A Contact-Free, Ballistocardiography-Based Monitoring System (Emfit QS) for Measuring Nocturnal Heart Rate and Heart Rate Variability: Validation Study

Ville Vesterinen*, PhD; Niina Rinkinen*, MSc; Ari Nummela*, PhD
KIHU - Research Institute for Olympic Sports, Jyväskylä, Finland
*all authors contributed equally

Corresponding Author:
Ville Vesterinen, PhD
KIHU - Research Institute for Olympic Sports
Rautpohjankatu 6
Jyväskylä, 40700
Finland
Phone: 358 505451049
Email: ville.vesterinen@kihu.fi

Abstract

Background: Heart rate (HR) and heart rate variability (HRV) measurements are widely used to monitor stress and recovery status in sedentary people and athletes. However, effective HRV monitoring should occur on a daily basis because sparse measurements do not allow for a complete view of the stress-recovery balance. Morning electrocardiography (ECG) measurements with HR straps are time-consuming and arduous to perform every day, and thus compliance with regular measurements is poor. Contact-free, ballistocardiography (BCG)-based Emfit QS is effortless for daily monitoring. However, to the best of our knowledge, there is no study on the accuracy of nocturnal HR and HRV measured via BCG under real-life conditions.

Objective: The aim of this study was to evaluate the accuracy of Emfit QS in measuring nocturnal HR and HRV.

Methods: Healthy participants (n=20) completed nocturnal HR and HRV recordings at home using Emfit QS and an ECG-based reference device (Firstbeat BG2) during sleep. Emfit QS measures BCG by a ferroelectret sensor installed under a bed mattress. HR and the root mean square of successive differences between RR intervals (RMSSD) were determined for 3-minute epochs and the sleep period mean.

Results: A trivial mean bias was observed in the mean HR (mean –0.8 bpm [beats per minute], SD 2.3 bpm, \( P=.15 \)) and Ln (natural logarithm) RMSSD (mean –0.05 ms, SD 0.25 ms, \( P=.33 \)) between Emfit QS and ECG. In addition, very large correlations were found in the mean values of HR (\( r=0.90, \ P<.001 \)) and Ln RMSSD (\( r=0.89, \ P<.001 \)) between the devices. A greater amount of erroneous or missing data (\( P<.001 \)) was observed in the Emfit QS measurements (28.3%, SD 14.4%) compared with the reference device (1.1%, SD 2.3%). The results showed that 5.0% of the mean HR and Ln RMSSD values were outside the limits of agreement.

Conclusions: Based on the present results, Emfit QS provides nocturnal HR and HRV data with an acceptable, small mean bias when calculating the mean of the sleep period. Thus, Emfit QS has the potential to be used for the long-term monitoring of nocturnal HR and HRV. However, further research is needed to assess reliability in HR and HRV detection.


KEYWORDS
wearable technology; cardiac autonomic regulation; monitoring; validity

Introduction

Technological development has brought forth numerous apps, gadgets, and high-tech solutions designed to enhance health, fitness, and performance. Based on a survey by the American College of Sports Medicine, wearable technology is the most popular fitness trend of 2019 [1]. Data from everyday life can be easily recorded by wearable solutions. Because of the growing use of wearable technologies, it has become important to determine their validity. However, more than half of the devices currently used to monitor and improve personal health
and sports performance have not been validated through independent research [2].

Over the past several decades, heart rate (HR) has been used to monitor physiological stress and workload during exercise [3]. Recently, heart rate variability (HRV) measurements have been growing in popularity. HRV reflects the activity of cardiac autonomic regulation and can be used in the monitoring of stress and recovery status [4-6]. Traditionally, HR has been determined through the process of electrocardiography (ECG) by measuring the electrical activity of the heart. A 12-lead ECG, with electrodes attached to the body surface, is widely used in medical examinations. Traditional HR monitors allow for real-time measurement, using a chest strap with wireless ECG sensors [3]. Nowadays, HR can also be measured optically by photoplethysmography (PPG) with wearables, such as watches and mobile phones [7,8]. In addition, HR can be determined by ballistocardiography (BCG), which measures ballistic forces on the heart arising from the sudden ejection of blood into the great vessels with each heart beat [9]. It is a long-established, noninvasive technique that uses several types of sensors like pressure sensors, film-type force sensors, microbend fiber optic BCG sensors, electromechanical film transducer sensors, piezoelectric film sensors, polyvinylidene fluoride sensors, strain gauges, and pneumatic and hydraulic sensors [10]. The novel sensor technologies may detect HR and HRV more accurately. In addition, several different algorithms are being used to detect BCG peaks, each with differing detecting abilities [11]. However, previous studies have been conducted under laboratory conditions. Thus, it is necessary to evaluate the validity of BCG measurements under real-life conditions.

