



Published in final edited form as:

Sleep Med. 2022 February ; 90: 65–73. doi:10.1016/j.sleep.2021.12.016.

Validation of actigraphy for sleep measurement in children with cerebral palsy

Bing Xue^{a,1}, Amy Licis^{b,c,*,1}, Jill Boyd^b, Catherine R. Hoyt^{b,d}, Yo-Ei S. Ju^{b,c,e}

^aSchool of Engineering, Washington University, Saint Louis, MO, USA

^bDepartment of Neurology, Washington University School of Medicine, Saint Louis, MO, USA

^cCenter on Biological Rhythms and Sleep (COBRAS), Washington University, Saint Louis, MO, USA

^dDepartment of Pediatrics, Washington University School of Medicine, Saint Louis, MO, USA

^eHope Center for Neurological Disorders, Saint Louis, MO, USA

Abstract

Objectives: Sleep issues are common in children with cerebral palsy (CP), although there are challenges in obtaining objective data about their sleep patterns. Actigraphs measure movement to quantify sleep but their accuracy in children with CP is unknown. Our goals were to validate actigraphy for sleep assessment in children with CP and to study their sleep patterns in a cross-sectional cohort study. **Methods:** We recruited children with (N = 13) and without (N = 13) CP aged 2–17 years (mean age 9 y 11mo [SD 4 y 10mo] range 4–17 y; 17 males, 9 females; 54% spastic quadriplegic, 23% spastic diplegic, 15% spastic hemiplegic, 8% unclassified CP). We obtained wrist and forehead actigraphy with concurrent polysomnography for one night, and home wrist actigraphy for one week. We developed actigraphy algorithms and evaluated their accuracy (agreement with polysomnography-determined sleep versus wake staging), sensitivity (sleep detection), and specificity (wake detection).

Results: Our actigraphy algorithms had median 72–80% accuracy, 87–91% sensitivity, and 60–71% specificity in children with CP and 86–89% accuracy, 88–92% sensitivity, and 70–75% specificity in children without CP, with similar accuracies in wrist and forehead locations. Our algorithms had increased specificity and accuracy compared to existing algorithms, facilitating detection of sleep disruption. Children with CP showed lower sleep efficiency and duration than children without CP.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

*Corresponding author. Campus Box 8111, Department of Neurology, Washington University School of Medicine, 660 South Euclid Ave, St Louis, MO, USA. Fax: +1 314 454 2523. licisa@wustl.edu (A. Licis).

¹Co-first authors and contributed equally to the manuscript. All authors of the manuscript agree to this designation.

Conflict of interest

The ICMJE Uniform Disclosure Form for Potential Conflicts of Interest associated with this article can be viewed by clicking on the following link: <https://doi.org/10.1016/j.sleep.2021.12.016>.

Competing interests

The authors have no competing interests to declare.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.sleep.2021.12.016>.

Conclusions: Actigraphy is a valid tool for sleep assessment in children with CP. Children with CP have worse sleep efficiency and duration.

Keywords

Actigraphy; Children; Cerebral palsy; Validation; Polysomnography

1. Introduction

Cerebral palsy (CP) is characterized by non-progressive, chronic motor deficits with onset during early childhood, with an estimated prevalence of 2.11 per 1000 live births [1]. CP comorbidities such as epilepsy, visual and cognitive impairment, hearing deficits, pain, dysmotility, and gastroesophageal reflux may contribute to sleep disturbances [2]. Sleep issues are reported in 23–40% of children with CP in questionnaire-based studies, with high rates of insomnia and symptoms of sleep-disordered breathing, versus 13–20% of typically developing children [3–7]. Poor sleep in children with CP correlates with problematic behaviors, decreased cognitive performance, and worse motor functioning [4]. Despite the high prevalence and negative impact of sleep disorders in CP, there are challenges in obtaining objective sleep data. Parent report may over-estimate sleep duration and underestimate night wakings compared to actigraphy for typically developing children as parents may be unaware of their children's wakings that occur during parental sleep, or children may not disclose or may forget their middle of the night wakings [8]. Parent-reported sleep information could be even less accurate for children with communication and movement impairments.

Actigraphy measures movement with non-invasive accelerometers typically worn on wrists to provide objective information about sleep–wake patterns. Algorithms transform actigraphy-detected activity data into sleep–wake staging for each epoch (ranging from 10 s to 1 min) of data. [9–11] Research in adults comparing actigraphy to polysomnography (PSG) demonstrated that actigraphy has a high sensitivity (ability to detect sleep) of 97–99%, but a low specificity (ability to detect wake) of 34–44% [12]. Similarly in children, actigraphy has shown a high sensitivity of 82–90% and relatively low specificity of 51–73%, with a tendency to overestimate sleep time and underestimate wake after sleep onset [13]. PSG is the gold standard for sleep–wake detection and assesses for a wide array of sleep disorders including obstructive sleep apnea, central sleep apnea, hypoventilation, hypoxemia, and periodic limb movement disorder whereas actigraphy can assess sleep duration, efficiency, and timing but not other sleep disorders. However, obtaining multiple night-PSG's to routinely assess sleep patterns is not feasible due to expense and limited availability; in comparison, actigraphy is portable, accessible in the home environment, useable over long periods of time, and requires less equipment and staff training compared to PSG.

To our knowledge, actigraphy has not been validated for sleep assessment in children with CP. The abnormal quality of movements in children with CP may reduce the accuracy of sleep-scoring algorithms. Variation in actigraphy placement on the body may affect results, especially in patients with movement disorders, and set protocols are not established. In

patients with Parkinson disease, wrist actigraphy was determined to be valid for sleep assessment if scoring parameters were adjusted [14], suggesting that actigraphy with modified algorithms may likewise be valid for sleep assessment in children with CP.

The goals of this study were to develop validated actigraphy analytic methods in children with CP, and to compare sleep in the home setting between children with and without CP. Our objectives were to 1) develop an accurate actigraphy sleep scoring algorithm for children with CP measured against concurrent gold-standard polysomnogram (PSG) staging; 2) determine optimal placement of actigraphs (wrist or forehead) in children with CP; and 3) apply these methods to assess sleep at home in children with and without CP using actigraphy.

2. Methods

2.1. Design

This cross-sectional cohort study included participants with CP and without CP who completed one night PSG while wearing wrist and forehead actigraphs, then wore wrist actigraphs for one week. The study was approved by the Human Research Protection office at Washington University School of Medicine.

