Performance Accuracy of Wrist-Worn Oximetry and Its Automated Output Parameters for Screening Obstructive Sleep Apnea in Children

Mon Ohn1,2,3,4, Kathleen J Maddison4,5,6, Julie Nguyen1,3,7, Daisy Evans3,4,7,8, Natasha Bear9,10, R. Nazim Khan3,4,11, Peter R Eastwood12, Britta S. von Ungern-Sternberg3,4,7,13, Andrew C Wilson12,14, Jennifer H Walsh4,5,6

1Department of Respiratory and Sleep Medicine, Perth Children’s Hospital, Nedlands, WA,
2Division of Pediatrics, Medical School, University of Western Australia, Perth, WA,
3Perioperative Medicine Team, Perioperative Care Program, Telethon Kids Institute, Nedlands, WA,
4Institute for Paediatric Perioperative Excellence, The University of Western Australia, Perth, WA,
5Centre for Sleep Science, School of Human Sciences, The University of Western Australia, Perth, WA,
6West Australian Sleep Disorders Research Institute, Department of Pulmonary Physiology & Sleep Medicine, Sir Charles Gairdner Hospital, Nedlands, WA,
7Department of Anaesthesia and Pain Medicine, Perth Children's Hospital, Nedlands, WA,
8School of Physics, Mathematics and Computing, The University of Western Australia, Perth, WA,
9Department of Clinical Research and Education, Child and Adolescent Health Service, Nedlands, WA,
10Institute for Health Research, Notre Dame University, Perth, WA,
11Department of Mathematics and Statistics, The University of Western Australia, Perth, WA,
12Health Futures Institute, Murdoch University, Perth, WA,
13Division of Emergency Medicine, Anaesthesia and Pain Medicine, Medical School, The University of Western Australia, Crawley, WA,
14Wal-yan Respiratory Research Centre, Telethon Kids Institute, Nedlands, WA, Australia

Received June 15, 2023
Revised August 2, 2023
Accepted August 18, 2023

Address for correspondence
Mon Ohn, FRACP
Department of Respiratory and Sleep Medicine,
Perth Children’s Hospital, 15 Hospital Avenue,
Locked Bag 2010, Nedlands, WA 6009, Australia
Tel: +61-8-6456-2222
Fax: +61-8-6456-2075
E-mail: mon.ohn@health.wa.gov.au

*Part of this study was presented by Dr. Mon Ohn in abstract form at the 2022 Sleep Down Under Conference, October 8-11, 2022, Brisbane, Australia.

Objectives: Obstructive sleep apnea (OSA) increases the risk of perioperative adverse events in children. While polysomnography (PSG) remains the reference standard for OSA diagnosis, oximetry is a valuable screening tool. The traditional practice is the manual analysis of desaturation clusters derived from a tabletop device using the McGill oximetry score. However, automated analysis of wearable oximetry data can be an alternative. This study investigated the accuracy of wrist-worn oximetry with automated analysis as a preoperative OSA screening tool. Methods: Healthy children scheduled for adenotonsillectomy underwent concurrent overnight PSG and wrist-worn oximetry. PSG determined the obstructive apnea-hypopnea index (OAHI). Oximetry data were auto-analyzed to determine 3% oxygen desaturation index (ODI3) and visually scored as per McGill criteria. The logistic regression model assessed the predictive performance of ODI3 for detecting the presence and severity of OSA after adjusting for covariates. Results: Seventy-six children (34 females), aged (mean±standard deviation) 5.7±1.6 years were classified, based on PSG-derived OAHI, as no OSA (n=31), mild (n=31), and moderate-severe OSA (n=14). Oximetric ODI3 was identified as the sole predictor of moderate-severe OSA (OAHI≥5 events/h) (odds ratio 1.38, 95% confidence interval 1.15, 1.65, p=0.001). The best diagnostic performance was at ODI3=5 events/h (78.6% sensitivity, 75.8% specificity [receiver operating characteristic-area under the curve {ROC-AUC}=0.857]). ODI3 was also more sensitive than the McGill oximetry score in diagnosing moderate-severe OSA (78.6% by ODI3 vs. 33.0% by McGill). The performance was suboptimal for any level of OSA (OAHI≥1 event/h) (75.6% sensitivity, 61.3% specificity [ROC-AUC=0.709]). Conclusions: Wrist-worn oximetry-derived automated ODI3 can reliably identify moderate-severe OSA in children undergoing adenotonsillectomy, making it a potentially useful preoperative OSA screening tool.

