Evaluation of pulse rate measurement with a wrist worn device during different tasks and physical activity

Rahel Gilgen-Ammann[1,*], Mark J. Buller[2], Jacqueline L. Bitterle[1], Simon K. Delves[3], Bertil J. Veenstra[4], Lilian Roos[1], Nadja Beeler[1] & Thomas Wyss[1]

1 Swiss Federal Institute of Sport Magglingen SFISM, Magglingen, Switzerland
2 U.S. Army Research Institute of Environmental Medicine, Natick, United States of Amerika
3 Institute of Naval Medicine, Hampshire, United Kingdom
4 Netherlands Ministry of Defence, Utrecht, Netherlands
* Corresponding author: Swiss Federal Institute of Sport Magglingen (SFISM), Federal Office of Sport, Hauptstrasse 247, CH-2532 Magglingen, Switzerland
Tel: +41 (0) 586 476 321
Email: rahel.gilgen@baspo.admin.ch

COMMENTARY

Article History:
Submitted 15th January 2018
Accepted 6th June 2018
Published 31st July 2018

Handling Editor:
Ernst-Joachim Hossner
University of Bern, Switzerland

Editor-in-Chief:
Martin Kopp
University of Innsbruck, Austria

Reviewers:
Reviewer 1: Kuno Hottenrott
Martin-Luther-Universität Halle-Wittenberg, Germany
Reviewer 2: Ilkka Korhonen
Tampere University of Technology, Finland

ABSTRACT

The purpose of this study was to evaluate the wrist-worn device Mio FUSE, which estimates heart rate (HR) based on photo-plethysmography, 1) in a large study group during a standardised activity, 2) in a small group during a variety of activities and 3) to investigate factors affecting HR accuracy in a real-world setting. First, 53 male participants (20 ±1 years; 1.79 ±0.07 m; 76.1 ±10.5 kg) completed a 35-km march wearing the Equivital EQ-02 as a criterion measure. Second, 5 participants (whereof 3 female; 29 ±5 years; 1.74 ±0.07 m; 67.8 ±11.1 kg) independently performed 25 activities, categorised as sitting passive, sitting active, standing, cyclic and anti-cyclic activities with the Polar H7 as a criterion device. Equivalence testing and Bland-and-Altman analyses were undertaken to assess the accuracy to the criterion devices. Third, confounders affecting HR accuracy were investigated using multiple backwards regression analyses. The Mio FUSE was equivalent to the respective criterion measures with only small systematic biases of -3.5 bpm (-2.6%) and -1.7 bpm (-1.3%) with limits of agreements of ±10.1 bpm and ±10.8 bpm during the 35-km march and during different activities, respectively. Confounding factors negatively affecting the accuracy of the Mio FUSE were found to include larger wrist size and intensified arm and/or wrist movement. The wrist-worn Mio FUSE can be recommended to estimate overall HR accurately for different types of activities in healthy adults. However, during sporting activities involving intensified arm and/or wrist movement or for detailed continuous analysis, a chest strap is preferred to the Mio FUSE to optimise HR estimation accuracy.

Keywords:
Fitness tracker – photo-plethysmography – measurement accuracy

Citation:
Gilgen-Ammann, R., Buller, M. J., Bitterle, J. L., Delves, S. K., Veenstra, B. J., Roos, L., Beeler, N., Wyss, T. (2018). Evaluation of pulse rate measurement with a wrist worn device during different tasks and physical activity. *Current Issues in Sport Science, 3:011*. doi: 10.15203/CISS_2018.011

Introduction

Heart rate (HR) is an informative physiological measurement that has long been utilised to assess work rate, energy expenditure and fitness (Achten & Jeukendrup, 2003). The HR derived from electrocardiograms (ECG) using wearable chest straps are a mature technology from about 1980 that provides reliable and valid measures when compared to standard medical ECG devices (Laukkanen & Virtanen, 1998; Leger & Thivierge, 1988). However, chest-belt HR systems can be problematic, especially for long-term wear (> 12 h). The belts tend to displace over time, which can negatively affect measurement accuracy and they can be seen as constrictive or can cause skin irritation, which all contribute to reducing wear compliance over time (Forrest et al., 2004; O'Toole, Douglas, & Hiller, 1998; Tharion et al., 2013). Furthermore, if electrodes are on dry skin, for example, during physical inactivity, the HR signal can be lost.