2.2. Participants

Participants with CP (“CP group”) and without CP (“control group”) aged 2–17 years receiving a polysomnogram (PSG) for clinical indications at St. Louis Children’s Hospital Sleep Center were recruited from January 2014 to July 2015. Children with extremity weakness were excluded from control group enrollment. Informed consent was provided by the participants’ parents or legal guardians. Participants aged 14–17 who were cognitively able signed an assent document. Additional methods are provided in the supplement.

2.3. Procedures

Actigraphs (Actical, Philips Respironics Inc., Murrysville PA, USA) were placed on the dominant wrist and forehead on the night of the polysomnogram. Actigraphs condensed activity data into 1-min epochs. A PSG with standard channels was performed on a SomnoStar (version 10.2, CareFusion, Yorba Linda, CA, USA) overnight, scored using standardized criteria [15] by a registered PSG technologist, and reviewed by a physician board-certified in sleep medicine. Following the overnight visit, participants were instructed to wear wrist actigraphs for one week. Caregivers were instructed to keep a sleep log to track bedtimes and wake times, corresponding to a time stamp button-push on the actigraph. Parents completed the Children’s Sleep Habits Questionnaire (CSHQ), which assesses for symptoms of sleep disorders with ratings based on symptom frequency, an abnormality cut-off of 41, and higher scores being indicative of more disturbed sleep [16]. Actigraphy data were downloaded from actigraphs using Respironics Actical software, and raw data were exported for further analysis.

2.4. Weighted logistic regression (WLR) model

All analyses were performed with custom scripts in MATLAB (version: 2018b, The MathWorks Inc, Natick, MA), which are available publicly (<https://github.com/xuebing1234/actigraphy>) [17]. Sleep–wake staging from PSG was transformed to 1-min epochs to match with actigraphy data using the following method: if one or both of the two 30-s epochs by PSG were scored as wake, the corresponding 1-min epoch was considered to be wake. Therefore, for the first night of actigraphy recording, we were able to match epoch-by-epoch sleep–wake staging by PSG, the gold standard.

Logistic regression was used to estimate the probability of sleep versus wake. Since the magnitudes of movement are different based on location and group, separate models were developed for the CP and control groups, and for forehead versus wrist actigraphy data. The model development process is depicted in Fig. S1. Since PSG data were collected overnight during the sleep period, the majority of epochs were sleep (82.7% for the control group and 63.5% for the CP group). Given the imbalance in the fraction of wake epochs in the CP and control groups, as well as a small minority of epochs being wake, a weighting method was incorporated into the model to improve accuracy and balance sensitivity and specificity [18,19]. Subsequently, we refer to our overall algorithm as weighted logistic regression (WLR).

In the personalized WLR (pWLR), a personalized model was “trained” using one individual’s PSG data, so that the feature coefficients were optimized for that participant’s unique movement patterns. In the combined WLR (cWLR), the model was trained using combined data from a group, so that we could identify optimized feature coefficients for future use in groups with shared characteristics (for example, children with CP) who do not have concurrent PSG data available.

2.5. Definition of model parameters and performance metrics

Activity count thresholds define the cutoffs for activity counts within an epoch or window (depending on the method) above which the activity level is determined to be low, medium, or high. “Activity count” in an epoch is an integration of both amplitude and number of movements within that epoch. Various algorithms then use the activity level to define whether sleep or wake stage is assigned to an epoch. Window size is the duration of time surrounding an epoch that is used for staging that particular epoch. Prior studies have defined often arbitrary thresholds to categorize low, medium, or high activity levels, as well as arbitrary window sizes [9,11,20]. We used a data-driven approach employing receiver operating characteristic (ROC) analyses and comparing area under the receiver operating curve (AUC) values to assess the effects of varying activity count thresholds and window sizes on model accuracy (Fig. S2).

2.6. Feature extraction and reduction

We extracted 12 features in our initial WLR algorithm, and reduced to 6 features to simplify the model for minimal generalization error. A stepwise feature selection method was used to rank and select the features (Table S1). The selected features were verified using univariate feature selection (Fig. S3), and both methods yielded the same feature set. The selected

features are: the ratio of epochs with zero activity counts within the window, standard deviation of activity count within the window, mean crossing rate (ratio of activity counts crossing the average activity level of a window), ratio of epochs with high activity counts (above high activity threshold) within a window, ratio of epochs with low activity accounts (below medium activity threshold) within a window, and maximum activity count within a window.

2.7. Performance testing

We tested the performance of our final WLR algorithms against two widely-used scoring algorithms: Cole–Kripke and Sadeh [9,11,20]. The Cole–Kripke and Sadeh algorithms identify wake from sleep by a weighted summation of activity during the previous epochs, current epoch, and following epochs. The difference between these two methods is explained by difference in the weighting of the respective components and the fact that the Sadeh algorithms employ a nonlinear function, rather than linear, in the computation [9,11,20]. We performed pWLR for each of the 46 actigraphy records, and cWLR for four groups: forehead actigraphy from the control group (control forehead), forehead actigraphy from the CP group (CP forehead), wrist actigraphy from the control group (control wrist), and wrist actigraphy from the CP group (CP wrist).

We determined the specificity, sensitivity, and overall accuracy for our WLR algorithms calculated against concurrent sleep–wake staging by polysomnogram. Accuracy was calculated as the percentage of epochs staged in agreement with polysomnography staging. Sensitivity was defined as the percentage of sleep epochs correctly identified as sleep, and specificity was defined as the percentage of wake epochs correctly identified as wake. Sensitivity, specificity, and accuracy are not normally distributed, therefore median with interquartile range are reported throughout.

2.8. Sleep metrics

Bedtime and waketime were scored based on sleep logs and timestamps (button pushes on the actigraphs) [21,22]. Time in bed was calculated as time between bedtime and waketime. To assess sleep at home, pWLR was applied to actigraphy data for all time in bed, excluding the first night (ie PSG night). Total sleep time (TST), sleep efficiency (SE), sleep onset latency (SOL) and wake after sleep onset (WASO) were calculated. TST is the sum of all epochs scored as sleep, and SE is TST divided by time in bed, expressed as a percentage. SOL is the duration from bedtime and sleep onset (SO), defined as the beginning of the first consecutive number τ of epochs containing no more than 1 min of wake. WASO is the sum of epochs scored as wake from sleep onset until waketime. Since varying τ affects SOL and WASO, we tested τ from 5 to 20.

