Classifying signals from a wearable accelerometer device to measure respiratory rate

Gordon B. Drummond, Darius Fischer, Margaret Lees, Andrew Bates, Janek Mann, D.K. Arvind

Classifying signals from a wearable accelerometer device to measure respiratory rate.

Gordon B Drummond[1]*, Darius Fischer[2], Margaret Lees[3], Andrew Bates[2], Janek Mann[2], and DK Arvind[4]

1. Honorary Senior Lecturer, Department of Anaesthesia, Critical Care, and Pain Medicine, University of Edinburgh, Edinburgh Royal Infirmary 51 Little France Crescent Edinburgh EH16 4SA
2. Research Associate, Centre for Speckled Computing, School of Informatics, University of Edinburgh, Room 1.33, Informatics Forum, 10 Crichton Street, Edinburgh EH8 9AB, UK
3. Research Nurse, Princess Margaret Eye Pavilion, 45 Chalmers Street Edinburgh EH3 9HA
4. Professor and Director, Centre for Speckled Computing, School of Informatics, University of Edinburgh, Room 2.18, Informatics Forum, 10 Crichton Street, Edinburgh EH8 9AB, UK

*Corresponding author: email g.b.drummond@ed.ac.uk

Summary:

We studied acutely ill patients wearing a small wireless motion sensor. Using machine learning, we developed a method to classify sections of the record and measure and record respiratory rate accurately and automatically, to allow automatic charting.

251 characters with spaces (max 256)

ABSTRACT

Background. Automatic measurement of respiratory rate in general hospital patients is difficult. Patient movement degrades the signal, and variation of the breathing cycle means that accurate observation for at least 60 seconds is needed for adequate precision.

Methods. We studied acutely ill patients recently admitted to a teaching hospital. Breath duration was measured from a tri-axial accelerometer attached to the chest wall, and compared with a signal from a nasal cannula. We randomly divided the patient records into a training (n=54) and a test set (n=7). We used machine learning to train a neural network to select reliable signals, automatically identifying signal features associated with accurate measurement of respiratory rate. We used the test records to assess the accuracy of the device, indicated by the median absolute difference between respiratory rates, provided by the accelerometer and by the nasal cannula.

Results. In the test set of patients, machine classification of the respiratory signal reduced the absolute difference from 1.25 (0.56 to 2.18) to 0.48 (0.30 to 0.78) breaths/minute (median, interquartile range). 50 % of the recording periods were rejected as unreliable, and in one patient, only 10% of the signal time was classified as reliable. However, even only 10% of observation time would allow accurate measurement for 6 minutes in an hour of recording, giving greater reliability than nurse charting, which is based on much less observation time.

Conclusion. Signals from a body-mounted accelerometer yield accurate measures of respiratory rate, which could improve automatic illness scoring in adult hospital patients.

Word count 250

Introduction

To assess acute illness, scoring systems using measurements such as respiratory rate, heart rate, systolic blood pressure, and oxygen saturation are often used in hospital settings.[1] Respiratory rate is a key observation in early recognition of acutely ill patients in hospital, and is important in illness scoring, such as in COVID 19 infection [2] where it has prognostic value. [3, 4] Respiratory rate is not commonly measured automatically, although clinicians now recognise the need for reliable wearable monitoring devices to assist management of outbreaks of infectious disease.[5]

Current measurements of respiratory rate need to be improved. Breaths are counted for a limited time,[6] values recorded by different observers often differ,[7] and charted values are not trustworthy.[8]

Accurate measurement of respiratory rate with a wearable device has practical difficulties. Artefacts caused by patient movement can cause repeated false alarms. In surveillance monitoring, rapid changes in patient condition are likely to be less than in critical illness, and management of artefacts may require a different strategy, such as signal selection.

A poorly recognised factor is that the breathing signal itself is not stable, which can affect comparisons of methods of measurement. Breath duration varies from breath to breath, and measurements of rate made using short periods of counting, such as 30 or 60 seconds, are imprecise. Repeating observations made over one minute gives results with an interquartile range of 3 breaths/minute.[6] When comparing methods, exactly the same breaths should be sampled, but previous studies failed to do this.[9, 10] Recent studies of automatic measurement of respiratory rate in low care and high dependency patients report similar limits of agreement, about ± 4 breath/minute.[11–13]

To monitor respiratory rate in a general hospital setting, we need sufficient reliable and accurate information to identify acute illness or deterioration.[14, 15] This should be possible by signal processing to select trustworthy sections of the signal. Although signal quality may sometimes be unreliable, a sample of less frequent but more trustworthy observations could provide adequate input values for a nursing chart and the illness scoring system, and more accurately identify acute illness or deterioration. [16, 17]

