Performance evaluation of a wrist-worn reflectance pulse oximeter during sleep

Hyunjun Jung, PhD\textsuperscript{a,\#}, Dongyeop Kim, MD\textsuperscript{b,\#}, Wonkyu Lee, PhD\textsuperscript{c}, Hyejung Seo, PhD\textsuperscript{c}, Jinwoo Seo, PhD\textsuperscript{c}, Jongmin Choi, PhD\textsuperscript{a,\***}, Eun Yeon Joo, MD, PhD\textsuperscript{a,*}

\textsuperscript{a} Samsung Electronics, Suwon, Republic of Korea
\textsuperscript{b} Department of Neurology, Seoul Hospital, Ewha Womans University College of Medicine, Seoul, Republic of Korea
\textsuperscript{c} Department of Neurology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea

A B S T R A C T

Objective: To characterize and evaluate the estimation of oxygen saturation measured by a wrist-worn reflectance pulse oximeter during sleep.

Methods: Ninety-seven adults with sleep disturbances were enrolled. Oxygen saturation was simultaneously measured using a reflectance pulse oximeter (Galaxy Watch 4 [GW4], Samsung, South Korea) and a transmittance pulse oximeter (polysonomography) as a reference. The performance of the device was evaluated using the root mean squared error (RMSE) and coverage rate. Additionally, GW4-derived oxygen desaturation index (ODI) was compared with the apnea-hypopnea index (AHI) derived from polysomnography.

Results: The GW4 had an overall RMSE of 2.3% and negligible bias of -0.2%. A Bland-Altman density plot showed good agreement between the GW4 and the reference pulse oximeter. RMSEs were 1.65 ± 0.57%, 1.76 ± 0.58%, 1.93 ± 0.54%, and 2.93 ± 1.71% for normal (n = 18), mild (n = 21), moderate (n = 23), and severe obstructive sleep apnea (n = 35), respectively. The data rejection rate was 26.5%, which was caused by the discarding of data less than 70% of saturation. A GW4-ODI ≥5/h had the highest ability to predict AHI ≥15/h with sensitivity, specificity, accuracy, and area under the curve of 89.7%, 64.1%, 79.4%, and 0.908, respectively.

Conclusions: This study evaluated the estimation of oxygen saturation by the GW4 during sleep. This device complies with both Food and Drug Administration and International Organization for Standardization standards. Further improvements in the algorithms of wearable devices are required to obtain more accurate and reliable information about oxygen saturation measurements.

© 2022 The Authors. Published by Elsevier Inc. on behalf of National Sleep Foundation. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)

Introduction

Pulse oximetry significantly contributes to clinical care by permitting noninvasive, convenient, and continuous measurement of oxygen saturation in peripheral blood (SpO\textsubscript{2}). Photoplethysmography (PPG) is an optical measurement method that enables wearable devices such as pulse oximeters to detect blood volume changes in the microvascular tissue bed; it has various clinical applications, including measuring the heart rate, blood pressure, and SpO\textsubscript{2}.\textsuperscript{1} This technique transmits light to the skin at 2 different wavelengths, red (660 nm) and infrared (940 nm), and measures reflected or transmitted light to calculate absorbance based on the Lambert-Beer law.\textsuperscript{2} Then, it calculates the absorbance ratio at different wavelengths, which reflects the ratio of oxygenated and deoxygenated hemoglobin levels. Several studies have attempted to measure SpO\textsubscript{2} at various body locations using various types of wearable devices.\textsuperscript{2-4} According to previous results, transmittance PPG on the fingertip was verified as one of the most accurate and stable methods for measuring SpO\textsubscript{2}. However, the fingertip-attached device can be obstructive and cumbersome to wear continuously because it prevents users from moving their hands freely; therefore, attempts have been made to develop and validate wrist-worn devices for measuring SpO\textsubscript{2}.\textsuperscript{5,7}

Obstructive sleep apnea (OSA) is a common sleep disorder, with a 9%–38% prevalence rate among the overall adult population.\textsuperscript{8} It is characterized by repetitive upper airway obstruction during sleep, thereby resulting in oxygen desaturation, frequent arousals, and increased sympathetic activity. In-laboratory polysomnography (PSG) is the gold standard for measuring the apnea-hypopnea index (AHI) and diagnosing sleep-related breathing disorders; however, it
has certain limitations, such as high costs, lengthy time requirements, and the need for an in-hospital setting. Furthermore, PSG has limitations that affect its functionality, including the first-night effect and inherent night-to-night variability. Therefore, accessible and relatively low-cost options are being developed to overcome the limitations of PSG so that daily sleep patterns can be monitored.