To overcome these drawbacks, the measurement of HR recently has been incorporated into wrist-worn activity trackers using photo-plethysmography (PPG). The approach of PPG is an optical measurement to assess blood volume changes in the microvascular bed of tissue (Challoner & Ramsay, 1974; Lemay et al., 2014). This can be used to determine oxygen saturation and pulse waves, through which HR can be derived. The few previous studies that evaluated wrist-based HR accuracy using PPG revealed conflicting results (Gorny, Liew, Tan, & Muller-Riemenschneider, 2017). Movement artefacts are thought to be a problem in the wrist-worn pulse rate measurements using PPG (Tamura, Maeda, Sekine, & Yoshida, 2014; Zhang, Pi, & Liu, 2015), especially during activities with increased intensity or particular arm and wrist movements (Gorny et al., 2017; Shimazaki, Hara, Okuhata, Nakamura, & Kawabata, 2014; Spierer, Rosen, Litman, & Fujii, 2015; Stahl, An, Dinkel, Noble, & Lee, 2016). Moreover, factors such as skin colour, skin temperature or body mass index (BMI) were shown potentially to affect measurement accuracy (Khan et al., 2015; Spierer et al., 2015; Valenti & Westerterp, 2013). Additionally, there are methodological concerns to be addressed as some previous research validated multiple devices that were simultaneously worn on one wrist. This is not recommended for two main reasons; firstly, by fitting more than one device on the wrist, a correct placement according to the manufacturer's instructions is not warranted and, secondly, devices touching each other tap together with every movement (Duking, Fuss, Holmberg, & Sperlich, 2018; Duking, Hotho, Holmberg, Fuss, & Sperlich, 2016; Malone, Lovell, Varley, & Coutts, 2017). Two concerns that might affect the measurement output. Furthermore, the majority of the previous validation studies was conducted in a laboratory setting. However, the wrist-worn pulse rate measurement systems are designed for monitoring general activity in various forms of exercise, intensity, positions, and while moving freely (Duking et al., 2018). For all the aforementioned reasons, there is still a lack of knowledge regarding the accuracy of the wrist-worn devices on the market that measure pulse rate using PPG and what exactly the influencing factors regarding accuracy are. Despite this fact, many manufacturers have launched wrist-worn devices that utilise the PPG approach to determine HR. Moreover, these devices are promoted as tools to monitor medical concerns or to enhance physical fitness and performance (Duking, Holmberg, & Sperlich, 2017). The annual survey of worldwide fitness trends revealed wearable technology as the number one fitness trend for 2017 (Thompson, 2016). Consequently, it is becoming more and more imperative that the data they supply are proven to be trustworthy by employing scientific evaluations (Sperlich & Holmberg, 2017).

The aim of the present study was therefore, to examine whether an example of a wrist-based PPG HR device of a foremost manufacturer (Mio FUSE, Mio Global Inc., Vancouver, CA) provides accurate HR data 1) for each subject of a large study group during a standardised activity (long-distance marching task), 2) for a small study group during a variety of activity tasks and 3) to detect factors affecting HR accuracy of a PPG HR device.

Methods

Experimental approach

The present study was conducted in a real-world setting with two independent study samples and designs. Part 1 consisted of data collection with a large study group (N=63) during a standardised activity: a 35-km military marching task. All participants had exactly the same conditions in terms of activity, environment, measurement time point, uniform and backpack (type and weight). This design allowed for the investigation of the influence of anthropometrical factors on measurement accuracy. Part 2 consisted of a variety of over 100 daily tasks and sporting activities with a smaller sample group (N=5). This design was chosen to examine the effect of environmental factors, such as versatile movement patterns, device placement and application of body lotion, on the wrist-based HR estimation.

Participants

Part 1 was performed with 63 male Swiss Army recruit volunteers of the Rescue Technician Training School (age 20 ±1 years; height 1.79 ±0.06 m; body mass 75.79 ±10.05 kg) and Part 2 was conducted with 5 civil participants (whereof 3 female; age 29 ±5 years; height 1.74 ±0.07 m; body mass 67.8 ±11.1 kg). Participants in both study parts signed informed consent forms that were approved by the ethics commission of the Canton Berne. All participants were healthy, neither took any HR affecting medication or were suffering from any known illnesses.

Apparatus

In both study parts, the investigated wrist-worn pulse rate device was the Mio FUSE (Firmware Version 01.20; Mio Global Inc., Vancouver, CA). The Mio FUSE has an integrated optical sensor based on PPG that uses a green light to detect differences in pulse waves caused by capillary blood flow. Most wrist-based devices use this PPG information to derive HR. Additionally, the Mio FUSE has an incorporated 3-axis accelerometer that is used for specific algorithms to cancel out artefacts and noise in the signal affected by, for instance, arm and/or wrist movements. All devices were set to collect data at 2-second intervals.

In Part 1, the participants wore a chest sensor belt electrocardiogram (Equivital EQ-02 LifeMonitor, Hidalgo Ltd., Cambridge, UK) as the criterion measure, which collected data at 15-second intervals. The validity and reliability of this method are presented elsewhere (Liu, Zhu, Wang, Ye, & Li, 2013). Wrist size, wrist hair, skin type and body size were recorded as possible confounders of accurate pulse rate measurement. Wrist size was quantified using a measuring tape. The hair on the forearm was categorised into two groups: 1=little or no hair and 2=moderate to a lot of hair. Skin type was assessed using the Fitzpatrick scale, from type 1=light, pale white to type 6=black, very dark skin (Fitzpatrick, 1988). Body size was quantified using the BMI. In Part 2, the criterion measure was the HR recording from the Polar H7 chest strap (Polar Electro Oy, Kempele, Finland) with a sampling rate of 1-second intervals (Cheatham, Kolber, & Ernst, 2015; Giles, Draper, & Neil, 2016). An additional sport watch (Polar V800) enabled data storage and exporting of the Polar H7 data. Investigated confounding factors of accurate HR estimation were self-classified arm and wrist movement on a Likert scale (1=little movement, relaxed to 3=much movement, like in a return play), device positioning (0=supinated; 1=pronated side of the wrist) and the application of a greasy body lotion right before the measurements (0=no lotion; 1=with lotion). Participants reported these factors on a predetermined digital questionnaire.