We used a non-invasive monitor (Respeck) to measure respiration by sensing chest wall movement, transmitting the data by a Bluetooth low energy connection to a bedside receiver. In patients after surgery, we obtained reliable measures of respiratory rate,[18] but these patients were "settled" and inactive. We now assessed

how the device performed when used in acutely ill medical patients. We developed an automatic classification process to detect periods when the signal was reliable, by identifying features that were present when the signal corresponded with a second, reliable, method of measuring breathing. Our comparison method was nasal cannula pressure, which measures gas flow at the nose.[19]

The research question was whether the device could be trained to measure respiratory rate *accurately* and *reliably* in patients with acute medical illness. An *accurate* process would give values that agreed with a standard measure, within ± 2 breaths/minute. Measurements would be *reliable* if they were sufficient for adequate input into a hospital early warning chart. Currently, national UK recommendations[1] suggest respiratory rate should be measured on admission to hospital and then 12 hourly, and increased up to hourly in patients with severe illness.

Methods

Ethical approval

We wished to recruit patients who had just been admitted to hospital with acute illness. The Scotland A Research Ethics Committee considered the research (ref 12/SS/0054) met the requirements of section 51 of the Adults with Incapacity (Scotland) Act with regard to research subjects who lack capacity to consent. We were thus not permitted to record patient data that could allow identification.

Patient recruitment.

A trained research nurse worked with nursing staff in the acute medical receiving unit of a 570 bed teaching hospital. Patients with medical illnesses are admitted for initial assessment in one ward, and moved to a second ward if further investigation and treatment is required. Most patients spend about 4-8 hours in the unit and no patient is in the unit for more than 24 hours.

Patients who might be approached by the research nurse were identified by the clinical nursing staff, and given an information leaflet. If the patient moved to the second ward, the study was explained to them, they were asked if they wished to take part, and recruited if they agreed. If possible, signals were to be recorded for one hour, or until the patient was prepared to leave the ward. Patients were allowed to discontinue wearing the devices if they wished.

Patient recordings

The Respeck sensor measures acceleration in three orthogonal axes. The accelerometer signals are digitised and transmitted by Bluetooth LE to a receiver

(iPod) for storage and later retrieval and analysis. The Respeck is enclosed in a plastic capsule 45×38×13 mm, and weighs 15.4 g. It is marked to allow a consistent orientation when attached to the patient. When attached as described below, a positive signal from the X axis indicates a force from right to left, on the Y axis from caudal to cranial, and on the Z axis from anterior to posterior. The device is sealed in a plastic bag, and the bag is taped to patient's skin using a conforming, perforated polyester fabric tape with acrylic adhesive (Mefix, Mölnlycke Health Care Limited, Dunstable, UK) to hold the capsule firmly on the body surface. The device is placed just under the rib margin, in the midclavicular line, and usually on the left side of the patient.

A single-use nasal cannula (Sleep Sense 15805-2, Medes Ltd., Radlett, Herts UK) was placed below the nostrils to provide a comparison signal. The cannula was connected via a bacterial filter to a battery-powered pressure transducer (PTAF2, Respironics, www.philips.com/respironics). This was placed beside the patient and connected to a base -station which transmitted the pressure signal wirelessly (Bluetooth LE) to the iPod receiver. All signals were digitised at 12.5 Hz.

Details of the reason for admission of the patient to hospital, and the year of birth, height, weight and gender of the patient were recorded, but no other patient data were kept. Recordings were transferred from the receiver to a secure computer for further analysis.

Signal analysis.

The nasal flow signal in each patient record was replayed, without displaying the motion sensor signals, using proprietary software (Spike2, version 5.19, CED, Cambridge, UK). The breathing pattern was inspected by a skilled observer unaware of the identity of the patient. Sections of each record that showed clear breathing signals were identified. Breath onset was timed when the nasal pressure became less than a threshold set to be just below atmospheric, and automatically marked using a "decrease through threshold" setting. The entire record was visually checked by the observer to avoid any false or missed breaths. After breath onsets had been measured, the patient records were randomly divided into two groups. The first group was used to develop and train a method to classify the Respeck signal (see below), and the second was used to test the classification model performance.

Respeck signal analysis

Gravity is the predominant component of acceleration sensed by the device. As the chest wall expands and changes shape, the angle of the device relative to gravity changes. The angle changes of the axis measurements are smoothed and combined. Previous studies showed a close correspondence between the Respeck signal and

nasal air flow.[18] By setting upper and lower threshold limits for the Respeck, we measure the timing of respiratory phases when the signal successively crosses these threshold limits. The limits are determined dynamically by the signal amplitude. Phases of breathing are identified by measuring successive half-breath cycles, with the duration of inspiration taken as the shorter.

Record synchronisation and comparison

Nasal cannula and Respeck signal values were time-stamped when stored on the iPod. The Respeck signal was shifted in time by 960ms to correct for delay introduced by the smoothing function. The respiratory rate measurements from the nasal cannula and the Respeck signals were compared in 20-second windows, with successive windows overlapping by 10 seconds. These would contain about 5-7 breaths, and allowed dynamic comparisons of the two signals. The instantaneous frequency of each breath, nasal cannula and Respeck, was calculated as 60 *[1/breath duration (sec)]. The breathing rate for the window was reported as the mean of the instantaneous rates of each complete breath cycle in the 20-second window. The accuracy of the Respeck measurement was defined as the modulus of the difference between nasal cannula and Respeck rates.

