory rate predicts cardiopulmonary arrest for internal medicine inpatients. *J Gen Intern Med* 1993; 8: 354-60.
7. *Cretikos M, Bellomo R, Hillman K, Chen J, Finfer S, Flabouris A*. Respiratory rate: the neglected vital sign. *Med J Aust* 2008; 188: 657-9.
8. *Odell M, Victor C, Oliver D*. Nurses' role in detecting deterioration in ward patients: systematic literature review. *J Adv Nurs* 2009; 65: 1992-2006.
9. *McBride J, Knight D, Piper J, Smith GB*. Long-term effect of introducing an early warning score on respiratory rate charting on general wards. *Resuscitation* 2005; 65: 41-4.
10. *Taenzler AH, Pyke JB, McGrath SP*. A review of current and emerging approaches to address failure-to-rescue. *Anesthesiology* 2011; 115: 421-31.
11. *van Loon K, van Zaane B, Bosch EJ, Kalkman CJ, Peelen LM*. Non-invasive continuous respiratory monitoring on general hospital wards: a systematic review. *PLoS One* 2015; 10: e0144626.
12. *World Medical Association*. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA* 2013; 310: 2191-4.
13. *van Loon K, Breteler MJ, van Wolfswinkel L, et al*. Wireless non-invasive continuous respiratory monitoring with FMCW radar: a clinical validation study. *J Clin Monit Comput* 2016; 30: 797-805.
14. *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.
15. *Ramsay MA, Usman M, Lagow E, Mendoza M, Untalan E, De Val 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.
16. *Bland JM, Altman DG*. Agreement between methods of measurement with multiple observations per individual. *J Biopharmaceutical Stat* 2007; 17: 571-82.
17. *Myles PS, Cui JJ*. Using the Bland-Altman method to measure agreement with repeated measures. *Br J Anaesth* 2007; 99: 309-11.
18. *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.
19. *Nassar BS, Schmidt GA*. Capnography during critical illness. *Chest* 2016; 149: 576-85.
20. *Casati A, Gallioli G, Scandroglio M, Passaretta R, Borghi B, Torri*. Accuracy of end-tidal carbon dioxide monitoring using the NBP-75 microstream capnometer. A study in intubated ventilated and spontaneously breathing nonintubated patients. *Eur J Anaesthesiol* 2000; 17: 622-6.