Keywords: Children; Oximetry; Oxygen desaturation index; Obstructive sleep apnea; Adenotonsillectomy.
INTRODUCTION

Childhood obstructive sleep apnea (OSA) is a common sleep-related breathing disorder with a prevalence of 3%–5%.\(^1\) OSA is characterized by recurrent, partial, or complete upper airway obstruction during sleep, causing sleep disruption and gas exchange abnormalities. OSA is a well-known risk factor for perioperative adverse events in children undergoing routine surgical procedures such as adenotonsillectomy.\(^2\) The risk could be as high as 50% in children with OSA.\(^3,4\) Thus, identifying OSA is vital for perioperative management.

While diagnostic polysomnography (PSG) is the reference standard for OSA, it is not readily available for most patients.\(^5,6\) Overnight continuous recording of peripheral oxygen saturation by oximetry has been proposed as an OSA screening tool.\(^7\) Tabletop oximetry is traditionally used in pediatric settings. Validated scoring tools, such as the McGill oximetry score, based on visual detection of ‘clusters’ of desaturations, are recommended for predicting the presence and severity of OSA.\(^5,7\)

Technologies have led to the development of wrist-worn, portable oximeters with automated analysis options. Unlike PSG, this technology is readily accessible, economical, offers easy data acquisition and interpretation, and requires minimal training. One such option is a wrist-worn oximeter (WristOx 3150, Nonin Medical, Plymouth, MN, USA). The oximetry report generated by its associated nVISION \(^8\) (version 6.5.1, Nonin Medical, Plymouth, MN, USA) oxygen saturation (SpO\(_2\)) analysis software provides a graphical summary of oxygen traces and various markers of hypoxemia: oxygen desaturation index (ODI); mean SpO\(_2\); the lowest value of oxygen saturation (SpO\(_2\) nadir); and the cumulative time in minutes spent SpO\(_2\) <90% (T90 [min]). Among those metrics, ODI is the most reported parameter used for screening OSA.\(^8,13\)

ODI is defined as the number of desaturation events per hour where the SpO\(_2\) decreases by a defined percentage from the baseline level. The desaturation criteria are typically ≥3% (ODI3) or ≥4% (ODI4). However, there is no consensus regarding using ODI3 or ODI4. To date, studies have examined the utility of wrist oximetry in pediatric OSA only using the ODI4 criteria.\(^14,15\) However, oxygen desaturation of ≥3% is in accordance with the standard polysomnographic scoring rule for hypopneas\(^16\) and may correlate better with ODI3. Therefore, ODI3 for OSA screening warrants further investigation.

Our study aims to investigate the potential of a wrist-worn oximeter with automated analysis as a tool for preoperative screening of OSA. Therefore, we assessed the efficacy of oximetric parameters in diagnosing OSA. Additionally, we investigated the relationship between traditional McGill scores and automated ODI3 in screening for OSA.

METHODS

Design and setting

This prospective, observational study was conducted at Perth Children’s Hospital, Western Australia. This study was a pre-planned sub-study of larger studies examining perioperative outcomes and alternative methods to PSG for perioperative screening for OSA; OSATS2 (The Obstructive Sleep Apnea Study: Making Tonsillectomies Safer) and NIGHT-OWL (New prediction strategies to help test for obstructive sleep apnea and correlate patient risk with the level of care).

Institutional ethics approvals were obtained before the commencement of the study (CAHS HREC Approval RGS 0000000014 & RGS00000003677 and UWA EC Notification RA/4/20/4012 & RA/4/20/6035). In addition, the trials were prospectively registered (ACTRN 12617001503314: ANZCTR - Registration and ACTRN 12620001190998: ANZCTR - Registration). Informed consent was obtained from the primary carer of each participant. All children simultaneously underwent PSG and wrist-worn oximetry (Fig. 1) in the sleep laboratory 2–4 weeks before the scheduled surgery.

Study participants

Eighty children (aged 1–8 years) undergoing adenotonsillectomy were prospectively included. Exclusion criteria were significant comorbidities such as significant respiratory disease (e.g., cystic fibrosis), cardiovascular disease or syndromic conditions, and less than 4 hours of total recording time of wrist-worn oximetry.

