Physiological
Measurement

PAPER • **OPEN ACCESS**

Bed sensor ballistocardiogram for non-invasive detection of atrial fibrillation: a comprehensive clinical study

To cite this article: Jonas Sandelin *et al* 2025 *Physiol. Meas.* **46** 035003

IOP Publishing Physiol. Meas. 46 (2025) 035003 https://doi.org/10.1088/1361-6579/adbb52

Physiological Measurement

IPEM Institute of Physics and Engineering in Medicine

OPEN ACCESS

RECEIVED
19 September 2024

REVISED
4 February 2025

ACCEPTED FOR PUBLICATION
27 February 2025

PUBLISHED
11 March 2025

PAPER

Bed sensor ballistocardiogram for non-invasive detection of atrial fibrillation: a comprehensive clinical study

Jonas Sandelin[1,*], O Lahdenoja[1], I Elnaggar[1], R Rekola[1], A Anzanpour[1], S Seifizarei[1], M Kaisti[1], T Koivisto[1], J Lehto[2], J Nuotio[2], J Jaakkola[2], A Relander[2], T Vasankari[2], J Airaksinen[2] and T Kiviniemi[2]

[1] Department of Computing, Digital Health Technology Group, University of Turku, Vesilinnantie 3, 20500 Turku, Finland
[2] Heart Center, Turku University Hospital and University of Turku, Vesilinnantie 3, 20500 Turku, Finland
[*] Author to whom any correspondence should be addressed.

E-mail: jojusan@utu.fi

Keywords: AFib, atrial fibrillation, ballistocardiogram, BCG, bed-sensor, bed, ECG

Abstract

Objective. Atrial fibrillation (AFib) is a common cardiac arrhythmia associated with high morbidity and mortality, making early detection and continuous monitoring essential to prevent complications like stroke. This study explores the potential of using a ballistocardiogram (BCG) based bed sensor for the detection of AFib. *Approach.* We conducted a comprehensive clinical study with night hospital recordings from 116 patients, divided into 72 training and 44 test subjects. The study employs established methods such as autocorrelation to identify AFib from BCG signals. Spot and continuous Holter ECG were used as reference methods for AFib detection against which BCG rhythm classifications were compared. *Results.* Our findings demonstrate the potential of BCG-based AFib detection, achieving 94% accuracy on the training set using a rule-based method. Furthermore, the machine learning model trained with the training set achieved an AUROC score of 97% on the test set. *Significance.* This innovative approach shows promise for accurate, non-invasive, and continuous monitoring of AFib, contributing to improved patient care and outcomes, particularly in the context of home-based or hospital settings.

1. Introduction

Atrial fibrillation (AFib) stands as a global health concern due to its increasing incidence and substantial impact on public health (Camm *et al* 2012, Kornej *et al* 2020). Traditional approaches for AFib detection rely on intermittent electrocardiogram (ECG) recordings obtained during medical visits or hospital stays, which may not effectively capture transient and often asymptomatic episodes (Camm *et al* 2012, Harris *et al* 2012). This limitation necessitates the exploration of novel health technologies (such as miniature Holter ECG) that offer continuous and non-intrusive monitoring solutions for enhanced detection and management of AFib. However, even with regular Holter recordings, screening for AFib can be a difficult task (Diederichsen *et al* 2020). The rationale for screening is based on data indicating that patients with screening-detected AFib have a risk of stroke and death that is similar to that of patients with asymptomatic AFib (Gibbs *et al* 2021).

This paper investigates the possibility of using a bed-based ballistocardiogram (BCG) sensor as a means to achieve continuous AFib detection. BCG is a non-invasive method to measure the micro movements of the ballistic forces generated by the heart (Inan *et al* 2014). In this study, we used a bed-based BCG. The primary objectives of this study were to develop a robust algorithm for reliable detection of AFib in BCG and to create accurate labeling in continuous monitoring of the data collected from the BCG sensor. In addition, since BCG signals are highly sensitive to even minor movements (Alivar *et al* 2017), it is crucial to detect and account for motion artifacts during the classification process (Enayati *et al* 2020).

In this paper, we use ECG as a reference signal in order to confirm the findings with the bed sensor. ECG is considered to be the gold standard for AFib detection (Babaeizadeh *et al* 2009). ECG is used to label the signals gathered in the study. There are generally two main approaches to detect AFib with ECG; by assessing

the periodicity of the signal or by finding abnormal or missing P-waves (Couceiro *et al* 2008). This study will only focus on finding the periodicity, since P-waves or their alternative are not currently reliably found in BCG.

One of the biggest benefits of a bed-based AFib detection is that it is easily accessible. The accessibility and ease of integration of a bed-based AFib detection system offer significant advantages over traditional monitoring methods. With the growing popularity of wearable and at-home health monitoring devices (Greiwe and Nyenhuis 2020), the implementation of a bed sensor provides an easily accessible continuous monitoring solution for AFib, reducing the need for frequent hospital visits and enabling timely intervention. Unlike traditional wearable devices, which patients need to wear and adjust constantly, the bed sensor offers a non-intrusive, non-contact and convenient solution for continuous monitoring of AFib during sleep. It offers a unique and user-friendly alternative that seamlessly integrates into patients' daily lives.

One of the challenges associated with the analysis of the BCG signal is its inherent variability (Nagura *et al* 2018, Sadek *et al* 2019). The morphology of the BCG signal varies not only between different subjects but also within individuals based on the posture and position (Sadek *et al* 2019), making it prone to environmental changes and artifacts. The positioning of the device itself can cause variations (Alivar *et al* 2017). These variabilities pose a significant concern for developing robust and accurate AFib detection algorithms using BCG signals.

Previous studies have explored the use of bed sensors to detect AFib (Brüser *et al* 2011, Zink *et al* 2015, Yu *et al* 2019, Su *et al* 2022, Feng *et al* 2023), but to our knowledge, none have yet been implemented on a similar end-to-end and comprehensive scale as ours. Some of these studies are also recorded during the day with short segments, while our study intends to use full-night recordings for a more realistic approach.

The most similar study to ours is a study by Su *et al* (2022), which has many similarities to ours. They have collected similar type of data to us with an piezoelectric sensor placed under the sheet. The placement is different from our study, which might make a difference to the quality of the signal. Another big difference is also the segment size. Our study uses multiple different segment size and was originally planned to detect paroxysmal AFib over a full night recording, while their study uses much shorter 10 s segments. Their approach was to use a residual convolutional neural network to detect AFib from 28 214 10 s segments of the BCG.