Measures

The data collection of Part 1 was accomplished in one day in March 2016. Participants were equipped with the measurement devices by a supervisor team. The Mio FUSE was placed on the non-dominant wrist and according to the user manual: snugly, about 1-3 inches above the wrist bone (mioglobal.com). Thereafter, participants were instructed not to displace or adjust the devices. Additionally, participants carried their usual military equipment (webbing, helmet and weapon) and a backpack, which overall weighted 25 kg. Wrist size, wrist hair, skin type and body size were recorded during device fitting. Mid-afternoon participants started the self-paced 35-km road march. The march was split into 5 sections with 4 extended breaks of approximately 20 minutes each. During these rests, participants were free to move or sit around and to eat and drink as desired. Part 2 was conducted in the 2015-2016 winter season. After the assessment of the anthropometrics, the participants were instructed and shown on how exactly to use and mount the Mio FUSE, the Polar H7 chest strap and the Polar sport watch. In case of pronated device positioning, the Mio FUSE was fasten snugly 1-3 inches above the wrist bone, whereas in case of supinated device positioning the Mio FUSE was worn slightly higher, since bending the wrist could affect the measurement (mioglobal.com). The Mio FUSE always was placed on the non-dominant wrist and the sport watch on the dominant wrist. Additionally, participants were briefed on how to choose and record their activities. In total, each participant recorded HR data simultaneously with all measurement devices during 25 activities categorised into sitting passive, sitting active, standing, cyclic and anti-cyclic activities tasks (Table I). Five tasks per category had to be accomplished individually, whereas the intensity, wearing position and application of a greasy body lotion were self-selected randomly without preconditioning. After initiation of a measurement, the activity had to be performed for a minimum of 30 minutes.

Table 1: Examples of some activity tasks in Part 2 and the corresponding activity category including the scaling of arm and wrist movement.

Activity task	Activity category	Arm movement [scale 1-3]	Wrist movement [scale 1-3]
Relaxing (reading, watching TV)	Sitting passive	1	1
Riding a train	Sitting passive	1	1
Driving a car	Sitting active	2	2
Office work seated	Sitting active	1	2
Office work standing	Standing	1	2
Household chores	Standing	2	2
Cooking	Standing	2	2
Walking / hiking	Cyclic activity	2	1
Jogging	Cyclic activity	3	1
Skiing	Anti-cyclic activity	2	3
Climbing	Anti-cyclic activity	3	3
Strength training	Anti-cyclic activity	2	3
Lifting weight	Anti-cyclic activity	3	3
Playing soccer	Anti-cyclic activity	3	2
Playing squash	Anti-cyclic activity	3	2

Data analysis

In the present study, the term continuous HR data was defined as HR values for each 1-minute time interval during an activity (Allen, 2007). While overall HR was defined as the mean HR of one subject over a single activity.

In Part 1, the HR for each device was averaged to 1-minute intervals to calculate continuous HR data, and then, an overall mean HR was calculated over the entire 35-km march, as determined by averaging all 1-minute intervals. Of 63 participants, the data of 53 were analysed, as data of 10 participants (16%) had to be excluded due to technical problems of either the concurrent or the criterion device. These included no data collected at all (3 times within both measurement systems) or flat battery (2 times within each measurement system). No outliers or artefacts were excluded. The potential confounding factors of wrist hair and skin type were included as dichotomous variables because the present study sample represented skin type I through IV only, whereof 85% represented skin type II and III. To ensure the same group sizes, skin type I and II were computed as 0=white fair skin type, and type III and IV as 1=medium white to moderate brown skin type.

In Part 2, the data of the Mio FUSE and the H7 chest strap were synchronised to the same starting second of each activity prior to computing continuous and overall HR data for each activity task. The first minute of each activity was discharged, and the following 29 minutes used for the analyses. Due to technical problems (as in Part 1) or inaccurate performance/reporting by the participants (e.g. mixing up activity categories or too short measurement periods), 11 activity tasks were excluded. Moreover, short-time data loss within the data files (4.5 ±10.6% of each file) were discarded from the analyses and replaced as missing data in both measurement systems.

Statistical analysis

For both study parts, the same statistical analyses were applied using the software Microsoft Excel (2011) and SPSS 24.0 (Inc., Armonk, NY, USA). The results were considered to be significant if $p \leq .05$. Equivalence testing was performed to determine whether the Mio FUSE measurements were significantly equivalent to the criterion measures (Dixon & Techmann, 2005; Walker & Nowacki, 2011). The HR obtained by the Mio FUSE were considered to be equivalent if the 95% confidence interval for the absolute mean error of the Mio FUSE measurement fell into the proposed equivalence zone (±5%) of the measured HR by the criterion measures. Bland-and-Altman plots with corresponding 95% limits of agreement (±1.96·SD) were used to calculate and visualise systematic differences in the HR estimations (Bland & Altman, 1986). Moreover, the root mean square errors (RMSE) and the Pearson correlation coefficients (r) were calculated. To investigate potential confounding effects on the HR accuracy during the march, a backward multiple linear regression was performed with the overall RMSE as the dependent variable and the skin type, wrist hair, wrist size and BMI as the independent variables. Similarly, potential predictors on HR accuracy estimations during different activities were evaluated using a backward multiple linear regression with the overall RMSE as the dependent variable and the position of the device on the wrist, arm movement, wrist movement and the application of body lotion as independent variables. In case of multicollinearity ($r \geq .80$) or non-significant prediction of the RMSE, the relevant variable was excluded from that particular regression analysis.