Polysomnography

Standard overnight PSG was performed at the Centre for Sleep Science, University of Western Australia. PSG data were acquired and displayed using a Compumedics Grael system...
Wrist Oximetry’s ODI3 for Childhood OSA

(Compumedics, Victoria, Australia). The following signals were recorded: electroencephalogram; electrooculogram; submental electromyogram; electrocardiogram; thoracoabdominal excursion (plethysmography); nasal airflow (pressure transducer and thermistor); snoring sound; SpO2 by an integrated pulse oximeter; transcutaneous carbon dioxide; leg movements; body position; and decibel level.

A sleep technician monitored the recordings and video for the study duration. PSG data were scored manually by experienced pediatric sleep technicians using Profusion software (Profusion PSG4, Compumedics, Victoria, Australia). The studies were scored according to the American Academy of Sleep Medicine Manual for the Scoring of Sleep and Associated Events and validated by the Royal Australasian College of Physicians accredited pediatric sleep physician.7

The presence and severity of childhood OSA were defined using the OAHII, which is calculated as the total number of apnea and hypopneas divided by the total sleep time in one hour. Based on OAHII, OSA was categorized as mild (1≤OAHII<5 events/h), moderate (5≤OAHII<10 events/h), and severe (OAHII≥10 events/h). PSG-derived ODI3 was defined as the number of oxygen desaturation events where the SpO2 decreased by ≥3% or more.

Wrist-worn oximetry

Participants wore a WristOx2 3150 (Nonin Medical, Plymouth, MN, USA) on the right or left wrist with a soft fingertip sensor (Nonin Soft Sensor 8000S, Nonin Medical, Plymouth, MN, USA) while simultaneously undergoing PSG. The recording settings included a 1-second sampling rate, <2–3 seconds averaging time (2–3 beat averaging), 0.1% SpO2 signal amplitude resolution, desaturation drop for an event of ≥3%, and minimum event duration of 3 seconds. These technical specifications met the recommendations of the American Academy of Sleep Medicine and the Australasian Sleep Association.7

Oximetry data were downloaded and automatically scored by custom software (nVISION data management software, version 6.5.1, Nonin Medical, Plymouth, MN, USA) which excluded areas of SpO2 artifact. Automated reports (Fig. 2) were obtained, displaying indices including ODI3, T90 (min), T90 (%), basal SpO2, nadir SpO2, and respiratory and graphical representations (SpO2, pulse rate, and signal quality). ODI3 was defined as the number of oxygen desaturation events ≥3 seconds where the SpO2 decreased by ≥3%. T90 (minutes) was the time in minutes spent with SpO2 <90%, and T90 (%) was the percentage of recording time with SpO2 <90%.

McGill oximetry scores

The study sleep physician graded the McGill scores by visually inspecting the Nonin wrist-worn oximetry tracing per Australasian Sleep Association recommendations.7 A “desaturation” was defined as a >4% fall in oxygen saturation from the preceding baseline. A cluster was defined as >5 desaturations within 30 minutes. McGill scores of 1–4 were assigned by the number of desaturation events below 90% (McGill score 1: <3 oxygen desaturations below 90%; McGill score 2: ≥3 desaturations below 90% but ≤3 desaturations below 85%; McGill score 3: >3 desaturation below 80% but ≤3 desaturations below 80%); McGill score 4: >3 desaturations less than 80%).19

McGill scores of 2–4 indicate abnormal or positive oximetry for predicting OSA; a score of 1 is interpreted as inconclusive. An abnormal McGill score (>2) corresponds to an OSA of at least moderate severity; however, a score of 1 does not necessarily exclude OSA.19

Clinical variables

The parents of the children completed a Paediatric Sleep Questionnaire,20 and demographic information was collected prior to the sleep study.

Statistical analysis

Descriptive statistics for demographic, PSG, and wrist-worn oximetry data were obtained. For continuous data, the mean±standard deviation was used if the distribution was symmetric and the median (interquartile range) if the distribution was skewed. Categorical data were presented as frequencies and percentages. The OSA predictive ability of wrist-worn oximetry was assessed by logistic regression analysis using the purposeful selection of variables method.21

Oximeter-derived automated parameters (ODI3, T90min, T90%, minimum SpO2, mean SpO2, mean event duration, and mean pulse rate) and demographic confounders (age, sex, and body mass index [BMI] z-score) were considered covariates. Insignificant covariates were removed one at a time. The final model reported estimates of model coefficients, mean, and 95% confidence interval (95% CI) of odds ratio (OR). Two models were assessed: the first, a binary presence or absence of OSA (PSG OAHII≥1 event/h) and the second, a binary moderate-severe OSA (PSG OAHII≥5 events/h) or normal-mild OSA. Statistical significance was assessed at p=0.05 (5%).