2.9. Statistical analysis

To assess for normal distribution, histograms were visually inspected and Kolmogorov–Smirnov tests applied to all continuous variables. Student’s *t* tests were used to compared normally-distributed continuous variables between groups, Wilcoxon-signed rank tests for non-normally-distributed continuous variables within an individual, and chi-squared tests for dichotomous variables. Statistical analyses were performed using MATLAB (version:

2018b) and Python. Bland Altman plots depict the correlation between the gold-standard polysomnogram and actigraphy.

3. Results

3.1. Demographics and overview

Twenty-six participants enrolled in the study, 13 each in the CP and control groups (Table 1). Parents of 5 children with CP and 2 children without CP declined enrollment for reasons including sensory sensitivity issues or concern about the potential for device loss. All 26 participants completed overnight PSG with concurrent overnight wrist and/or forehead actigraphy (Fig. S1). We developed novel WLR algorithms to analyze actigraphy data in children with and without CP using concurrent PSG staging as a gold standard. We used window sizes of 9 min for the control group and 15 min for the CP group, and activity thresholds of 60 for high activity and 50 for medium activity.

3.2. Power analysis

We have conducted a power analysis. Given that this is a pilot study, we have chosen a significance level at 0.05, an effect size (Cohen's d) at 1.2, and statistical power at 0.80. For the unpaired t test, the required sample size is 11.942. Since we have 13 subjects in each group, we have satisfied the required sample size.

3.3. Algorithm performance

We compared the performance of our algorithms against existing algorithms [9,12,20,23,24]. Overall, our methods for all locations and groups showed better accuracy (88%, IQR 76%–91% for pWLR and 79%, IQR 69%–88% for cWLR) and greatly improved specificity (73%, IQR 60%–84% for pWLR and 64%, IQR 39%–87% for cWLR), compared to the Cole–Kripke method (accuracy 81%, IQR 63%–89% and specificity 9%, IQR 2%–23%) and the Sadeh method (accuracy 82%, IQR 60%–89% and specificity 16%, IQR 4%–45%), without sacrificing much on sensitivity (91%, IQR 87%–95% for pWLR and 88%, IQR 77%–95% for cWLR) (Fig. 1, Table 2, Table S2). Additionally, pWLR, in which each individual's polysomnography data was used to determine model coefficients, outperformed cWLR in which the model coefficients were used for a combined group (Fig. 1, Table 2, Table S2). The ROC curves of both algorithms are plotted (Fig. S4). Model interpretations of pWLR and cWLR were visualized using Shapley values. Shapley values examine the contributions of individual components of the model to cause differences between a model's prediction and an average baseline (Fig. S5).

Accuracy was higher for all methods in the control group compared to the CP group (Fig. 1, Table 2, Table S2). We observed minimal differences in accuracy, sensitivity, and specificity between wrist and forehead locations (Fig. 1, Table 2, Table S2). In dividing our participants into three age groups (2–5 years, 6–12 years and 13–17 years), we observed only slight variations among sensitivity, specificity, and accuracy values (Table S3).

Bland–Altman plots demonstrate excellent agreement between polysomnogram staging and sleep–wake scoring of actigraphy by the pWLR method (Fig. 2). Average differences

between polysomnogram- and actigraphy-determined sleep variables for pWLR were 3 min for SOL, 7 min for WASO, 1% for SE, and 5 min for TST for both CP and control groups regardless of placement. In general, wrist placement had better concordance with polysomnography compared to forehead placement. As the mean WASO increased, the magnitude of underestimation of WASO by actigraphy increased; indicating that the longer the WASO, the more actigraphy underestimated WASO.

3.4. Home sleep actigraphy

Finally, we compared home sleep between CP and control groups as measured by wrist actigraphy, applying pWLR. Among the 26 participants, 6 participants did not return the actigraph and 4 declined to wear actigraphs at home due to concerns about tolerance of wearing actigraphy for several days; therefore, a total of 16 participants completed home sleep actigraphy: 8 in the CP group for a total of 42 nights and 8 in the control group for a total of 46 nights. We found that the CP group had longer WASO compared to the control group (181 vs 113 min, $p = 0.04$), resulting in lower decreased TST in the CP group. A shorter SOL in the CP group attenuated the difference, so the overall difference in TST was not statistically significant (355 vs 417 min, $p = 0.18$) (Table 1).

3.5. Test for normality

We conducted the Kolmogorov–Smirnov (KS) test for normality and did not reject the hypothesis that data is normally distributed, and performed paired t tests (Table S5). While multiple paired t tests could theoretically inflate p -values, this condition does not affect our conclusion, as our p values are smaller in magnitude. For example, using the Bonferroni correction, given that we have 3 repeated t tests for each measure, the p value to reject the null hypothesis would be $p/3$. At 5% confidence level, the correct $p^* = 0.017$. In our Table S2, most of the p values are less than 0.005, so our conclusion remains the same.

3.6. Drop-out effects

We have conducted a sensitivity analysis for the drop out effects. To evaluate the model performance with and without the drop-out participants, we have trained cWLR and pWLR and measured the area under the curve (ROC) in each case. This analysis revealed that the effect of drop-out participants to the overall model performance was minimal, and in fact the elimination of drop-out subjects increased model performance in many cases (Table S6)

4. Discussion

Our study is the first to evaluate the accuracy of actigraphy for sleep assessment in children with CP. We developed a weighted logistic regression (WLR) based sleep–wake scoring method for actigraphy data validated against concurrent polysomnography. We found that actigraphy is a valid tool for assessing sleep in children with CP using these methods, offering a novel approach to understand sleep in this population. Sleep issues are frequently reported but difficult to assess in children with CP, and development of instruments is needed to facilitate the routine investigation of their sleep [25]. Our study demonstrates that actigraphy provides a meaningful way to understand the sleep patterns of children with CP

in their home environment. We anticipate that these findings will inform healthcare practice to improve clinical outcomes and quality of life for children with CP and their caregivers.