Results

If not noted differently, all results presented are related to the overall HR data.

In Part 1, the mean marching time for the 35 km (including 80-minute resting time) was 493 ±47 min, which corresponded to a mean moving speed of 1.4 $m \cdot s^{-1}$. Over the duration of the march, the air temperature was 4.0 ±1.7°C and relative humidity 71 ±12%. The overall mean HR (Table II) recorded by the EQ-02 was 130.1 bpm, of which 5% (±6.5 bpm) was used to determine the interval of tolerable difference. The Mio FUSE recorded a mean HR of 126.6 bpm and a systematic bias from the criterion of -3.5 bpm (-2.6%) with limits of agreement of ±10.1 bpm (Figure 1). Since the reported 95% confidence intervals (-4.9, -2.0) for the difference between the Mio FUSE and the EQ-02 were completely within the interval of tolerable difference (-6.5, +6.5), the HR estimations of the two devices can be declared equivalent. The Bland-and-Altman plot for continuous HR data derived an absolute difference from the criterion of -3.6 bpm (-2.7%) with limits of agreement of ±31.1 bpm (Figure 2). Participants' skin types were 51% fair white and 49% medium white to moderate brown. In terms of wrist hair, 57% had no to little hair and 43% moderate to a lot of hair. The mean wrist size was 20.1 cm (ranging from 18.0-23.0 cm) and the BMI

Table 2: Summary of the heart rate data obtained during the 35-km march in Part 1.

	Mio FUSE ±SD [bpm]	**EQ-02 ±SD** [bpm]	**Absolute error ±SD** [bpm]	**RMSE ±SD** [bpm]	**Mean *r*¥** (*p*-value)
35-km march	126.6 ±10.1	130.1 ±10.0	-3.5 ±5.2	4.3 ±4.5	.867 (< .001)

Note: ¥correlation coefficient of each participants' overall heart rate obtained with the wristband Mio FUSE and the chest sensor belt Equivital EQ-02. SD=standard deviation; RMSE=root mean square error.

was 23.9 kg·m^{-2} (ranging from 19.2-33.6 kg·m^{-2}). A multiple linear regression was calculated to predict measurement errors based on the aforementioned factors. A significant regression equation was found ($F_{4, 52}$ = 3.336, p = .018), with an R^2 of .225 (Table III). The RMSE of the Mio FUSE as the dependent variable increased with a larger wrist circumference (p = .006). Whereas, less wrist hair (p = .033), a lighter skin type (p = .047) and a higher BMI (p = .053) decreased the RMSE.

Table 3: Results of the multiple linear regression analyses (backward method) with root mean square error as dependent variable assessed during (a) the 35-km march and (b) the different daily activities.

(a)	*OR*	*P-value*	*95% CI*
Wrist size	1.932	.006	0.583; 3.282
Wrist hair[1]	-3.230	.033	-6.191; -0.268
Skin type[1]	3.012	.047	0.036; 5.989
Body mass index	-0.483	.053	-0.972; 0.007
(b)			
Arm movement[2]	4.628	.000	2.701; 6.555
Wrist movement[2]	2.852	.004	0.936; 4.769

Note: [1] 0=no to little wrist hair, 1=moderate to a lot wrist hair; 0=fair skin type, 1=medium white to moderate brown skin type;

[2] 1=no to little arm/wrist movement, 2=moderate arm/wrist movement; 3=much arm/wrist movement.

OR=odds ratio; CI=confidence interval

In Part 2, the data of 114 activity tasks were analysed (Table IV). Overall, the mean HR obtained by the chest strap was 91.0 bpm. The Mio FUSE estimated a mean HR of 89.2 bpm and a systematic bias from the criterion of -1.7 bpm (-1.3%) with ±10.8 bpm limits of agreement (Figure 1). The HR estimations of the Mio FUSE was equivalent to the chest strap as the 95% confidence intervals (-2.8, -0.7) for the difference between the two devices were completely within the interval of tolerable difference (-4.5, +4.5). The Bland-and-Altman plot for continuous HR data derived an absolute difference from the criterion of -1.8 bpm (-2.0%) with limits of agreement of ±23.6 bpm (Figure 2).

The five different activity tasks were performed with an equal frequency (Table IV). Altogether, 67 (59%) measurements were recorded with the Mio FUSE placed on the pronated wrist side and 47 (41%) placed on the supinated side. Eighty-eight (77%) measurements were recorded without and 26 (23%) with applying a body lotion prior the activity initiation. The distribution of the intensity in arm and wrist movements were reported with a similar frequency of approximately 40% of little, 45% of moderate and 15% of much movement. To investigate the impact of the activity tasks and the aforementioned factors on the overall RMSE, a multiple linear regression analysis was performed. Due to its multicollinearity with arm movement (r = .834, p < .001), the variable activity task had to be excluded. A significant regression equation was revealed ($F_{2, 113}$ = 27.451, p < .001), with R^2 of .331 (Table III). The more arm movement, the larger the error was in the HR estimation (p < .001). Likewise, the more activity-induced wrist movement, the larger the measurement error was (p = .004). No significant effects were detected for the wearing position or the application of body lotion.