Receiver operating characteristic (ROC) curves and the area under the curve (AUC) were used to assess the diagnostic ability of the models. AUC of <0.5 indicated no discrimination; 0.7–0.8 was acceptable; 0.8–0.9 was excellent; >0.9 was perfect.22 Sensitivity, specificity, positive predictive value, and negative predictive value were reported for different probability cutoffs. The preferred cutoff was determined considering the need for high sensitivity and balanced specificity and was
verified by maximizing the Youden Index.\textsuperscript{23} The maximum value of the Youden Index is 1 (perfect test), and the minimum is 0, which denotes no diagnostic value.\textsuperscript{24} The distributions of oximetry-derived automated ODI3 values across different McGill oximetry score groups were compared graphically using box and whisker plots. No previous pediatric data on wrist-worn oximetry-derived automated ODI3 exists; therefore, a formal sample size calculation was not attempted, and a convenience sample size was selected. Stata Standard Edition (College Station, TX, USA) and R version 4.2 (R Foundation for Statistical Computing, Vienna, Austria; https://www.R-project.org/) were used for the statistical analysis.

RESULTS

Participant characteristics

This prospective study included a total of 80 participants scheduled to undergo adenotonsillectomy. Four participants were excluded because of insufficient data captured (less than 4 hours). Of the remaining 76 children (age mean±standard
deviation] 5.7±1.6 years, 34 females), 31 (41%) had no OSA (OAHI<1 event/h), 31 (40.8%) had mild OSA (1≤OAHI<5 events/h) and 14 (18.4%) moderate-severe OSA (OAHI≥5 events/h). Participant characteristics and PSG parameters for each OSA group are summarized in Table 1.

Predictive ability of the wrist-worn oximeter for detecting moderate-severe OSA (OAHI≥5 events/h)

The wrist-worn oximetry-derived automated ODI3 and sex were the only statistically significant covariates for moderate-severe OSA in the multiple logistic regression. However, the number of moderate-severe OSA cases by sex was small (nine females and five males), so sex was excluded from the model. Moreover, although sex was significant in the multiple logistic regression model, the finding was likely to be a spurious effect.25,26 Therefore, the final model included ODI3 (OR 1.38, 95% CI 1.15, 1.65, p=0.001) as the sole predictor for moderate-severe OSA.

Table 2 shows the sensitivity and specificity at different ODI3 cutoff values. At a cutoff of ODI3=1 event/h, 100% sensitivity was achieved, but a low specificity of 9.7%. At a cutoff of ODI3=5 events/h, 78.6% sensitivity and 75.8% specificity were achieved. The maximum Youden Index (0.618) reached an ODI3 cutoff of 6.4 events/h, showing 71.4% sensitivity and 90.3% specificity. Fig. 3 (red line) shows the ROC curve for this moderate-severe OSA analysis (AUC=0.857). Fig. 4 shows the probability of a moderate-severe OSA diagnosis as assigned by the model for each patient.

Predictive ability of the wrist-worn oximeter for detecting any OSA (OAHI≥1 event/h)

Wrist-worn oximetry-derived ODI3 (OR 1.19, 95% CI 1.01, 1.41, p=0.039) and mean SpO₂ (OR 2.31, 95% CI 1.10, 4.82, p=0.026) were significant covariates for identifying the pres-

Table 1. Patient characteristics and summary of sleep and breathing-related parameters from polysomnography

<table>
<thead>
<tr>
<th>Participant characteristics</th>
<th>All (n=76)</th>
<th>No OSA (OAHI&lt;1) (n=31)</th>
<th>Mild OSA (1≤OAHI&lt;5) (n=31)</th>
<th>Moderate-Severe OSA (OAHI≥5) (n=14)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, yr</td>
<td>5.7±1.6</td>
<td>6.2±1.6</td>
<td>5.1±1.6</td>
<td>5.7±1.4</td>
</tr>
<tr>
<td>Female, n (%)</td>
<td>34 (45)</td>
<td>11 (35)</td>
<td>14 (45)</td>
<td>9 (64)</td>
</tr>
<tr>
<td>BMI z-score</td>
<td>0.36±1.52</td>
<td>0.53±1.30</td>
<td>0.20±1.66</td>
<td>0.36±1.75</td>
</tr>
<tr>
<td>PSQ</td>
<td>0.59±0.21</td>
<td>0.54±0.24</td>
<td>0.63±0.17</td>
<td>0.61±0.19</td>
</tr>
</tbody>
</table>

