Monitoring of heart rate and inter-beat intervals with wrist plethysmography in patients with atrial fibrillation

Monitoring of heart rate and inter-beat-intervals with wrist plethysmography in patients with atrial fibrillation

Jarkko Harju[1], Adrian Tarniceriu[2], Jakub Parak[3, 4], Antti Vehkaoja[3, 4], Arvi Yli-Hankala[1, 5] and Ilkka Korhonen[3, 4]

[1]Department of Anesthesia, Tampere University Hospital, Tampere, Finland

[2]PulseOn SA, Neuchâtel, Switzerland

[3]Tampere University of Technology, Tampere, Finland

[4]Pulseon Oy, Espoo, Finland

[5]Medical School, University of Tampere, Tampere, Finland

Corresponding author:
Jarkko Harju
Jarkko.harju@fimnet.fi
+358 3 311 6111

Abstract:

Atrial fibrillation (AF) causes marked risk for patients, while silent fibrillation may remain unnoticed if not suspected and screened. Development of comfortable yet accurate beat-to-beat heart rate (HR) monitoring with good AF detection sensitivity would facilitate screening and improve treatment. The purpose of this study was to evaluate whether a wrist-worn photoplethysmography (PPG) device can be used to monitor beat-to-beat HR accurately during postoperative treatment in patients suffering from AF and whether wrist-PPG can be used to distinguish AF from sinus rhythm (SR).

29 patients (14 with AF, 15 with SR, mean age 71.5y) with multiple comorbidities were monitored during routine postoperative treatment. The monitoring included standard ECG, finger PPG monitoring and a wrist-worn PPG monitor with green and infrared light sources. The HR from PPG sensors was compared against ECG derived HR.

The wrist PPG technology had very good HR and beat detection accuracy when using green light. For the SR group, the mean absolute error (MAE) for HR was 1.50 bpm, and for the inter-beat-intervals (IBI), the MAE was 7.64 ms. For the AF group, the MAE for HR was 4.28 bpm and for IBI, the MAE was 14.67 ms. Accuracy for the infrared (IR) channel was worse. Finger PPG provided similar accuracy for HR and better accuracy for the IBI. AF detection sensitivity using green light was 99.0% and the specificity was 93.0%. Performance can be improved by discarding unreliable IBI periods.

Results suggest that wrist PPG measurement allows accurate HR and beat-to-beat HR monitoring also in AF patients, and could be used for differentiating between SR and AF with very good sensitivity.

Keywords: Heart Rate, Pulse Rate, Photoplethysmography, Atrial fibrillation, perioperative monitoring

1 Introduction

Atrial fibrillation (AF) is the most common heart rhythm abnormality and it is associated with a markedly increased risk for acute stroke, heart failure and reduced survival (Rietbrock 2008, Rob A Vermond et al. 2015). Although the risk for morbidity is increased, a marked percentage of AF may remain unnoticed as silent fibrillation if not suspected or searched (Savelieva, Camm 2000).

The gold standard for heart rhythm analysis is electrocardiogram (ECG), but the prevalence of rhythm abnormalities has also been recorded successfully using photoplethysmography (PPG) at the finger (McManus et al. 2016). Long-term monitoring, often required for detecting seldom occurring AF episodes with either method is however not convenient. ECG recording requires placement of typically three to twelve electrodes (Feild et al. 2008). Placement of sensors at fingers is similarly cumbersome and suspect to movement artefacts (Shelley 2007). Thus, more optimal placement options are needed. Traditionally ECG devices have been used for beat-to-beat interval detection because of their high accuracy (Dash et al. 2009, Weippert et al. 2010), but some reports also exist on using PPG sensors at different sites (Parak et al. 2015, Conroy et al. 2017, McManus et al. 2013) or even video plethysmography (Hernandez Guzman 2015).

The analysis of the PPG measurement differs slightly from that of ECG analysis. The recognition of AF is based on two characteristic phenomena's in the ECG: the absence of p-wave and irregularity of the heartbeat intervals. The analysis is typically classified into three categories, first time domain methods, second frequency domain methods and third non-linear methods (Sahoo et al. 2011). In contrast, in the analysis of PPG, the p-wave cannot be seen or analyzed. Thus, the decision on the rhythm is based on regularity or irregularity of the heart contractions and assessed from peripheral pulse wave. As a limitation to the technology, rhythm abnormalities with regular beat such as junctional rhythm cannot be detected using PPG waveform (Nemati et al. 2016). If the heart rate (HR) is irregular pulse amplitude changes constantly especially when the HR is high (Stanton et al. 2008) causing challenges for the measurement.

Recent development in wrist-based fitness monitors and their improved measurement capability also during movements makes it tempting to target their use to hospital patients

(El-Amrawy, Nounou 2015). Some reports exist on the use of such devices on hospitalized patients (Harju et al. 2017), but the evidence on their reliability and accuracy in association to rhythm abnormalities is still very limited. The performance of sensors based on the use of green wavelengths seems most promising during movement (Matsumura et al. 2014, Lee, Matsumura et al. 2013). These optical HR monitors could be used to monitor HR and beat-to-beat HR variability to detect AF (Nemati et al. 2016) or testing of rhythm devices (Fridman et al. 2016). Therefore, we wanted to test the feasibility of the technology during postoperative treatment in patients with AF, the accuracy in both AF and sinus rhythm (SR) patients and whether the technology could be used to distinguish between AF and SR rhythms.

2 Materials and methods

An ethical approval was obtained from the local ethical committee (ETL R17024), and the study was approved by Finnish National Supervisory Authority of Health and Welfare. The study was registered at clinicaltrials.gov (NCT03081793) before study initiation. All patients were treated at the postoperative care unit of Tampere University Hospital after a surgical procedure between May and August 2017. The purpose was to obtain a heterogeneous sample of patients for preliminary evaluation of the feasibility of the novel monitoring method. Prior to any study measurements, all patients gave a written informed consent. The patients were allocated into one of two groups having either AF or sinus rhythm (SR). Equal number of patients were initially recruited into both groups. One patient in AF group was later discarded. The rhythm was prescreened based on preoperative ECG, but the group allocation was decided based on current rhythm at the beginning of each measurement from the reference monitor by an anesthesiologist (JH).

The study population consisted of patients subjected to surgery requiring treatment at post-operative care unit and willing to participate in the study. The patients were mostly lying down but were allowed to move hands freely. They were also subjected to normal nursery during their treatment. Patients with marked shivering at the time of approval such as patients with Parkinson disease were excluded from the study. Similarly, patients with cardiac pacemakers were not included.

The tested variables were HR, beat detection, inter-beat-interval (IBI) estimation and AF detection. ECG recorded with clinical patient monitor was used as a reference. For the wrist PPG sensor, we tested the performance for both green (PPG_{green}) and IR (PPG_{IR}) wavelengths. To have a better understanding of the accuracy of this device, also fingertip PPG signals (PPG_{finger}) were analyzed.