The notable improvement in specificity (wake detection) favors our algorithms over existing algorithms for sleep assessment in children with CP. Their sleep is often disrupted, and algorithms with low specificity are not well-suited to assess disrupted sleep. Achieving good specificity is a difficult task in sleep–wake staging by actigraphy and good specificity is an asset of our algorithms. In addition, our WLR methods have comparable sensitivity (sleep detection), enhanced accuracy (agreement with PSG for children with and without CP). We have extended the methodology of the Cole–Kripke and Sadeh algorithms with modern feature extraction and machine learning strategies into cWLR; we start with extracting the statistical features in the sliding window of epochs, followed by feature reduction using three evaluation metrics (Fig. S3): Shapley values, mutual information measuring the dependence between two variables, and ranks of importance in sleep scoring. By doing so, we have reduced feature dimensions for better generalizability and robustness, as shown in the results section. To further improve the model performance, we have taken the individual’s different characteristics into the building of scoring algorithms (pWLR), with the guidance of PSG data. Bland–Altman plots demonstrated high correlation between our algorithms and polysomnography, including average differences among all conditions of 3 min for SOL, 7 min for WASO, and 1% for SE and 5 min for TST. Our methods can be applied on a group (using cWLR) or individual basis (using pWLR or cWLR), and pWLR improved accuracy even further.

The best window size for actigraphy analysis was longer for children with CP than without CP, to facilitate detection of relatively fewer movements. Forehead placement had similar accuracy as wrist placement and could be considered when wrist placement is impractical. Head actigraphy has been studied in patients with quadriplegia resulting from spinal cord lesions and validated against polysomnography in a general sample of adult patients [26,27]. The dominant wrist was chosen because the frequency of nondominant wrists movements may be insufficient in some children with CP. The dominant and nondominant wrists provided equally valid results when tested against PSG data in adults and adolescents without movement disorders [20]. We note that Actical devices are designed for physical activity estimation rather than sleep–wake measurement, however, these devices were selected due to their suitability for assessment of non-wrist locations. Acticals are similar to wrist-worn actigraphs intended for sleep measurement in that they use a piezoelectric accelerometer to record a digitally integrated measurement of gross motor activity, produce voltage based on movement, and express activity counts for each epoch [28]. Furthermore, Actical devices have been validated for sleep assessment compared to PSG, have been shown to be as reliable for sleep monitoring as wrist-worn actigraphs built to measure sleep, and are used at some institutions for the purpose of sleep assessment [28,29]. Our algorithms are not specific for Actical devices and can be applied with other types of actigraphs.

Our study is the first to use algorithms validated for the assessment of home sleep in children with CP. Prior home actigraphy studies using data reported collectively for children with neurodevelopmental disorders, including CP, showed relatively prolonged sleep latency and reduced sleep efficiency [30,31]. We assessed home sleep actigraphy data in children

with CP distinctly and in comparison to a control group. We found that children with CP had higher WASO and shorter TST than children without CP, perhaps due the synergistic sleep-disrupting effects of issues such as pain, obstructive sleep apnea, epilepsy, gastroesophageal reflux, and medication side effects [2,32,33].

Strengths of our study include validation of actigraphy usage for children with CP against the gold-standard polysomnography, comparison of our algorithm with established actigraphy algorithms, and concurrent data collection in a control group of children. When PSG data is available, pWLR allows personalization to each individual's unique movement patterns, a feature that is especially useful for participants with CP. In the absence of concurrent PSG data, cWLR can be used to accurately assess sleep in children with or without CP, using variables and scripts that we have publicly shared. Also, the accuracy of our algorithms were maintained in both wrist and forehead locations and across young childhood, middle childhood, and adolescence. Other strengths include collection of sleep data in the home environment using actigraphy algorithms validated for children with CP. Due to their relatively high specificity, our algorithms are well-suited for use in children with CP or others with disrupted sleep, and improve determination of common actigraphy measures affected by specificity such as wake after WASO, SE, and TST. In addition, our algorithms provide an option for sleep assessment in typically-developing children. Actigraphy provides a practical and convenient method of obtaining objective sleep data for children without significant sensory sensitivities and when cost or access is not a barrier.

Limitations include the small sample size and the possibility of overfitting of the scoring algorithms, although cross-validation was used to avoid overfitting or selection bias. With feature selection, we minimized generalization error and observed similar performance (measured by both AUC and area under precision–recall curve (AUPRC) between training and testing datasets (Table S4 and Fig. S4 in the appendix). Since the comparison was based on the average of 5 random shuffles of 5-fold cross validation, this strongly validates the robustness and consistency of our proposed pWLR and cWLR methods. While the Bland–Altman plots did not show overall bias, when the CP group and control group are examined separately, most of the data points from the CP group were above the WASO bias line (and below the SE bias line) and most of the data points from the control group were below the bias line (and above the SE bias line). Therefore, the mean biases in WASO and SE may be greater if only one of these groups were examined. In addition, pWLR algorithms require access to PSG data, which may not be available. However, in the case of a child with sleep problems for whom actigraphy data is sought, PSG is often performed to evaluate for sleep disorders. Concurrent actigraphy could be considered so that response to interventions could be assessed using pWLR.

Study limitations also include other practical considerations. A “first night” effect has been described, in which reduced sleep time and sleep efficiency occur during the first night spent in a sleep lab [34]. However, multiple-night polysomnograms were not feasible in this study due to resource limitations and are not routine in actigraphy validation studies or clinical practice [9,11,12,14,28]. Similarly, one week was selected for the duration of home actigraphy due to resource limitations. An additional limitation is that both the control group and the group of children with CP consisted of children undergoing polysomnography

for clinical reasons, since the expense of research-only polysomnograms for actigraphy validation in children without sleep issues was not feasible for this study. However, this limitation may provide an advantage in that our control group includes typically developing children with symptoms of sleep disturbance, another population for whom actigraphy monitoring may be clinically useful. In our sample 10/13 children with CP and 7/13 children without CP had OSA diagnosed by PSG; however, OSA did not affect actigraphy accuracy in prior research validating actigraphy against PSG in children [35]. Also, our study may not directly comment on the comparison of home sleep efficiency and duration in children with and without CP who are not pre-selected for sleep issues. Device retention may be problematic in actigraphy studies, and six participants did not return the device in our study. While the signal quality was interpretable for all participants who returned the device, it is possible that those with a higher likelihood of having signal quality issues may have declined participation or not returned the device. A barrier to any actigraphy research or clinical application is that actigraphy devices and software can be expensive if not already available. Also, some children with CP have significant sensory sensitivities and five parents withheld participation for that reason.

