

Article

Wrist-Based Photoplethysmography Assessment of Heart Rate and Heart Rate Variability: Validation of WHOOP

Clint R. Bellenger ^{1,2,*}, Dean Miller ³, Shona L. Halson ⁴, Greg Roach ³ and Charli Sargent ³¹ Alliance for Research in Exercise, Nutrition and Activity (ARENA), Allied Health and Human Performance, University of South Australia, Adelaide 5001, Australia² South Australian Sports Institute, Adelaide 5001, Australia³ The Appleton Institute for Behavioural Science, Central Queensland University, Adelaide 5034, Australia; d.j.miller@cqu.edu.au (D.M.); greg.roach@cqu.edu.au (G.R.); charli.sargent@cqu.edu.au (C.S.)⁴ School of Behavioural and Health Sciences, Australian Catholic University, Brisbane 4014, Australia; Shona.Halson@acu.edu.au

* Correspondence: clint.bellenger@unisa.edu.au; Tel.: +618-8302-2060

Abstract: Heart rate (HR) and HR variability (HRV) infer readiness to perform exercise in athletic populations. Technological advancements have facilitated HR and HRV quantification via photoplethysmography (PPG). This study evaluated the validity of WHOOP's PPG-derived HR and HRV against electrocardiogram-derived (ECG) measures. HR and HRV were assessed via WHOOP and ECG over 15 opportunities. WHOOP-derived pulse-to-pulse (PP) intervals were edited with WHOOP's proprietary filter, in addition to various filter strengths via Kubios HRV software. HR and HRV (Ln RMSSD) were quantified for each filter strength. Agreement was assessed via bias and limits of agreement (LOA), and contextualised using smallest worthwhile change (SWC) and coefficient of variation (CV). Regardless of filter strength, bias ($\leq 0.39 \pm 0.38\%$) and LOA ($\leq 1.56\%$) in HR were lower than the CV (10–11%) and SWC (5–5.5%) for this parameter. For Ln RMSSD, bias ($1.66 \pm 1.80\%$) and LOA ($\pm 5.93\%$) were lowest for a 200 ms filter and WHOOP's proprietary filter, which approached or exceeded the CV (3–13%) and SWC (1.5–6.5%) for this parameter. Acceptable agreement was found between WHOOP- and ECG-derived HR. Bias and LOA in Ln RMSSD approached or exceeded the SWC/CV for this variable and should be interpreted against its own level of bias precision.

Citation: Bellenger, C.R.; Miller, D.; Halson, S.L.; Roach, G.; Sargent, C. Wrist-Based Photoplethysmography Assessment of Heart Rate and Heart Rate Variability: Validation of WHOOP. *Sensors* **2021**, *21*, 3571. <https://doi.org/10.3390/s21103571>

Academic Editor: Marco Altini

Keywords: autonomic nervous system; agreement; electrocardiogram

Received: 9 April 2021

Accepted: 14 May 2021

Published: 20 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Resting heart rate (HR) quantification and monitoring have been common in exercise physiology research and practice for centuries [1]. Additionally, HR variability (HRV), a sophisticated derivative of HR, has been quantified to provide insight into cardiac modulation by the parasympathetic and sympathetic divisions of the autonomic nervous system (ANS) [2]. Given the integral role of the ANS to all physiological function, including those related to exercise and training [3], the body's ability to tolerate or adapt to an exercise stimulus may be inferred by examining ANS responsiveness [3]. Consequently, HRV has been used to infer training tolerance or readiness to perform exercise in athletes [4,5].

Advancements in HR monitor technology, namely the first wireless HR monitor [6], have facilitated frequent and accurate HR quantification. However, this technology is reliant on the wireless communication of the heart's electrical activity from an elastic electrode chest strap to relevant receivers, and such reliance on chest straps can be inconvenient and problematic. Regarding HR and HRV assessment for the day-to-day monitoring of readiness to perform specifically, compliance is challenged by wearing a chest strap

during daily recordings [7]. Consequently, HR and HRV assessments that occur without a chest strap, such as photoplethysmography (PPG), are advantageous.

PPG detects changes in pulsatile blood flow between the heart's systole and diastole via LED-emitted light at the wrist, fingertip or earlobe [8]. The LED light illuminates the skin, while a photodetector quantifies the intensity of the light reflected back from the skin [7]. Since blood volume is acutely increased following cardiac systole (which obstructs the LED light and reduces the intensity of the reflected light) and decreased during cardiac diastole (increasing the intensity of reflected light), the heart's rhythm can be detected via PPG [7].

Calculation of HR and HRV from pulse-to-pulse (PP) intervals quantified via PPG is not a novel concept, having been utilised in research as historically as 1938 [9]. However, its contemporary application has been facilitated by technological advancements in commercially available HR monitors. PPG validation studies demonstrate acceptable agreement in HR quantification at rest [8,10] and during sleep [11]. Similarly, PPG-derived resting HRV (and specifically the root mean square of successive beat to beat (BB) interval differences; RMSSD) demonstrates acceptable agreement assessed at the earlobe and fingertip at rest [7,12–17] and during sleep [18].

WHOOP is a commercially available unit that quantifies HR and HRV (in the form of RMSSD) via wrist-based PPG. Uniquely, however, WHOOP quantifies these measures during slow-wave sleep (SWS) [19], which it is able to determine with moderate accuracy [20]. Given that SWS is thought to be important for physiological recovery from exercise [21–23], and HRV is considered a marker of physiological recovery [4,5], HRV assessment during SWS may quantify the degree of physiological recovery facilitated by this sleep stage, and thus may be used in the day-to-day monitoring of training status by practitioners. Additionally, WHOOP subsequently utilises HR and HRV measures (along with sleep duration) in an algorithm to predict a "Recovery Score" out of 100% [19]. This Recovery Score may be used to individually guide training prescription as a measure of readiness to perform. Given the novelty of WHOOP for assessing wrist-based PPG-derived HR and HRV, in addition to its unique dependence on measuring SWS in its quantification of HR and HRV, this study aimed to evaluate the agreement between WHOOP-derived HR and HRV and gold-standard assessment via ECG during SWS episodes.

2. Materials and Methods

2.1. Participants

Six healthy, young adults (male: $n = 3$; female: $n = 3$; age: 22.9 ± 3.4 years) participated in this study. Participants were excluded if they reported any existing medical conditions or sleep disorders, or had a recent history of shift work and/or transmeridian travel. This study was approved by the Central Queensland University Human Research Ethics Committee.