Similar type of solutions have also been used in previous studies using seismocardiogram (SCG) signals (Hurnanen *et al* 2016). SCG as a signal is similar to BCG in multiple ways, which makes it a relevant reference point to our study. The autocorrelation approach for detection used in this study for BCG highly resembles the ones in our previous studies for SCG. This study and few others (Pänkäälä *et al* 2016, Mora *et al* 2020) have showed great promise on AFib detection with SCG. However, the drawback with SCG is that the device usually needs to be in contact to the patient, which means that co-operation from the user is needed. The primary advantage of BCG lies in its non-contact and non-intrusive nature. Unlike most SCG solutions, it does not require daily adjustments or re-installation. These characteristics make BCG particularly well-suited for long-term monitoring applications.

Figure 1 represents the basic structure of the study. BCG and ECG signals were gathered from a person lying in bed. These signals were then processed separately, getting the label from the reference ECG signal and then comparing it to the prediction of heart rhythm made based on the BCG signal. The exact same data gathering setup was used in both training and test phases of the study. With the training data, we created the classifiers based on data analysis and adjusted the preprocessing parts to work as good as possible. For the test set these previously created steps were left untouched and the classifier models created with the training data were used to get the predictions based on the features, creating the setup we have in figure 1.

In this paper, we present the methodology, data analysis, and results of the clinical study, highlighting the potential of the bed-based BCG sensor for continuous AFib monitoring and its implications for improving patient care and outcomes.

2. Materials and methods

2.1. Data

Data used in the study was collected from the Turku University Hospital (TYKS), Finland. The dataset includes 116 patients, described in table 1. Inclusion criteria for the study were the following: patient has been informed about the study and agrees to it, patient is hospitalized in TYKS, patient is over 18 years old and do not suffer of acute heart failure (HF) or moderate heart valve disease.

The data collection was done by using two sensors, Emfit QS Bed Sensor and Bittium Faros Holter. From each patient, simultaneous BCG and a reference ECG signals were recorded with these sensors while the patient was in supine position in the bed. The length of these recordings varied from 2 h to 24 h. The average length of a recording was 14 h. The quality of the signals also vary a lot, since the patient could have been

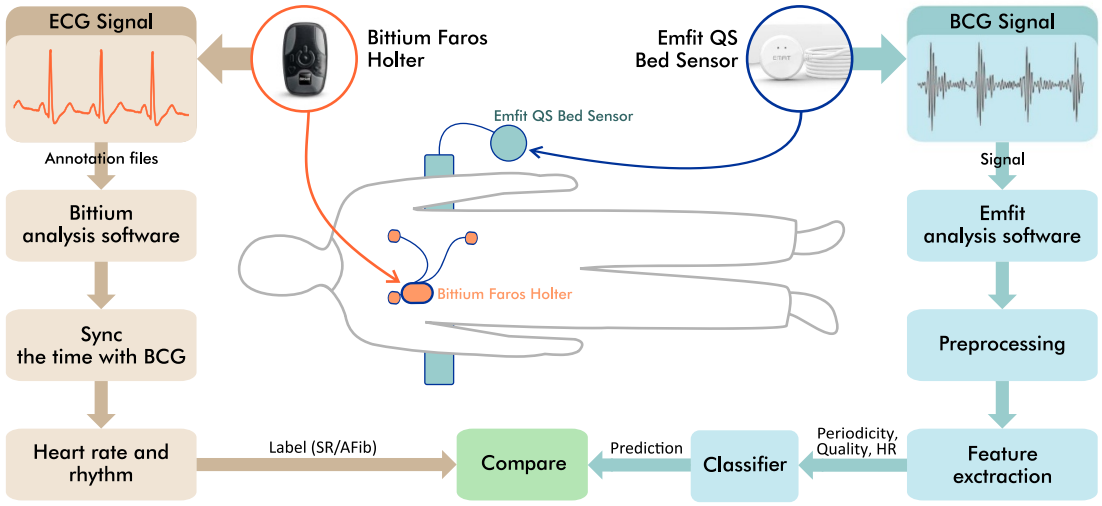

Figure 1. The pipeline for the study. The setup comprises of two different sensors, one for BCG and one for ECG reference.

Table 1. Data sets.

	Train set	Test set
Quantity	72	44
Quantity 1 h	951	476
Quantity 30 min	1880	945
Quantity 10 min	4970	2532
Females	32 (44%)	23 (52%)
Males	40 (56%)	21 (48%)
AFib	32 (44%)	22 (50%)
SR	40 (56%)	22 (50%)
Age	68 ± 12	70 ± 13
Body mass index (BMI)	30 ± 6	30 ± 6
Left ventricular ejection fracture (LVEF)	47 ± 15	49 ± 13
Heart failure	22 (31%)	12 (27%)
Coronary artery disease	32 (45%)	22 (50%)
Previous myocardial infarction	20 (27%)	6 (14%)
Previous stroke	4 (5.6%)	1 (2.3%)
Chronic AFib	12 (17%)	13 (30%)

awake during the recordings and thus move a lot. We tried to increase the quality by filtering the signals and detecting clear artifacts. The data was collected by respecting the declaration of Helsinki. Informed consent was obtained from all participants, and rigorous measures were taken to ensure data privacy and confidentiality.

The dataset was partitioned into two subsets. The training set includes 72 patients and was used in the development of the algorithm. The test set includes 44 patients and was used blindly for testing after the algorithm was developed. The decision for this specific split was based on the time when the algorithm was finalized and 'locked' to ensure consistency during testing.

As seen in table 1, sinus rhythm (SR) and AFib were split quite evenly in both data sets. The data sets were also very similar compared to each other making them viable for this type of testing. Chronic AFib however was split a bit unevenly between the test and training set.

The datasets had combined 34 patients with HF. Five of these HFs were HFs with mid-range ejection fracture, 17 of them were with reduced ejection fraction (HFrEF) and the last 12 were with preserved ejection fraction (HFpEF).