5. Conclusions

In conclusion, we developed novel sleep–wake scoring actigraphy algorithms for children with and without CP. Our WLR algorithms showed improved performance compared to existing methods, especially with regard to specificity, and high agreement with polysomnographically-determined sleep variables. Future directions include implementation of our algorithms in other actigraphy devices, in a larger sample size of children with or without CP including children with dyskinetic CP, in individuals with other etiologies of movement impairment, and incorporating multiple-night polysomnograms to mitigate the possible first night effect. In general, validation of sleep actigraphy in special populations of children is an important endeavor and requires further attention [36]. Our algorithms can help advance understanding of sleep disturbances in children with CP and monitor response to treatment interventions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful to the Washington University Institute of Clinical and Translational Sciences Research Forum-Child Health Initiative for project feedback, to Jennifer S. McLeland, MSW, MA, CCRC, ACRP-PM, RPSGT and Cristina D. Toedebusch, CCRC, RPSGT for their technical assistance and input, and to the participants and their families for their contributions.

The WLR algorithm is publicly shared on github: <https://github.com/xuebing1234/actigraphy>. The algorithm is not proprietary [17].

Funding

This research was supported by the Washington University Institute of Clinical and Translational Sciences Pilot Award “Actigraphy Validation in Children with Spastic Quadriplegic Cerebral Palsy.” (CTSA 604; Licis, PI, NIH CTSA UL1TR002345). Research reported in this publication was supported by National Institutes of Health awards

K23-NS089922, R01-AG059507, UL1RR024992 SubAward KL2-TR000450, and the Washington University Institute of Clinical and Translational Sciences grant UL1TR000448 from the National Center for Advancing Translational Sciences (NCATS) of the National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official view of the National Institutes of Health.

Abbreviations:

AUC	area under the receiver operating curve
AUPRC	area under precision–recall curve
CP	cerebral palsy
CSHQ	Children’s Sleep Habits Questionnaire
cWLR	combined WLR
PSG	polysomnogram
pWLR	personalized WLR
ROC	receiver operating characteristic
SE	sleep efficiency
SO	sleep onset latency
TST	total sleep time
WASO	wake after sleep onset
WLR	weighted logistic regression

References

- [1]. Oskoui M, Coutinho F, Dykeman J, et al. An update on the prevalence of cerebral palsy: a systematic review and meta-analysis. *Dev Med Child Neurol* 2013;55:509–19. 10.1111/dmcn.12080. [PubMed: 23346889]
- [2]. Pruitt DW, Tsai T. Common medical comorbidities associated with cerebral palsy. *Phys Med Rehabil Clin N Am* 2009;20:453–67. 10.1016/j.pmr.2009.06.002. [PubMed: 19643347]
- [3]. Newman CJ, O’Regan M, Hensey O. Sleep disorders in children with cerebral palsy. *Dev Med Child Neurol* 2006;48:564–8. 10.1017/S0012162206001198. [PubMed: 16780625]
- [4]. Romeo DM, Brogna C, Quintiliani M, et al. Sleep disorders in children with cerebral palsy: neurodevelopmental and behavioral correlates. *Sleep Med* 2014;15:213–8. 10.1016/j.sleep.2013.08.793. [PubMed: 24424102]
- [5]. Elsayed RM, Hasanein BM, Sayyah HE, et al. Sleep assessment of children with cerebral palsy: using validated sleep questionnaire. *Ann Indian Acad Neurol* 2013;16:62–5. 10.4103/0972-2327.107708. [PubMed: 23661965]
- [6]. Calhoun SL, Fernandez-Mendoza J, Vgontzas AN, et al. Prevalence of insomnia symptoms in a general population sample of young children and preadolescents: gender effects. *Sleep Med* 2014;15:91–5. 10.1016/j.sleep.2013.08.787. [PubMed: 24333223]
- [7]. Kim DS, Lee CL, Ahn YM. Sleep problems in children and adolescents at pediatric clinics. *Korean J Pediatr* 2017;60:158–65. 10.3345/kjp.2017.60.5.158. [PubMed: 28592979]
- [8]. Iwasaki M, Iwata S, Iemura A, et al. Utility of subjective sleep assessment tools for healthy preschool children: a comparative study between sleep logs, questionnaires, and actigraphy. *J Epidemiol* 2010;20:143–9. 10.2188/jea.je20090054. [PubMed: 20139658]