Reflectance PPG sensors are embedded in most commercially available wearable devices for collecting bio-signals. Nevertheless, several challenges associated with wrist-worn reflectance PPG sensors have been encountered because these devices are known to have a relatively low signal-to-noise ratio and are prone to artifacts; such drawbacks have prevented their widespread use in clinical practice.

Previous studies have validated the accuracy and reliability of wrist-worn reflectance PPG sensors by detecting SpO2 at different hypoxic levels during wake states in laboratory environments. However, detecting wrist-worn reflectance PPG signals during sleep is different because the individuals are unconscious and cannot adjust the watch position. If the measurements during sleep are not properly performed for a long duration, then the test cannot be interrupted to correct for artifacts, resulting in substantially more errors compared to when measurements are obtained during awake states. Only one published study has evaluated the performance of the wrist-worn reflectance PPG device during sleep.

This study aimed to evaluate the SpO2 during sleep as obtained by Galaxy Watch 4, which permits continuous measurement every second. In addition, we attempted to compare the oxygen desaturation index (ODI) obtained using wrist-worn reflectance pulse oximetry with the AHI during sleep.

Participants and methods

Participants

A total of 97 adults with sleep disturbances (age 44.4 ± 13.0 years; 74 men and 23 women) who visited the sleep laboratory at Samsung Medical Center, Seoul, South Korea, to undergo PSG were enrolled in this study. All participants were Asian. The following individuals were excluded: those with parasomnia because of aggressive limb movements during sleep; those with hypoxemia (SpO2 < 90% during waking hours) caused by an underlying pulmonary disease; those with neurological or cardiovascular diseases, including peripheral artery disease; and those with mental illness who were unable to comply with study procedures. All participants provided written informed consent, and the study protocol was approved by the institutional review board of Samsung Medical Center (IRB no. 2021-04-166).

Tested device

Our target device was a GW4 series (model SM-R890N or SM-R860N, Samsung Electronics Co., Seoul, South Korea), which is a watch-type of wearable device that includes a reflectance pulse oximeter module on its underside that is worn against the skin (Fig. 1). This module comprises a series of closely located light-emitting diodes for each wavelength (λIR = 940 nm, λRed = 660 nm) at the center. The 8 photodiodes that sense reflected light are located radially, with an average distance of 4.5 mm from the center. This device captured the PPG signal at a sampling frequency of 25 Hz for each wavelength, and it calculated the SpO2 every second. Output SpO2 data from the GW4 (WristO2) were presented as integers, and this device covered 70%–100% range of saturation because of its lack of accuracy for saturation levels less than 70%. The GW4 continuously monitored the similarity of repetitive pulse waveforms to filter out inconclusive conditions. In that case, the PPG signal quality was considered low, and the output value was not analyzed.

Performance evaluation of oxygen saturation obtained using GW4

All participants underwent level 1 PSG (Embla N7000 PSG System, Medcare Flaga, Reykjavik, Iceland) while wearing the GW4. To compare the accuracy of its reflectance pulse oximeter, we simultaneously measured SpO2 using both the GW4 and a transmittance pulse oximetry system placed on the fingertip as a reference (SpO2ref). Participants wore the GW4 on the left wrist and were instructed to tighten the device against the skin to obtain a high-quality PPG signal. Participants were classified into 4 groups based on the PSG results as follows: normal (AHI <5/h), mild (5 ≤ AHI <15/h), moderate (15 ≤ AHI <30/h), and severe (AHI ≥30/h) OSA. Accordingly, we employed a scoring metric following the guidelines published by the American Academy of Sleep Medicine (AASM).14

SpO2 calculation method

A well-defined alternating component (AC)/direct component (DC) method was used to calculate SpO2. Briefly, noise in the PPG waveform was eliminated using a low-pass filter (fc = 12.5 Hz). Then, the AC and DC were measured based on the recent PPG waveform.
Performance evaluation of WristO2 and the association between AHI and GW4-ODI