Features associated with accurate estimates of respiratory rate.

To allow training of the selection method, and subsequent testing, satisfactory records were blinded, randomly numbered, and sorted for length. To ensure that the sample used for testing was adequate, the shortest four records were temporarily removed. The remainder were then separated into a sample of 50 for training the selection method, and a sample of 7 for final testing. The short records were then added back to the training group.

In each window of the Respeck signal, features were identified that could show whether the measurements could be trusted. These features were used to classify the Respeck signal windows and only windows that gave a close estimate of the breathing rate were selected.[20] A trusted period was defined as one in which the difference between nasal cannula and Respeck rate was no greater than a specified threshold value. There is an inverse relationship between this threshold, and agreement of the estimate with the breathing rate, and also the proportion of the record that could be selected. The threshold value was initially set to 1.0 breaths/minute. For different values of threshold less than 1, features of the signals in all the windows were compared for each patient record in the training set. We inspected the distribution of these features to generate a list of those that would possibly classify Respeck signal windows as those with a breathing rate close to the nasal cannula breathing rate. (see table 1)

Table 1

Initial features considered for testing in the model. These are derived from the waveform of the Respeck signal, the accelerometer signals, and the breath timing measurements

Factors	Comment or explanation
Fourier transform of Respeck signal	Normalised as (Fourier factors in range 9 - 24 breaths/minute)/ (sum of all frequency factors)
Mean, median, and Standard deviation (SD) of individual breathing rates	Extreme values are less likely to be reliable. A smaller SD reflects more regular breathing
Maximum – minimum rate	A large range of rates within the sampling window would be less trustworthy
Mean and median of differences	Mean and median of frequency differences between one breath and the next, within the window.
SD of axis acceleration values	A measure of large changes in axis signal, often associated with posture changes
SD, median and skew of activity levels	Activity level is computed as the length of the vector differences, using the X Y and Z coordinates
SD of Respeck signal	A large SD indicates probable posture change
Number of valid breaths	A small number of detected breaths suggests the signal is unreliable
SD of Respeck amplitude peaks.	Greater variation in the size of breaths in a window suggests unstable breathing.
Mean and median of peak differences	Large differences between breath amplitudes suggest unstable breathing
Mean/median of peak intervals	Times between successive peaks in the window should be similar

The most effective features were selected and ranked using recursive feature elimination and logistic regression. A limited number of features would save on computation, allowing implementation to run on the Respeck device itself. A neural network with two hidden layers was trained to select acceptable windows using a linear rectifier as the activation function.[21, 22] Using leave-one-out cross-validation for each patient, the model was trained for the rest of the patients in the set and then tested on the chosen patient. This prevents the model learning subject-specific patterns before testing. The features, ranked in descending order of effectiveness, were: the SD of the breathing rates, the mean difference of the breathing rates, the number of rate values in the window, the mean of the peak intervals, the SD of the breathing signal and the difference between the maximum and minimum rates.

Statistical analysis.

We used the modulus of the difference between nasal cannula and Respeck rates to assess agreement. For comparison with other reports, we compared the Respeck and

nasal cannula measures, before and after selection of reliable signals, using the method of Bland and Altman for repeated measures in multiple subjects.[23]

Results

Clinical information

We approached 96 patients for the study. Of these, 78 agreed to participate. In 12 data sets, the nasal pressure records were absent or of poor quality, and in 5, the recording was too short to be useful (less than 10 minutes). Thus 61 records were used for analysis. The age of these patients was 61 (18) years (mean, SD), their BMI was 28 (7) kg.m^{-2}, and 35 were female. The most frequent presenting illness was respiratory disease, either exacerbation of chronic obstructive pulmonary disease or asthma (16 patients). Ischaemic heart disease or venous thromboembolism was suspected in 14 patients, and 13 had either sepsis or urinary tract infection. The remainder had neurological conditions such as stroke (9), gastrointestinal disease (6), and other conditions (4). Many patients had comorbid conditions: these were cardiovascular in 32, respiratory in 20, diabetes in 11, and renal impairment in 12.

The total number of 20-second windows analysed was 23,278. Table 2 shows how these periods were used in the study.

Numbers are sample windows (% total)	Sample	
	Used for training	Used to test
Total windows	21055	2223
No cannula or Respeck rate	603 (3)	29 (1)
No Respeck rate	618 (3)	67 (3)
No cannula rate	4844 (23)	363 (16)
Discarded due to undefined feature	575 (3)	94 (4)
Used for analysis	14415 (68)	1670 (75)

Table 2.