Emfit QS is a BCG-based commercial device for monitoring sleep and recovery. An EMFi sensor (6 cm x 55 cm in size) installed under a bed mattress can detect HR, HRV, breathing, and other body movements (Figure 1 [12]). Emfit QS measurement starts automatically shortly after the user goes to bed and stops recording once they have left the bed in the morning. Data are transferred via a Wi-Fi or 3G network to the internet, and the results can be accessed from a smartphone, tablet, or computer shortly after awakening. Thus, it is a contact-free, effortless, and user-friendly method with the capacity to improve user compliance on a daily basis. However, to the best of our knowledge, there is no previously published data on measuring BCG-based nocturnal HR and HRV under real-life conditions. Thus, the aim of this study was to evaluate the accuracy of Emfit QS in measuring HR and HRV during sleep, alongside an ECG-based device as a reference.

**Figure 1.** Correct placement of the Emfit QS sensor beneath the mattress or mattress topper under the chest area. Source: Emfit Ltd [12].

### Methods

#### Participants

A total of 20 participants were recruited to the study. Women (n=11; mean age 34 yrs, SD 7 yrs; mean height 1.69 m, SD 0.05 m; mean weight 67 kg, SD 10 kg) and men (n=9; mean age 42 yrs, SD 8 yrs; mean height 1.80 m, SD 0.06 m; mean weight 78 kg, SD 3 kg) were healthy (eg, no disease) and nonsmokers, and did not take medication on a regular basis. The participants were fully informed about the study design and the use of measurement devices before signing an informed consent document. The study complies with the standards set by the Ethics Committee of the University of Jyväskylä, Finland.

#### Data Collection

Nocturnal recordings were taken at home during sleep. Recordings began shortly after participants went to bed and
stopped once they left their bed in the morning. Before the first recording, Emfit QS’s own proprietary cellular ferroelectret sensor was placed beneath the mattress or mattress topper under chest area (Figure 1). The reference RR interval (RRI) data were recorded with Firstbeat Bodyguard 2 (BG2), an ECG-based recorder with two disposable electrodes and a sampling frequency of 1000 Hz. BG2 and its electrodes were set up on a participant’s body according to instructions in the user manual. The accuracy of BG2 was previously evaluated in laboratory protocol studies by Parak et al [13] and Bogdány et al [14], and in our unpublished study, which showed perfect agreement in the detection of RRRs (r=1.00) during a 30-minute rest period with Custo Cardio 100BT, a 12-channel ECG device (Custo med GmbH). The participants recorded the data over 3 consecutive nights, and the last recording was used in the analysis. Before measurements were taken, time synchronization was performed on the devices.

Data Analysis

Emfit QS provides HR, the vagal-related HRV index, and the root mean square of successive differences (RMSSD) in RRs throughout the sleep period in 3-minute epochs. If heart beat detection is disturbed due to a poor signal or artifacts any time during this period, data are not collected. The RRI data of BG2 were analyzed using the Firstbeat Sports software (Firstbeat Technologies Ltd). RRs were checked by an artifact detection filter of the Firstbeat Sports software [15] and subsequently excluded all falsely detected, missed, and premature heart beats caused by movement artifacts or any other artifacts of unknown origin. HR and RMSSD values were calculated for each 3-minute epoch throughout the measurement. Averages of the whole night period, which were used in the analysis, were calculated from those 3-minute values. Emfit QS and BG2 data were synchronized according to the time stamp. In addition to HR and RMSSD data, the amount of missing data was calculated.