The performance of the target device was evaluated using the RMSE and coverage rate. The RMSE was calculated using the squared error (\( \text{SpO}_2\text{Ref} - \text{WristO2} \)) of all valid sample points when both the \( \text{SpO}_2\text{Ref} \) and WristO2 data were available. The coverage rate was defined as the percentage of valid sample points out of the total time asleep. The data rejection rate was calculated as (1−coverage rate). To evaluate the performance, the RMSE and coverage rate were calculated using the pooled \( \text{SpO}_2 \) errors. The RMSE and coverage rate were also calculated for each participant to assess the dependency on the OSA severity. The GW4 oxygen desaturation index (GW4-ODI) was derived from the WristO2, which was defined as the number of desaturation episodes divided by the total sleep time. The number of desaturation events was calculated as the number of drops with a difference of a certain criteria (2%, 3%, or 4%) or more between the maximum \( \text{SpO}_2 \) and minimum \( \text{SpO}_2 \) of a consecutive slope. The total sleep time was estimated using the algorithm of the watch, and the rejected data points were not included.

Statistical analysis

Demographics and PSG parameters were analyzed using descriptive statistical analysis to calculate the means and standard deviations of individual variables. MATLAB was used to analyze and visualize the RMSE and coverage rate of the WristO2. Bland-Altman density plots were used to show agreement between the WristO2 and \( \text{SpO}_2\text{Ref} \). Moreover, these plots were used along with correlation plots for correlation and agreement analyses.

An analysis of the receiver-operating characteristic (ROC) curve was performed to compare the diagnostic performance of different ODI thresholds. All calculated \( P \) values were 2-tailed, and statistical significance was defined as \( P<0.05 \), using SPSS version 18.0 (SPSS Inc., Chicago, IL, USA).

Results

The demographics and polysomnographic findings of the participants are summarized in Table 1.

Oxygen saturation

Fig. 2 shows a representative image of the WristO2 with its reference during sleep. The WristO2 of a normal participant (AHI, 2.3/h) was distributed in the normal range of 90%–100% (Fig. 2A). Transient false recordings and baseline drift were observed intermittently. The erroneous readings (Fig. 2A) were assumed to be related to positional changes or fluctuations in contact pressure between the sensor and the skin, both of which have been reported previously. Multiple true desaturation events associated with the WristO2 were observed in one patient with severe OSA (AHI, 91/h) (Fig. 2B).

Performance evaluation

To evaluate the WristO2, we included valid data points for which the WristO2 and \( \text{SpO}_2\text{Ref} \) were both available, and these points were used for analysis. Of 672.3 hours of recorded sleep time, 3.4 hours were excluded due to inconclusive data from \( \text{SpO}_2\text{Ref} \). 171.7 hours from WristO2 and 6.4 hours from both WristO2 and \( \text{SpO}_2\text{Ref} \). Thus, a considerable amount of data was missing from the WristO2 calculations. After excluding these missing data, 490.8 hours of recorded sleep time were used for the analysis.

The coverage rate of the GW4 was 73.5%, and its data rejection rate was 26.5%. The RMSE was 2.3%, which was calculated during the entire valid period; therefore, it met the requirements of the FDA and ISO standards (Table 2). The Bland-Altman density plot showed good agreement between the 2 measurements, with a mean bias of -0.16% (Fig. 3).

Performance based on the severity of OSA

To characterize the performance of WristO2 among participants with different OSA severity, the RMSE and coverage rate of each participant were compared according to AHI and OSA severity groups. Almost all participants had an AHI ≤ 60, and the performance of the device met FDA and ISO standards (Fig. 4A). However, when participants had a AHI > 60, the estimation error increased and the WristO2 did not meet the FDA and ISO standards.
The RMSEs were 1.65 ± 0.57%, 1.76 ± 0.65%, 1.93 ± 0.54%, and 2.93 ± 1.71% for normal, mild, moderate, and severe OSA, respectively. There was no significant difference in the RMSEs of the different groups; however, the severe OSA group had a trend of poor performance of WristO2 (Fig. 4C). Conversely, the coverage rates were 74.30 ± 18.16%, 79.26 ± 14.13%, 77.94 ± 15.48%, and 64.29 ± 19.08% for normal, mild, moderate, and severe OSA, respectively; these were variable regardless of the AHI (Fig. 4B, D).