2.1 Data collection

Wrist PPG signals were recorded at 25Hz using the PulseOn Optical Heart Rate (OHR) monitor prototype (PulseOn Ltd, Espoo, Finland), presented in Figure 1. Similar technology has been previously tested on healthy subjects and achieved high reliability and accuracy on HR estimation during sport, daily and rest activities (Delgado-Gonzalo et al. 2015) and on inter-beat interval (IBI) detection (Parak et al. 2015). Prior to any study measurements the device was tested and approved by the hospital for patient. The device was placed around the wrist about one finger width proximally from the capitulum radii as instructed by the manufacturer. The wristband was tightened firmly but comfortably to ensure adequate skin contact. A mobile phone was used to control the device and the recorded data were downloaded from the device after each measurement period. The device was capable to measure PPG with green (535 nm) and IR (940 nm) light. The device did not have a display for data verification simultaneously to the measurement. The PPG signals were processed with the PulseOn proprietary algorithms afterwards to obtain the HR and IBI and to make the AF/SR assessment.

Figure 1 PulseOn OHR monitor

The standard monitoring was carried out using GE Carescape™ B850 monitor (GE Healthcare Oy, Helsinki, Finland) including measurement of ECG, PPG (TruSignal finger blood oxygen saturation sensor) and skin temperature. The recordings from standard monitoring were collected using the S5Collect software (GE Healthcare Oy, Helsinki, Finland); the ECG and finger PPG as 300 Hz waveform collection and other parameters (e.g. HR and PPG pulse rate) as 10 seconds average values.-RR intervals (RRI) were obtained from the ECG signal using the Kubios HRV software, version 2.2 (Kubios Ltd, Kuopio, Finland). The ECG waveforms were also visually inspected to ensure that no R-waves were missed, and to correct any R-wave detection errors.

The recording was done during normal postoperative care treatment. The duration of the measurement was between one and two hours depending on patient availability and the data storage limitation of the current version of study device. An example of ECG, fingertip and wrist (green and infrared) PPG waveforms is given in Figure 2.

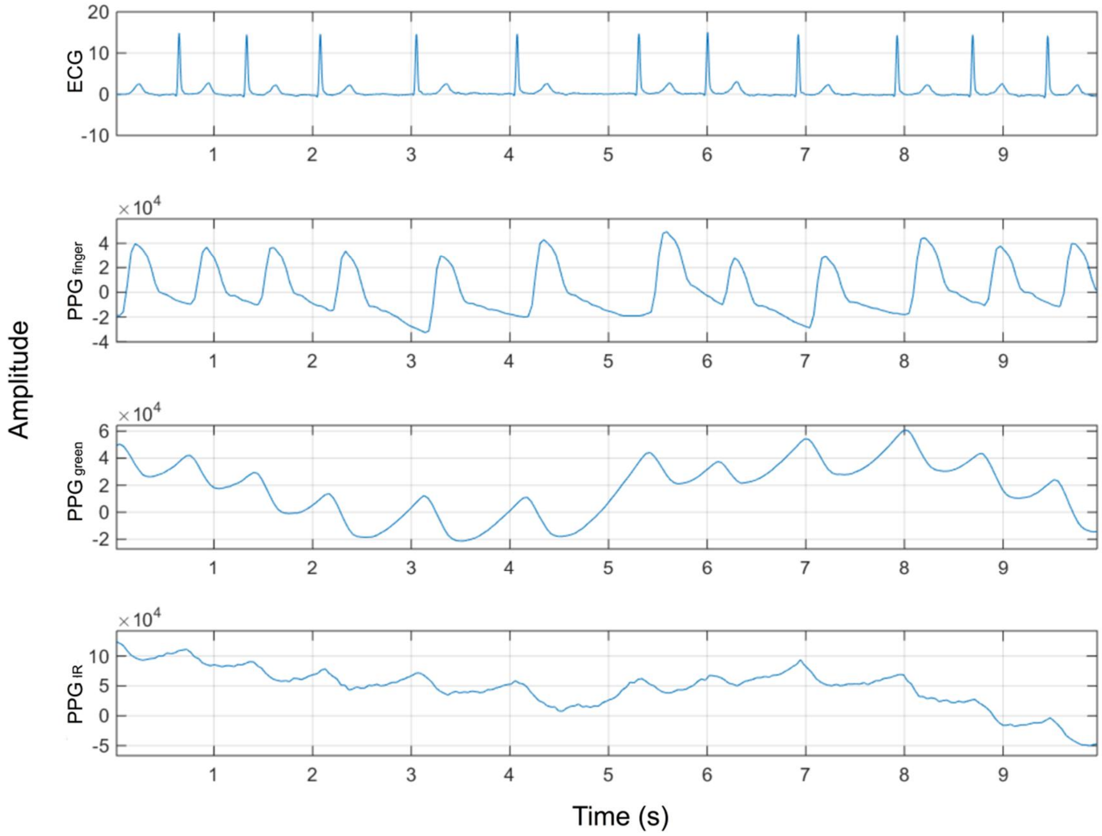

Figure 2. An example of ECG, fingertip and wrist (green and infrared) PPG waveforms.

2.2 Analysis of monitoring performance

PPG HR, beat detection, IBI estimation, and AF detection were all obtained starting from the raw PPG signals using the PulseOn algorithms. As the processing steps are not the same for these parameters, we will describe them separately.

2.2.1 HR estimation

Firstly, the PPG HR and the reference HR obtained from the ECG were interpolated to the same sampling frequency (100 Hz) and synchronized by minimizing the mean absolute error between them. Afterwards, both PPG and ECG HR data were averaged over 5 second windows. The HR estimation performance was evaluated by the following parameters:

- Mean Error (ME): average of the difference between the reference and estimated HR
- Mean Absolute Error (MAE): average of the absolute difference between the reference and estimated HR
- Root Mean Square Error (RMSE): square root of the average of squared errors between the reference and estimated HR
- Reliability: percentage of time when the absolute error is below 10 beats per minute (bpm)

2.2.2 IBI detection and estimation

The IBIs are estimated from the PPG signals as described in (Renevey et al. 2013). As a first processing step, the derivative of the PPG signal is computed by taking the difference of consecutive samples. Afterwards, IBIs are computed as the time-distance between consecutive local minima of the smoothed PPG derivative. At this point, the resolution is given by the sampling frequency of 25 Hz, corresponding to a sampling period of 40 ms, and is not sufficient for HRV analysis or AF detection. However, this limitation is overcome by one more processing step, consisting of second-order polynomial interpolation around the detected minima points. Similarly, as for HR estimation, the vectors containing IBI from PPG signals and reference RRI from ECG signals, were synchronized by aligning their time axis so that the MAE between the time series was minimized. Afterwards, to compensate for eventual time drifts between the PulseOn and Carescape B850 clocks, we split the data in intervals of one minute and performed a new synchronization for each interval.

After synchronization, for each one-minute interval, we determined the percentage of correctly detected beats, extra beats, and missed beats with respect to the ECG reference. This was done with a method similar to the one proposed in (Parak et al. 2015). For every PPG beat, we checked how many reference beats were detected in the interval $[t - 0.5l, t + 0.5l]$, where t is the time when the beat was detected and l is the length of the corresponding IBI. If there was only one reference beat, the beat was marked as correctly detected. If there were no corresponding reference beats, then an extra beat had been detected. The reference beats with no corresponding PPG-detected beat were considered as missed beats. Note that the extra and missing beats are mostly a consequence of movement or other types of artefacts. Thus, they could be used as an indicator for the PPG signal quality. Also, as we are currently targeting only SR or AF scenarios, ectopic beats were discarded from the analysis.