The quantity of instances increases when segmenting the signals into smaller segments. However, some instances had to be cut from the datasets since some of the data entries were found to be corrupt, missing, or otherwise unusable, necessitating their removal from the analysis. For example these removed entries include instances where the signal was found to be full of only artifact or absence. This could happen if the patient for example left the bed or moved a lot during the specific segment. For the 10 min variant, we were able to

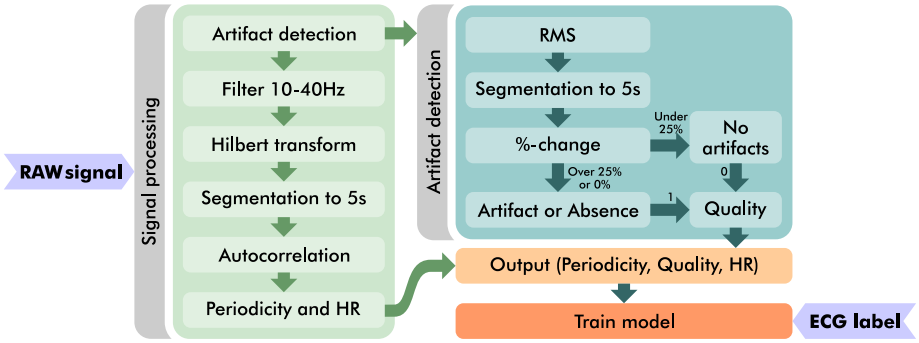

Figure 2. The flow chart of the algorithm used for BCG.

use 83% of the data available, but for 30 min and 1 h coverage is 95%. For full overnight recordings the coverage is 100%.

2.2. Biosignal hardware

The Emfit Quantified Sleep (QS) system was used as the bed-based BCG sensor in this study (Emfit homepage 2025). The Emfit QS uses a pressure-sensitive mat integrated into the mattress, which captures subtle mechanical vibrations induced by cardiac activities during sleep. In our study the sensor records data at 200 Hz sampling rate to ensure accurate detection and analysis of the BCG signal. The raw BCG signals can be extracted from the UI called Emfit Navigator which we have used in this study.

Bittium Faros was used as a reference ECG for the study (Bittium faros homepage 2025). The Faros device records the patient's ECG data, which was used as a reference signal. Faros Holter patch measures single-lead ECG signal from the chest and is attached via FastFix electrode. The data collected with Faros can then be examined with Bittium Cardiac Navigator. From this UI, we can extract the annotations about the heart rates(HRs), RR-intervals and the label for the heart rhythm. With these annotations we can then create a label for each segment based on the majority label within each segment. This means that the labeling for ECG was done with the annotations provided by Bittium Faros. When examining a 10 min signal segment for example, we designate the label as AFib if the majority (over 50%) of annotations were AFib. Conversely, if the annotations predominantly indicate SR, the label assigned was SR.

2.3. Algorithm pipeline

A customized algorithm pipeline was developed to process and analyze the BCG data collected from the Emfit QS sensor. The pipeline involved several stages, including preprocessing, feature extraction and classification. This pipeline is described in figure 2.

The first step of the algorithm was to detect the artifacts and absences from the raw BCG signal. This way the computation time used for the analysis process can be minimized, since no calculations were done to the parts containing artifacts or absences. After this, the signal was filtered with a 3rd order Butterworth band-pass filter with 10–40 Hz cut-off frequencies. Last part of the preprocessing was to envelope the signal with Hilbert transform (Ulrich 2006). Earlier, Hilbert transform has been used in Tadi *et al* (2016) for SCG, but we observed it to be useful for BCG as well as it makes the autocorrelation peaks that we utilize more reliable.

After the preprocessing part, the preprocessed signal was fed through a auto-correlation algorithm. From this algorithm we can acquire a feature called 'periodicity', which describes how regular the signal was. This algorithm also gives us an 'HR' value, which is the estimated HR. From the artifact and absences values we can determine the 'quality' of the signal, which describes how much of the signal was corrupted and cannot be used in the auto-correlation algorithm.

The classification can then be done with these features. To classify the particular signals to be either SR or AFib, we used machine learning methods and also a rule based method. A rule based method simply looks at the periodicity value and if the periodicity was over 10%, then it can be classified to be SR. This value was based on the data analysis done to the full length signals in the training set. The machine learning classification was done for the training set by using leave-one-subject-out (LOGO) cross-validation (CV) method and for the test set by creating the model with the training set and then classifying the signals in the test set.

In this study we chose to analyze different segment lengths to see how it affects the accuracy of classification. The lengths we chose were 10 min, 30 min, 1 h and full length recordings.

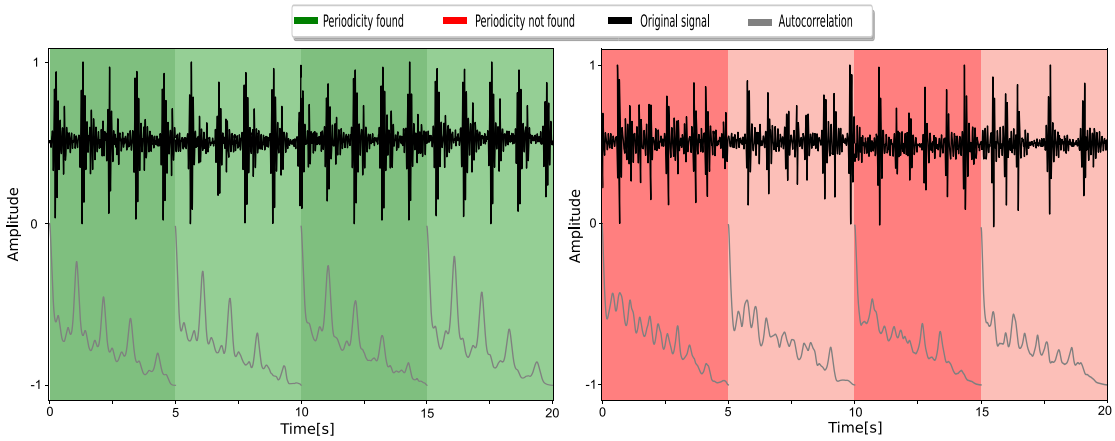

Figure 3. Autocorrelation of SR and AFib.