**Table 2**

<table>
<thead>
<tr>
<th>Specification</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>2.28%</td>
</tr>
<tr>
<td>Bias</td>
<td>-0.16%</td>
</tr>
<tr>
<td>95% lower limit of bias</td>
<td>-4.63%</td>
</tr>
<tr>
<td>95% upper limit of bias</td>
<td>4.31%</td>
</tr>
<tr>
<td>Total data duration</td>
<td>672.3 hours</td>
</tr>
<tr>
<td>Valid data duration</td>
<td>490.8 hours</td>
</tr>
<tr>
<td>Data rejection rate</td>
<td>26.5%</td>
</tr>
</tbody>
</table>

GW4, Galaxy Watch 4; RMSE, root mean square error.

**Fig. 2.** Illustrations of overnight WristO2 traces of normal participants (A) and patients with severe OSA (B). Transient false recordings were observed intermittently (blue arrows). The black trace indicates the WristO2, and the gray-shaded area indicates the SpO2Ref with ±4% error. The gray arrows indicate examples of motion-related transient drop in the WristO2, and the black arrow points to the baseline shift and slow recovery to the baseline. Magnified views of red rectangular areas in (B) are shown in (C) and (D); these represent periodic desaturation events caused by apnea or hypopnea and relatively small changes in oxygen saturation (C) and large changes with the lowest saturation of 71% (D). WristO2, oxygen saturation derived from a wrist-worn re fle ctance pulse oximeter; OSA, obstructive sleep apnea; SpO2Ref, oxygen saturation according to the finger-attached transmittance pulse oximeter used as a reference.

### Association between AHI ≥15/h and GW4-ODI

ROC analysis to predict an AHI ≥15/h was performed to compare the GW4-ODI based on different thresholds, specifically 2%, 3%, and 4% (Fig. 5). The areas under the ROC curve (AUCs) were 0.890 (95% confidence interval [CI], 0.829-0.952), 0.908 (95% CI, 0.852-0.963), and 0.893 (95% CI, 0.832-0.953) for the thresholds of 2%, 3%, and 4%, respectively. Therefore, the GW4-ODI was defined as a 3% decrease in the WristO2 from baseline per hour.

Strong positive correlations were observed between the AHI and PSG-ODI (Pearson correlation, \( r = 0.951 \)) and between the AHI and GW4-ODI (\( r = 0.918 \)) (Fig. 6A, B). Bland-Altman plots were used to show agreement between the AHI and ODI obtained from PSG and GW4 (Fig. 6C, D). The biases were similar under both conditions, and the AHI was higher than both the PSG-ODI (mean, 9.28; SD, 10.26) and GW4-ODI (mean, 9.32; SD, 11.46). Participants with a difference of 10 or more between the AHI and GW4-ODI had a higher total hypopnea index (26.18 ± 12.54 vs. 8.82 ± 7.53, \( P < .001 \)) and lower saturation (81.84 ± 6.52 vs. 85.36 ± 9.35, \( P = .03 \)) than those with a difference less than 10.
Fig. 3. Bland-Altman density plot showing good agreement between the WristO2 and SpO2Ref (A). The bar indicates the number of samples. The black dashed lines indicate the upper and lower 95% limits of agreement calculated from the estimated error, ranging from -4.63% to 4.31%. When the saturation value is low, a large estimation error is more likely. Probability density functions of the estimation error (B) and averaged oxygen saturation (C) are shown. The bias of the WristO2 is -0.16%, with a standard deviation of 2.28%. WristO2, oxygen saturation derived from a wrist-worn reflectance pulse oximeter; SpO2Ref, oxygen saturation according to the finger-attached transmittance pulse oximeter used as a reference.
Parameters for predicting AHI ≥15/h were estimated according to the different cutoff values associated with the GW4-ODI \(a_i\). The highest sensitivity was observed with a cutoff value of GW4-ODI ≥5/h. A GW4-ODI ≥15/h and ≥20/h provided 100% specificity.

**Discussion**

Wearable devices are increasingly being used in healthcare to facilitate continuous real-life monitoring of users’ data.\(^{17,18}\) In this study, we evaluated the estimated SpO₂ obtained from a commercially available watch-type wearable device. Previous studies have shown that reflectance pulse oximetry can relatively accurately detect desaturation, except when measurements are performed at high altitudes.\(^{10-12}\) However, wrist-worn reflectance pulse oximetry has more limitations than transmittance pulse oximetry.\(^{19}\) Because they are susceptible to motion and noise artifacts, commercially available wearable smartwatches have their own algorithms to enhance PPG measurements.\(^{19-21}\)