The breathing rates of the patients, measured from the nasal pressure records, varied considerably. The breathing rate for each patient was summarised as a median duration for all the breaths measured in that patient's record. For all patients, breath duration was 2.80 (2.40 to 3.44) seconds (median, interquartile range). These represent breathing rates of 21 (17 to 25) breaths/minute. Figure 1 shows a representative trace from the study. It shows a period of quiet breathing, transiently affected by a change most prominent in one axis, which disturbs the Respeck signal and would affect the matching between the Respeck rate and the nasal cannula rate.

Figure 2 shows the results obtained when the selection process was applied to the test data. Before selection, the median difference between Respeck and cannula rates was 1.25 breaths/minute. After selection, the median difference was reduced to 0.46 breaths/minute. Although this required removal of a median 58% of the time periods recorded, in all but one of the test subjects the remaining time available for analysis would have been sufficient to give one minute of respiratory rate measurements in every ten minutes of data. Figure 3 shows the change in performance after selection. The limits of agreement ranged from -2.01 to 1.9 breaths/minute and the confidence intervals of these limits were small. This contrasts with the greater limits found before signal selection.

Figure 4 compares two plots of rate estimates for patients A and G in the test sample (see figure 2). The record for patient G includes only 10% of acceptable windows, but there remain sufficient estimates to record the respiratory rates in this 15-minute period.

Discussion

A machine learning-based approach to Respeck data analysis has shown that useful respiratory rate information can be extracted from noisy Respeck signals in a challenging clinical setting. In the test data set, the respiratory frequency matched the reference values closely (median absolute difference of 0.46 breath/min). The experimental setup to gather time-synchronised Respeck and reference data ensured comparisons over exactly contemporaneous periods. A previous analysis of the nasal cannula signals obtained from this study showed that repeated estimates of respiratory rate, made from the same record, but at different times, could vary substantially.[6] The performance after signal selection is superior to results reported in previous studies.[11–13, 24, 25]

The study has several strengths: a medically appropriate cohort of patients, studied in clinical circumstances, using an appropriate and accurate comparison signal. The respiratory signal was evaluated before the signal analysis was undertaken, and with the other signal records concealed. The nasal cannula signal allows the exact timing of the breathing cycle to be identified easily and precisely, making it an attractive choice as a "gold standard" in comparison to other devices which give a less faithful measure of respiration.

We had some technical difficulties in this study. We had to set up a new system for recording nasal pressure, and this caused some early failures to record. Some of the nasal pressure records were of poor quality, probably caused by fluid within the cannula tubing. This could not be detected because the signal could not be checked as the recording was being made. Some records may have been degraded by

predominant mouth breathing, and speech, coughing, cannula displacement or disconnection to allow ambulation can also interrupt nasal pressure measurements.

Signal analysis for model development was conducted without reference to the nasal pressure waveform, although onset times of breaths were used to measure the breath periods. We noted substantial heterogeneity in the breathing of the patients whose records were used to provide the training dataset. We considered that this would be helpful in training the classification process, because the model incorporates these variations when considering the test data and future subjects.

The final model performed well on the test data. To allow comparison with other studies, we present this in a Bland and Altman analysis. (figure 3) This showed a median rate difference of only 0.84 breaths/min, 95% limits of agreement from -2.01 to 1.90, and in 90% of the patients, the rate differences were less than 1.27 breaths/min. This degree of precision should be clinically acceptable for a scoring system. Very few reports assessing respiratory rate measuring devices define acceptable agreement explicitly, but one study suggested that a difference less than 2 breaths/minute indicates equivalence.[26]

Previous work has been done to recognise respiratory signals that are "reliable". For example, some cardiac monitors measure respiration from an impedance signal measured using ECG chest leads, and body motion frequently makes this signal unreliable. A computer system of decision rules was developed to score reliability of this signal, using a database of recordings from trauma victims.[27] Subsequent re-analysis of data using the selected signals improved prediction of outcome events.[28] Our work uses the same approach, but has the strength of incorporating a concurrent reference measure, rather than depending on subjective identification of "trusted" sections of the original record.

Although vital signs monitoring in general wards appears to improve early detection of illness,[29, 30] monitoring using continuous acquisition of data is not clearly more effective than when using intermittent chart-based systems.[31] Detecting deterioration may be more effective when the time course of values is used, rather than using threshold values.[32, 33] If so, regular and accurate measurement of respiratory rate may be particularly useful. More precise measurements may be more effective for scoring systems with narrow cutoff ranges (20, 21-24, 25).

Ideally, an automatic summary measure, at 10-minute intervals, could display features such as a progressive increase in rate in a patient with sepsis, or a decrease associated with excess opioid. Machine learning allowed us to select trustworthy signals from our monitor to do this, giving results that differ by less than 2

breath/minute from a reference device. This promising approach would allow automatic charting and provide input to illness scoring systems.