2.4. Artifact and absences detection and bed status monitoring

To ensure the accuracy and reliability of the BCG signal, an artifact detection technique was implemented. The artifact detection system developed was based on a paper published in 2010 (Wiard *et al* 2011). On top of this we added a way to detect absences, since the patients could leave their beds. First, the signal was segmented into 5 s long windows. For each 5 s window, root-mean-square (RMS) value is calculated. After this we check the percentage change within the segments. If the RMS value of a segment is 75% higher than the RMS value of the previous segment, we determine the segment as an artifact. Absences on the other hand were detected by seeing if the percentage change is 0% compared to the previous segment. If this was the case, we determine the segment as absent. For each 5 s segment, a binary label vector was created for each bed status. If a segment contained an artifact, the artifact label was set to 1; otherwise, it was labeled as 0. The validity of this method was evaluated using ten manually annotated signals obtained from sources external to the main datasets. This data was manually annotated by one of the authors and was collected outside of the study.

2.5. Autocorrelation, rhythm detection and periodicity

To estimate the periodicity of a specific signal, we compute the autocorrelation signal for each 5 s segment. For each autocorrelation segment, we then calculate the side peaks and the time between them. After this we calculate the difference of these times between the side peaks. In each segment, we aim to find 4 side peaks. By looking at the smallest difference, we can determine whether the specific segment was periodic or not. The algorithm takes in filtered signal, artifacts and absences, the segment length and the sampling frequency. After this the signal was sliced into 5 s segments and if that specific segment has either artifact or absence label, the process was stopped, and we move on to the next segment. The algorithm returns an output which includes vector array which has periodicity value for each segment of the signal. If it is periodic, the segment has the value 1, if not then 0. The overall mean value of the array consisting of (0, 1 or NaN) values was used as the final periodicity value. If an artifact or absence was detected this value was NaN. In figure 3 we can see auto-correlation of a SR patient and of an AFib patient and their original signal.

HR can be estimated from the autocorrelation calculating the mean time difference between the side peaks of the autocorrelation. As we see in figure 3, peaks can sometimes be unobtainable especially when signal is non-periodic, meaning that the estimated HR will not be accurate. HR will still be estimated even if it would not be periodic. Similar method to detect HR from BCG was used in a previous study (Elnaggar *et al* 2022).

3. Results

Figure 4 shows a comparison between the HR derived from the BCG and the reference ECG. The figure shows all the HR's that have the frequency above 10%. The figure is mainly for SR patients, and we can see that the HR estimate is accurate for most of them. There are also a few AFib instances (in red) while for some of them the HR estimate is correct and for majority not. This is mainly due to the fact that the HR is derived from autocorrelation. If there are irregular beats, it might not be able to find the correct HR. For AFib detection this is not a huge problem as the impact of HR to the overall classification is limited (see table 2). However, for HR estimation, it is obviously a problem. A better solution needs to be found in the future.

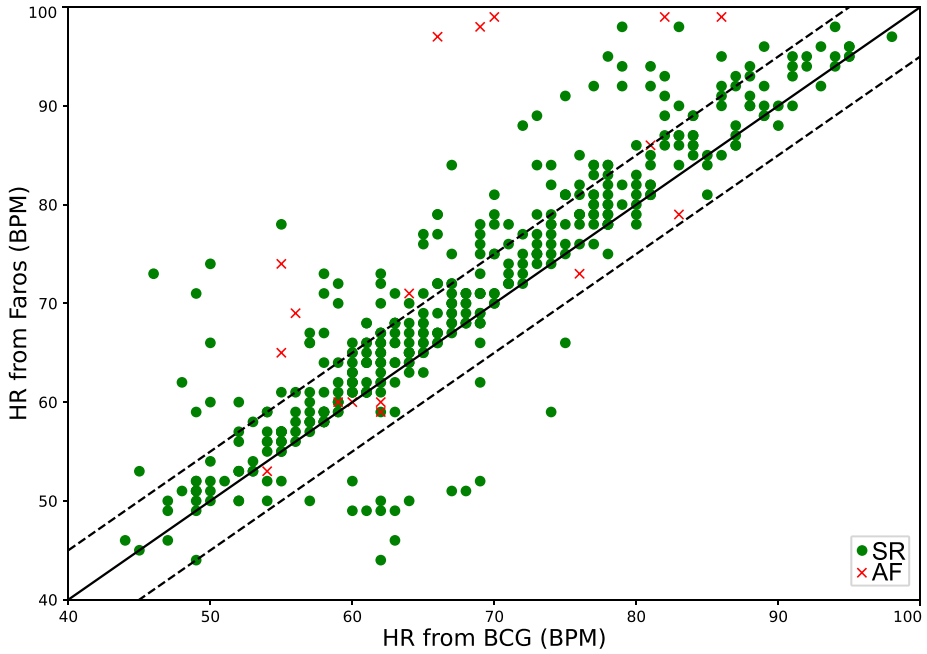

Figure 4. Heart rate from BCG compared to ECG reference. Dash lines represent ±5 bpm. Each dot represents the 1 h mean of heart rate. Each patient can have multiple dots. R^2 value of the results is 0.75, when only the instances with periodicity above the 10% cut-off are plotted and calculated.

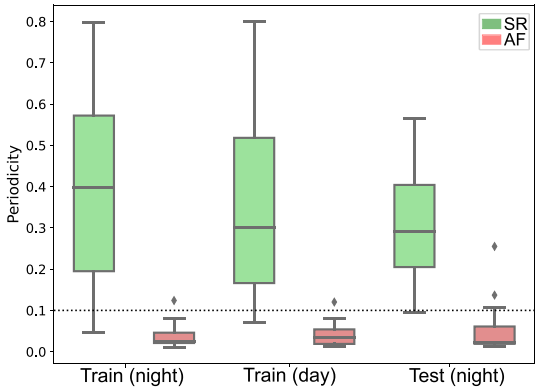

Figure 5. Box plot of full recordings periodicities and the rule-based cut-off line of 10%. The data includes train and test set recorded during the day and separate data from train set patients during the day.

Figure 5 illustrates the notable difference between SR and AFib periodicities. We initially used these values to determine the cut-off for our rule-based method to be 10%. With this cut-off, only 4 people out of 72 in the training set were misinterpreted. With the same cut-off, the test set achieves similar results, 4 people out of 44 were misclassified. We also included a day dataset within this comparison. Most of the patients had multiple recordings while in the hospital. The train and test sets were chosen to only include overnight recordings, but on the basis of the figure AFib can be separated from SR even during the day. The day data was selected from the patients within the train set. It includes 40 patients with 168 h of data, equaling approximately 4.2 h per patient. The missing patients did not have enough data, so we left them out of the analysis. The datasets seem to be behaving similarly, but there are some differences between the median values of SR periodicities.