Sleep architecture

Total sleep time, min       | 501 (441–537) | 503 (432–525) | 501 (456–555) | 497 (420–543) |
Sleep onset latency, min    | 34.5 (12–63)  | 45 (12–66)     | 32 (9–51)     | 39.8 (14–80)  |
Sleep efficiency, %         | 85.8 (79.2–89.3) | 85.2 (75.9–87.5) | 88.1 (82.4–91.2) | 81 (78.7–86.8) |
WASO, min                  | 43.8 (24.8–68.5) | 46.5 (24.5–82) | 34.5 (24–61.5) | 48.8 (31–64)  |
Arousal index, events/hour  | 17.6 (12.9–23) | 13.4 (11.7–17.8) | 18.1 (13.7–22.3) | 24.3 (23–31.8) |

Breathing-related

AHI, events/hour            | 2.85 (1.25–5.2) | 1 (0.8–2.8) | 3.4 (2.1–4.8) | 12.1 (9.4–16.3) |
OAHI, events/hour           | 1.27 (0.6–3.55) | 0.50 (0.12–0.75) | 2.16 (1.3–3.1) | 10.05 (5.9–14.8) |
CAHI, events/hour           | 0.9 (0.4–2.2)  | 0.7 (0.3–2.1) | 0.7 (0.3–2.2) | 1.8 (1.1–3)     |
Mean SpO₂, %                | 98 (97–98)     | 98 (97–98)   | 98 (97–98)    | 97 (96–98)      |
Nadir oxygen, %             | 91 (89–93.5)   | 92 (90–94)   | 91 (90–94)    | 88.5 (84–91)    |
Lowest SpO₂ in REM, %       | 93.5 (90–95)   | 94 (91–96)   | 94 (90–95)    | 89 (84–93)      |
Lowest SpO₂ in NREM, %      | 92 (91–94)     | 92 (91–94)   | 93 (91–94)    | 92 (89–92)      |
ODI3, events/hour           | 0.7 (0.3–2.3)  | 0.5 (0.2–1.4) | 0.6 (0.2–1.65) | 2.9 (1.6–10.4)  |
T90, %                     | 0 (0–0)        | 0 (0–0)      | 0 (0–0)       | 0 (0–0)         |
T90, min                   | 0 (0–0.1)      | 0 (0–0)      | 0 (0–0.1)     | 0.15 (0–1.8)    |
Average CO₂ in sleep, mmHg | 44.8 (41.7–47.5) | 44 (38.9–46.9) | 45.05 (42.4–47.6) | 44.85 (42.2–48.7) |

Values are presented as mean±standard deviation, n (%), or median (interquartile range). OSA, obstructive sleep apnea; OAHI, obstructive apnea-hypopnea index; BMI, body mass index; PSQ, Pediatric Sleep questionnaire; WASO, wake after sleep onset; AHI, apnea-hypopnea index; CAHI, central apnea-hypopnea index; SpO₂, saturation of peripheral oxygen; REM, rapid eye movement; NREM, non-rapid eye movement; ODI3, the number of oxygen desaturation events where the SpO₂ decreases by 3%; T90 (%), cumulative time percentage with SpO₂<90%; T90 (min), cumulative time in minutes spent SpO₂<90%; CO₂, carbon dioxide
ence of OSA (mild, moderate or severe) in the multiple regression model. A sensitivity of 75.6% and a specificity of 61.3% were achieved at a probability cutoff of 0.55 (Youden Index=0.368). Fig. 3 (blue line) shows the ROC curve for this model (AUC=0.709).

**Relationship between wrist-worn oximeter-derived automated ODI3 and McGill oximetry score**

The ODI3 values for the McGill score groups 1 and 2–4 were (median [IQR] 2.75 [1.95, 4.00]) and (median [IQR] 6.55 [4.80, 13.60]), respectively (Fig. 5). There was a significant ODI3 overlap between the McGill score 1 and 2 groups; consequently, no clear ODI3 cutoffs were identified for different McGill score oximetry groups. However, 75% of the McGill score 1 group cases fell below ODI3=4 events/h, and over 75% of the McGill score 2–4 group cases fell above this threshold. In addition, there was an apparent increase in ODI3 for higher McGill scores. All patients in the McGill score 3 and 4 groups had ODI3>4 events/h; however, the number was small, with only six cases in groups 3 and 4 combined. Therefore, these groups were combined with group 2 in the summary.