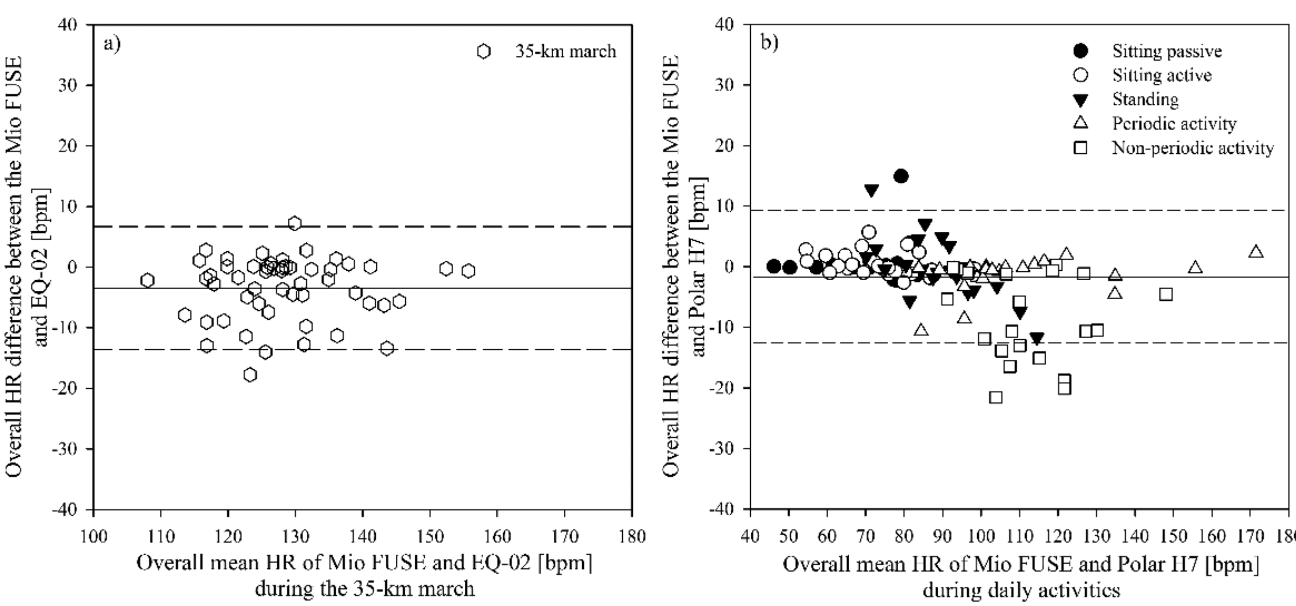

Figure 1: Bland-and-Altman plots of the overall mean heart rate (HR) obtained during a) a 35-km march (N=53) and b) during different activities (114 measurements). The solid lines represent the systematic bias; the dashed lines represent the limits of agreement (systematic bias ±1.96·standard deviation).

Table 4: Summary of the heart rate data obtained during the different activity tasks in Part 2.

Activity category	Amount of activity tasks (%)	Mio FUSE ±SD [bpm]	Polar H7 ±SD [bpm]	Absolute error ±SD [bpm]	RMSE ±SD [bpm]	Mean $r^{¥}$ (*p*-value)
Sitting passive	23 (22.8)	71.2 ±10.3	70.8 ±9.9	0.4 ±3.4	5.5 ±6.4	.945 (< .001)
Sitting active	26 (20.2)	73.7 ±10.5	73.3 ±11.2	0.4 ±1.9	5.7 ±3.8	.986 (< .001)
Standing	21 (18.4)	88.8 ±10.3	89.0 ±13.9	-0.2 ±5.3	11.0 ±5.7	.948 (< .001)
Cyclic activity	23 (20.2)	108.3 ±22.6	109.7 ±21.6	-1.4 ±3.0	7.2 ±6.8	.992 (< .001)
Anti-cyclic activity	21 (18.4)	107.8 ±14.3	116.5 ±15.1	-8.7 ±7.3	16.7 ±9.4	.879 (< .001)
Total	114	89.2 ±21.4	91.0 ±23.7	-1.7 ±5.5	9.0 ±7.7	.975 (< .001)

Note: ¥correlation coefficient of each participants' overall heart rate obtained with the wristband Mio FUSE and the chest strap Polar H7. SD=standard deviation; RMSE=root mean square error.