Word count 2999

References

1. National Institute for Clinical Excellence. Acutely ill patients in hospital. Recognition of and response to acute illness in adults in hospital. 2007.Available from: https://www.nice.org.uk/guidance/CG50. Last checked but not updated January 2020, accessed 7 September 2020

2. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, Zhang L, Fan G, Xu J, Gu X, Cheng Z, Yu T, Xia J, Wei Y, Wu W, Xie X, Yin W, Li H, Liu M, Xiao Y, Gao H, Guo L, Xie J, Wang G, Jiang R, Gao Z, Jin Q, Wang J, Cao B. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 2020; 395: 497–506.

3. Cecconi M, Piovani D, Brunetta E, Aghemo A, Greco M, Ciccarelli M, Angelini C, Voza A, Omodei P, Vespa E, Pugliese N, Parigi TL, Folci M, Danese S, Bonovas S. Early Predictors of Clinical Deterioration in a Cohort of 239 Patients Hospitalized for Covid-19 Infection in Lombardy, Italy. *J. Clin. Med.* 2020; 9.

4. Burian E, Jungmann F, Kaissis GA, Lohöfer FK, Spinner CD, Lahmer T, Treiber M, Dommasch M, Schneider G, Geisler F, Huber W, Protzer U, Schmid RM, Schwaiger M, Makowski MR, Braren RF. Intensive Care Risk Estimation in COVID-19 Pneumonia Based on Clinical and Imaging Parameters: Experiences from the Munich Cohort. *J. Clin. Med.* 2020; 9.

5. Ming DK, Sorawat S, Chanh HQ, Nhat PTH, Yacoub S, Georgiou P, Holmes AH. Continuous physiological monitoring using wearable technology to inform individual management of infectious diseases, public health and outbreak responses. *Int. J. Infect. Dis.* 2020; 96: 648–654.

6. Drummond GB, Fischer D, Arvind D. Current clinical methods of measurement of respiratory rate give imprecise values. *Eur. Respir. J. Open* 2020; .*in press.*

7. Lim WS, Carty SM, Macfarlane JT, Anthony RE, Christian J, Dakin KS, Dennis PM. Respiratory rate measurement in adults - How reliable is it? *Respir. Med.* 2002; 96: 31–33.

8. 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: 832–836.

9. Bianchi W, Dugas AF, Hsieh YH, Saheed M, Hill P, Lindauer C, Terzis A, Rothman RE. Revitalizing a vital sign: Improving detection of tachypnea at primary triage. *Ann. Emerg. Med.* 61: 37–43.

10. van Loon K, Breteler MJM, van Wolfwinkel L, Rheineck Leyssius AT, Kossen

S, Kalkman CJ, van Zaane B, Peelen LM. Wireless non-invasive continuous respiratory monitoring with FMCW radar: a clinical validation study. *J. Clin. Monit. Comput.* 2016; 30.

11.   Hernandez-Silveira M, Ahmed K, Ang S-S, 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. Last updated 19 May 2015, accessed 7 September 2020

12.   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; 65: 1324–1332.

13.   Breteler MJM, KleinJan EJ, Dohmen DAJ, Leenen LPH, van Hillegersberg R, Ruurda JP, van Loon K, Blokhuis TJ, Kalkman CJ. Vital Signs Monitoring with Wearable Sensors in High-risk Surgical Patients: A Clinical Validation Study. *Anesthesiology* 2020; 132: 424–439.

14.   Buist M, Bernard S, Nguyen T V, Moore G, Anderson J. Association between clinically abnormal observations and subsequent in-hospital mortality: a prospective study. *Resuscitation* 2004; 62: 137–141.

15.   McGaughey J, Alderdice F, Fowler R, Kapila A, Mayhew A, Moutray M. Outreach and Early Warning Systems (EWS) for the prevention of intensive care admission and death of critically ill adult patients on general hospital wards. *Cochrane.Database.Syst.Rev.* 2007. last amended 22 April 2008, accessed 7 September 2020

16.   Odell M, Victor C, Oliver D. Nurses' role in detecting deterioration in ward patients: Systematic literature review. *J. Adv. Nurs.* 2009; 65: 1992–2006.

17.   Tirkkonen J, Ylä-Mattila J, Olkkola KT, Huhtala H, Tenhunen J, Hoppu S. Factors associated with delayed activation of medical emergency team and excess mortality: An Utstein-style analysis. *Resuscitation* 2013; 84: 173–178.

18.   Drummond GB, Bates A, Mann J, Arvind DK. Validation of a new non-invasive automatic monitor of respiratory rate for postoperative subjects. *Br. J. Anaesth.* 2011; 107: 462–469.

19.   Thurnheer R, Xie X, Bloch KE. Accuracy of nasal cannula pressure recordings for assessment of ventilation during sleep. *Am. J. Respir. Crit. Care Med.* 2001; 164: 1914–1919.

20.   Bishop CM. Pattern Recognition and Machine Learning. New York: Springer; 2006.

21.   Pedregosa F, Varoquaux G, Gramfort A, Michel V, Bertrand T, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay É. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* 2011; 12: 2825–2830.