2.2. Experimental Overview

Data collection occurred concurrently with a larger pre-existing sleep study which has not yet been published. Data were collected over three consecutive sleep opportunities at the Appleton Institute of Behavioural Science, Central Queensland University, which contains two co-located, purpose-built accommodation suites that are sound attenuated, free from external environmental cues and can simultaneously house a total of six participants with private bedrooms and bathrooms. Participants wore a WHOOP unit (CB Rank, Greater Boston, New England) on their non-dominant wrist during sleep opportunities at the end of day 1 (2300–0800), end of day 2 (0300–1200) and during day 4 (1430–2130). Agreement between WHOOP and ECG-derived HR and HRV was evaluated through four time- and SWS stage-matched analyses (Table 1). Firstly, agreement between time-matched WHOOP- and ECG-derived HR and HRV was assessed during the final WHOOP-derived SWS episode in line with the technology's ecological use, and across a

range of data editing filters to determine the extent to which WHOOP-derived PP intervals need to be edited for erroneous PP intervals, including WHOOP's proprietary filter. Secondly, to determine whether the accurate identification of the final SWS episode impacts upon the agreement between WHOOP- and ECG-derived HR and HRV (given that WHOOP has only moderate sensitivity to accurately identifying sleep stage [20]), WHOOP- and ECG-derived HR and HRV were also quantified during the final polysomnography (PSG)-derived SWS episode for comparison. Additionally, to provide insight into the impact of misrepresentation of true SWS periods by WHOOP on HR and HRV, two SWS stage-matched analyses were conducted. Consequently, the third analysis assessed agreement between ECG-derived HR/HRV (i.e., true HR/HRV) during PSG-derived SWS episodes (i.e., true SWS) and WHOOP-derived HR/HRV during WHOOP-derived SWS episodes. The fourth analysis assessed the agreement between WHOOP-derived HR/HRV during PSG-derived SWS episodes and WHOOP-derived HR/HRV during WHOOP-derived SWS episodes as a means of determining the impact of SWS misrepresentation on WHOOP-derived measures.

Table 1. Analyses of agreement overview.

| SWS period derived by... | HR/HRV derived by... | | |
|--------------------------|----------------------|-------|-----|
| | WHOOP | 1,3,4 | 1 |
| | PSG | 2,4 | 2,3 |

HR, heart rate; HRV, heart rate variability; PSG, polysomnography; SWS, slow-wave sleep; 1, time-matched analysis of WHOOP-derived HR/HRV vs. ECG-derived HR/HRV during WHOOP-derived SWS; 2, time-matched analysis of WHOOP-derived HR/HRV vs. ECG-derived HR/HRV during PSG-derived SWS; 3, sleep stage-matched analysis of WHOOP-derived HR/HRV during WHOOP-derived SWS vs. ECG-derived HR/HRV during PSG-derived SWS; 4, sleep stage-matched analysis of WHOOP-derived HR/HRV during WHOOP-derived SWS vs. WHOOP-derived HR/HRV during PSG-derived SWS.

2.3. Sleep Stage Identification

To acquire WHOOP strap sleep data, researchers manually entered the start and end times of each sleep opportunity into the WHOOP smart phone application. The manufacturer then provided data in 30 s epochs for wake, light sleep, SWS and rapid eye movement (REM) sleep. PSG data were recorded directly to data acquisition, storage and analysis systems (Grael, Compumedics; Victoria, Australia). Brain, eye and muscle activity were quantified from electrodes attached to the face and scalp of participants, including three electroencephalography electrodes (i.e., C4-M1, F4-M1, O2-M1), two electro-oculograms (i.e., left/right outer canthus) and a submental electromyogram. PSG records were manually scored (in 30 s epochs) by a registered and experienced polysomnographic technician in compliance with standard criteria [24]. Time in bed during each sleep opportunity was arranged into wake, non-rapid eye movement sleep (non-REM; stage 1 [S1], stage 2 [S2] and SWS) and rapid eye movement (REM) sleep. Cardiac activity was assessed via two ECG electrodes (left-positive and right-negative) recorded using the aforementioned Grael PSG system. The negative electrode was placed three centimetres below the right clavicle, positioned on the torso parallel to the right leg. The positive electrode was positioned on the left side of the torso parallel to the left hip and leg, between either the fifth, sixth, or seventh intercostal spaces on the lower left side of the rib cage.

2.4. Heart Rate and Heart Rate Variability Calculation

Using both the WHOOP- and PSG-derived sleep staging data for each sleep opportunity, the final five minutes of the final SWS episodes were identified. If the final sleep stage was less than five minutes in duration, the preceding stages were identified until a

five-minute stage was found. Subsequently, time-matched PP and RR intervals were extracted from WHOOP- and ECG-derived files, respectively, for each SWS episode and analysed using WHOOP's proprietary filter and HRV analysis software (Kubios HRV Analysis, version 2.0 beta 1, Biomedical Signals Analysis Group, University of Kuopio, Finland). To determine the degree of PP interval editing required to facilitate the best agreement between WHOOP- and ECG-derived measures, Kubios's default filters (i.e., "Low", "Medium", "Strong" and "Very Strong"—equivalent to 400, 300, 200 and 100 ms editing thresholds, respectively), in addition to no filter (i.e., "None") and WHOOP's proprietary filter, were separately applied to WHOOP-derived PP intervals and recorded for analysis. For HRV analysis, RMSSD and its natural logarithm transformation (i.e., Ln RMSSD) were recorded for analysis.

2.5. Statistical Analysis

Data were analysed using SPSS (IBM Corp. IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY) and presented as the mean \pm 95% confidence intervals. Agreement between WHOOP- and ECG-derived measures of HR and HRV was determined through absolute and percentage mean bias (WHOOP minus ECG), absolute and percentage limits of agreement (LOA) and intra-class correlation (ICC). ICCs were evaluated as: 0.0–0.1, trivial; 0.1–0.3, small; 0.3–0.5, moderate; 0.5–0.7, large; 0.7–0.9, very large; 0.9–1.0, nearly perfect [25].