Statistical Analysis

Values are expressed as mean (SD). Averages of the whole night period were calculated from the 3-minute epochs. The Gaussian distribution of the data was assessed with the Shapiro-Wilk goodness-of-fit test. Ln (natural logarithm) transformation was applied to the RMSSD data in order to meet the assumptions of the parametric statistical analysis. The accuracy of HR and RMSSD measured by Emfit QS was evaluated by determining the amount of missing data, the mean bias (absolute and percentage), and the root mean square error (RMSE) compared with the reference (BG2). Statistical difference between the measurements of BG2 and Emfit QS was analyzed using a paired Student t test. In addition, the magnitude of the differences was expressed as effect size (ES). The difference was considered trivial when ES ≤ 0.2, small when ES ≤ 0.6, moderate when ES ≤ 1.2, large when ES ≤ 2.0, and very large when ES > 2.0. A Pearson product-moment correlation and a Bland-Altman plot were used to analyze agreement between the reference and Emfit QS data. A Spearman rank correlation coefficient was calculated to investigate the correlation between the absolute differences (reference minus Emfit QS) and the average of the devices for HR and Ln RMSSD. In addition to the measures of statistical significance, the following criteria were adopted to interpret the magnitude of the correlation between measurement variables: < 0.10 (trivial), 0.11-0.30 (small), 0.31-0.50 (moderate), 0.51-0.70 (large), 0.71-0.90 (very large), and 0.91-1.0 (almost perfect) [16]. Significance was accepted as P < 0.05. Data were analyzed using SPSS Statistics 25 software (IBM Corp).

Results

No significant differences were found in HR (mean difference –1.7%, SD 4.6%, P =.15) and Ln RMSSD (mean difference 2.0%, SD 6.5%, P =.33) between the measurements by ECG (BG2) and BCG (Emfit QS) (Figure 2 and Table 1). A greater amount of erroneous or missing data (P <.001) was found in the 3-minute values of Ln RMSSD by Emfit QS (mean 28.3%, SD 14.4%) compared with the reference device (mean 1.1%, SD 2.3%). Very large correlations were found in HR and Ln RMSSD between the devices (Figure 3). Bland-Altman plots detailing the differences in the mean HR and Ln RMSSD between the reference and Emfit QS are shown in Figure 4. The Spearman rank correlation coefficient between the absolute differences and the average of the devices was r = 0.39 (P = .09) for HR and r = 0.53 (P = .02) for Ln RMSSD.

No differences were found in the 3-minute averaged HR values between the recordings of the reference device and Emfit QS (Figure 5). Significant differences in Ln RMSSD were found between the devices at 4 time points: 15 minutes (P =.01), 57 minutes (P =.05), 171 minutes (P =.04), and 450 minutes (P =.03). A very large correlation (r = 0.72, P < .001) was found in the 3-minute averaged HR, and a large correlation (r = 0.58, P < .001) was found in Ln RMSSD between the reference device and Emfit QS.
Figure 2. (A) Mean nocturnal heart rate (HR) and (B) the natural logarithm of the root mean square of successive differences between RR intervals (Ln RMSSD), measured by electrocardiogram (the reference) and ballistocardiogram (Emfit QS) (n=20). Box plots represent median values (solid line), 50th percentile values (box outline), and minimal and maximal values (whiskers).

Table 1. Comparison between the mean (SD) values of Emfit QS and the reference device.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Bias&lt;sup&gt;a&lt;/sup&gt; (absolute), mean (SD)</th>
<th>Bias&lt;sup&gt;b&lt;/sup&gt; (%), mean (SD)</th>
<th>RMSE&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>HR&lt;sup&gt;c&lt;/sup&gt; (bpm)</td>
<td>-0.8 (2.3)</td>
<td>-1.7 (4.6)</td>
<td>2.4</td>
<td>0.16 (trivial)</td>
</tr>
<tr>
<td>Ln RMSSD&lt;sup&gt;d&lt;/sup&gt; (ms)</td>
<td>-0.05 (0.25)</td>
<td>-2.0 (6.5)</td>
<td>0.25</td>
<td>0.14 (trivial)</td>
</tr>
</tbody>
</table>

<sup>a</sup>Bias: the difference between the reference and Emfit QS.

<sup>b</sup>RMSE: the root mean square error.

<sup>c</sup>HR: heart rate.