22.   sklearn.neural_network.MLPClassifier.Available from: http://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPClassifier.html#sklearn.neural_network.MLPClassifier. No update: version scikit-learn 0.23.2, accessed 7 September 2020

23.   Bland JM, Altman DG. Agreement between methods of measurement with multiple observations per individual. *J. Biopharm. Stat.* 2007; 17: 571–582.

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: 1444–1452.

25.   Breteler MJM, Huizinga E, Loon K Van, 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: 1–9.

26.   Smith I, MacKay J, Fahrid N, Krucheck D. Respiratory rate measurement: a comparison of methods. *Br. J. Healthc. Assist.* 2011; 5: 18–23.

27.   Chen L, McKenna T, Reisner A, Reifman J. Algorithms to qualify respiratory data collected during the transport of trauma patients. *Physiol. Meas.* 2006; 27: 797–816.

28.   Chen L, Reisner AT, Gribok A, McKenna TM, Reifman J. Can we improve the clinical utility of respiratory rate as a monitored vital sign? *Shock* 2009; 31: 574–580.

29.   Smith GB, Prytherch DR, Schmidt P, Featherstone PI, Knight D, Clements G, Mohammed MA. Hospital-wide physiological surveillance–A new approach to the early identification and management of the sick patient. *Resuscitation* 2006; 71: 19–28.

30.   Schmidt PE, Meredith P, Prytherch DR, Watson D, Watson V, Killen RM, Greengross P, Mohammed MA, Smith GB. Impact of introducing an electronic physiological surveillance system on hospital mortality. *BMJ Qual. Saf.* 2015; 24: 10–20.

31.   Cardona-Morrell M, Prgomet M, Turner RM, Nicholson M, Hillman K. Effectiveness of continuous or intermittent vital signs monitoring in preventing adverse events on general wards: a systematic review and meta-analysis. *Int. J. Clin. Pract.* 2016; 70: 806–824.

32. 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: 628–633.

33. Kellett J, Wang F, Woodworth S, Huang W. Changes and their prognostic implications in the abbreviated VitalPAC Early Warning Score (ViEWS) after admission to hospital of 18,827 surgical patients. *Resuscitation*. 2013; 84: 471–476.

Figure 1 An example of a section of recording from a single subject Top trace: nasal cannula pressure. A downward deflection (pressure decrease) indicates inspiration Bottom three traces: accelerometer forces. These are smoothed and combined into the Respeck signal. The arrow indicates the start of a patient movement, seen in one of the axis measures, that causes a large deflection of the Respeck signal.

Figure 2 Effects of applying selection process to the test data. The six test patients are identified by letters, A – F. The test patients are ordered according to proportion of retained measurement windows. Left hand axis: Differences between the rate measurement pairs made by nasal cannula and Respeck, in each subject. Open columns, before window selection, grey columns after window selection. Values are median, error bars are quartiles. Right hand axis, the proportion of retained measurement windows for each patient is shown (black filled columns). Patient G has a low proportion of windows retained.

Figure 3 Bland and Altman plots of (nasal – Respeck) differences before and after selection of suitable signals, plotted in relation to the mean respiratory rate ([nasal + Respeck]/2). The data are from all six validation subjects: each dot indicates one pair of compared values, nasal and Respeck measures, for a single observation window, in one validation subject's data. Plots show bias (solid line), and limits of agreement (outer dashed lines) surrounded by their 95% confidence ranges (shaded). To allow comparable scales to be used in both panels, 31 of the 1670 data points from the "before selection" panel are not plotted.