Figure 6 represents the logistic regression classification (LR) boundary between SR and AFib recordings. Looking at the figure, it is clear that when the quality of the signal decreases, the periodicity also seems to be decreasing. Based on figure 6, it appears that the decision boundary created by the LR model is effective. The amount of false negatives is low, but then again the amount of false positives is significantly higher. The LR model used in the study is also presented in table 2. It confirms that periodicity is the most significant feature, but quality and HR do also matter. One thing to note is that periodicity and quality were normalized, but HR is not. Figure 6 and table 2 can be explained as such: When quality increases, the periodicity

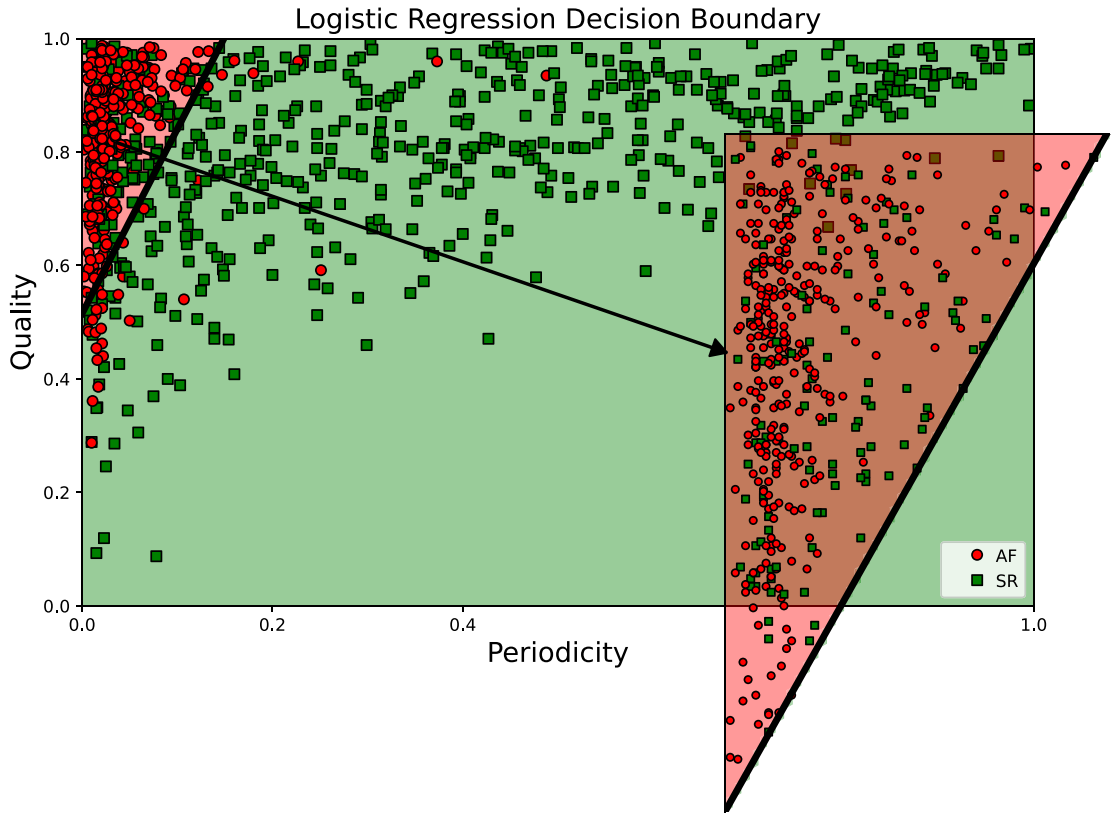

Figure 6. Periodicity compared to the quality and the logistic regression decision boundary. Each dot represents 1 h mean of periodicity and quality. Each patient can have multiple dots based on the length of the signal.

Table 2. Logistic regression model.

	Coef	Std err	z
`Periodicity`	20.3355	2.086	9.750
`Quality`	−6.6554	0.655	10.164
`HR`	0.0610	0.008	7.491

threshold that determines whether the signal is AFib or SR increases. This means that if the quality is good, then the model is able to loosen the decision boundary.

In table 3 we can see the classification results. Overall, the results were good, but upon further analysis, it becomes evident that more instances were misclassified when decreasing the time interval. Especially for the rule-based method, precision and specificity drops significantly. The machine learning models have higher accuracies when the time frame was segmented compared to the rule-based method, meanwhile rule-based method was better when using the full night recording. It seems that increasing the time interval diminishes the risk of false positives. All of the different classifiers seem to work well, but random forest and rule-based method seem to stand out. Interestingly, however, with the shortest time frame the LR model was better than the others. One thing to note is that the rule-based method was originally designed for full night analysis. This explains that it seems to become very unbalanced when lowering the segment time. This does not happen as significantly for the machine learning models.

We also conducted a test using a 10-fold CV. The obtained results are represented in table 4. For these results we tightened the way we label segments with Faros ECG labels. Earlier we determined the segment label by majority, now we used only the segments that had at least 95% SR or AFib within the segment. On top of this we tightened the artifact detection from 75% to 25%. These changes lowered the coverage, but also increased the performance significantly. The results also seem to be more balanced compared to table 3.

4. Discussion

In this study, an algorithm for AFib screening using a mattress sensor was introduced, highlighting the advantages and challenges of the method, as well as methods for handling issues such as noise and artifacts.

Table 3. Statistics with different times and classifiers. The training set was used to train the models while test set was used to get these results. Best statistics for each segment times are highlighted.