- [9]. Cole RJ, Kripke DF, Gruen W, et al. Automatic sleep/wake identification from wrist activity. *Sleep* 1992;15:461e9. 10.1093/sleep/15.5.461.
- [10]. Long X, Fonseca P, Foussier J, et al. Sleep and wake classification with actigraphy and respiratory effort using dynamic warping. *IEEE J Biomed Health Inform* 2014;18:1272–84. 10.1109/JBHI.2013.2284610. [PubMed: 24108754]
- [11]. Sazonov E, Sazonova N, Schuckers S, et al. Activity-based sleep–wake identification in infants. *Physiol Meas* 2004;25:1291–304. 10.1088/0967-3334/25/5/018. [PubMed: 15535193]
- [12]. de Souza L, Benedito-Silva AA, Nogueira Pires ML, et al. Further validation of actigraphy for sleep studies. *Sleep* 2003;26:81–5. 10.1093/sleep/26.1.81. [PubMed: 12627737]
- [13]. Meltzer LJ, Montgomery-Downs HE, Insana SP, et al. Use of actigraphy for assessment in pediatric sleep research. *Sleep Med Rev* 2012;16:463–75. 10.1016/j.smrv.2011.10.002. [PubMed: 22424706]
- [14]. Maglione JE, Liu L, Neikrug AB, et al. Actigraphy for the assessment of sleep measures in Parkinson’s disease. *Sleep* 2013;36:1209–17. 10.5665/sleep.2888. [PubMed: 23904681]
- [15]. Berry RB, Brooks R, Gamaldo CE, et al. For the American academy of sleep medicine. In: *The AASM manual for the scoring of sleep and associated events: rules, Terminology and technical specifications, version 2.2*; 2015. Darien, Illinois, www.aasmnet.org.
- [16]. Owens JA, Spirito A, McGuinn M. The Children’s Sleep Habits Questionnaire (CSHQ): psychometric properties of a survey instrument for school-aged children. *Sleep* 2000;23:1043–51. [PubMed: 11145319]
- [17]. [dataset] Xue B, Licis A, Boyd J, et al. Validation of actigraphy for sleep measurement in children with cerebral palsy: sleep training/staging algorithm for actigraphy data. Github; 2020. vol. 1, <https://github.com/xuebing1234/actigraphy>.
- [18]. Manski CF, Lerman SR. The estimation of choice probabilities from choice based samples. *Econom J Econom Soc* 1977;45:1977–88. 10.2307/1914121.
- [19]. King G, Zeng L. Logistic regression in rare events data. *Political Anal* 2001;9: 137–63.
- [20]. Sadeh A, Sharkey M, Carskadon MA. Activity-based sleep-wake identification: an empirical test of methodological issues. *Sleep* 1994;17:201–7. 10.1093/sleep/17.3.201. [PubMed: 7939118]
- [21]. Ju YS, McLeland JS, Toedebusch CD, et al. Sleep quality and preclinical Alzheimer disease. *JAMA Neurol* 2013;70:587–93. 10.1001/jamaneurol.2013.2334. [PubMed: 23479184]
- [22]. Ju YS, Ooms SJ, Sutphen C, et al. Slow wave sleep disruption increases cerebrospinal fluid amyloid- β levels. *Brain* 2017;140:2104–11. 10.1093/brain/awx148. [PubMed: 28899014]
- [23]. Sitnick SL, Goodlin-Jones BL, Anders TF. The use of actigraphy to study sleep disorders in preschoolers: some concerns about detection of nighttime awakenings. *Sleep* 2008;31:395–401. 10.1093/sleep/31.3.395. [PubMed: 18363316]
- [24]. Marino M, Li Y, Rueschman MN, et al. Measuring sleep: accuracy, sensitivity, and specificity of wrist actigraphy compared to polysomnography. *Sleep* 2013;36:1747–55. 10.5665/sleep.3142. [PubMed: 24179309]
- [25]. Hulst R, Willem Gorter J, Voorman J, et al. Sleep problems in children with cerebral palsy and their parents. *Dev Med Child Neurol* 2021;63:1344–50. 10.1111/dmcn.14920. [PubMed: 33990937]
- [26]. Spivak E, Oksenberg A, Catz A. The feasibility of sleep assessment by actigraph in patients with tetraplegia. *Spinal Cord* 2007;45:765–70. 10.1038/sj.sc.3102040. [PubMed: 17339889]
- [27]. Hummel R, Bradley TD, Fernie GR, et al. Estimation of sleep status in sleep apnea patients using a novel head actigraphy technique. *Conf Proc IEEE Eng Med Biol Soc* 2015:5416–9. 10.1109/EMBC.2015.7319616. 2015.
- [28]. Weiss AR, Johnson NL, Berger NA, et al. Validity of activity-based devices to estimate sleep. *J Clin Sleep Med* 2010;6:336–42. [PubMed: 20726281]
- [29]. Smith C, Galland B, Taylor R, et al. ActiGraph GT3Xp and actical wrist and hip worn accelerometers for sleep and wake indices in young children using an automated algorithm: validation with polysomnography. *Front Psychiatry* 2020;10:958. 10.3389/fpsy.2019.00958. [PubMed: 31992999]

- [30]. Chen X, Velez JC, Barbosa C, et al. Evaluation of actigraphy-measured sleep patterns among children with disabilities and associations with caregivers' educational attainment: results from a cross-sectional study. *BMJ Open* 2015;5:e008589. 10.1136/bmjopen-2015-008589.
- [31]. Wasdell MB, Jan JE, Bomben MM, et al. A randomized, placebo-controlled trial of controlled release melatonin treatment of delayed sleep phase syndrome and impaired sleep maintenance in children with neurodevelopmental disabilities. *J Pineal Res* 2008;44:57–64. 10.1111/j.1600-079X.2007.00528.x. [PubMed: 18078449]
- [32]. Garcia J, Wical B, Wical W, et al. Obstructive sleep apnea in children with cerebral palsy and epilepsy. *Dev Med Child Neurol* 2016;58:1057–62. 10.1111/dmcn.13091. [PubMed: 26991829]
- [33]. Penner M, Xie WY, Binopal N, et al. Characteristics of pain in children and youth with cerebral palsy. *Pediatrics* 2013;132:e407–13. 10.1542/peds.2013-0224. [PubMed: 23858420]
- [34]. Agnew HW, Webb WB, Williams RL. The first night effect, an EEG study of sleep. *Psychophysiology* 1966;2:263–6. 10.1111/j.14698986.1966.tb02650.x. [PubMed: 5903579]
- [35]. Hyde M, O'Driscoll D, Binette S, et al. Validation of actigraphy for determining sleep and wake in children with sleep disordered breathing. *J Sleep Res* 2007;16:213–6. 10.1111/j.1365-2869.2007.00588.x. [PubMed: 17542951]
- [36]. Sadaka Y, Sadeh A, Bradbury L, et al. Validation of actigraphy with continuous video-electroencephalography in children with epilepsy. *Sleep Med* 2014;15: 1075–81. 10.1016/j.sleep.2014.04.021. [PubMed: 24974198]