For the parameters of HR, RMSSD and Ln RMSSD, separate two-way (filter strength \times SWS quantification method [i.e., WHOOP vs. PSG]) repeated measures ANOVAs determined statistically significant differences in ECG and WHOOP-derived values, and in bias between filter strengths for each SWS quantification method. As it was not possible to statistically compare LOA between filters and SWS quantification method, mean residuals were calculated as a measure of variability about the bias and compared via separate two-way (filter strength \times SWS quantification method) repeated measures ANOVAs for HR, RMSSD and Ln RMSSD. Individual residuals were calculated as the square root of the squared difference between the individual value and the mean value for both absolute and percent bias. Statistical significance was set at $p < 0.05$.

For values of HRV, the filter strength resulting in the smallest bias and smallest LOA, in addition to WHOOP's proprietary filter, was subsequently used to compare differences in percent bias and percent residuals between analytical method (i.e., RMSSD vs. Ln RMSSD) via two-way ANOVA.

Effect sizes ([ES] with 95% confidence intervals) between variables of interest were calculated using pooled standard deviation. Threshold values for ES were ≤ 0.2 (trivial), > 0.2 (small), > 0.6 (moderate), > 1.2 (large), > 2.0 (very large), and > 4.0 (extremely large) [25].

3. Results

Of the 18 opportunities for data collection, HR and HRV data from three sleep opportunities were lost due to equipment malfunction and/or experimenter error. Thus, data from 15 sleep opportunities were available for comparison.

3.1. Filter Analysis

With regard to WHOOP's proprietary filter, trivial but statistically significant biases were found between WHOOP- and ECG-derived Ln RMSSD across both SWS quantification methods (percent bias $\leq 3.25 \pm 1.53\%$; bias as ES $\geq 0.18 \pm 0.16$; $p \leq 0.04$; LOA $\leq 6.59\%$; Figure 1a,b), and RMSSD during PSG-derived SWS (percent bias = $12.74 \pm 6.53\%$; bias as ES = 0.14 ± 0.09 ; $p = 0.005$; LOA = 25.31% ; Figure 1d), but not RMSSD during WHOOP-derived SWS (percent bias = $8.54 \pm 6.65\%$; bias as ES = 0.13 ± 0.15 ; $p = 0.10$; LOA = 25.75% ; Figure 1c).

Regarding the BB editing filters applied in Kubios, a Strong filter (equivalent to a 200 ms editing threshold) applied to WHOOP-derived PP intervals resulted in the smallest

percent bias and smallest percent LOA across both SWS quantification methods for RMSSD (percent bias $\leq 8.39 \pm 6.70\%$; bias as ES $\leq 0.08 \pm 0.15$; LOA $\leq 27.26\%$) and Ln RMSSD (percent bias $\leq 2.30 \pm 1.69\%$; bias as ES $\leq 0.14 \pm 0.17$; LOA $\leq 6.98\%$) (Supplementary Tables S1 and S2).

In comparison to Kubios's Strong filter, the bias between WHOOP and ECG values was greater when edited using WHOOP's proprietary filter during PSG-derived SWS across both RMSSD (ES = 0.51 ± 0.24 ; $p = 0.04$) and Ln RMSSD values (ES = 0.34 ± 0.14 ; $p = 0.03$). Percent LOAs were similar when WHOOP-derived PP intervals were edited using Kubios's Strong filter compared to WHOOP's proprietary filter across both SWS quantification methods for RMSSD and Ln RMSSD, and analysis of residuals indicated no statistical differences (ES $\leq 0.50 \pm 0.59$; $p \geq 0.12$).

With regard to HR, there was a trivial but statistically significant bias between WHOOP-derived HR edited by WHOOP's proprietary filter and ECG-derived HR during both WHOOP- and PSG-derived SWS (ES $\geq -0.03 \pm 0.02$; percent bias $\geq -0.39 \pm 0.38\%$; $p \leq 0.04$; Figure 1e,f). For all other filters, there were trivial and non-statistically significant biases between WHOOP- and ECG-derived HR (ES $\leq 0.006 \pm 0.023$; $p \geq 0.59$). LOA for all filters across both SWS quantification methods were $\leq 1.56\%$ and ICCs were almost perfect ($r = 1.00$ [95% confidence interval range 0.99 to 1.00]; Supplementary Table S3).

3.2. Analytical Method Analysis

For both Kubios's Strong filter and WHOOP's proprietary filter, percent bias for measures of Ln RMSSD were smaller than measures of RMSSD (ES $\leq -0.73 \pm 0.61$; $p \leq 0.03$) with the exception of Kubios's Strong filter during WHOOP-derived SWS which trended toward statistical significance (ES = -0.56 ± 0.60 ; $p = 0.09$). Similarly, percent residuals for measures of Ln RMSSD were smaller than measures of RMSSD (ES $\leq -1.23 \pm 0.60$; $p \leq 0.001$).

3.3. SWS Stage-Matched Analyses

For the analysis of agreement between ECG-derived HR/HRV during PSG-derived SWS and WHOOP-derived HR/HRV during WHOOP-derived SWS, there were trivial to small biases in HR (ES = 0.13 ± 0.17 ; $1.85 \pm 2.51\%$; $p = 0.15$; Figure 2a) and HRV (ES = 0.03 ± 0.29 ; $20.08 \pm 28.42\%$; $p = 0.86$ for RMSSD [Figure 2b] and ES = 0.21 ± 0.39 ; $3.74 \pm 5.76\%$; $p = 0.29$ for Ln RMSSD [Figure 2c]). There were moderate to large LOAs for HR (9.71%; ES = 0.67), RMSSD (110.07%; ES = 1.12) and Ln RMSSD (22.31%; ES = 1.50).

For the analysis of agreement between WHOOP-derived HR/HRV during PSG-derived SWS and WHOOP-derived HR/HRV during WHOOP-derived SWS, there were trivial biases in HR (ES = 0.11 ± 0.17 ; $1.57 \pm 2.45\%$; $p = 0.21$; Figure 3a) and HRV (ES = 0.16 ± 0.31 ; $5.75 \pm 23.16\%$; $p = 0.34$ for RMSSD [Figure 3b] and ES = 0.01 ± 0.38 ; $0.68 \pm 5.13\%$; $p = 0.94$ for Ln RMSSD [Figure 3c]). There were moderate to large LOAs for HR (9.48%; ES = 0.66), RMSSD (89.70%; ES = 1.21) and Ln RMSSD (19.85%; ES = 1.47).

Figure 4 depicts the time differential between WHOOP-derived SWS episodes and PSG-derived SWS episodes per sleep opportunity; there was a small bias (WHOOP minus PSG) of 35.1 ± 47.2 min (ES = 0.31 ± 0.42 ; $p = 0.17$) with large LOA of 182.7 min (ES = 1.62).