**Accuracy of SpO₂ measurements during sleep**

Reflectance pulse oximetry of the GW4 showed an overall RMSE of 2.28% and a negligible bias of -0.16% during sleep. This performance is comparable to that of reported by previous studies of wrist-worn wearable devices monitoring SpO₂ that included an in-house device, the Withings ScanWatch and Garmin-branded devices.\(^{5,10-12}\) Individuals move and change their position unconsciously throughout the sleep cycle, which prevents proper measurement of SpO₂ because of a poor sensor-skin interface. Furthermore, in the clinical setting, such as a sleep laboratory, sleep apnea induces oxygen desaturation and subsequent arousals; which are usually followed by body or limb movements. False recordings of SpO₂ frequently occur because of venous pulsations,\(^{22}\) position changes,\(^{15}\) respirations,\(^{23}\) or issues with the sensor-skin interface.\(^{24}\) Empirically, we found that recorded SpO₂ measurements were much more accurate when

---

**Fig. 4.** The obtained WristO₂ according to different AHI values (A, B) and OSA severity (C, D) on scatter plots and bar plots. Each dot represents one participant. This image shows a non-linear correlation between the RMSE and the AHI (A). There was no correlation found between the coverage rate and the AHI (B). WristO₂, oxygen saturation derived from a wrist-worn reflectance pulse oximeter; AHI, apnea-hypopnea index; OSA, obstructive sleep apnea; RMSE, root mean square error.

**Fig. 5.** Receiver-operating characteristic curves of the ODI with different thresholds for predicting an AHI ≥15/h. ODI: oxygen desaturation index; AUC, area under the curve; AHI: apnea-hypopnea index.
participants adhered to the manufacturer’s instructions to wear the device as tightly as possible above the wrist bone.

Although most sampled data from previous studies had an SpO2 range of 90%–100%, this study included participants presenting a wide range of SpO2, with the lowest SpO2Ref of 54%. The Bland-Altman density plot showed good agreement between the WristO2 and SpO2ref, with an acceptable error in the range of 70%–100% of the average saturation (Fig. 3A). Poor performance of WristO2 with profound desaturation may be caused by arousal from sleep due to apnea or hypopnea. It is possible that some motion-derived artifacts in PPG waveforms coincide with the changes in the accelerometric parameters. However, we found that the coverage rate was variable regardless of the OSA severity; that is, low saturation itself is related to increases in error. This result is consistent with that of previous reportings showing a decrease in accuracy with a lower saturation level.25,26 Although the GW4 showed a relatively imprecise estimation of SpO2 for participants with very severe OSA, it demonstrated generally good performance for the other 3 groups.

Table 3

<table>
<thead>
<tr>
<th>ODI cutoff value</th>
<th>Sensitivity, %</th>
<th>Specificity, %</th>
<th>PPV, %</th>
<th>NPV, %</th>
<th>Accuracy, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>≥5/h</td>
<td>89.7</td>
<td>64.1</td>
<td>78.8</td>
<td>80.6</td>
<td>79.4</td>
</tr>
<tr>
<td>≥10/h</td>
<td>74.1</td>
<td>89.7</td>
<td>91.5</td>
<td>70.0</td>
<td>80.4</td>
</tr>
<tr>
<td>≥15/h</td>
<td>56.9</td>
<td>100</td>
<td>100</td>
<td>60.9</td>
<td>74.2</td>
</tr>
<tr>
<td>≥20/h</td>
<td>48.3</td>
<td>100</td>
<td>100</td>
<td>56.5</td>
<td>69.1</td>
</tr>
</tbody>
</table>

GW4, Galaxy Watch 4; ODI, oxygen desaturation index; PPV, positive predictive value; NPV, negative predictive value.

The AASM Sleep Apnea Definitions Task Force recommends the use of ≥3% (instead of ≥4%) as the oxygen desaturation criterion to define hypopnea.27 However, these guidelines do not specify a definition of the ODI. Using the ODI for screening OSA could be an alternative technique for measuring the AHI, but it is difficult to directly compare the diagnostic performance of the ODI across studies because different definitions and software have been used.9,28 To the best of our knowledge, no previous study has attempted to calculate the ODI using wrist-worn reflectance pulse oximetry during sleep. Instead, wrist-worn wearable devices with transmittance pulse oximetry attached to the finger, such as WatchPAT or Pulsox-300i, have been verified with outstanding discrimination.29-31 The ROC analysis in our study found that a 3% decrease in the WristO2 from baseline showed outstanding discrimination, with an AUC of 0.908 to
predict an AHI ≥ 15/h. The GW4-ODI ≥ 5/h was chosen as a predictor for AHI ≥ 15/h because of its high sensitivity. However, it should be noted that GW4 is not a screening device for OSA due to its high rejection rate.