<sup>d</sup>Ln RMSSD: natural logarithm of the root mean square of successive differences.

Figure 3. Correlation of the mean (A) heart rate (HR) and (B) the natural logarithm of the root mean square of successive differences (Ln RMSSD) between RR intervals between electrocardiogram (the reference) and ballistocardiogram (Emfit QS). The dashed line represents the line of equivalence (r=1.0).
Figure 4. A Bland-Altman plot comparing the reference mean to that of Emfit QS: (A) heart rate (HR) and (B) the natural logarithm of the root mean square of successive differences (Ln RMSSD) between RR intervals.

Figure 5. Nocturnal (A) heart rate (HR) (n=17-20) and (B) the natural logarithm of the root mean square of successive differences (Ln RMSSD) between RR intervals (n=9-19) in 3-minute epochs during sleep between recordings of electrocardiogram (the reference) and ballistocardiogram (Emfit QS). Significant difference between the devices noted (P<.05).
Discussion

Principal Results
To the best of our knowledge, this study is the first to evaluate the accuracy of BCG-based nocturnal HR and HRV measurements under real-life conditions. The main finding of this study is that BCG-based Emfit QS showed a trivial mean bias in nocturnal mean HR (1.7%, ES=0.16) and RMSSD (2.0%, ES=0.14) compared with the ECG-based reference device. In addition, the correlation coefficient showed good agreement (r=0.89-0.90) between the devices. Terbizan et al [17] suggested a minimum correlation of 0.9 for heart monitors to be clinically reliable. However, it needs to be acknowledged that a high correlation coefficient does not solely represent good agreement in all cases [18].

Comparison With Prior Work
Choe and Cho [19] found an RMSE of 1.8 bpm (beats per minute) for HR measured over a 15-minute rest period by a piezoelectric sensor (BCG) under laboratory conditions. The error was slightly smaller compared with that in our study (2.4 bpm). In addition, Xie et al [20] reported a 0.90 bpm mean absolute error in HR detection by a BCG sensor placed under a chair during 15-minute recordings and a greater correlation (r=0.98) compared with the present study. In previous studies, the participants were instructed to avoid movement since muscular activity may cause measurement errors in BCG. In this study, the measurements were carried out at home under real-life conditions. During sleep, movement can have interfering effects on HR and HRV detection, which can explain the slightly higher error and the amount of missing data in the present study.

For valid HRV determination, consecutive heart beats must be detected with a high degree of accuracy. Shin et al [21] observed a 5% relative error and a strong correlation (r=0.97) in a time-domain analysis of HRV by BCG and ECG. Wang et al [22] observed perfect correlations (r=0.99-1.00) in HRV variables between BCG and ECG. The authors concluded that HRV can be measured reliably with BCG. In this study, the mean bias in RMSSD was smaller (2.0%) than that in the study by Shin et al [21], but the correlation was weaker (r=0.89).

The Bland-Altman plots showed that 5% (1/20) of the mean HR and Ln RMSSD values were outside the limits of agreement (LoA). Bland and Altman [23] recommended that 95% of the data points should lie within the mean difference (SD 1.96). Our results showed a proportional error in the mean Ln RMSSD determined by Emfit QS. Based on the Bland-Altman plot, a larger error can be found in smaller and larger Ln RMSSD values; it seems that Emfit QS underestimates Ln RMSSD at high HRV levels and overestimates Ln RMSSD at low HRV levels. The LoA for the mean HR is relatively narrow (SD 4.6 bpm, ~8%). It is important that LoA would be narrower than daily changes in long-term monitoring. Al Haddad et al [24] reported a ~12% day-to-day variation in resting Ln RMSSD. In the present study, the LoA for the mean Ln RMSSD was ~12%. Thus, it can be concluded that it is not greater than the day-to-day variation in Ln RMSSD; the LoA is barely acceptable. Unfortunately, this research did not answer whether the biases are stable in repetitive measurements. Reliability is crucial for long-term monitoring of stress and recovery states; if the bias varies within an individual in repetitive measurements, the true changes in daily HR and HRV cannot be detected, and thus, an accurate interpretation of cardiac autonomic regulation will be compromised. In addition, it is good to standardize sleeping and measurement conditions (eg, bed, mattress) in long-term monitoring for reducing biases.