Of the moderate-severe OSA cases, 78.6% were correctly diagnosed by an ODI3 cutoff of ≥5, compared with 66.7% by the McGill oximetry score 2–4 group. Table 3 shows the number of OSA cases diagnosed by wrist-worn oximetry-derived automated ODI3 and the McGill oximetry score. McGill scores could not be determined in ten participants because of the presence of artifactual data for the recording period.

**DISCUSSION**

This study assessed the ability of a wrist-worn oximeter and its automated parameters to identify OSA in children during the preoperative period. The key finding was that the wrist-worn oximetry-derived automated ODI3 effectively identified children with moderate-severe OSA. Additionally, we did not find any clear ODI3 cutoffs for different McGill oximetry groups.

ODI3 was the only automated oximetry parameter significantly associated with diagnosing moderate-severe OSA. At an ODI3 cutoff of 1 event/hr (i.e., the presence of OSA of any severity), OSA cases could be correctly identified with 100% sensitivity and no false negatives. However, specificity was low at 9.7%, leading to many false-positive results. At a lower ODI3 cutoff, the predictive performance was suboptimal. The best diagnostic performance was at the ODI3 cutoff of 5 events/hr, obtaining a balanced sensitivity and specificity above 75% and a high negative predictive value (94.0%). In addition, ROC-AUC for the moderate-severe OSA model was 0.857, which indicates an excellent ability to identify moderate-severe OSA cases.

When we included mild OSA cases in the analysis to identify the presence of any OSA (mild, moderate, or severe), ODI3 and mean SpO2 were found to be significant. Although we did not observe adequate diagnostic efficacy, ROC-AUC showed acceptable discriminative power at 0.709. Our results align

**Table 2. A model for moderate-severe obstructive sleep apnea diagnosis by wrist-worn oximetry-derived automated ODI3**

<table>
<thead>
<tr>
<th>ODI3</th>
<th>Sensitivity, %</th>
<th>Specificity, %</th>
<th>Accuracy, %</th>
<th>PPV, %</th>
<th>NPV, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100</td>
<td>9.7</td>
<td>26.3</td>
<td>20.0</td>
<td>100</td>
</tr>
<tr>
<td>2</td>
<td>92.9</td>
<td>21.0</td>
<td>34.2</td>
<td>21.0</td>
<td>92.9</td>
</tr>
<tr>
<td>3</td>
<td>92.9</td>
<td>50.0</td>
<td>57.9</td>
<td>29.6</td>
<td>96.9</td>
</tr>
<tr>
<td>4</td>
<td>85.7</td>
<td>61.3</td>
<td>65.8</td>
<td>33.3</td>
<td>95.0</td>
</tr>
<tr>
<td>5</td>
<td>78.6</td>
<td>75.8</td>
<td>76.3</td>
<td>42.3</td>
<td>94.0</td>
</tr>
<tr>
<td>6.4*</td>
<td>71.4</td>
<td>90.3</td>
<td>86.8</td>
<td>62.5</td>
<td>93.3</td>
</tr>
</tbody>
</table>

*Youden index maximum. ODI3, oxygen desaturation event, the number of oxygen desaturation events where the SpO2 decreases by 3%; PPV, positive predictive value; NPV, negative predictive value

**Figure 3.** Receiver operating characteristic (ROC) curves. ROC curve for moderate-severe obstructive sleep apnea predicted by wrist-worn oximetry-derived automated ODI3 (red line). ROC curve for any obstructive sleep apnea predicted by wrist-worn oximetry-derived automated ODI3 and mean SpO2 (combined predictors) (blue line). PSG, polysomnography; OAH, obstructive apnea-hypopneea index, events/h; AUC, area under the curve; ODI3, 3% oxygen desaturation index; SpO2, oxygen saturation.
Wrist Oximetry’s ODI3 for Childhood OSA

with previous literature, consistently showing that positive oximetry identifies moderate-severe OSA but does not reliably rule out mild OSA. In children, episodes of upper airway obstruction during sleep, characteristic of OSA, may lead to brief arousal from sleep without desaturations. Therefore, oximetry is known to have a reduced capacity for detecting OSA in children with high arousal indices without desaturation.