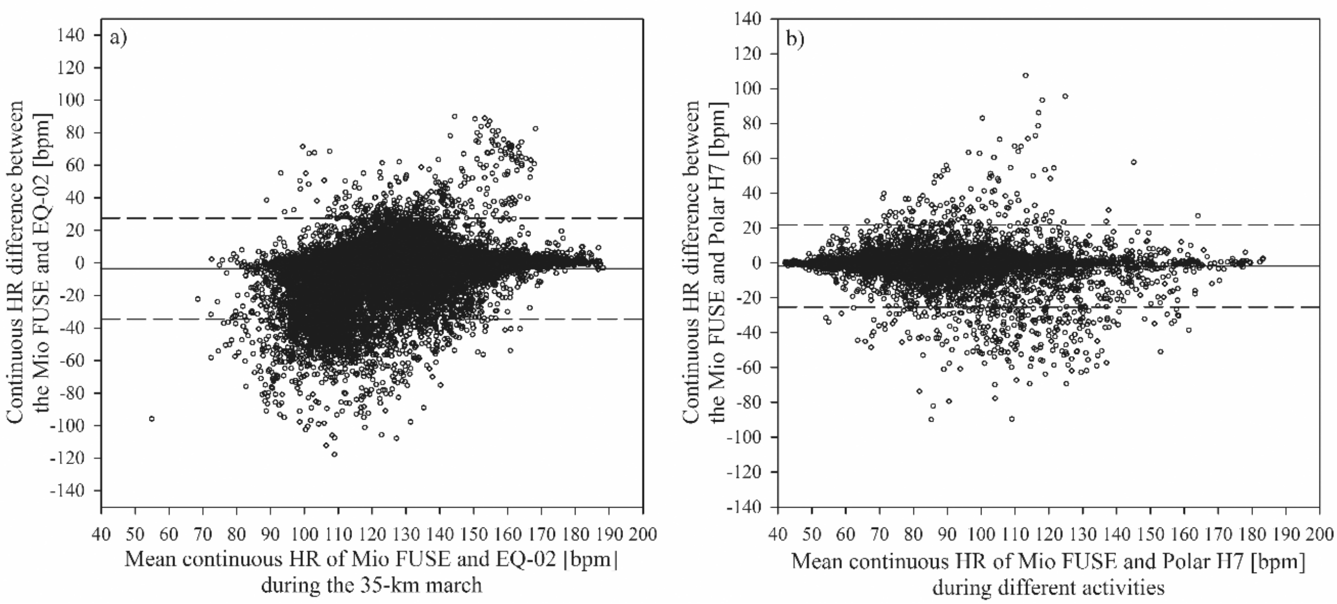

Figure 2: Bland-and-Altman plots of the continuous heart rate (HR) obtained during a) a 35-km march and b) during different activities. The solid lines represent the systematic bias; the dashed lines represent the limits of agreement (systematic bias ±1.96·standard deviation).

Discussion

In this study, a commercially available wrist-based PPG pulse rate monitor was evaluated under real-world conditions. The accuracy of the Mio FUSE was evaluated, first, during a standardised activity within a large study group and, second, during a variety of activities within a small group, by comparing the HR estimations with that of a mobile ECG device and a chest strap. The findings suggest high overall concordance between the methods in both study parts. With relative deviations from the criterion measures of -2.6 ±3.9% during the 35-km march and -1.3 ±5.5% during the different activities, the Mio FUSE was equivalent to the respective criterion devices. The revealed confounding factors affecting the HR accuracy of the Mio FUSE were larger wrist size, intensified arm and/or wrist movement, denser wrist hair, darker skin type and lower BMI. Generally, the results of the present study confirm and extend the investigations with comparable methodologies of wrist-worn devices, with a constant underestimation of the HR during different activities (Gorny et al., 2017; Spierer et al., 2015).

However, mean overall HR computed across the trials masks the poorer performance of the Mio FUSE regarding continuous HR data as indicated by larger limits of agreements in Figure 2 compared to Figure 1. It appears that the Mio FUSE is less appropriate to monitor continuous HR during single activities, rather for overall training or 24/7 HR assessment. The accuracy of the Mio FUSE seems to be person- and/or task-dependent, which was in line with previous observations using similar devices (Gorny et al., 2017; Spierer et al., 2015; Stahl et al., 2016). The present study revealed significant confounding factors affecting the quality of the HR assessment. In terms of anthropometrics (person-dependent), larger wrist size was the most relevant predictor for decreased measurement accuracy. Previously, it has been highlighted that the wristband must not be too small, as PPG signal quality would be hampered by modified pressure between the sensor and the skin (Lemay et al., 2014). Therefore, a perfect fit of the wristband to the individuals' anthropometry plays a paramount role in order to achieve more accurate HR data. Nevertheless, the wearing position of the Mio FUSE on the pronated or supinated wrist side did not affect the measurement error. Considering environmental factors (task-dependent), activities with more arm and/or wrist movement were shown to impair the HR assessment significantly. Such movement patterns were prevalent during cyclic (e.g. hiking) and anti-cyclic (e.g. serve/return plays) activities. In particular, during anti-cyclic activities, HR accuracy was lowered with a mean relative error of 8%. This was in accordance with previously published findings, showing poorer measurement quality caused by movement artefacts (Shimazaki et al., 2014; Spierer et al., 2015; Stahl et al., 2016; Tamura et al., 2014; Zhang et al., 2015). As such, weightlifting led to decreased measurement accuracy within the wrist-worn pulse rate monitor (Spierer et al., 2015). Similar, in a study investigating 7 different device positions between the upper arm and a finger, the smallest movement artefacts in the PPG signal were detected when placed as proximal as possible (Maeda, Sekine, & Tamura, 2011). Measurement errors were smaller when placed on the upper arm than on the wrist or finger. Furthermore, in line with previous findings was the factor that participants with darker skin types indicated enlarged measurement errors, presumably due to lowered translucence of the skin (Spierer et al., 2015). Potentially related to this, denser wrist hair was also a significant predictor for enlarged measurement errors. Interestingly, higher BMI values tended to improve measurement quality. However, this most likely does not mean the larger the user the more accurate the HR estimation. The BMI is not the most expressive measure in terms of body size, as for example no information like muscle or fat mass is provided. Hence, the impact of BMI on measurement accuracy should not be over-interpreted. The application of any kind of body lotion did not influence the HR assessment. Hence, it is assumed that a greasy layer on the skin, potentially changing conductivity or increasing sweat, does not affect the PPG signal.