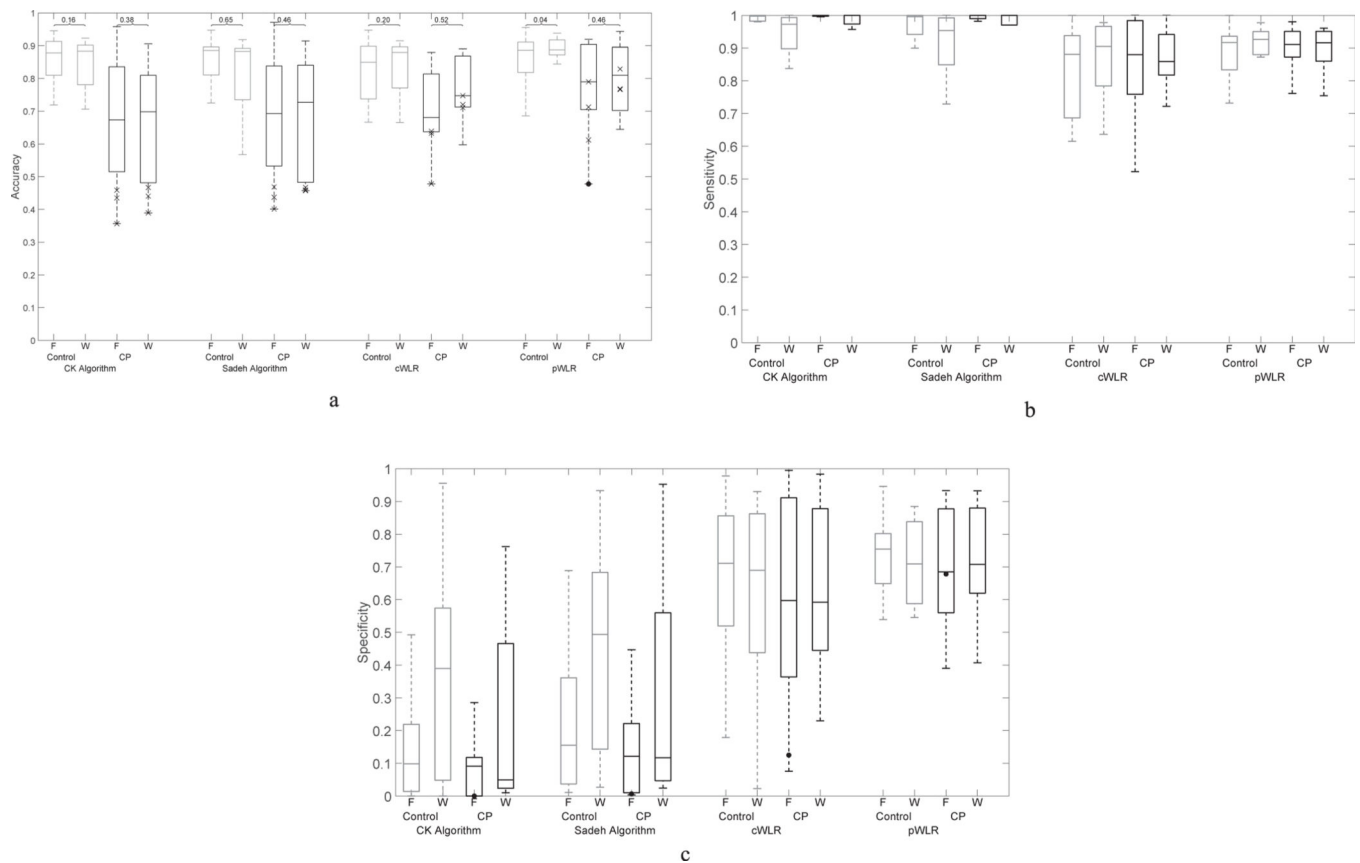


Fig. 1. Performance of weighted logistic regression, Boxplots of a) accuracy, b) sensitivity, and c) specificity of the Cole–Kripke (CK) algorithm, Sadeh algorithm, combined weighted logistic regression (cWLR) algorithm and personalized weighted logistic regression (pWLR) algorithm are shown. The central horizontal line indicates the median, the box the interquartile range, and the whiskers the maximum and minimum. Black boxes represent the cerebral palsy (CP) group and gray boxes represent the control group. F – forehead actigraphy; W – wrist actigraphy. The *p* values of Wilcoxon signed rank tests between forehead and wrist datasets are shown above the boxplots in Fig. 2a. “X” symbols show three outliers with poor (<50%) accuracy in the CP group using the Cole–Kripke and Sadeh algorithms. These outliers were able to be “rescued” with pWLR to an average accuracy of 70% using forehead actigraphy and 78% using wrist actigraphy using the pWLR model, and 58% using forehead actigraphy and 73% using wrist actigraphy using the cWLR model. Although there was a new outlier in pWLR model using forehead actigraphy (black circle), the specificity is much higher (74%) than with the Cole–Kripke algorithm (0%) and Sadeh algorithm (0.01%) as illustrated in Fig. 2c.