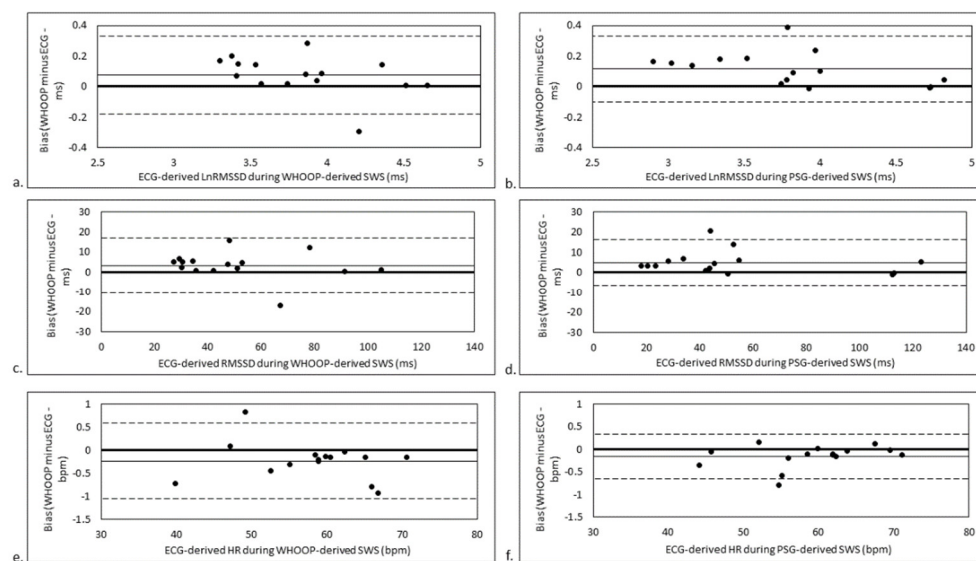


Figure 1. Agreement between (a) ECG-derived Ln RMSSD and WHOOP-derived Ln RMSSD during WHOOP-derived SWS, (b) ECG-derived Ln RMSSD and WHOOP-derived Ln RMSSD during PSG-derived SWS, (c) ECG-derived RMSSD and WHOOP-derived RMSSD during WHOOP-derived SWS, (d) ECG-derived RMSSD and WHOOP-derived RMSSD during PSG-derived SWS, (e) ECG-derived HR and WHOOP-derived HR during WHOOP-derived SWS and (f) ECG-derived HR and WHOOP-derived HR during PSG-derived SWS. Thin continuous line represents mean bias. Dashed lines represent mean bias \pm limits of agreement. bpm, beats per minute; ECG, electrocardiogram; HR, heart rate; Ln RMSSD, natural logarithm of the root mean square of successive BB interval differences; ms, milliseconds; PSG, polysomnography; RMSSD, root mean square of successive BB interval differences; SWS, slow-wave sleep.

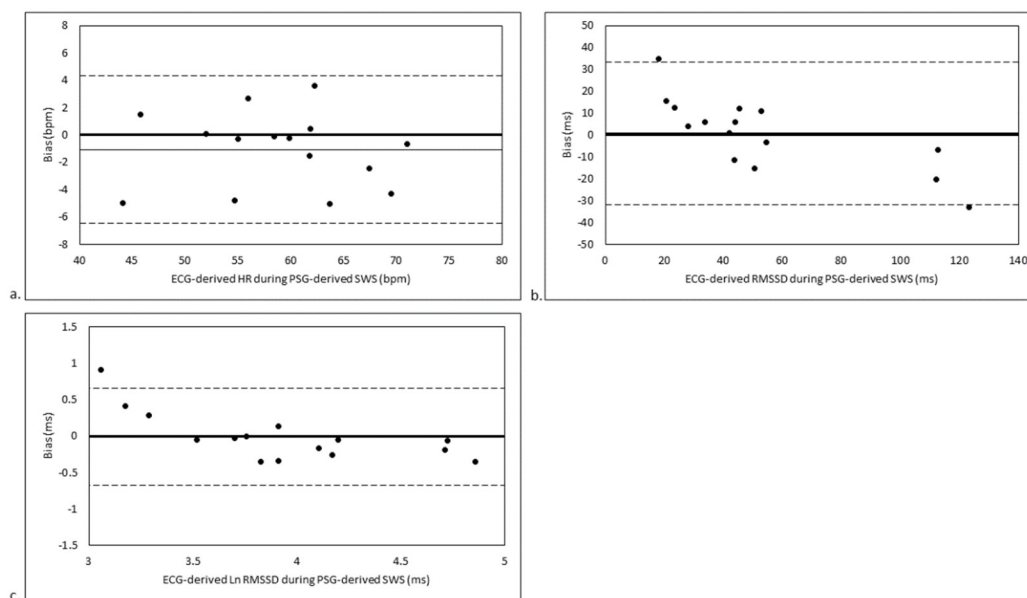


Figure 2. Agreement between (a) ECG-derived HR during PSG-derived SWS and WHOOP-derived HR during WHOOP-derived SWS, (b) ECG-derived RMSSD during PSG-derived SWS and WHOOP-derived RMSSD during WHOOP-derived SWS, and (c) ECG-derived Ln RMSSD during PSG-derived SWS and WHOOP-derived Ln RMSSD during WHOOP-derived SWS. Bias is calculated as WHOOP-derived variable during WHOOP-derived SWS minus ECG-derived variable during PSG-derived SWS. Thin continuous line represents mean bias. Dashed lines represent mean bias \pm limits of agreement. bpm, beats per minute; ECG, electrocardiogram; HR, heart rate; Ln RMSSD, natural logarithm of the root mean square of successive BB interval differences; ms, milliseconds; PSG, polysomnography; RMSSD, root mean square of successive BB interval differences; SWS, slow-wave sleep.