The IBI detection and estimation performance was evaluated by the following parameters:

- Percentage of correctly detected beats
- Percentage of missed beats
- Percentage of extra beats
- ME: average of the difference between the reference RRI and estimated IBI
- MAE: average of the absolute difference between the reference RRI and estimated IBI
- RMSE: square root of the average of squared errors between the reference RRI and estimated IBI

The ME, MAE, and RMSE were estimated using four different approaches.

- Using only beats from one-minute intervals where no extra or missed beats were detected. As the goal of the paper is to evaluate IBI estimation during rest, this ensures that no beats affected by motion are considered in the statistics. This corresponds to theoretical maximum performance but might ignore too much of the collected data.
- By discarding +/- 5 beats around each extra or missed beat. This eliminates most motion contaminated periods from the data.

- By discarding +/- 2 beats around each extra or missed beat. This includes more data for the statistics, but also increases the risk of considering motion-affected IBI.
- Considering all IBI with a one-to-one correspondence to a reference RRI. This includes all motion-affected IBI in the statistics.

2.2.3 AF Detection

We developed an AF detection algorithm that discriminates AF from SR based on processing the values of successive IBI. For each input IBI, the algorithm returns the most likely condition, AF or SR. The AF detection algorithm is based on the work of (Moody & Mark 1983). The IBI series are modelled as a Markov process, where the states are given by the IBI length. The transition matrices containing the transition probabilities between states have been computed using the MIT/BIH Arrhythmia Database, MIT/BIH Atrial Fibrillation Database, and MIT/BIH Supraventricular Arrhythmia Database (www.physionet.org, Goldberger et al. 2000), for both SR and AF data. Given these two matrices, one can determine the most likely condition, SR or AF, by maximizing the conditional probability for an observed sequence of beats. In our implementation, we used sliding windows of 20 consecutive beats. This is large enough to provide a stable output, but still provides a fast decision (the maximum delay is 20 beats).

For ease of implementation, the log-likelihood matrix was computed from the two transition matrices by dividing them element-wise and taking the logarithm. Thus, instead of computing two conditional probabilities for an observed IBI sequence, one only needs to sum the log-likelihoods, and take the decision based on the sign of the result. In addition, as the observed IBI intervals are not identical to the Markov states, the entries of the log-likelihood matrix are interpolated based on the IBI values to reduce the quantization error. Further smoothing of the result increases the method’s stability.

To have an estimate of the influence of IBI reliability on AF detection, we evaluated four different cases, as in 2.2.2. Considering all IBI from one-minute intervals with no extra or missed beats, discarding +/- 5 beats around each extra or missed beat, discarding +/- 2 beats around each extra or missed beat, and considering all IBI that have a one-to-one correspondence to an ECG RRI. For all cases, we estimate:

- Sensitivity: the percentage of AF data classified as AF
- Specificity: the percentage of SR data classified as SR

2.3 Statistical analysis

The data were analyzed using IBM SPSS statistics version 24 (IBM, IL, USA). The data are described as mean and standard deviation (SD) or number and percentage. The IBI comparisons are described as numbers of extra detected and missed beats, as Bland–Altman plots (Bland, Altman 2007), ME, MAE, and RMSE. The statistical test used for p-value calculation is independent samples t-test for normally distributed values and Man-Whitney U test for nominal values. P-value<0.05 was considered significant. As an exception, wrist diameter and skin temperature were divided into three equal quantiles and compared using ANOVA.

3 Results

A total of thirty patients were recruited. One patient in the AF group had initially AF, which was reversed to SR soon after the beginning of the measurement. Therefore, the patient was discarded from the final analysis. The mean age of the study group was 71.5 (SD 10.4) years. Most of the patients (n=21) had paroxysmal or current AF rhythm, although the current rhythm was SR in 15 of the patients. The patient characteristics and background variables are described in Table 1.

Table 1 Patient characteristic and selected background variables for sinus rhythm (SR) and atrial fibrillation (AF) groups. Described as mean (SD) or frequency (%).

Characteristics	SR (n=15)	AF (n=14)	p-value
Age	67.5 (10.7)	74.8 (8.3)	0.053
Height (cm)	169.6 (8.1)	168.2 (7.8)	0.644
Weight (kg)	79.9 (13.7)	71.9 (18.7)	0.199
Wrist diameter (cm)	18.4 (1.4)	18.1 (2.0)	0.707
Sex (n)			0.652
Male (n)	8 (53.3%)	6 (42.9%)	
Female (n)	7 (46.7%)	8 (57.1%)	
Diabetes (n)	3 (20.0%)	2 (14.3%)	0.813
Peripheral artery disease (n)	2 (13.3)	1 (7.1%)	0.780
Coronary artery disease (n)	2 (13.3%)	7 (50%)	0.093
Atrial fibrillation (n)	7 (46.7%)	14 (100%)	0.014

3.1 HR estimation

The HR estimation results are provided in Table 2. For both SR and AF groups, the performance in terms of MAE is relatively close for PPG_{green} and PPG_{finger}. PPG_{IR} has a much lower accuracy than the PPG_{green} and PPG_{finger} measurements. The error statistics are better for the SR group than for the AF group.

Table 2 HR analysis describing differences between ECG, and photoplethysmographic measurement with each sensor and the average time of heart rate error under 10 bpm. (SR patients n=15, AF n=14). Described as mean (SD).

	PPG_{green}			PPG_{IR}			PPG_{finger}		
	SR	AF	p-value	SR	AF	p-value	SR	AF	p-value
ME [bpm]	-0.63 (1.1)	0.07(3.6)	0.149	1.67 (2.8)	2.62(8.1)	0.670	-0.69(0.6)	3.22(4.8)	0.004
MAE [bpm]	1.50(1.5)	4.28 (2.0)	<0.001	3.35 (3.0)	8.77 (4.7)	0.001	1.35 (0.5)	5.83 (2.9)	<0.001
RMSE [bpm]	2.22(1.0)	5.76 (2.4)	<0.001	5.89(4.9)	12.50 (8.2)	0.005	1.92(0.8)	7.32 (3.2)	<0.001
Reliability	99.21(1.4)	91.06 (10.8)	0.009	92.77(11.3)	72.39 (20.4)	0.002	99.42(1.0)	82.10 (16.8)	<0.001

3.2 IBI estimation analysis

The beat detection performance is presented in **Error! Reference source not found.**. The ability to recognize correct beats was very good in both fingertip and green wrist data, and lower for IR wrist data. As extra/missed beats are a good indicator of the presence of motion or other noise, we can conclude that the IR signals are more sensitive to noise. Figure 3 describes an example of the waves used for IBI detection in both SR and AF.