**Figure 4.** Probability of moderate-severe obstructive sleep apnea (OSA) diagnosis based on the logistic regression model using wrist-worn oximetry-derived automated ODI3 (color) plotted with polysomnography-derived OAHI. The vertical line shows the chosen threshold for moderate-severe OSA diagnosis (OAHI≥5 events/h); participants to the right of this line are true positives. The horizontal line shows a chosen wrist-worn oximetry-derived automated ODI3 cutoff (ODI3=5 events/h); participants above this line are assigned positive by the model. Adjusting the position of the horizontal and vertical lines will adjust the sensitivity and specificity of the model. Data in the bottom right quadrant of the figure are false negatives; all were moderate OSA cases which were not captured at the ODI3=5 events/h cutoff. Data in the top left quadrant are false positives—all were no OSA or mild OSA cases, which were falsely assigned as positive at the ODI3=5 events/h cut point. PSG, polysomnography; OAHI, obstructive apnea-hypopnea index, events/h; ODI3, the number of oxygen desaturation events where the SpO2 decreases by 3%.

**Table 3.** Polysomnography confirmed OSA cases diagnosed by wrist-worn oximetry-derived automated ODI3 and McGill oximetry score

<table>
<thead>
<tr>
<th>McGill oximetry score</th>
<th>No-Mild OSA (n=62)</th>
<th>Moderate-Severe OSA (n=14)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score 1</td>
<td>40 (74.1)</td>
<td>4 (33.3)</td>
</tr>
<tr>
<td>Score 2–4</td>
<td>14 (25.9)</td>
<td>8 (66.7)</td>
</tr>
<tr>
<td>NA, n</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>ODI3 cut-off</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ODI3&lt;5</td>
<td>47 (75.8)</td>
<td>3 (21.4)</td>
</tr>
<tr>
<td>ODI3≥5</td>
<td>15 (24.2)</td>
<td>11 (78.6)</td>
</tr>
</tbody>
</table>

Values are presented as n (%) or n. OSA, obstructive sleep apnea; ODI3, oxygen desaturation event, the number of oxygen desaturation events where the SpO2 decreases by 3%; NA, not available.

**Figure 5.** Distribution of wrist-worn oximetry-derived automated ODI3 values grouped by McGill oximetry score 1 and 2-4. McGill oximetry score Groups 3 and 4 were merged with Group 2 since they have less than five observations and were not informative separately. ODI3, the number of oxygen desaturation events where the SpO2 decreases by 3%.
we may not have enough observations to draw valid conclusions about possible OSA predictors other than ODI3.

The current study is the first to explore the utility of wrist-worn oximetry-derived automated ODI3 for detecting OSA in children. Two previous studies have examined the diagnostic performance of wrist-worn oximetry, exploring the utility of an automated ODI4. Ma et al.14 assessed ODI4 in diagnosing OSA in 32 children. The authors reported the best diagnostic ability at a notably high apnea-hypopnea index (AHI) >20 events/h: 83.3% sensitivity, 92.3% specificity, and AUC 0.929. In addition, AUC for standard AHI cutoffs >1, 5, 10 events/h were <0.80 and sensitivities <65%.14 A more recent study by Thavagnanam et al.15 evaluated the value of Nonin WristOx2 (Nonin Medical) oximetry-derived automated ODI4 in 162 children. The authors reported high specificity of 96.7% but low sensitivity of 50.0% in detecting moderate-severe OSA cases (ROC-AUC=0.852 at ODI4≥2 events/h).15

In contrast, this study analyzed all automated parameters and possible covariates. Consistent with prior studies,14,30 we report similar accuracy issues in detecting mild OSA cases. Nonetheless, children with mild OSA have only slightly increased risk in the perioperative setting. However, it is essential to reliably identify children with moderate-severe OSA because such a diagnosis may change the perioperative decision pathway (e.g., choice of regional vs. tertiary hospital).3

We also explored the relationship between the wrist-worn oximetry-derived automated ODI3 and the McGill oximetry score. Our findings indicated that an ODI3≥5 events/h was more sensitive than the McGill oximetry score in identifying moderate-severe OSA cases (78.6% by automated ODI3 vs. 66.7% by the McGill oximetry score 2–4 group). However, the comparison between the two methods could not be made for all groups because of the small sample size in groups 3 and 4.