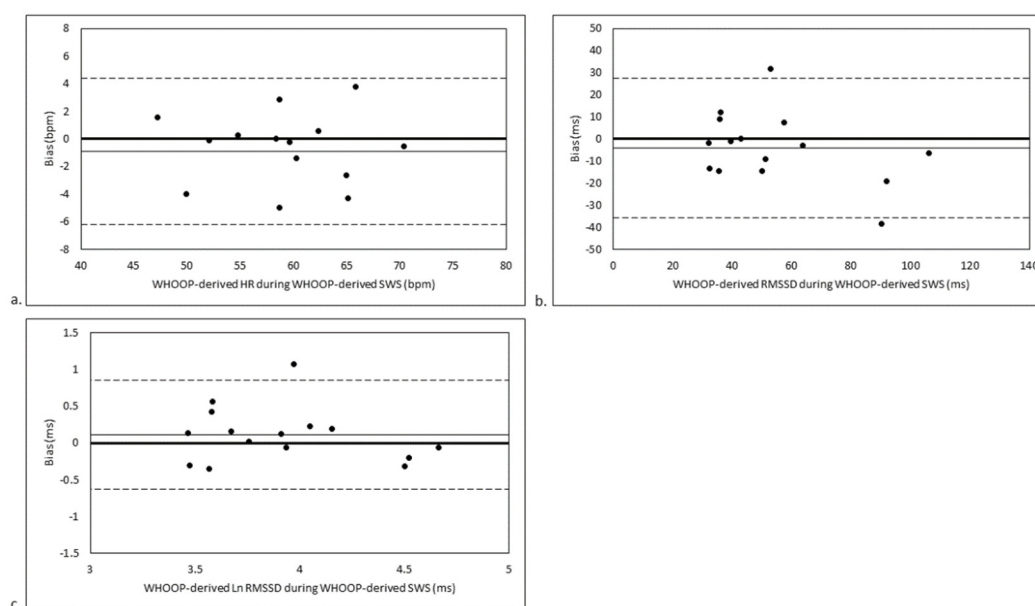


Figure 3. Agreement between (a) WHOOP-derived HR during PSG-derived SWS and WHOOP-derived HR during WHOOP-derived SWS, (b) WHOOP-derived RMSSD during PSG-derived SWS and WHOOP-derived RMSSD during WHOOP-derived SWS, and (c) WHOOP-derived Ln RMSSD during PSG-derived SWS and WHOOP-derived Ln RMSSD during WHOOP-derived SWS. Bias is calculated as WHOOP-derived variable during WHOOP-derived SWS minus WHOOP-derived variable during PSG-derived SWS. Thin continuous line represents mean bias. Dashed lines represent mean bias \pm limits of agreement. bpm, beats per minute; HR, heart rate; Ln RMSSD, natural logarithm of the root mean square of successive BB interval differences; ms, milliseconds; PSG, polysomnography; RMSSD, root mean square of successive BB interval differences; SWS, slow-wave sleep.

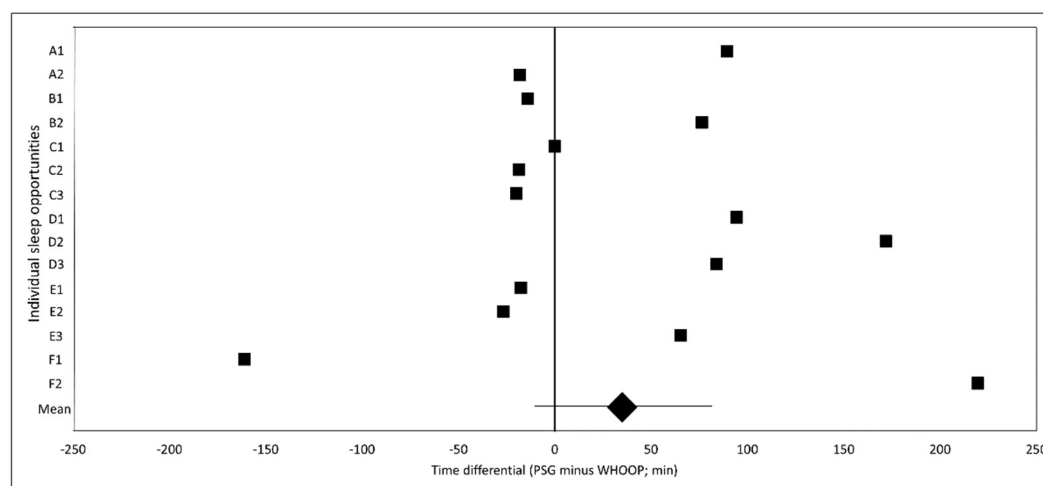


Figure 4. Time differential (min) for SWS periods derived by PSG and WHOOP. Data are presented as WHOOP time of day minus PSG time of day such that positive values indicate WHOOP-derived SWS periods occurred after PSG-derived SWS periods. Diamond marker represents mean bias \pm 95% confidence interval. Y-axis letters represent individual participants. Y-axis numerals represent individual participants' sleep opportunities. min, minutes; PSG, polysomnography; SWS, slow-wave sleep.

4. Discussion

This study evaluated agreement between PPG assessment of HR and HRV by a commercially available wrist-worn activity monitor (WHOOP) and gold-standard assessment via ECG. The primary findings were that of trivial bias ($ES \leq 0.03$) and LOA ($ES \leq 0.10$) for time-matched HR assessment, and trivial bias ($ES \leq 0.19$) and small LOA ($ES \leq 0.59$) for time-matched HRV assessment when either a Strong filter or WHOOP's proprietary filter was applied to WHOOP-derived PP interval data and analysed as Ln RMSSD. SWS stage-

matched WHOOP-derived HR and HRV demonstrated trivial bias ($ES \leq 0.11$) and moderate to large LOA ($ES = 0.66\text{--}1.47$).

The present study identified that WHOOP-derived PP intervals need to be filtered prior to HRV calculation to facilitate optimal agreement with ECG-derived HRV. The methodological consideration of BB interval editing is not a novel concept and has been advocated in HRV analysis [2,5]. Indeed, Buchheit [5] demonstrated that a single erroneous BB interval over a five-minute recording substantially altered RMSSD calculation. Consequently, it is important to edit BB intervals prior to analysis to ensure a true reflection of ANS status. While visual inspection and manual editing of BB intervals is ideal, it is an unrealistic practice in the field where a multitude of files are recorded, and instantaneous feedback is required to guide athletic training. Accordingly, automatic BB interval editing within manufacture software is commonplace [5]. In the present study, bias and LOA was minimised, while ICC was maximised, as filter strength increased to a Strong level (and also with WHOOP's proprietary filter). However, a Very Strong filter resulted in poorer agreement between WHOOP- and ECG-derived HRV, indicating this filter was too aggressive, and excessively altered the true BB interval patterning. Thus, the present study indicates that while erroneous WHOOP-derived PP intervals have little impact on HR calculation, these intervals have a small to moderate impact on RMSSD/Ln RMSSD calculation, as evidenced by the small biases ($ES = 0.21\text{--}0.32$) and small to moderate LOAs ($ES = 0.56\text{--}0.78$) in unfiltered RMSSD and Ln RMSSD across both SWS quantification methods, in comparison to the trivial biases ($ES = 0.04\text{--}0.19$) and small LOAs ($ES = 0.33\text{--}0.59$) in filtered values (i.e., Kubios's Strong filter and WHOOP's proprietary filter).