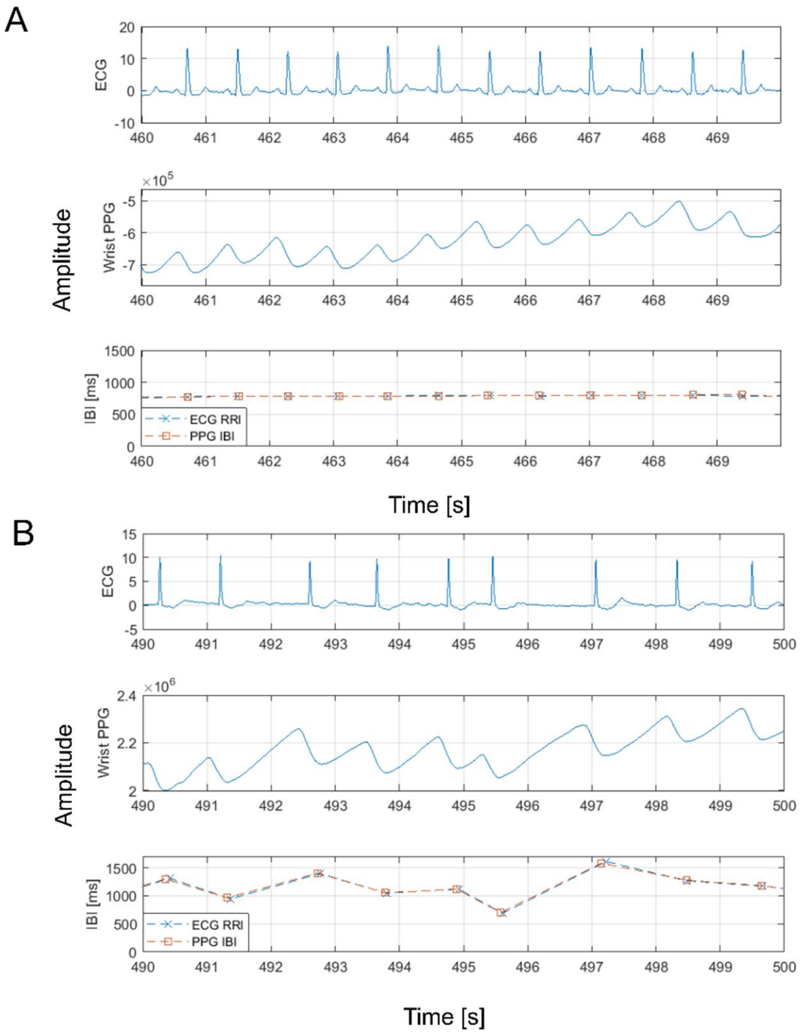

Figure 3 An example of signals and IBI detection during sinus rhythm (A) and atrial fibrillation (B)

The beat estimation performance is presented in Table 4. Besides ME, MAE and RMSE, also the percentage of beats used in the analysis is shown.

For both SR and AF sets, the statistics are better for the fingertip collected signals. This is not surprising, because the fingertip is better vascularized than the wrist. As for HR estimation, PPG_{IR} has the worst performances. The statistics are better for the SR group than for the AF group.

Table 3 Beat detection analysis. Comparison of values obtained from ECG to finger and wrist measurements in patients with sinus rhythm (SR) or atrial fibrillation (AF).-Described as mean(SD). (SR patients n=15, AF n=14)

	PPG $_{green}$			PPG $_{IR}$			PPG $_{finger}$		
	SR	AF	p-value	SR	AF	p-value	SR	AF	p-value
Extra beats %	2.4 (1.8)	3.6 (3.5)	0.262	8.9 (5.2)	10.3(8.9)	0.465	1.9 (2.5)	4.1 (5.1)	0.184
Missing beats %	1.2 (1.3)	3.4 (4.3)	0.088	4.5 (4.8)	6.2 (7.1)	0.483	0.3 (0.5)	1.7 (4.4)	0.216
Total beats [n]	87793	86519		81143	83365		85283	89035	

Table 4 Beat accuracy analysis. Comparison of values obtained from ECG to finger and wrist measurements in patients with sinus rhythm (SR) or atrial fibrillation (AF) compared to ECG. The numbers are calculated for clean one-minute intervals without extra or missed beats, when discarding +/- 5 beats around each extra or missed beat, when discarding +/- 2 beats around each extra or missed beat, and considering all IBI that have a one-to-one correspondence to an ECG RRI. Described as mean (SD) (SR patients n=15, AF n=14).

		PPG $_{green}$			PPG $_{IR}$			PPG $_{finger}$		
		SR	AF	p-value	SR	AF	p-value	SR	AF	p-value
Beats from clean intervals	beat %	44.2 (22.8)	31.4 (27.4)	0.133	13.03 (18.3)	13.46 (16.2)	0.652	59.68(26.8)	26.9 (22.8)	0.002
	ME [ms]	-0.1(2.0)	-0.5(0.7)	0.213	-0.7(-5.2)	-0.2(0.4)	0.272	-0.5(0.5)	-0.5(0.4)	0.414
	MAE [ms]	7.6(5.0)	14.7(7.1)	0.013	12.4(28.7)	19.1(11.9)	0.067	6.0(4.4)	8.7(4.0)	0.221
	RMSE [ms]	18.92	23.20		22.61	27.84		15.75	17.23	
Ignore +/-5 beats	beat % (min, max)	69.43 (18.8)	61.4 (24.4)	0.332	37.8 (22.9)	38.0 (28.3)	0.892	74.8(22.8)	62.46 (26.1)	0.126
	ME [ms]	-0.5(0.5)	-0.6(1.3)	0.738	-0.4(2.6)	-0.5(7.2)	0.585	-0.6(0.7)	-1.0(5.6)	0.284
	MAE [ms]	10.4(5.8)	17.8(6.1)	0.004	20.7(28.7)	24.4(31.4)	0.406	7.3(5.8)	10.9(9.3)	0.073
	RMSE [ms]	29.57	34.69		50.05	48.85		21.28	33.74	
Ignore +/- 2 beats	beat % (min, max)	80.84 (12.7)	74.8 (17.9)	0.296	54.0 (21.6)	52.0 (27.7)	0.702	83.65 (16.4)	77.7 (20.2)	0.225
	ME [ms]	-0.6(0.6)	-0.5(1.8)	0.881	-0.5(4.3)	-1.2(5.3)	0.652	-0.6(0.8)	-1.1(3.6)	0.353
	MAE [ms]	12.4(6.4)	20.5(8.4)	0.008	29.2(29.5)	31.8(32.7)	0.506	8.4(7.1)	12.1(9.3)	0.163
	RMSE [ms]	35.76	42.42		67.58	65.03		26.08	37.18	
All beats with a corresponding reference	beat % (min, max)	94.01(5.2)	93.84(4.3)	0.989	85.33(9.4)	84.04(11.0)	0.585	93.94(7.8))	94.37(5.6))	0.905
	ME [ms]	-1.67(7.0)	-4.48(27.0)	0.473	-16.09(24.2)	-32.50(65.6)	0.220	-7.28(12.8)	-20.16(47.3)	0.186
	MAE [ms]	26.55(15.6)	50.95(29.0)	0.008	83.10(48.3)	104.78(73.1)	0.199	18.75(19.4)	43.69(50.3)	0.034
	RMSE [ms]	88.97	138.25		174.91	215.44		79.08	168.10	

The IBIs from wrist and finger sensors were compared with reference RRIs also through Bland-Altman plots, shown in 4 and 5 for SR and AF groups, respectively. In the SR group the mean difference was 0.22 ms between ECG and PPG_{green}, 0.67 ms between ECG and PPG_{IR}, and 0.45 ms between ECG and PPG_{finger}. In the AF group the mean difference was 0.48 ms between ECG and wrist PPG_{green}, 0.24ms between ECG and PPG_{IR} and 0.52 ms between ECG and PPG_{finger}.