	10 min	30 min	1 h	Full
Rule-based				
Accuracy	76.97%	80.72%	80.25%	**90.51%**
Specificity	63.19%	70.78%	70.03%	**95.23%**
Sensitivity	**92.29%**	**96.21%**	**95.76%**	85.71%
Precision	63.19%	67.88%	67.79%	**94.74%**
F1 score	75.02%	79.59%	79.39%	**90.00%**
Logistic regression				
Accuracy	**79.61%**	79.56%	79.41%	80.95%
AUROC	**84.60%**	86.72%	85.50%	85.03%
Specificity	**81.82%**	**82.09%**	**81.88%**	80.95%
Sensitivity	75.05%	78.85%	78.77%	80.95%
Precision	**77.05%**	**78.51%**	78.48%	80.95%
F1 score	**77.12%**	**78.66%**	**78.61%**	80.95%
Random forest				
Accuracy	75.62%	78.18%	**80.25%**	85.71%
AUROC	82.92%	**87.79%**	**87.69%**	97.16%
Specificity	74.08%	76.55%	78.84%	80.95%
Sensitivity	80.31%	84.00%	85.71%	**95.24%**
Precision	73.87%	77.20%	**79.56%**	87.06%
F1 score	73.99%	76.83%	79.14%	85.58%
XGBoost				
Accuracy	74.91%	**78.28%**	79.20%	88.10%
AUROC	83.37%	86.82%	84.88%	**97.28%**
Specificity	73.63%	76.93%	78.37%	80.95%
Sensitivity	78.73%	83.13%	85.02%	**95.24%**
Precision	73.28%	77.24%	78.45%	88.89%
F1 score	73.44%	77.07%	78.01%	88.03%

Table 4. Results with 10-fold CV and with less coverage. Inside the brackets are the 90% confidence intervals (CI). The best performance for each metric is highlighted. RF = Random forest, LogReg = Logistic regression, XGBoost = Extreme gradient boosting.

Model	10 min	30 min	1 h
RF			
Accuracy	84.10% [83.25, 84.96]	88.95% [87.99, 89.91]	89.85% [87.40, 92.30]
AUROC	92.19% [91.55, 92.83]	95.68% [95.00, 96.36]	96.75% [95.68, 97.81]
Specificity	79.62% [78.10, 81.15]	**87.81%** [86.18, 89.44]	86.99% [81.16, 92.82]
Sensitivity	86.67% [85.27, 88.07]	89.77% [88.13, 91.42]	90.64% [88.48, 92.81]
Precision	88.17% [87.36, 88.98]	92.32% [90.93, 93.71]	**93.80%** [91.25, 96.36]
LogReg			
Accuracy	82.91% [81.80, 84.02]	85.13% [83.61, 86.65]	84.79% [81.66, 87.92]
AUROC	90.94% [90.31, 91.57]	92.66% [91.79, 93.54]	92.73% [90.52, 94.94]
Specificity	83.66% [82.47, 84.86]	83.74% [81.25, 86.24]	75.79% [69.22, 82.35]
Sensitivity	82.48% [80.71, 84.24]	85.83% [83.61, 88.06]	89.63% [87.34, 91.93]
Precision	89.82% [88.93, 90.61]	89.75% [88.09, 91.41]	87.12% [84.00, 90.25]
XGBoost			
Accuracy	84.15% [83.08, 85.22]	88.38% [86.74, 90.03]	**91.31%** [89.33, 93.29]
AUROC	92.36% [91.50, 93.23]	94.94% [94.04, 95.85]	**96.82%** [95.93, 97.70]
Specificity	81.07% [79.74, 82.39]	87.38% [85.05, 89.71]	85.85% [80.70, 90.99]
Sensitivity	85.92% [84.32, 87.52]	89.01% [87.18, 90.84]	**93.75%** [92.11, 95.38]
Precision	88.79% [87.94, 89.64]	91.97% [90.16, 93.77]	92.90% [91.06, 94.73]

The study results offer valuable insights into the strengths and limitations of this technology, with a specific emphasis on the bed sensor and its potential applications.

The motion artifacts can have a significant effect on the estimated periodicity value. This was why a quality assessment and detecting movement was necessary. By implementing a robust motion artifact detection system, the study was able to reduce the impact of motion-related disturbances on the BCG signal, enhancing the algorithm's performance and overall effectiveness in AFib detection. The most notable motion artifacts usually happen within the first few hours of recordings when the patients were still awake. We noticed a clear decrease between the periodicities of the few first hours of the recording compared to the middle and end parts of the recording. The same can be stated for the quality. In order to get reliable approximation of the periodicity, the signal needs to be clean for some time rather than just few seconds. However, when comparing the overnight and day recordings, there does not seem to be a huge difference in periodicity as figure 5 shows. The average length of the recordings is shorter which may have an effect. The authors also observed that the quality is lower with the day recordings but not significantly. The statistics however for the day results remained similar with the rule-based method, with accuracy being 95%, specificity 95.5%, precision, sensitivity and F1 score all being 94.4%. For both classes there was one misclassification, making it 2 subject out of the 40 incorrectly classified.

One of the problems of the study was the amount of false positives. This might be related to the nature of BCG. It is a very vulnerable signal which is affected by multiple variables. Different postures, minor movement, snoring and multiple other factors can affect it. These all can affect the result of autocorrelation, but might not be enough to alarm the artifact detection system, leading to false classifications. When looking at the BCG, sometimes it was impossible to find the heart beating or even remotely close to it. Previous studies have noted similar findings (Perez-Macias *et al* 2022). Even though it might be a clear SR in the reference ECG, the BCG signal can just look like white-noise due to low signal to noise ratio. This is something to be addressed in future studies. Proper pre-processing steps such as filtering and enveloping was important. During the algorithm developing process we found that enveloping the signal with Hilbert transform improved the classification rate of out system significantly (approximately 5%). Different band-widths for the Butterworth filter were also tested, but 10–40 Hz was found to be the best based on LOGO-CV results with the training set. However, different signals did better with different band-widths. Some sort of adjustable frequency band selector would most likely be an optimal solution and could decrease the amount of false positives. One solution might also be to detect amplitude changes within the signal. When a person changes position on the bed, the strength of the signal changes, changing the amplitude. By removing the lower amplitude segments, it might be possible to filter out the segments without detectable HR. However, this would most likely remove multiple hours within the recordings and as stated before, the variability between person to person in signals might be too high in order to make this type of solution.

One limitation of this study is the use of a single-channel ECG. While ECG is considered the gold standard for heart rhythm monitoring, a single-channel ECG is less comprehensive than a 12-lead ECG and is also more prone to motion artifacts. The authors observed that Faros ECG occasionally failed to correctly identify R-peaks when the signal was affected by motion artifacts.