A further methodological consideration in HRV determination is the natural logarithm transformation of RMSSD (i.e., Ln RMSSD). Natural logarithm transformation reduces bias from non-uniformity of error [25], and has become standard practice for the longitudinal monitoring of training status via HRV [5]. In the present study, natural logarithm transformation of RMSSD resulted in small to moderate ($ES = 0.56\text{--}1.19$) reductions in percent bias, and large to very large ($ES = 1.23\text{--}2.15$) reductions in percent residuals in comparison to raw RMSSD.

While some statistically significant differences in bias were found between Kubios's Strong filter and WHOOP's proprietary filter, agreement statistics (i.e., bias and LOA) in WHOOP-derived HR and HRV may also be contextualised using the natural day-to-day variability in these variables. Some variation exists in the literature with regard to day-to-day variability in HR (10–11% coefficient of variation [5,26]) and Ln RMSSD (3–13% coefficient of variation [5,26–32]) which is likely dependent on timing of assessment (i.e., morning waking versus nocturnal) and posture (i.e., supine versus sitting versus standing). With specific regard to nocturnally collected Ln RMSSD, Costa et al. [32] demonstrated a coefficient of variation of 4–6%. Additionally, given that a smallest worthwhile change in HR/HRV has been proposed to be calculated as 0.5 multiplied by coefficient of variation [25], the smallest worthwhile change is 5–5.5% for HR and 1.5–6.5% for Ln RMSSD (and 2–3% for nocturnally derived Ln RMSSD). Consequently, since the bias ($<0.5\%$) and LOA (1–1.5%) in WHOOP-derived HR were less than both the smallest worthwhile change and coefficient of variation in HR, it may be concluded that WHOOP's proprietary filter provides suitable editing of PP intervals. However, since the bias (2–3.5%) and LOA (6–6.5%) in WHOOP-derived Ln RMSSD (when edited with WHOOP's proprietary filter) approaches the upper limit for the smallest worthwhile change in Ln RMSSD (and exceeds both the coefficient of variation and smallest worthwhile change for nocturnally derived Ln RMSSD), WHOOP-derived measures of Ln RMSSD may need to be interpreted against their own level of bias precision (i.e., LOA of 6–6.5%).

The physiological determinants of the trivial biases with small LOAs for agreement between ECG- and PPG-derived Ln RMSSD demonstrated in the present study potentially lie within the pulse travel time. Specifically, the electrical activity of the heart is followed by spread of the pulsatile wave of blood to the periphery [12]. While this pulse travel time demonstrates BB fluctuations of only a few milliseconds [33–35], this does indicate that

BB intervals derived from ECG and PPG will rarely be *exactly* the same, which intuitively indicates that variability in BB intervals (i.e., RMSSD/Ln RMSSD) will rarely be *exactly* the same also.

WHOOP quantify HR and HRV in the final SWS episode of each sleep opportunity, which is intuitive from a physiological perspective. SWS is thought to be important for physiological recovery from exercise [21], a hypothesis supported by the synchronisation between SWS periods and growth hormone release in humans (suggesting that sleep periods provide optimal anabolic conditions) and findings of SWS duration being proportional to preceding wakefulness [22]. Since HRV is a measure of ANS status [2], which is in turn considered a marker of physiological readiness to perform exercise [3], it is intuitive that HRV be assessed during SWS, since this may be used to quantify the magnitude of physiological recovery facilitated by SWS (although future research is required to confirm this). However, given that WHOOP has only moderate sensitivity in accurately identifying sleep stage [20], which is supported by the findings of the present study indicating a small time differential (ES = 0.31) in the determination of SWS periods by WHOOP in comparison to PSG (and a large [ES = 1.62] LOA about that time differential), the misrepresentation of true SWS periods by WHOOP may have an impact on WHOOP-derived HR and HRV since autonomic HR modulation is physiologically impacted upon by sleep stage [36–39]. Indeed, while the bias in stage-matched WHOOP-derived HR (ES = 0.11; 1.57%) and Ln RMSSD (ES = 0.01; 0.68%) during WHOOP-derived SWS episodes compared to PSG-derived SWS episodes was trivial, the LOA for WHOOP-derived HR (9.48%) exceeded the smallest worthwhile change (5–5.5%) and approached the coefficient of variation (10–11%) in this variable, while the LOA for WHOOP-derived Ln RMSSD (19.85%) exceeded both the smallest worthwhile change (1.5–6.5%) and coefficient of variation (3–13%) for this variable. Thus, while the exploratory analysis performed in the present study does not quantify the true day-to-day variability in WHOOP-derived HR and HRV, it does indicate that a misrepresentation of SWS periods impacts upon the precision of bias in WHOOP-derived HR and HRV, which may in turn impact upon the day-to-day variability in WHOOP-derived HR and HRV. Consequently, future research should evaluate the day-to-day variability in WHOOP-derived HR and HRV.

The agreement between PPG- and ECG-derived HR shown in this study (i.e., bias = 0.02–0.23 bpm [ES = 0.01–0.03]; LOA = 0.72–0.81 bpm; $r = 1.00$ across both SWS quantification methods) is similar to that previously shown in other wearable devices during quiet rest and sleep (bias = 0.09–1.00 bpm [ES = 0.01–0.09]; LOA = 2.67–5.29 bpm; $r = 0.99$) [8,10,11]. The agreement between wrist-based PPG- and ECG-derived HRV demonstrated in the present study (i.e., bias = 1.33–4.90 ms [ES = 0.04–0.14]; LOA = 11.26–15.96 ms; $r = 0.98$ – 0.99 for RMSSD) is within the range found using earlobe and fingertip PPG-derived HRV (bias = 0.00–2.71 ms [ES = 0.00–0.11]; LOA = 1.40–14.50 ms; $r = 0.99$ – 1.00) during quiet rest and sleep [7,12–18].