It is worth noting that the discrimination between the GW4-ODI and AHI of individuals requires different criteria. The GW4-ODI underestimated the AHI by 9.32 ± 1.146 in our study. Furthermore, a similar degree of bias between the AHI and ODI, defined as desaturation of 4% or more, was reported in a previous study. Although normal participants showed good agreement between the AHI and GW4-ODI, more severe OSA was related to increased bias and greater dispersions in the Bland-Altman plot (Figure 6D). The bias and variability drastically increased in the severe OSA group. When participants were divided into 2 groups according to the difference between the AHI and GW4-ODI, the group with a difference of ≥10/h had a higher total hypopnea index. This suggests that the GW4-ODI underestimates hypopnea because the ODI is unable to detect hypopneas without desaturation. This group with a difference of ≥10/h also included participants with a lower SpO2 nadir, indicating that patients with more severe OSA may be more susceptible to prediction error.

In the previous studies that predict OSA using a finger-attached pulse oximeter, the sensitivity and specificity varied with various ODI definitions and target severities. The present study showed lower accuracy when predicting an AHI ≥ 15/h compared to prior studies using an ODI of 3% and reported sensitivity ranging from 86.1% to 96% and specificity from 89% to 94%. The relatively low specificity of the GW4-ODI could be attributed to the vulnerability to artifacts compared to the ODI derived from a firmly attached transmittance pulse oximeter.

**Limitations**

This study had several limitations that require particular attention. First, the error in the WristO2 was larger than the true SpO2 level. During this investigation, transmittance pulse oximetry was used as a reference instead of co-oximetry, as we were not able to easily perform invasive co-oximetry measurements during sleep in this study. Second, the WristO2 is inherently vulnerable to artifacts. An algorithm will enable the correction of the afore-mentioned errors, such as a sudden and transient drop or baseline shift with the slow in contact pressure caused by non-supine body positions, movements, and the discarding of data less than 70% of saturation (Supplementary Fig. 1). Therefore, it is important to minimize such artifacts by following strict guidelines or an error-correction algorithm. Improvement of the algorithm will enable the correction of the afore-mentioned errors, such as a sudden and transient drop or baseline shift with the slow recovery of the WristO2, thus allowing improvement of the accuracy or coverage rate. Third, the weakness of wrist-worn reflectance pulse oximetry is its declining accuracy at low saturation levels. Although the accuracy of the device with a low SpO2 range was relatively well-preserved in a previous study, the difference from our study was that participants were in the semi-supine position and remained still during measurements. When SpO2 decreased during sleep, the WristO2 tended to have a lower value than the reference, thus underestimated the true SpO2.

**Conclusions and future work**

This study characterized the performance of the GW4 when measuring the estimated SpO2 during sleep. The accuracy of the GW4 complies with the FDA and ISO standards. The major strength of this study is that it evaluated the performance of wrist-worn reflectance pulse oximetry when measuring the SpO2 of patients with sleep apnea and of some patients with profound desaturation.

In the field of sleep medicine, the development of precise and convenient SpO2 measurements should be associated with the screening and monitoring of OSA. With improvements in the accuracy and error correction algorithms of wrist-worn wearable devices, it is expected that future devices will provide more accurate and reliable information to patients and clinicians alike.

**Authors’ contribution**

HJ, DK, and EYJ conceived and designed the study. WL, HS, and JS participated in data curation. HJ and DK conducted data analyses and drafted the manuscript. EYJ and JC reviewed and edited the manuscript. All authors approved the final version of the manuscript.

**Declaration of conflict of interest**

This study was funded by Samsung Electronics. Some authors (HJ, WL, HS, JS, JC, and EYJ) are affiliated with Samsung Electronics. Open access to the data was provided and the research was monitored independently by the 2 institutions.

**Funding**

This study was supported by a grant from Samsung Medical Center (grant OTC1190671).

**Supplementary materials**

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.sleh.2022.04.003.

**References**


Please cite this article as: H. Jung et al., Performance evaluation of a wrist-worn reflectance pulse oximeter during sleep, Sleep Health (2022), https://doi.org/10.1016/j.sleh.2022.04.003


https://doi.org/10.1016/j.sleh.2022.04.003