Overall, according to the present findings, the Mio FUSE showed good practical accuracy in the HR estimation obtained on the wrist using PPG. The Mio FUSE can be recommended as a fitness tracker in order to obtain 24/7 HR data. There are two main advantages compared to a chest strap. First, the method based on PPG also works on dry skin during longer periods of inactivity, and second, the wristband has a high wearing comfort. When well fitted to the users' wrist size, it appears particularly recommendable for accurate overall HR estimations during sitting, standing and cyclic activities with low-intensity and little arm and/or wrist movements. Meanwhile, during high-intensity or anti-cyclic movements the Mio FUSE demonstrated impaired measurement accuracy by a mean value of 8%. Consequently, during sporting activities or training sessions, particularly incorporating more arm movement, a chest strap is superior to the Mio FUSE. This supports the findings by Parak and Korhonen (2014) showing that measurement systems based on ECG outperformed wrist-worn systems on more rigorous tasks. Also, the average temperature during the 35-km march was rather cold, a condition that was shown to reduce the PPG signal (Khan et al., 2015). Hence, one may expect even smaller errors in the HR assessment in warmer conditions.

In Part 2, the participants independently accomplished all different activities. Therefore, participants had to decide about the intensity and where to place the Mio FUSE and when to apply body lotion. There were markedly more assessments with the Mio FUSE on the pronated than the supinated wrist placement, and many more without than with the application of a body lotion. Nevertheless, as wearing the Mio FUSE on the pronated wrist side and not applying any body lotion right before an assessment are assumed as the common approach, this is not perceived as a limitation of the presented results. However, in further research wrist-worn pulse rate monitors should be evaluated in larger study samples incorporating equally both gender and a broad range of ages and skin types.

Conclusions

The Mio FUSE wristband based on PPG can estimate overall HR data accurately during a variety of activity tasks. Particularly for a 24/7 monitoring, when well fitted to the wrist size, it can be recommended as an accurate HR estimation device, overcoming the disadvantages of a chest strap. Whereas, during activities incorporating intensified arm and/or wrist movement or the detailed continuous HR analysis of single tasks, a chest strap is still preferred to the Mio FUSE in order to optimise HR measurement accuracy.

Funding

The authors have no funding or support to report.

Competing Interests

This study is a collaboration between Switzerland, the Netherlands, United States of America and the United Kingdom through NATO Panel HFM-260. The authors declare that they have no conflict of interest. The views expressed in this paper are those of the authors and do not reflect the official policy of their employers or respective governments.

Data Availability Statement

All relevant data are within the paper.

References

Achten, J., & Jeukendrup, A. E. (2003). Heart rate monitoring: applications and limitations. *Sports Med, 33*(7), 517-538.

Allen, J. (2007). Photoplethysmography and its application in clinical physiological measurement. *Physiol Meas, 28*(3), R1-39. doi:10.1088/0967-3334/28/3/R01

Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet, 1*(8476), 307-310.

Challoner, A. V., & Ramsay, C. A. (1974). A photoelectric plethysmograph for the measurement of cutaneous blood flow. *Phys Med Biol, 19*(3), 317-328.

Cheatham, S. W., Kolber, M. J., & Ernst, M. P. (2015). Concurrent validity of resting pulse-rate measurements: a comparison of 2 smartphone applications, the polar H7 belt monitor, and a pulse oximeter with bluetooth. *J Sport Rehabil, 24*(2), 171-178.

Dixon, P. M., & Techmann, J. H. K. (2005). A statistical test to show negligible trend. *Ecology, 86*, 1751 – 1756.

Duking, P., Fuss, F. K., Holmberg, H. C., & Sperlich, B. (2018). Recommendations for Assessment of the Reliability, Sensitivity, and Validity of Data Provided by Wearable Sensors Designed for Monitoring Physical Activity. *JMIR Mhealth Uhealth, 6*(4), e102. doi:10.2196/mhealth.9341

Duking, P., Holmberg, H. C., & Sperlich, B. (2017). Instant Biofeedback Provided by Wearable Sensor Technology Can Help to Optimize Exercise and Prevent Injury and Overuse. *Front Physiol, 8*, 167. doi:10.3389/fphys.2017.00167

Duking, P., Hotho, A., Holmberg, H. C., Fuss, F. K., & Sperlich, B. (2016). Comparison of Non-Invasive Individual Monitoring of the Training and Health of Athletes with Commercially Available Wearable Technologies. *Front Physiol, 7*, 71. doi:10.3389/fphys.2016.00071

Fitzpatrick, T. B. (1988). The validity and practicality of sun-reactive skin types I through VI. *Arch Dermatol, 124*(6), 869-871.

Forrest, A. J., Blanton, C. A., Gale, B. J., Gustafson-Storms M., Horn, W. F., Keim, N., Kretsch, M. J. (2004). Subject compliance using digital activity log, accelerometer and heart rate monitor to record physical activity. *FASEB Journal, 18*(4-5), 106-107.