The present study used repeat measures in only a small number of unique individuals, which may be considered a limitation in certain scientific contexts. However, the authors propose this is not the case in the present study where simple statistical agreement is the focus. Specifically, to suitably “challenge” the WHOOP unit for accurately assessing PPG-derived HR and HRV, an appropriate range and variability in BB intervals is required. The current dataset provides HR in the range of 40–75 bpm and RMSSD in the range of 15–125 ms, and it is hypothesised that these ranges cover the typical range seen in both the general and athletic population.

While the WHOOP-derived HR and HRV variables assessed in the present study feed into WHOOP’s Recovery Score, validation of this Recovery Score itself was beyond the scope of this study, and thus should be validated in future research.

5. Conclusions

A wrist-worn, commercially available activity monitor (WHOOP) demonstrated acceptable agreement in HR via PPG assessment compared with gold-standard assessment

via ECG. Regarding HRV assessment however, bias and LOA in Ln RMSSD approached or exceeded the smallest worthwhile change/coefficient of variation for this variable, and thus should be interpreted against its own level of bias precision when suitably edited to remove erroneous PP intervals and analysed as Ln RMSSD. SWS stage-matched assessment of HR and HRV indicated that misrepresentation of SWS periods impacted upon the precision of bias in WHOOP-derived HR and HRV, which may in turn have an impact on the day-to-day variability in WHOOP-derived HR and HRV.

Supplementary Materials: The following are available online at www.mdpi.com/1424-8220/21/10/3571/s1, Table S1. Agreement between ECG- and WHOOP-derived Heart Rate (bpm), Table S2. Agreement between ECG- and WHOOP-derived RMSSD (ms), Table S3. Agreement between ECG- and WHOOP-derived Ln RMSSD (ms).

Author Contributions: Conceptualization, C.R.B. and S.L.H.; methodology, C.R.B., D.M., S.L.H., G.R. and C.S.; software, C.R.B. and D.M.; formal analysis, C.R.B.; writing—original draft preparation, C.R.B.; writing—review and editing, C.R.B., D.M., S.L.H., G.R. and C.S.; visualization, C.R.B.; project administration, C.R.B. and D.M.; funding acquisition, G.R. and C.S. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by the Australian Research Council under Grant (DP160104909) and the Australian Institute of Sport. The funders and suppliers had no input in the design of the study and interpretation of results. The results of the current study do not constitute endorsement of the product by the Australian Institute of Sport, authors or the journal.

Institutional Review Board Statement: This study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board (or Ethics Committee) of Central Queensland University (protocol approval number: H16/06-168; approved on August 7 2018).

Informed Consent Statement: Informed consent was obtained from all subjects involved in this study.

Data Availability Statement: The datasets generated from the current study are available from the corresponding author on reasonable request.

Conflicts of Interest: Dean Miller's position as a Research Officer at Central Queensland University is currently sponsored by WHOOP Inc.—the company that produces the wearable monitors that were used in this study. However, this sponsorship arrangement was initiated after the data were collected for this study. Nevertheless, this represents a potential conflict of interest for Dean Miller and for the co-authors who work in the same research group as Dean Miller (i.e., Greg Roach and Charli Sargent). The funders of this study had no role in the design of this study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Achten, J.; Jeukendrup, A.E. Heart rate monitoring: Applications and limitations. *Sports Med.* **2003**, *33*, 517–538.
2. Malik, M.; Bigger, J.T.; Camm, A.J.; Kleiger, R.E.; Malliani, A.; Moss, A.J.; Schwartz, P.J. Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. *Eur. Heart J.* **1996**, *17*, 354–381, doi:10.1161/01.CIR.93.5.1043.
3. Borresen, J.; Lambert, M.I. Autonomic control of heart rate during and after exercise: Measurements and implications for monitoring training status. *Sports Med.* **2008**, *38*, 633–646.
4. Bellenger, C.R.; Fuller, J.T.; Thomson, R.L.; Davison, K.; Robertson, E.Y.; Buckley, J.D. Monitoring athletic training status through autonomic heart rate regulation: A systematic review and meta-analysis. *Sports Med.* **2016**, *46*, 1461–1486.
5. Buchheit, M. Monitoring training status with HR measures: Do all roads lead to Rome? *Front Physiol.* **2014**, *5*, 1–19.
6. Laukkanen, R.; Virtanen, P.K. Heart rate monitors: State of the art. *J Sports Sci.* **1998**, *16*, S3–S7.
7. Plews, D.; Scott, B.; Altini, M.; Wood, M.; Kilding, A.; Laursen, P. Comparison of heart-rate-variability recording with smartphone photoplethysmography, Polar H7 chest strap, and electrocardiography. *Int. J. Sport Physiol. Perf.* **2017**, *12*, 1324–1328.
8. Støve, M.P.; Haucke, E.; Nymann, M.L.; Sigurdsson, T.; Larsen, B.T. Accuracy of the wearable activity tracker Garmin Forerunner 235 for the assessment of heart rate during rest and activity. *J. Sports Sci.* **2019**, *37*, 895–901.
9. Hertzman, A.B. The blood supply of various skin areas as estimated by the photoelectric plethysmograph. *Am. J. Physiol.* **1938**, *124*, 328–340.