The 3-minute averaged data showed no differences in HR at any time points during the night and few differences in Ln RMSSD (Figure 5). The differences can be explained by the greater amount of missing data of Emfit QS; it is possible that a 3-minute epoch can include only, for example, 15 seconds of Emfit QS data; yet, the system provides a value for the 3-minute epoch. Comparatively, the reference data could include the entire 3-minute data, which could explain the differences in the 3-minute values. Thus, it seems that HR and HRV values for the entire night are more adequate. The overall view of nocturnal HRV is also more important than a single 3-minute value because HRV fluctuates during sleep according to sleep phases [25,26].

Although the bias and LoA analyses appear promising, future development in the accuracy of HRV measurement is needed to decrease the amount of missing data and incorrect 3-minute values. Because of the relatively large amount of missing data in HRV by Emfit QS compared to the reference data (28% vs 1%), the HRV data do not reflect values of the whole night for every individual. Suliman et al [11] observed large differences (77%-95%) in the ability to detect BCG peaks between different algorithms during a short resting period. Thus, the development of these algorithms may decrease missing data. It could be speculated that individuals with a high amount of missing data could also have more differences between the reference and Emfit QS, but our results showed no correlation between the amount of missing data and the mean bias of the measurements. Furthermore, future studies should clarify reliability over time for measuring HR and HRV by Emfit QS.

Practical Applications
Morning ECG measurements with HR straps are time-consuming and arduous to perform every day, and thus, compliance with regular measurements is poor [27]. Nocturnal HRV is a time-efficient method compared to morning HRV measurements taken at waking [28]. Previous studies have mainly focused on BCG measurements under laboratory conditions. This study showed the potential of Emfit QS to serve as a tool for everyday use at home to measure nocturnal HRV. Contact-free and fully automatic analysis by Emfit QS facilitates effortless daily monitoring of stress and recovery status among athletes. Furthermore, it does not require attaching electrodes on the body surface and thus does not disturb participants’ typical sleep behaviors, which is advantageous over ECG. Based on our results, Emfit QS provides HR and HRV data with an acceptable, small mean bias compared with ECG. However, due to some large errors in detecting HR and HRV in some individuals, it would be best practice to ensure accuracy by comparing initial results from Emfit QS with ECG.
Conclusions
This study evaluated the accuracy of BCG-based Emfit QS for measuring HR and HRV. Our results showed that Emfit QS provides HR and HRV data with an acceptable, small mean bias compared with ECG. Thus, Emfit QS can be a potential tool for the long-term monitoring of HR and HRV. However, further research is needed to evaluate the reliability of HR and HRV detected by Emfit QS.

Acknowledgments
Funding for this study was provided by Emfit Ltd. The study was conducted by KIHU – Research Institute for Olympic Sports.

Conflicts of Interest
None declared.

References


Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCG</td>
<td>ballistocardiography</td>
</tr>
<tr>
<td>BG2</td>
<td>Bodyguard 2</td>
</tr>
<tr>
<td>bpm</td>
<td>beats per minute</td>
</tr>
<tr>
<td>ECG</td>
<td>electrocardiography</td>
</tr>
<tr>
<td>ES</td>
<td>effect size</td>
</tr>
<tr>
<td>HR</td>
<td>heart rate</td>
</tr>
<tr>
<td>HRV</td>
<td>heart rate variability</td>
</tr>
<tr>
<td>Ln</td>
<td>natural logarithm</td>
</tr>
<tr>
<td>LoA</td>
<td>limits of agreement</td>
</tr>
<tr>
<td>PPG</td>
<td>photoplethysmography</td>
</tr>
<tr>
<td>RMSE</td>
<td>root mean square error</td>
</tr>
<tr>
<td>RMSSD</td>
<td>root mean square of successive differences</td>
</tr>
<tr>
<td>RRI</td>
<td>RR interval</td>
</tr>
</tbody>
</table>

©Ville Vesterinen, Niina Rinkinen, Ari Nummela. Originally published in JMIR Biomedical Engineering (http://biomedeng.jmir.org), 23.04.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Biomedical Engineering, is properly cited. The complete bibliographic information, a link to the original publication on http://biomedeng.jmir.org/, as well as this copyright and license information must be included.