Although our findings did not fully support automated ODI3 as a replacement for the McGill scoring system, it may still be a practical alternative in the preoperative setting. Graphical interpretation using McGill scoring criteria can be time-consuming and require training and practice for perioperative physicians. In contrast, wrist-worn oximetry-derived ODI3 provides an automated analysis that can derive a score despite periods of missing data due to artifacts. In addition, this automated nature may minimize potential clinician bias in interpretation, resulting in reliable and reproducible reports.

Our study has few limitations. Firstly, the study only included otherwise healthy children scheduled for adenotonsillectomy, which may limit the generalizability of our findings to other procedures. Secondly, while all children underwent PSG and wrist-worn oximetry simultaneously in the sleep laboratory, the two devices might not have measured the same periods. The wrist-worn oximeter automatically filtered out periods where a sensor may have fallen off, which could impact the results. Thirdly, further validation studies in larger samples with severe OSA cases are required to strengthen our findings. Finally, it is best practice to split a dataset into training and test sets when building predictive models. However, because of the small sample size of moderate-severe OSA cases, we used the whole dataset for model building, which is a limitation. Therefore, future studies should test the model on an independent sample with the same inclusion criteria to enhance the reliability and validity of our predictive model.

Further direction
A reliable simplified screening tool for mild OSA is still unavailable. ODI3, in combination with other oximetric indices, clinical information, and questionnaires, may be better able to identify mild OSA.

Conclusion
Our findings demonstrate that a wrist-worn oximetry-derived automated ODI3 effectively identified children with moderate-severe OSA. In addition, its easy acquisition and interpretation can assist in triaging the urgency of adenotonsillectomy and can help to streamline preoperative decision-making.

Availability of Data and Material
The datasets generated or analyzed during the study are available from the corresponding author on reasonable request.

Conflicts of Interest
The authors have no potential conflicts of interest to disclose.

ORCID iDs
Mon Ohn https://orcid.org/0000-0001-9652-8827
Kathleen J Maddison https://orcid.org/0000-0002-6827-0993
Julie Nguyen https://orcid.org/0000-0003-3300-7033
Daisy Evans https://orcid.org/0000-0001-8262-5706
Natasha Bear https://orcid.org/0000-0002-9311-4386
R. Nazim Khan https://orcid.org/0000-0003-3349-5006
Peter R Eastwood https://orcid.org/0000-0002-4490-4138
Britta S. von Ungern-Sternberg https://orcid.org/0000-0002-8043-8541
Andrew C Wilson https://orcid.org/0000-0003-0875-2736
Jennifer H Walsh https://orcid.org/0000-0003-3950-6079

Author Contributions
Conceptualization: Mon Ohn, Kathleen J Maddison, Peter R Eastwood, Britta S. von Ungern-Sternberg, Andrew C Wilson, Jennifer H Walsh. Data curation: Mon Ohn, Kathleen J Maddison, Julie Nguyen, Jennifer H Walsh. Formal analysis: Mon Ohn, Daisy Evans, Natasha Bear, R. Nazim Khan. Funding acquisition: Mon Ohn, Britta S. von Ungern-Sternberg. Investigation: all authors. Methodology: all authors. Project administration: Mon Ohn, Julie Nguyen, Britta S. von Ungern-Sternberg. Resources: all authors. Software: Mon Ohn, Julie Nguyen, Daisy Evans, Natasha Bear, R. Nazim Khan. Supervision: Mon Ohn, Kathleen J Maddison, Peter R Eastwood, Britta S. von Ungern-Sternberg, Andrew C Wilson, Jennifer H Walsh. Validation: all authors. Visualization: all authors. Writing—original draft: Mon Ohn. Writing—review & editing: all authors.
**Funding Statement**

We are grateful for the support of grants from various sources: Dr Mon Ohn was awarded the Telegen Research Fellowship 2018 by Telegen Channel 7 for “The Obstructive Sleep Apnea: Study: Making Tonsillectomies Safer” study; Dr Mon Ohn is a recipient of the Raine Clinician Research Fellowship (CRF14-R9, 2021-2023) from the WA Department of Health and the Raine Medical Research Foundation for the “NIGHT OWL” study; Prof. Britta S. von Ungern-Sternberg is partly funded by the Stan Perron Charitable Foundation partly funded the study equipment. We also received in-kind support from our institutions and departments.

**Acknowledgments**

The authors thank the children and their families for participating in this research project. We gratefully acknowledge the collaboration and contribution of theater staff and Ear, Nose, Throat (ENT) colleagues at Perth Children’s Hospital.

**REFERENCES**