10. Horton, F.J.; Stergiou, S.P.; Fung, S.T.; Katz, S.L. Comparison of Polar M600 optical heart rate and ECG heart rate during exercise. *Med. Sci. Sports Exerc.* **2017**, *49*, 2600–2607.
11. Haghayegh, S.; Khoshnevis, S.; Smolensky, M.H.; Diller, K.R. Accuracy of PurePulse photoplethysmography technology of Fitbit Charge 2 for assessment of heart rate during sleep. *Chronobiol. Int.* **2019**, *36*, 927–933.
12. Selvaraj, N.; Jaryal, A.; Santhosh, J.; Deepak, K.K.; Anand, S. Assessment of heart rate variability derived from finger-tip photoplethysmography as compared to electrocardiography. *J. Med. Eng. Technol.* **2008**, *32*, 479–484.
13. Heathers, J.A.J. Smartphone-enabled pulse rate variability: An alternative methodology for the collection of heart rate variability in psychophysiological research. *Int. J. Psychophysiol.* **2013**, *89*, 297–304.
14. Gil, E.; Orini, M.; Bailón, R.; Vergara, J.M.; Mainardi, L.; Laguna, P. Photoplethysmography pulse rate variability as a surrogate measurement of heart rate variability during non-stationary conditions. *Physiol. Meas.* **2010**, *31*, 1271–1290.
15. Pernice, R.; Javorka, M.; Krohova, J.; Czippelova, B.; Turianikova, Z.; Busacca, A.; Faes, L. Comparison of short-term heart rate variability indexes evaluated through electrocardiographic and continuous blood pressure monitoring. *Med. Biol. Eng. Comput.* **2019**, *57*, 1247–1263.
16. Charlot, K.; Cornolo, J.; Brugniaux, J.V.; Richalet, J.P.; Pichon, A. Interchangeability between heart rate and photoplethysmography variabilities during sympathetic stimulations. *Physiol. Meas.* **2009**, *30*, 1357–1369.
17. Vescio, B.; Salsone, M.; Gambardella, A.; Quattrone, A. Comparison between electrocardiographic and earlobe pulse photoplethysmographic detection for evaluating heart rate variability in healthy subjects in short- and long-term recordings. *Sensors* **2018**, *18*, 844.
18. Flatt, A.A.; Esco, M.R. Validity of the ithlete™ smart phone application for determining ultra-short-term heart rate variability. *J. Hum. Kinet.* **2013**, *39*, 85–92.
19. WHOOP. Experience WHOOP Recovery. Available online: <https://www.whoop.com/experience/> (accessed on 30 September 2019).
20. Miller, D.J.; Lastella, M.; Scanlan, A.T.; Bellenger, C.; Halson, S.L.; Roach, G.D.; Sargent, C. A validation of the WHOOP strap against polysomnography to assess sleep. *J. Sports Sci.* **2020**, doi10.1080/02640414.2020.1797448.
21. Halson, S. Sleep in elite athletes and nutritional interventions to enhance sleep. *Sports Med.* **2014**, *44* (Suppl. 1), 13–23.
22. Shapiro, C.M.; Bortz, R.; Mitchell, D.; Bartel, P.; Jooste, P. Slow-wave sleep: A recovery period after exercise. *Science* **1981**, *214*, 1253–1254.
23. Dijk, D.J. Slow-wave sleep deficiency and enhancement: Implications for insomnia and its management. *World J. Biol. Psychiatry* **2010**, *11*, 22–28.
24. Berry, R.B.; Brooks, R.; Gamaldo, C.; Harding, S.M.; Lloyd, R.M.; Quan, S.F.; Troester, M.T.; Vaughn, B.V. AASM Scoring Manual Updates for 2017 (Version 2.4). *J Clin Sleep Med.* **2017**, *13*, 665–666.
25. Hopkins, W.G.; Marshall, S.W.; Batterham, A.M.; Hanin, J. Progressive statistics for studies in sports medicine and exercise science. *Med. Sci. Sports Exerc.* **2009**, *41*, 3–12.
26. Al Haddad, H.; Laursen, P.; Chollet, D.; Ahmaidi, S.; Buchheit, M. Reliability of resting and postexercise heart rate measures. *Int. J. Sports Med.* **2011**, *32*, 598–605.
27. Flatt, A.A.; Howells, D. Effects of varying training load on heart rate variability and running performance among an Olympic rugby sevens team. *J. Sci. Med. Sport* **2019**, *22*, 222–226.
28. Plews, D.J.; Laursen, P.B.; Le Meur, Y.; Hausswirth, C.; Kilding, A.E.; Buchheit, M. Monitoring training with heart rate-variability: How much compliance is needed for valid assessment? *Int. J. Sport Physiol. Perf.* **2014**, *9*, 783.
29. Chen, Y.S.; Clemente, F.M.; Bezerra, P.; Lu, Y. Ultra-short-term and short-term heart rate variability recording during training camps and an international tournament in U-20 national futsal players. *Int. J. Environ. Res. Public Health* **2020**, *17*, 775.
30. Nakamura, F.Y.; Antunes, P.; Nunes, C.; Costa, J.A.; Esco, M.R.; Travassos, B. Heart rate variability changes from traditional vs. ultra-short-term recordings in relation to preseason training load and performance in Futsal players. *J. Strength Cond Res.* **2020**, *34*, 2974–2981.
31. Flatt, A.A.; Esco, M.R. Smartphone-derived heart-rate variability and training load in a women's soccer team. *Int. J. Sports Physiol. Perform.* **2015**, *10*, 994–1000.
32. Costa, J.A.; Brito, J.; Nakamura, F.Y.; Oliveira, E.M.; Rebelo, A.N. Effects of late-night training on “slow-wave sleep episode” and hour-by-hour-derived nocturnal cardiac autonomic activity in female soccer players. *Int. J. Sports Physiol. Perform.* **2018**, *13*, 638–644.
33. Drinnan, M.J.; Allen, J.; Murray, A. Relation between heart rate and pulse transit time during paced respiration. *Physiol. Meas.* **2001**, *22*, 425–432.
34. Johansson, A.; Ahlstrom, C.; Lanne, T.; Ask, P. Pulse wave transit time for monitoring respiration rate. *Med. Biol. Eng. Comput.* **2006**, *44*, 471–478.
35. Porta, A.; Gasperi, C.; Nollo, G.; Lucini, D.; Pizzinelli, P.; Antolini, R.; Pagani, M. Global versus local linear beat-to-beat analysis of the relationship between arterial pressure and pulse transit time during dynamic exercise. *Med. Biol. Eng. Comput.* **2006**, *44*, 331–337.
36. Somers, V.K.; Dyken, M.E.; Mark, A.L.; Abboud, F.M. Sympathetic-nerve activity during sleep in normal subjects. *N. Engl. J. Med.* **1993**, *328*, 303–307.

-
37. Burgess, H.J.; Penev, P.D.; Schneider, R.; Van Cauter, E. Estimating cardiac autonomic activity during sleep: Impedance cardiography, spectral analysis, and Poincaré plots. *Clin. Neurophysiol.* **2004**, *115*, 19–28.
 38. Gronfier, C.; Simon, C.; Piquard, F.; Ehrhart, J.; Brandenberger, G. Neuroendocrine processes underlying ultradian sleep regulation in man. *J. Clin. Endocrinol. Metab.* **1999**, *84*, 2686–2690.
 39. Cabiddu, R.; Cerutti, S.; Viardot, G.; Werner, S.; Bianchi, A.M. Modulation of the sympatho-vagal balance during sleep: Frequency domain study of heart rate variability and respiration. *Front. Physiol.* **2012**, *3*, 45–45.