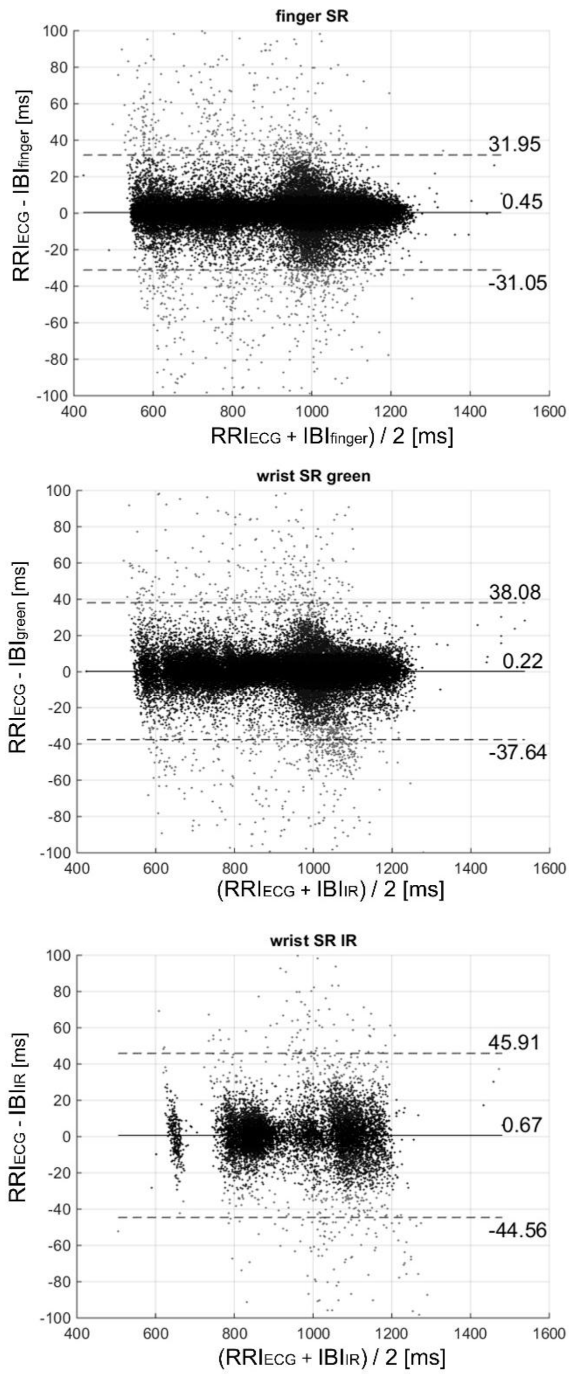

Figure 4 A Bland Altman Plot showing comparison of ECG and finger or ECG and wrist inter-beat interval measurement in patients with sinus rhythm. The numbers indicate mean error and limits of agreement.

.

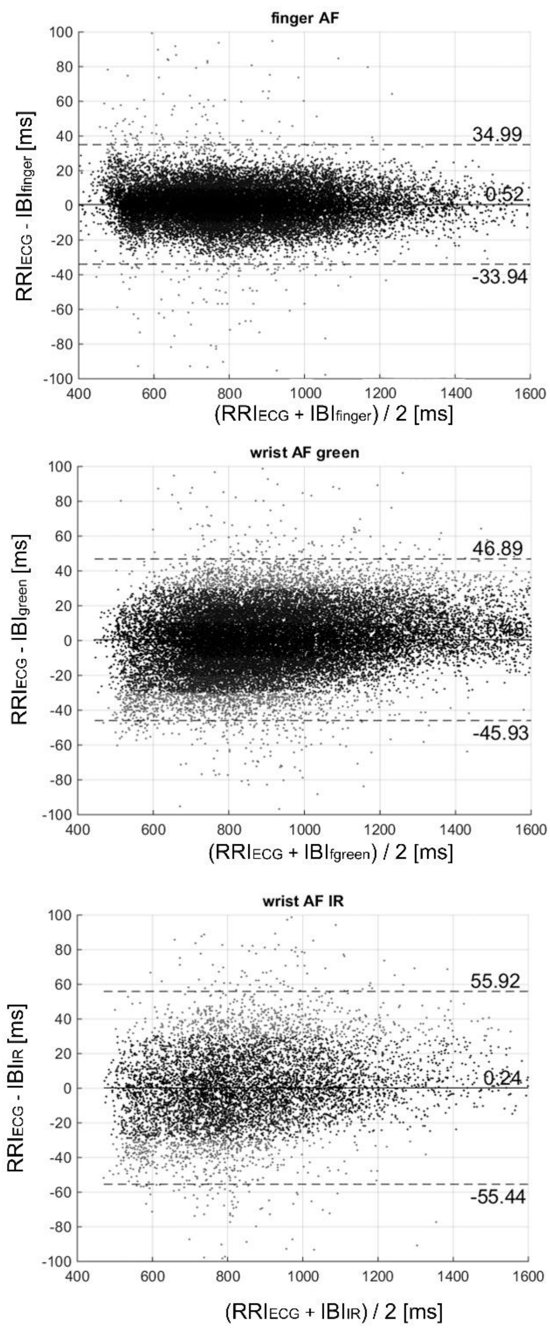

Figure 5 A Bland Altman Plot showing comparison of ECG and finger or ECG and wrist inter-beat interval measurement in patients with atrial fibrillation. The numbers indicate mean error and limits of agreement.

Selected background variables potentially affecting the IBI estimation accuracy using clean 1-minute intervals are compared in Table 5 . There were no statistically significant differences caused by any of the background variables with respect to ME or MAE.

Table 5 The comparison of selected background variables to mean error (ME) and mean absolute error (MAE) of IBI estimation accuracy. Described as mean (standard deviation).

Variable	n	ME (ms)	p-value	MAE (ms)	p value
Location			0.154		0.065
PPG_{finger}	29	-0.46(0.42)		7.93 (4.15)	
PPG_{green}	29	-0.23(1.54)		10.50(6.08)	
PPG_{IR}	29	-1.39(3.83)		16.15(22.15)	
AF Yes/No	14 / 15	-0.57(0.67) / 0.09(2.01)	0.256	11.43(7.10) / 9.64(5.03)	0.439
MCC Yes/No	9 / 20	-0.17(1.25) / -0.256(1.61)	0.891	12.82(5.66) / 9.46(6.1)	0.174
ASO Yes/No	3 / 26	0.753 (2.22) / -0.34(1.46)	0.250	13.11(7.71) / 10.18(5.96)	0.405
Diabetes yes / no	5 / 24	0.86(2.94) / -0.45(1.03)	0.081	11.62(3.28) / 10.28 (6.53)	0.662
Wrist diameter			0.230		0.347
Q1	10	-0.65(0.74)		9.44(5.4)	
Q2	11	-0.37(1.36)		12.67(7.06)	
Q3	10 / 11 / 7	//0.64(2.4)		8.78(5.56)	
Skin temperature			0.664		0.075
Q1	9	-0.41(0.54)		6.26(4.44)	
Q2	8	0.10(2.53)		11.62(6.62)	
Q3	9	-0.48(0.21)		11.87(5.44)	

3.3 AF detection

The AF detection results are presented in Table 6. The sensitivity was found very good with all sensors whereas the specificity was varying. The lowest specificity was associated to the PPG_{IR} sensor and when all beats were accounted into the analysis. This is the case when most pulse artefacts are considered.