The length of the segment also influences the rhythm annotations. For instance, a 10 min segment might contain 4 min of AFib and 6 min of SR. Such a segment would be labeled as SR in majority voting, but would affect the periodicity value of the segment and might lead to a misclassification. However, since the study was designed to detect also paroxysmal AFib, and the patients' labels were provided by physicians, this approach was deemed acceptable. This reasoning also justifies the use of full-night recordings, as these types of inconsistencies would balance out over longer periods. However, the 10-fold CV results with lower coverage based on a tighter Faros label, demonstrate that the majority vote is causing some misclassifications.

Comparing our results with previous studies is not straightforward due to the major differences between the studies. However, the accuracy, precision and sensitivity are compared to other studies in table 5. Other studies seem to have similar amounts of false positives and false negatives making their systems more balanced compared to independent test set results. When comparing our CV results, they seem to be within line with other studies.

One difference between our study and the others is the length of the segment. Before the study a full night segment length was established for our study, while the others used very short segment sizes. The small amount of data also provided the possibility to manually label the segments. Both (Bruser *et al* 2012, Yu *et al* 2019) report that they manually identified the segments and then picked bad quality parts out, while we used a fully automated algorithm in an end-to-end manner. This may affect the results quite significantly because of the unstable nature of BCG.

The data collected thus far strongly supports the need for further algorithm development to establish a cost-effective AFib detection ecosystem. However, it is important to notice that this data was collected within a hospital environment. This could affect the result in different ways. On the other hand, it can enhance the

Table 5. Study result comparison.

Study	(Bruser *et al* 2012)	(Yu *et al* 2019)	(Su *et al* 2022)	Our study	Our study
Patients	10	12	45	116	116
Data (hours)	7.1 h	6.7 h	78.4 h	1250.8 h (+168 h)	552 h
Avg length	0.71 h	0.56 h	1.72 h	10.78 h	4.76 h
Segment size	30 s	30 s	10 s	Full night	1 h
Coverage[a]	95%	7.4%	18.4%	100%	44%
Classes	3	2	3	2	2
Testing	10-fold CV	30% hold-out	10-fold CV	Test set	10-fold CV
Precision	90.7%	94.6%	96.8%	94.7%	92.9%
Sensitivity	93.8%	97.0%	93.7%	85.7%	93.8%
Accuracy	92.1%	94.4%	96.8%	90.5%	91.3%

[a] Coverage calculated from the original data reported and the data actually used.

results when the setup was optimal and the beds are similar for every patient. However, it can also affect it negatively. The hospital environment, sleepless nights due to unfamiliar bed and the stress caused by the heart diseases can all lower the quality of the signal and thus harm the result. The first few hours being more polluted with noise indicates that more relaxed environment could improve the signal quality. In future studies it would be beneficial to have data collected at home, since it will give a more of a realistic setting for the study. The database itself was challenging. Most patients recorded have some arrhythmia's or other heart problems. Even though the patients were labeled as SR and AFib by doctors and the Faros ECG, these other heart issues can potentially affect the BCG signal.

For this type of study and for the future of these type of devices it was crucial that the data was collected at the hospital and had a comprehensive dataset. If the algorithm is too sensitive for detecting AFib, it can produce a lot of false positives which can then overload the healthcare system. That was why the datas*et al*so included other arrhythmia's since these were the one's that in theory should produce most of the false alarms. If the algorithm works with this dataset, it should work with healthy patients as well. The fact that we used the first 72 patients to create the algorithm and then have a big independent test set, assured us that we did not over fit the algorithm.

The study design offered several advantages and provided valuable insights. We for example noticed that the median periodicity was lower within female patients for both SR and AFib. The difference was not significant and would not change the rule-based results, but it was noticable. Age or BMI did not seem to have an affect on periodicity.

The results indicate promising possibilities for the development of a home monitoring device for AFib detection. This system could possibly used as a early warning system, as an indication, that could be potentially confirmed by using already proven methods such as ECG. It does have potential to be a detection system for AFib, but at the moment the amount of false positives is currently too high especially if assessing the shorter segments. This might possibly be improved by having some type of posture detection algorithm or other solutions. The orientation of the patients can change throughout the night thus changing the basic nature of the signal and potentially decreasing the periodicity or making it harder to reliably detect the rhythm. However, the full night signals do provide much better results. If suffering from AFib most of the night, this system will detect it quite reliably. This however can lead to the problem where short segments of AFib might be undetected.

5. Conclusion

In conclusion, the findings of this study highlight the potential of the bed-based BCG sensor for AFib detection, although further algorithm development is needed. The results provide insights into the strengths and limitations of the bed sensor, emphasizing its role as a component of AFib detection ecosystem. The proposed technology holds promise for enhancing AFib diagnosis, management, and overall patient care. Future studies should focus more on home monitoring and finding out how to limit the amount of false positives. Bed sensor based AFib monitoring does not require co-operation from the patient and our method was shown to perform well on hospitalized elderly population. Therefore, these elderly people could be the potential target group for BCG based screening as the likelihood of AFib is higher for them as well.

Data availability statement

The data cannot be made publicly available upon publication because they contain sensitive personal information. The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgment

This project was funded by Moore4Medical project which received funding from the ECSEL JU and Business Finland, under the Grant Agreement H2020-ECSEL-2019-IA-876190 and 7215/31/2019.

Ethical statement

The study was approved by Ethics Committee at of The Hospital District of Southwest Finland and the data was collected by respecting the declaration of Helsinki. All patients were over 18 and all provided written informed consent for participation.

ORCID iDs

Jonas Sandelin https://orcid.org/0009-0009-7666-0068
O Lahdenoja https://orcid.org/0000-0003-2081-600X
I Elnaggar https://orcid.org/0000-0002-2230-983X
R Rekola https://orcid.org/0000-0001-6578-5281
A Anzanpour https://orcid.org/0000-0002-7614-0232
S Seifizarei https://orcid.org/0009-0008-8379-0859
M Kaisti https://orcid.org/0000-0001-8118-2851
T Koivisto https://orcid.org/0000-0003-3792-8999
J Lehto https://orcid.org/0000-0002-6420-8607
J Nuotio https://orcid.org/0000-0003-3890-7081
J Jaakkola https://orcid.org/0000-0002-3398-4283
A Relander https://orcid.org/0000-0002-2177-5900
J Airaksinen https://orcid.org/0000-0002-0193-568X
T Kiviniemi https://orcid.org/0000-0002-0908-3741