Giles, D., Draper, N., & Neil, W. (2016). Validity of the Polar V800 heart rate monitor to measure RR intervals at rest. *Eur J Appl Physiol, 116*(3), 563-571. doi:10.1007/s00421-015-3303-9

Gorny, A. W., Liew, S. J., Tan, C. S., & Muller-Riemenschneider, F. (2017). Fitbit Charge HR Wireless Heart Rate Monitor: Validation Study Conducted Under Free-Living Conditions. *JMIR Mhealth Uhealth, 5*(10), e157.

Khan, M., Pretty, C. G., Amies, A. C., Elliott, R., Shaw, G. M., & Chase, J. G. (2015). Investigating the Effects of Temperature on Photoplethysmography. *IFAC-PapersOnLine, 48*(20), 360-365.

Laukkanen, R. M., & Virtanen, P. K. (1998). Heart rate monitors: state of the art. *J Sports Sci, 16 Suppl*, S3-7.

Leger, L., & Thivierge, M. (1988). Heart Rate Monitors: Validity, Stability, and Functionality. *Phys Sportsmed, 16*(5), 143-151. doi:10.1080/00913847.1988.11709511

Lemay, M., Bertschi, M., Sola, J., Renevey, P., Parak, J., & Korhonen, I. (2014). Application of Optical Heart Rate Monitoring. In S. Edward & N. Michael (Eds.), *Wearable Sensors* (pp. 105-129). San Diego (CA): Elsevier Academic Press.

Liu, Y., Zhu, S. H., Wang, G. H., Ye, F., & Li, P. Z. (2013). Validity and reliability of multiparameter physiological measurements recorded by the Equivital LifeMonitor during activities of various intensities. *J Occup Environ Hyg, 10*(2), 78-85.

Maeda, Y., Sekine, M., & Tamura, T. (2011). Relationship between measurement site and motion artifacts in wearable reflected photoplethysmography. *J Med Syst, 35*(5), 969-976.

Malone, J. J., Lovell, R., Varley, M. C., & Coutts, A. J. (2017). Unpacking the Black Box: Applications and Considerations for Using GPS Devices in Sport. *Int J Sports Physiol Perform, 12*(Suppl 2), S218-S226. doi:10.1123/ijspp.2016-0236

mioglobal.com. *Mio FUSE complete user guide*. Retrieved from https://www.mioglobal.com/docs/mio_fuse_complete-user-guide_en.pdf

O'Toole, M. L., Douglas, P. S., & Hiller, W. D. (1998). Use of heart rate monitors by endurance athletes: lessons from triathletes. *J Sports Med Phys Fitness, 38*(3), 181-187.

Parak, J., & Korhonen, I. (2014, 26-30 Aug. 2014). *Evaluation of wearable consumer heart rate monitors based on photopletysmography.* Paper presented at the 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society.

Shimazaki, T., Hara, S., Okuhata, H., Nakamura, H., & Kawabata, T. (2014). *Cancellation of motion artifact induced by exercise for PPG-based heart rate sensing*. Paper presented at the Conf Proc IEEE Eng Med Biol Soc, Chicago. http://www.ncbi.nlm.nih.gov/pubmed/25570675

Sperlich, B., & Holmberg, H. C. (2017). Wearable, yes, but able...?: it is time for evidence-based marketing claims! *Br J Sports Med, 51*(16), 1240. doi:10.1136/bjsports-2016-097295

Spierer, D. K., Rosen, Z., Litman, L. L., & Fujii, K. (2015). Validation of photoplethysmography as a method to detect heart rate during rest and exercise. *J Med Eng Technol, 39*(5), 264-271.

Stahl, S. E., An, H. S., Dinkel, D. M., Noble, J. M., & Lee, J. M. (2016). How accurate are the wrist-based heart rate monitors during walking and running activities? Are they accurate enough? *BMJ Open Sport Exerc Med, 2*(1), e000106. doi:10.1136/bmjsem-2015-000106

Tamura, T., Maeda, Y., Sekine, M., & Yoshida, M. (2014). Wearable Photoplethysmographic Sensors – Past and Present. *Electronics*(3), 282-302.

Tharion, W. J., Buller, M. J., Potter, A. W., Karis, A. J., Goetz, V., & Hoyt, R. W. (2013). Acceptability and Usability of an Ambulatory Health Monitoring System for Use by Military Personnel. *IIE Transactions on Occupational Ergonomics and Human Factors, 1*(4), 203-214.

Thompson, W. R. (2016). WORLDWIDE SURVEY OF FITNESS TRENDS FOR 2017. *ACSM's Health & Fitness Journal, 20*(6), 8-17.

Valenti, G., & Westerterp, K. (2013). *Optical heart rate monitoring module validation study*. Paper presented at the International Conference on Consumer Electronics (ICCE), Las Vegas.

Walker, E., & Nowacki, A. S. (2011). Understanding Equivalence and Noninferiority Testing. *Journal of General Internal Medicine, 26*(2), 192-196.

Zhang, Z., Pi, Z., & Liu, B. (2015). TROIKA: a general framework for heart rate monitoring using wrist-type photoplethysmographic signals during intensive physical exercise. *IEEE Trans Biomed Eng, 62*(2), 522-531. doi:10.1109/TBME.2014.2359372