Table 6 AF detection analysis for the SR and AF groups. The results are calculated for clean one-minute intervals without extra or missed beats, when discarding +/- 5 beats around each extra or missed beat, when discarding +/- 2 beats around each extra or missed beat, and considering all IBI that have a one-to-one correspondence to an ECG RRI. Described as percentage. (SR patients n=15, AF n=14)

	Sensitivity (%)			Specificity (%)		
	PPG_{green}	PPG_{IR}	PPG_{finger}	PPG_{green}	PPG_{IR}	PPG_{finger}
Beats from clean intervals	99.88	100.00	94.50	99.92	99.40	99.87
Ignore +/-5 beats	98.31	99.76	96.61	99.83	96.64	99.90
Ignore +/- 2 beats	98.55	99.76	97.12	99.56	91.54	99.76
All beats with a corresponding reference	99.00	99.70	97.70	92.96	53.85	92.73

4 Discussion

The main finding in our study is that reflective PPG technology was able to correctly measure HR and detect AF very well. As the HR is an average value obtained from a longer segment of PPG signal it was not that much affected by artifacts in the pulse wave as the IBIs and AF detection. This is clearly shown in the detailed analysis of both IBI estimation and AF, where the performance or the specificity dropped markedly if all IBI were taken into analysis. The PPG_{green} measurements were less prone to artifacts than the PPG_{IR} measurements.

HR estimation

The HR based on wrist PPG, was highly accurate when using green light. The MAE and reliability for HR estimation was found to be similar to fingertip measurements. The error relative to ECG readings was very small for both SR and AF permitting beat-to-beat comparison.

Previous studies have compared HR estimation to ECG with varying results. The measurement at finger has shown high accuracy and precision (Chuang et al. 2015, Coppetti et al. 2017). However, the performance of wrist-based measurements has been studied less. In a study by Kroll et al., the mean difference between HR from ECG and personal fitness tracker was -1.14 bpm (limits of agreement 24 bpm) during intensive care treatment. However, in that study, the HR was recorded only every 5 minutes (Kroll, Boyd & Maslove 2016). In a study by Fukushima et al., the mean SD was 8.7 bpm when an acceleration sensor was used to enhance the performance of a wrist sensor during movement in healthy subjects (Fukushima et al. 2012). In another study, the use of multiple artefact rejection showed 2.34 bpm average absolute error of HR estimation (Zhang, Zhou & Zeng 2017). Commercially available wrist based sensors have also been compared in some studies with healthy subjects but the results have been varying and thus cannot be applied directly to hospitalized patients (Terbizan, Dolezal & Albano 2002, Parak & Korhonen 2014, Wang et al. 2017, Hendrikx et al. 2017, Shcherbina et al. 2017). In the present study, the very good accuracy previously found in healthy subjects (Delgado-Gonzalo et al. 2015) was repeated in patients with marked comorbidities and cardiac rhythm irregularities.

The existence of cardiac rhythm irregularities makes the estimation of HR from peripheral pulse more challenging. The irregularity in heart contractions changes the cardiac output volume causing changes in pulse amplitude especially at high HR (Stanton et al. 2008). This was also seen in our study with a slightly lower accuracy in AF patients when compared to SR. In the estimation of HR, averaging or different type of signal processing methods, such as spectral analysis is often used to lower the effect of interference in the signal. However, spectral techniques are compromised when dealing with highly irregular rhythm. Further, if the signal is used for more detailed analysis such as heart rate variability (Gil et al. 2010) or AF detection (Weippert et al. 2010), estimation of beat-to-beat intervals is needed.

IBI estimation

In the absence of motion, the beat-to-beat estimation accuracy reported in this study is comparable to studies using ECG signals and automated beat-to-beat detection (Dash et al. 2009, Stoyan Tanev 2012, Phukpattaranont 2015, Cvikl, Zemva 2010). This proves that wrist PPG signals can provide accurate IBI estimation for elderly subjects, suffering of arrhythmias. As a result, the used technology has a potential to be used for arrhythmia detection.

As it was expected, the best performance is obtained when considering only beats from the one-minute intervals with no extra or missed beats. However, this approach is too restrictive, the amount of used IBI being below 50% in most cases. Discarding only +/- 5 beats or +/- 2 beats around the extra or missed beats enables more data to be used when computing the statistics. The performance decreases when considering more IBI, because the risk of considering motion-affected IBI increases. The bottom part of Table 4 shows the statistics when using all IBI with a one-to-one correspondence to a reference RRI. This also includes IBI that are affected by motion. The statistics are clearly worse than the other cases because IBI estimation is not accurate during motion periods. However, the main goal of this paper is to evaluate IBI estimation during rest periods, and the first two sections (“Beats from clean intervals” and “Ignore +/- 5 beats”) of Table 4 are representative for this.

AF detection

Previous evidence on the performance of wrist based sensors in AF patients is very limited. Nemati et al. reported the development of an AF detecting algorithm using a multiple sensor smartwatch and obtained 94% specificity and 97 % sensitivity in detecting AF (Nemati et al. 2016). A database was used to develop the algorithm, but the accuracy of beat-to-beat interval estimation was not reported. Other studies measuring the reflection PPG at finger have shown slightly poorer performance (Lee, Reyes et al. 2013). In our study when considering all beats, there are many cases when the IBI are not correctly estimated because of the presence of motion or other artifacts. This is shown in our study, when only data from one-minute intervals with no extra or missed beats were accepted. With a less sensitive approach more data was included in the analysis causing reduction in the accuracy. Still, the overall performance of the AF detection algorithm was comparable to previous findings. The next step is to develop an automatic way of rejecting poor data, which would enable improving the specificity of AF detection.

The sensitivity values are generally high, being above 98% for each dataset with one exception. It might seem surprising that the sensitivity is the lowest for the fingertip measurements, when considering only beats from clean intervals, when the IBI estimation is the most accurate. This is explained by the fact that the AF detection algorithm is based on IBI values, and for one set there are periods with stationary IBI, similar to SR. However, in the inspection of the ECG waveform p-waves were missing and the rhythm was defined as AF. These periods of regular IBI were more accurately estimated from finger PPG, and, because the classification method is based on IBI variability, classified as SR. For PPG_{green} and PPG_{IR}, there are more missed or extra beats, and more estimation errors. As a result, the IBI irregularity increases, and more data is classified as AF.

The specificity decreases as we consider more beats for the estimation, because more wrongly estimated beats are considered in the analysis. This leads to more irregular IBI sequences, which are more representative for AF than for SR. The PPGIR performance is considerably lower than PPGgreen or PPGfinger.

Sensor location and wavelength comparison

The performance was similar for the wrist green light and the fingertip sensor for HR and IBI estimation, and for AF detection. However, the wrist PPG represents a more comfortable solution and does not limit the subject's movement or ability to use hands.