Figure 4 Plots of rate estimates from two representative patients shown for a 17 minute time period (Patient records are from A and G from the test sample, see figure 2). These are patients with extreme proportions of measurement windows that were chosen as accurate, from the whole window dataset for that patient. Even with a 10% acceptance rate, in patient G, these samples are sufficient to give a reasonable measure of the overall respiratory rate, over the time period shown.
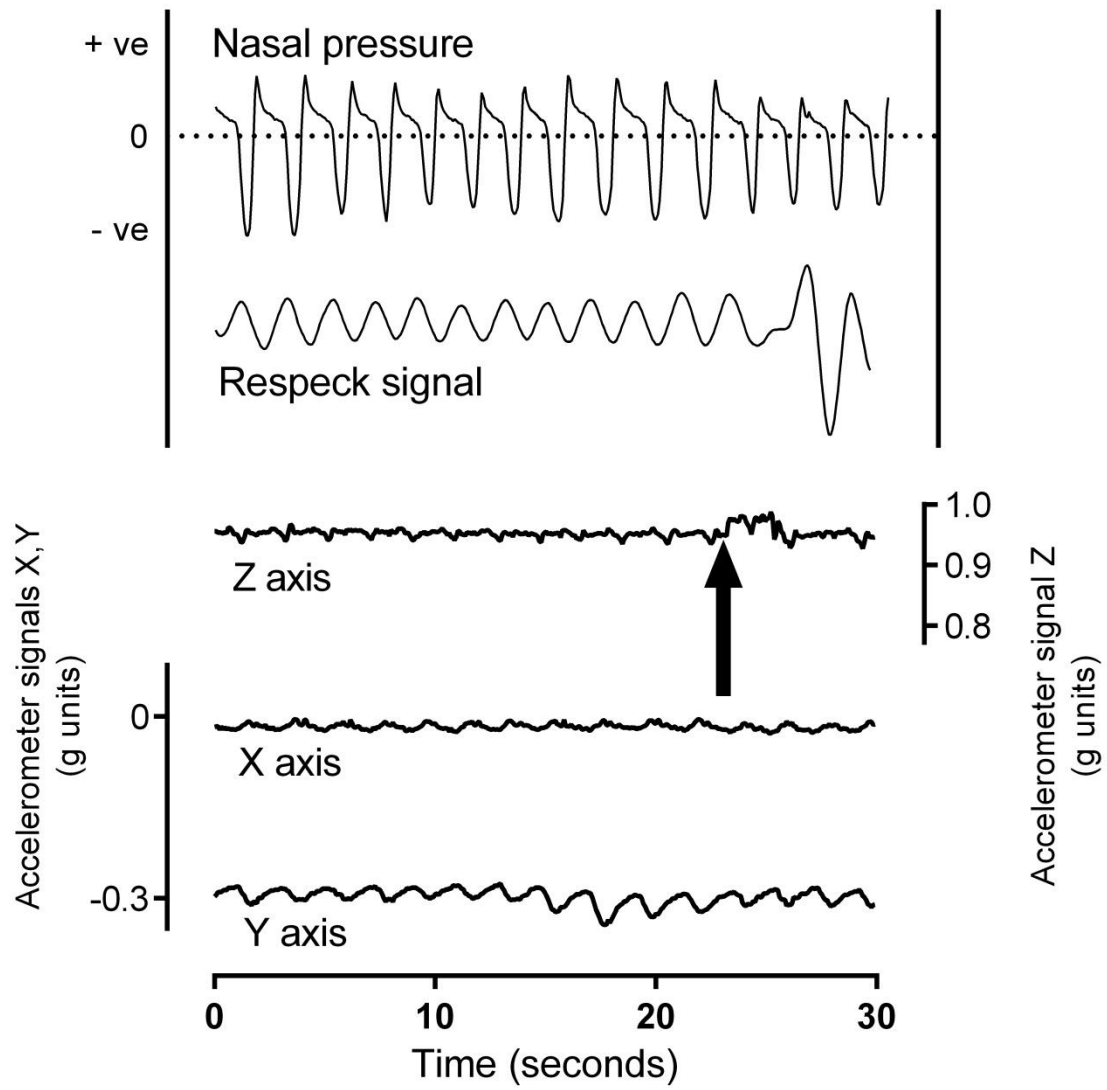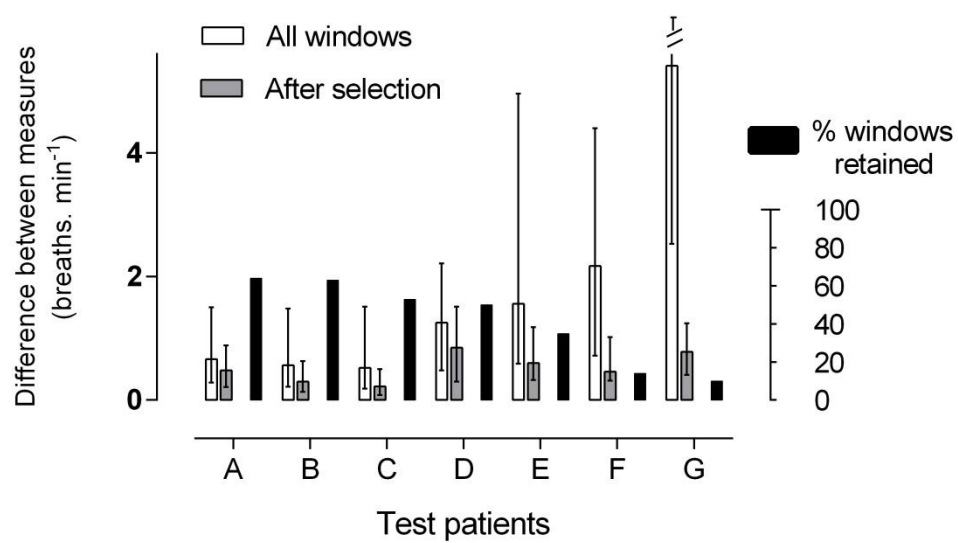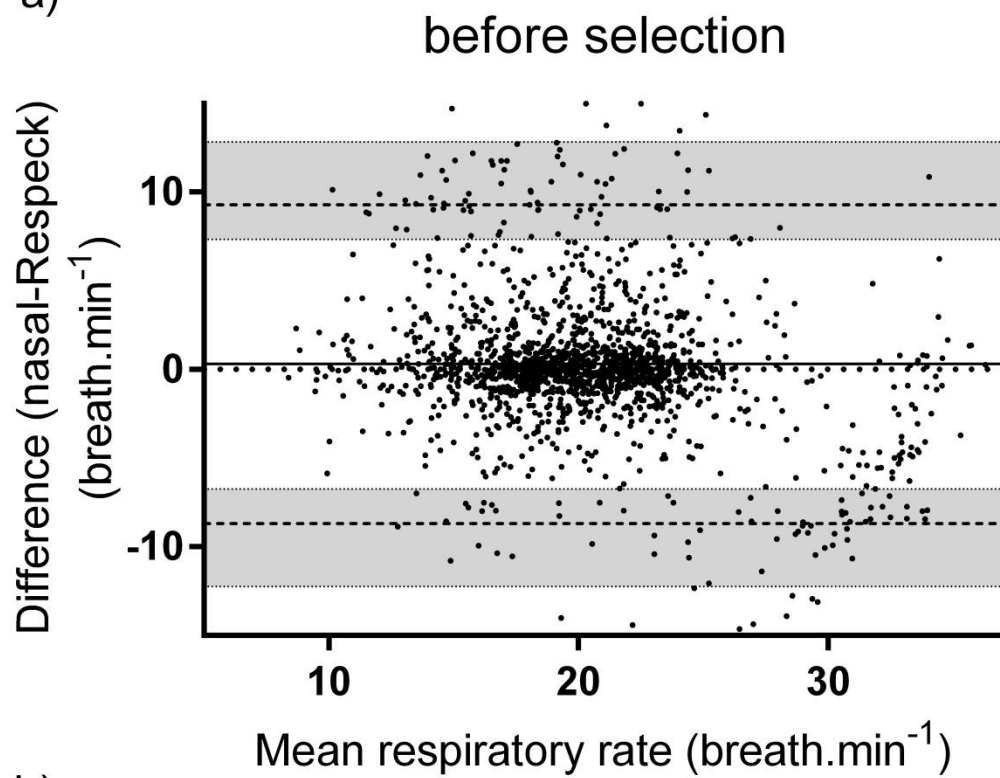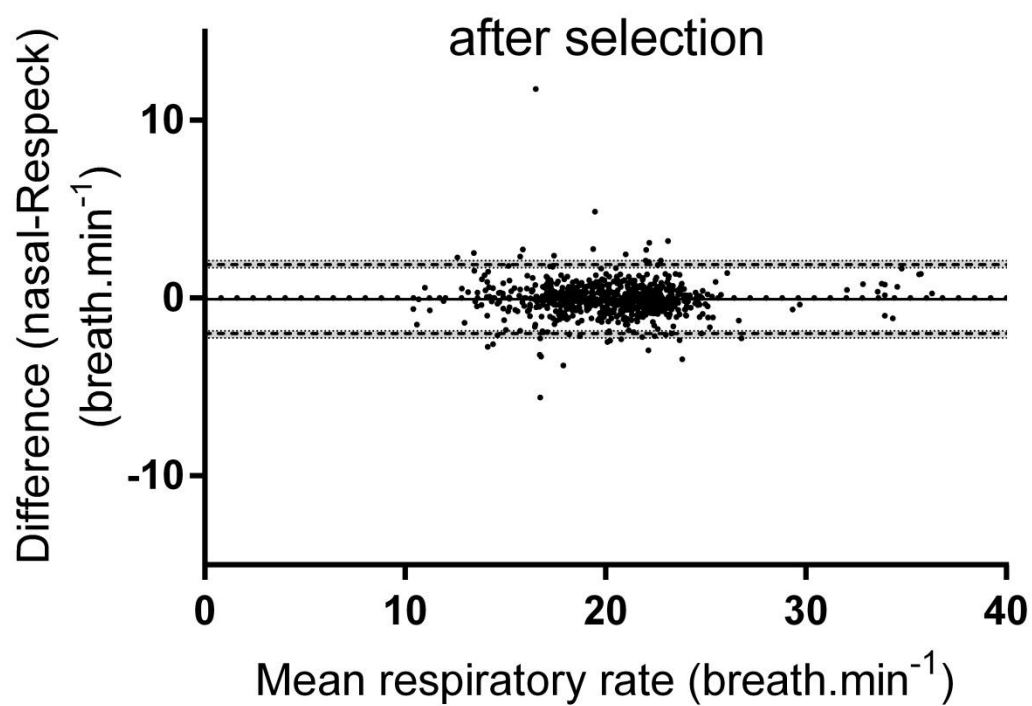
Fig 1

Fig 2

Fig 3
a)

before selection

b)

after selection

Fig 4