References

Alivar A, Carlson C, Suliman A, Warren S, Prakash P, Thompson D and Natarajan B 2017 Motion detection in bed-based ballistocardiogram to quantify sleep quality *GLOBECOM 2017-2017 IEEE Global Communications Conf.* (IEEE) pp 1–6

Babaeizadeh S, Gregg R E, Helfenbein E D, Lindauer J M and Zhou S H 2009 Improvements in atrial fibrillation detection for real-time monitoring *J. Electrocardiol.* **42** 522–6

Bittium Inc. 2025 (https://www.bittium.com/medical/bittium-faros)

Bruser C, Diesel J, Zink M D, Winter S, Schauerte P and Leonhardt S 2012 Automatic detection of atrial fibrillation in cardiac vibration signals *IEEE J. Biomed. Health Inf.* **17** 162–71

Brüser C, Zink M D, Winter S, Schauerte P and Leonhardt S 2011 A feasibility study on the automatic detection of atrial fibrillation using an unobtrusive bed-mounted sensor *2011 Computing in Cardiology* (IEEE) pp 13–16

Camm A J, Corbucci G and Padeletti L 2012 Usefulness of continuous electrocardiographic monitoring for atrial fibrillation *Am. J. Cardiol.* **110** 270–6

Couceiro R, Carvalho P, Henriques J, Antunes M, Harris M and Habetha J 2008 Detection of atrial fibrillation using model-based ecg analysis *2008 19th Int. Conference on Pattern Recognition* (IEEE) pp 1–5

Diederichsen S Z *et al* 2020 Comprehensive evaluation of rhythm monitoring strategies in screening for atrial fibrillation: insights from patients at risk monitored long term with an implantable loop recorder *Circulation* **141** 1510–22

Elnaggar I, Hurnanen T, Sandelin J, Lahdenoja O, Airola A, Kaisti M and Koivisto T 2022 Multichannel bed based ballistocardiography heart rate estimation using continuous wavelet transforms and autocorrelation *2022 Computing in Cardiology (CinC) vol 498* (IEEE) pp 1–4

EmFit Ltd 2025 (https://emfit.com/)

Enayati M, Farahani N Z and Skubic M 2020 Machine learning approach for motion artifact detection in ballistocardiogram signals *Proc. 14th EAI Int. Conf. on Pervasive Computing Technologies for Healthcare* pp 406–10

Feng J, Huang W, Jiang J, Wang Y, Zhang X, Li Q and Jiao X 2023 Non-invasive monitoring of cardiac function through Ballistocardiogram: an algorithm integrating short-time Fourier transform and ensemble empirical mode decomposition *Frontiers Physiol.* **14** 1201722

Gibbs H *et al* 2021 Clinical outcomes in asymptomatic and symptomatic atrial fibrillation presentations in GARFIELD-AF: implications for AF screening *Am. J. Med.* **134** 893–901

Greiwe J and Nyenhuis S M 2020 Wearable technology and how this can be implemented into clinical practice *Curr. Allergy Asthma Rep.* **20** 1–10

Harris K, Edwards D and Mant J 2012 How can we best detect atrial fibrillation? *J. R. Coll. Physicians Edinburgh* **42** 5–22

Hurnanen T, Lehtonen E, Tadi M J, Kuusela T, Kiviniemi T, Saraste A, Vasankari T, Airaksinen J, Koivisto T and Pänkäälä M 2016 Automated detection of atrial fibrillation based on time–frequency analysis of seismocardiograms *IEEE J. Biomed. Health Inf.* **21** 1233–41

Inan O T *et al* 2014 Ballistocardiography and seismocardiography: a review of recent advances *IEEE J. Biomed. Health Inf.* **19** 1414–27

Kornej J, Börschel C S, Benjamin E J and Schnabel R B 2020 Epidemiology of atrial fibrillation in the 21st century: novel methods and new insights *Circ. Res.* **127** 4–20

Mora N, Cocconcelli F, Matrella G and Ciampolini P 2020 A unified methodology for heartbeats detection in seismocardiogram and ballistocardiogram signals *Computers* **9** 41

Nagura M, Mitsukura Y, Kishimoto T and Mimura M 2018 A practical bcg measuring system with bed sensors and algorithm for heartbeat detection *2018 IEEE 15th Int. Workshop on Advanced Motion Control (AMC)* (IEEE) pp 317–21

Pänkäälä M, Koivisto T, Lahdenoja O, Kiviniemi T, Saraste A, Vasankari T and Airaksinen J 2016 Detection of atrial fibrillation with seismocardiography *2016 38th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)* (IEEE) pp 4369–74

Perez-Macias J M, Värri A, Himanen S L, Tenhunen M and Viik J 2022 Heart pulse demodulation from emfit mattress sensor using spectral and source separation techniques *2022 Computing in Cardiology (CinC) vol 498* (IEEE) pp 1–4

Sadek I, Biswas J and Abdulrazak B 2019 Ballistocardiogram signal processing: a review *Health Inf. Sci. Syst.* **7** 1–23

Su Q, Huang Y, Wu X, Zhang B, Lu P and Lyu T 2022 Atrial fibrillation detection based on a residual CNN using BCG signals *Electronics* **11** 2974

Tadi M J, Lehtonen E, Hurnanen T, Koskinen J, Eriksson J, Pänkäälä M, Teräs M and Koivisto T 2016 A real-time approach for heart rate monitoring using a Hilbert transform in seismocardiograms *Physiol. Meas.* **37** 1885

Ulrich T 2006 Envelope calculation from the Hilbert transform *Technical Report* (Los Alamos National Laboratory)

Wiard R M, Inan O T, Argyres B, Etemadi M, Kovacs G T and Giovangrandi L 2011 Automatic detection of motion artifacts in the ballistocardiogram measured on a modified bathroom scale *Med. Biol. Eng. Comput.* **49** 213–20

Yu B, Zhang B, Xu L, Fang P and Hu J 2019 Automatic detection of atrial fibrillation from ballistocardiogram (bcg) using wavelet features and machine learning *2019 41st Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)* (IEEE) pp 4322–5

Zink M D, Brüser C, Winnersbach P, Napp A, Leonhardt S, Marx N, Schauerte P and Mischke K 2015 Heartbeat cycle length detection by a ballistocardiographic sensor in atrial fibrillation and sinus rhythm *Biomed. Res. Int.* **2015** 840356