IR wavelength was also tested for the wrist location. It showed constantly lower performance than the green light. This is in agreement with previous studies showing that although HR can be measured from wrist region using IR wavelength (Harju et al. 2017), it is associated with increased effect of motion artefacts and lower accuracy when compared to green light (Matsumura et al. 2014, Lee et al. 2013, Maeda, Sekine & Tamura 2011). Interestingly, one study has reported lowest artifact rate by using a differential channel that combined the recordings of both red and green wavelengths (Zhou et al. 2016).

4.1 Limitations

While obtaining good results, there are some limitations in the current study. All patients were mostly stationary allowing generally good conditions for the measurement. The hemodynamics were also stable and no adrenergic drugs were used during the study period although that was not an exclusion criteria in the study. Furthermore, all but one patient had a stable heart rhythm that did not convert from AF to SR or vice versa during study measurement. Data recorded during changing of the rhythm would be needed to evaluate how the algorithm behaves in such circumstances, since the change doesn't necessarily happen immediately. Although the results are promising, more evidence on the performance in non-stationary patients and the performance in long term follow up at ward is needed.

5 Conclusions

The wrist PPG was able to estimate HR with high accuracy for both SR and AF patients when green light was used. In the absence of motion, also the beat-to-beat interval estimation was accurate. The accuracy was lower for patients with AF when compared to patients having SR, but in both cases the IBI were in close agreement with the RRI reference, enabling reliable AF detection. Thus, this technology could be used to monitor HR in hospitalized patients and has high potential in recognizing arrhythmias. In addition, its location makes it easy to wear and is therefore suitable for long term monitoring.

Acknowledgements The authors wish to thank the voluntary subjects for their participation and PulseOn Ltd (Espoo, Finland) for providing the study device.

Conflicts of interest: JH and AYH declare no conflict of interest. AT, JP and AV are employed workers of PulseOn Ltd. IK is a member of advisory board of PulseOn Ltd.

Ethical approval This study was approved by the Pirkanmaa Hospital district ethics committee. All procedures involving human participants were performed in accordance with the ethical standards of the institutional and/or national research committee, as well as with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

ORCID iDs

J Harju https://orcid.org/0000-0002-3633-0314

References

Bland, J.M. & Altman, D.G. 2007, "Agreement Between Methods of Measurement with Multiple Observations Per Individual", *Journal of Biopharmaceutical Statistics,* vol. 17, no. 4, pp. 571-582.

Chuang, C., Ye, J., Lin, W., Lee, K. & Tai, Y. 2015, "Photoplethysmography variability as an alternative approach to obtain heart rate variability information in chronic pain patient", *Journal of Clinical Monitoring and Computing,* vol. 29, no. 6, pp. 801-806.

Conroy, T., Hernandez Guzman, J., Hall, B., Tsouri, G. & Couderc, J.P. 2017, "Detection of atrial fibrillation using an earlobe photoplethysmographic sensor", *Physiological Measurement,* .

Coppetti, T., Brauchlin, A., Müggler, S., Attinger-Toller, A., Templin, C., Schönrath, F., Hellermann, J., Lüscher, T.,F., Biaggi, P. & Wyss, C.A. 2017, "Accuracy of smartphone apps for heart rate measurement", *Eur J Prev Cardiolog,* vol. 24, no. 12, pp. 1287-1293.

Cvikl, M. & Zemva, A. 2010, "FPGA-oriented HW/SW implementation of ECG beat detection and classification algorithm", *Digital Signal Processing,* vol. 20, no. 1, pp. 238-248.

Dash, S., Chon, K.H., Lu, S. & Raeder, E.A. 2009, "Automatic real time detection of atrial fibrillation", *Annals of biomedical engineering,* vol. 37, no. 9, pp. 1701-1709.
Delgado-Gonzalo, R., Parak, J., Tarniceriu, A., Renevey, P., Bertschi, M. &

Korhonen, I. 2015, "Evaluation of accuracy and reliability of PulseOn optical heart rate monitoring device", *Conference proceedings : ...Annual International Conference of the IEEE Engineering in Medicine and Biology Society.IEEE Engineering in Medicine and Biology Society.Annual Conference,* vol. 2015, pp. 430-433.

El-Amrawy, F. & Nounou, M.I. 2015, "Are Currently Available Wearable Devices for Activity Tracking and Heart Rate Monitoring Accurate, Precise, and Medically Beneficial?", *Healthcare informatics research,* vol. 21, no. 4, pp. 315-320.

Feild, D.Q., Zhou, S.H., Helfenbein, E.D., Gregg, R.E. & Lindauer, J.M. 2008, "Technical challenges and future directions in lead reconstruction for reduced-lead systems", *Journal of electrocardiology,* vol. 41, no. 6, pp. 466-473.

Fridman, V., Saponieri, C., El-Sherif, N. & Turitto, G. 2016, "Cardiac rhythm device threshold testing via pulse oxymetry", *Journal of Atrial Fibrillation,* vol. 8, no. 6, pp. 29-31.

Fukushima, H., Kawanaka, H., Bhuiyan, M.S. & Oguri, K. 2012, "Estimating heart rate using wrist-type Photoplethysmography and acceleration sensor while running", IEEE, United States, pp. 2901.

Gil, E., Orini, M., Bailón, R., Vergara, J.M., Mainardi, L. & Laguna, P. 2010, "Photoplethysmography pulse rate variability as a surrogate measurement of heart rate variability during non-stationary conditions", *Physiological Measurement,* vol. 31, no. 9, pp. 1271-1290.

Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng C-K, Stanley HE. 2000, ”PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals.” *Circulation* vol **101 no 23 pp. e**215-e220

Harju, J., Vehkaoja, A., Lindroos, V., Kumpulainen, P., Liuhanen, S., Yli-Hankala, A. & Oksala, N. 2017, "Determination of saturation, heart rate, and respiratory rate at forearm using a Nellcor™ forehead SpO2-saturation sensor", *Journal of Clinical Monitoring and Computing,* vol. 31, no. 5, pp. 1019-1026.

Hendrikx, J., Ruijs, S.L., Cox, G.L., Lemmens, M.P., Schuijers, G.E. & Goris, H.A. 2017, "Clinical Evaluation of the Measurement Performance of the Philips Health Watch: A Within-Person Comparative Study", *JMIR Mhealth Uhealth,* vol. 5, no. 2, pp. e10.

Hernandez Guzman, J. 2015, *Cardiac Inter Beat Interval and Atrial Fibrillation Detection using Video Plethysmography*, RIT Scholar Works.

Kroll, R.R., Boyd, J.G. & Maslove, D.M. 2016, "Accuracy of a Wrist-Worn Wearable Device for Monitoring Heart Rates in Hospital Inpatients: A Prospective Observational Study", *Journal of medical Internet research,* vol. 18, no. 9, pp. e253.

Lee, J., Matsumura, K., Yamakoshi, K., Rolfe, P., Tanaka, S. & Yamakoshi, T. 2013, "Comparison between red, green and blue light reflection photoplethysmography for heart rate monitoring during motion", *Conference proceedings : ...Annual International Conference of the IEEE Engineering in Medicine and Biology Society.IEEE Engineering in Medicine and Biology Society.Annual Conference,* vol. 2013, pp. 1724-1727.

Lee, J., Reyes, B.A., McManus, D.D., Maitas, O. & Chon, K.H. 2013, "Atrial fibrillation detection using an iPhone 4S", *IEEE transactions on bio-medical engineering,* vol. 60, no. 1, pp. 203-206.

Maeda, Y., Sekine, M. & Tamura, T. 2011, "The Advantages of Wearable Green Reflected Photoplethysmography", *Journal of medical systems,* vol. 35, no. 5, pp. 829-834.

Matsumura, K., Rolfe, P., Lee, J. & Yamakoshi, T. 2014, "iPhone 4s photoplethysmography: which light color yields the most accurate heart rate and normalized pulse volume using the iPhysioMeter Application in the presence of motion artifact?", *PloS one,* vol. 9, no. 3, pp. e91205.

McMANUS, D.D., Chong, J.W., Soni, A., Saczynski, J.S., Esa, N., Napolitano, C., Darling, C.E., Boyer, E., Rosen, R.K., Floyd, K.C. & Chon, K.H. 2016, "PULSE-SMART: Pulse-Based Arrhythmia Discrimination Using a Novel Smartphone Application", *Journal of cardiovascular electrophysiology,* vol. 27, no. 1, pp. 51.

McManus, D.D., Lee, J., Maitas, O., Esa, N., Pidikiti, R., Carlucci, A., Harrington, J., Mick, E. & Chon, K.H. 2013, "A novel application for the detection of an irregular pulse using an iPhone 4S in patients with atrial fibrillation", *Heart rhythm : the official journal of the Heart Rhythm Society,* vol. 10, no. 3, pp. 315-319.

Moody, G.B. & Mark, R.R. 1983, "New method for detecting atrial fibrillation using R-R intervals ", Computers in Cardiol, pp. 227-230.

Nemati, S., Ghassemi, M.M., Ambai, V., Isakadze, N., Levantsevych, O., Shah, A. & Clifford, G.D. 2016, "Monitoring and detecting atrial fibrillation using wearable technology", *Conference proceedings : ...Annual International Conference of the IEEE Engineering in Medicine and Biology Society.IEEE Engineering in Medicine and Biology Society.Annual Conference,* vol. 2016, pp. 3394-3397.

Parak, J. & Korhonen, I. 2014, "Evaluation of wearable consumer heart rate monitors based on photopletysmography", *Conference proceedings : ...Annual International Conference of the IEEE Engineering in Medicine and Biology Society.IEEE Engineering in Medicine and Biology Society.Annual Conference,* vol. 2014, pp. 3670-3673.

Parak, J., Tarniceriu, A., Renevey, P., Bertschi, M., Delgado-Gonzalo, R. & Korhonen, I. 2015, "Evaluation of the beat-to-beat detection accuracy of PulseOn wearable optical heart rate monitor", *Conference proceedings : ...Annual International Conference of the IEEE Engineering in Medicine and Biology Society.IEEE Engineering in Medicine and Biology Society.Annual Conference,* vol. 2015, pp. 8099-8102.

Phukpattaranont, P. 2015, "QRS detection algorithm based on the quadratic filter", *Expert Systems with Applications,* vol. 42, no. 11, pp. 4867-4877.

Renevey, P., Sola, J., Theurillat, P., Bertschi, M., Krauss, J., Andries, D. & Sartori, C. 2013, "Validation of a wrist monitor for accurate estimation of RR intervals during sleep", *Conference proceedings : ...Annual International Conference of the IEEE Engineering in Medicine and Biology Society.IEEE Engineering in Medicine and Biology Society.Annual Conference,* vol. 2013, pp. 5493-5496.

Rietbrock, S. 2008, "Chronic atrial fibrillation: Incidence, prevalence, and prediction of stroke using the Congestive heart failure, Hypertension, Age >75, Diabetes mellitus, and prior Stroke or transient ischemic attack (CHADS2) risk stratification scheme", *Am Heart J,* vol. 156, no. 1, pp. 57-64.

Rob A Vermond, Bastiaan Geelhoed, Niek Verweij, Robert G Tieleman, Pim van der Harst, Hans L Hillege, Wiek H van Gilst, Isabelle C van Gelder & Michiel Rienstra 2015, "Incidence of Atrial Fibrillation and Relationship With Cardiovascular

Events, Heart Failure, and Mortality", *Journal of the American College of Cardiology,* vol. 66, no. 9, pp. 1000.

Sahoo, S.K., Lu, W., Teddy, S.D., Kim, D. & Feng, M. 2011, "Detection of Atrial fibrillation from non-episodic ECG data: A review of methods", United States, pp. 4992.

Savelieva, I. & Camm, A.J. 2000, "Clinical Relevance of Silent Atrial Fibrillation: Prevalence, Prognosis, Quality of Life, and Management", *Journal of Interventional Cardiac Electrophysiology,* vol. 4, no. 2, pp. 369-382.

Shcherbina, A., Mattsson, C.M., Waggott, D., Salisbury, H., Christle, J.W., Hastie, T., Wheeler, M.T. & Ashley, E.A. 2017, "Accuracy in Wrist-Worn, Sensor-Based Measurements of Heart Rate and Energy Expenditure in a Diverse Cohort", *Journal of personalized medicine,* vol. 7, no. 2, pp. 10.3390/jpm7020003.

Shelley, K.H. 2007, "Photoplethysmography: Beyond the Calculation of Arterial Oxygen Saturation and Heart Rate", *Anesthesia & Analgesia,* vol. 105, no. 6) (OLine Supplement, pp. S36.

Stanton, T., Hawkins, N.M., Hogg, K.J., Goodfield, N.E.R., Petrie, M.C. & McMurray, J.J.V. 2008, "How should we optimize cardiac resynchronization therapy?", *European heart journal,* vol. 29, no. 20, pp. 2458-2472.

Stoyan Tanev 2012, "Ventricular Beat Detection and Classification in Long Term ECG Recordings", *International Journal Bioautomation,* vol. 16, no. 4, pp. 273-290.

Terbizan, D.J., Dolezal, B.A. & Albano, C. 2002, "Validity of Seven Commercially Available Heart Rate Monitors", *Measurement in Physical Education and Exercise Science,* vol. 6, no. 4, pp. 243-247.

Wang, R., Blackburn, G., Desai, M., Phelan, D., Gillinov, L., Houghtaling, P. & Gillinov, M. 2017, "Accuracy of Wrist-Worn Heart Rate Monitors", *JAMA Cardiology,* vol. 2, no. 1, pp. 104-106.

Weippert, M., Kumar, M., Kreuzfeld, S., Arndt, D., Rieger, A. & Stoll, R. 2010, "Comparison of three mobile devices for measuring R–R intervals and heart rate variability: Polar S810i, Suunto t6 and an ambulatory ECG system", *European Journal of Applied Physiology,* vol. 109, no. 4, pp. 779-786.

Zhang, Q.X., Zhou, D. & Zeng, X. 2017, "Highly wearable cuff-less blood pressure and heart rate monitoring with single-arm electrocardiogram and photoplethysmogram signals", *Biomedical engineering online,* vol. 16, no. 1.

Zhou, C., Feng, J., Hu, J. & Ye, X. 2016, "Study of Artifact-Resistive Technology Based on a Novel Dual Photoplethysmography Method for Wearable Pulse Rate Monitors", *Journal of medical systems,* vol. 40, no. 3, pp. 1-10.