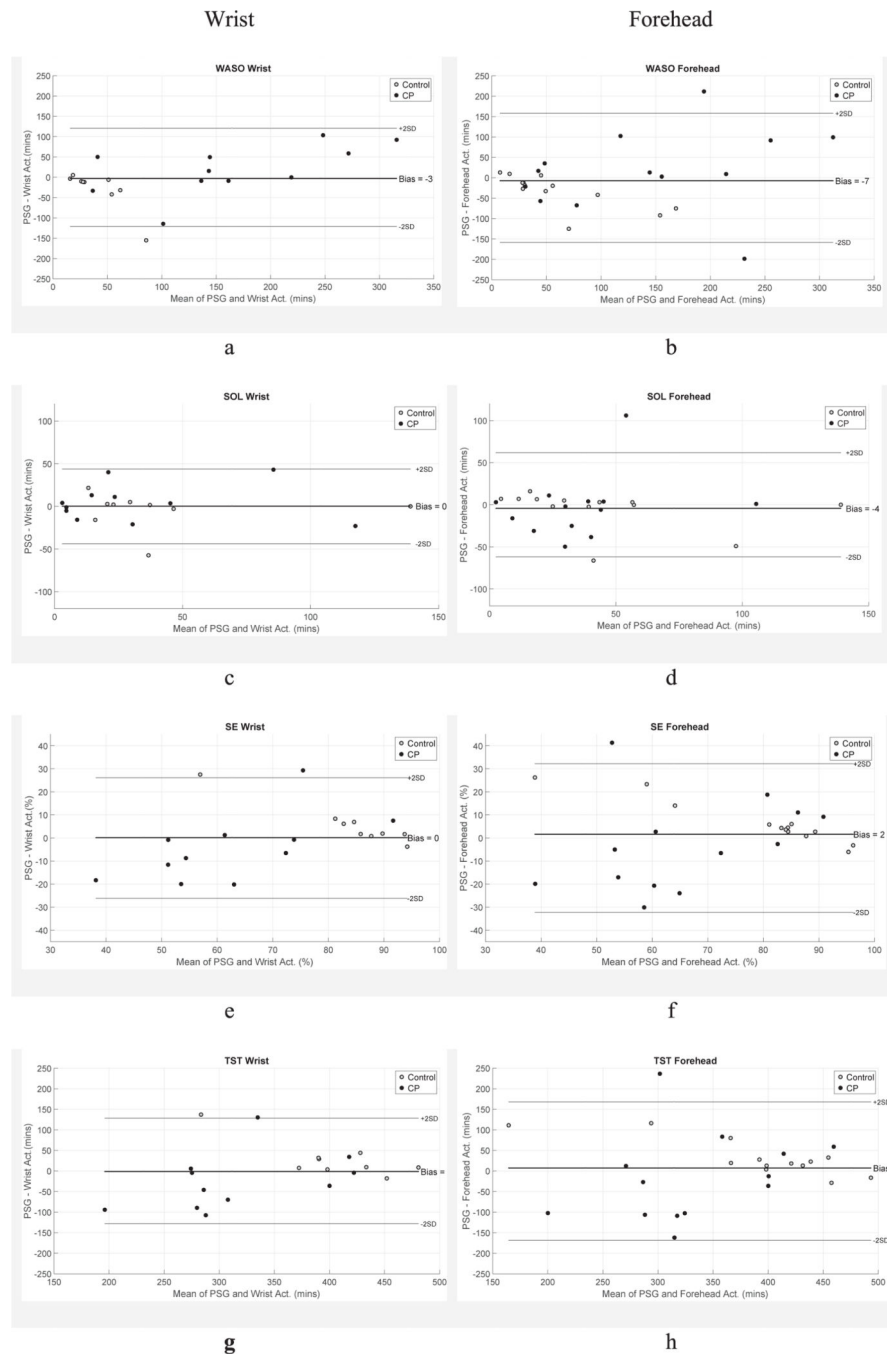


Fig. 2. Bland–Altman plots of weighted logistic regression. Bland–Altman plots of sleep variables determined by personalized weighted logistic regression (pWLR) using actigraphy and polysomnography (PSG). Plots for wrist actigraphy are shown on the left, while plots for forehead actigraphy are shown on the right. In each plot, the horizontal bias line represents the mean difference between PSG and actigraphy scoring. WASO- wake after sleep onset (min); SOL-sleep onset latency (min), SE-sleep efficiency (%); TST-total sleep time (min).

Table 1

Participant characteristics.

	All	CP	Control	<i>p</i> values
N (%)	26	13 (50)	13 (50)	–
Age (years, months)	9 y 11 mo [4 y 10 mo] range 4–17 y	10 y 8 mo [4 y 11 mo] range 4–17 y	9 y 1 mo [4 y 10 mo] range 2–16 y	0.43
Sex				
Male N (%)	17 (65)	20 (77)	14 (54)	0.22
Female N (%)	9 (35)	6 (23)	12 (46)	
Children's sleep habits score	47 [11] (N = 21)	46 [11] (N = 11)	48 [12] (N = 10)	0.68
Polysomnogram night sleep characteristics				
Total sleep time (minutes)	366 [93]	325 [100]	406 [67]	0.02
Total time in bed (minutes)	501 [49]	512 [52]	491 [45]	0.28
Sleep efficiency (%)	73 [19]	64 [21]	83 [12]	0.01
Sleep onset latency (minutes)	38 [35]	35 [35]	42 [35]	0.62
Total sleep time (minutes)	366 [93]	325 [100]	406 [67]	0.02
Obstructive sleep apnea diagnosis N (%)	17 (65)	10 (77)	7 (54)	0.94
Actigraphically-measured sleep at home				
No. of participants	16	8	8	–
No. of nights	88	42	46	–
Total sleep time (minutes)	–	355 [139]	417 [104]	0.18
Time in bed (minutes)	–	586 [68]	595 [54]	0.78
SE (%)	–	60 [20]	70 [16]	0.17
SOL (minutes)	–	57 [58]	81 [74]	0.52
WASO (minutes)	–	181 [75]	113 [61]	0.04

p values were calculated using chi-squared tests for categorical rows, and *t*-tests for rows with normal distribution (normality assessed by Kolmogorov–Smirnov tests). Mean [SD]. CP = cerebral palsy; CSHQ = Children's Sleep Habits Questionnaire; SE = sleep efficiency (%); SOL = sleep onset latency (min); WASO = wake after sleep onset (min).

Table 2

Accuracy, sensitivity, and specificity for pWLR, cWLR, Cole-Kripke, and Sadeh methods. Reported in (median %, IQR-interquartile range %).

Overall	pWLR	cWLR	Cole-Kripke	Sadeh
Accuracy (%)	88, IQR 76–91	79, IQR 69–88 ^a	81, IQR 63–89 ^a	82, IQR 60–89 ^a
Sensitivity (%)	91, IQR 87–95	88, IQR 77–95	99, IQR 98–100 ^a	99, IQR 95–100 ^a
Specificity (%)	73, IQR 60–84	64, IQR 39–87 ^a	9, IQR 2–23 ^a	16, IQR 4–45 ^a
Control group				
Accuracy (%)	89, IQR 85–91	86, IQR 77–90	88, IQR 81–91	88, IQR 76–89
Sensitivity (%)	92, IQR 88–95	88, IQR 76–94	99, IQR 95–100 ^a	99, IQR 92–100
Specificity (%)	75, IQR 61–81	70, IQR 56–86	13, IQR 3–43 ^a	24, IQR 6–52 ^a
CP group				
Accuracy (%)	80, IQR 70–90	72, IQR 67–84 ^a	68, IQR 51–81 ^a	71, IQR 51–83 ^a
Sensitivity (%)	91, IQR 87–95	87, IQR 80–96	100, IQR 99–100 ^a	100, IQR 98–100 ^a
Specificity (%)	71, IQR 61–87	60, IQR 37–91 ^a	7, IQR 1–15 ^a	12, IQR 3–28 ^a
Wrist				
Accuracy (%)	88, IQR 77–91	79, IQR 72–88	80, IQR 66–89 ^a	78, IQR 56–88 ^a
Sensitivity (%)	92, IQR 88–95	89, IQR 81–95	99, IQR 95–99 ^a	97, IQR 91–99
Specificity (%)	71, IQR 61–85	62, IQR 46–86	14, IQR 3–57 ^a	24, IQR 6–67 ^a
Forehead				
Accuracy (%)	88, IQR 73–90	78, IQR 68–87 ^a	83, IQR 63–90 ^a	83, IQR 66–89
Sensitivity (%)	91, IQR 87–94	88, IQR 75–95	100, IQR 99–100 ^a	100, IQR 98–100 ^a
Specificity (%)	74, IQR 61–81	67, IQR 37–87 ^a	9, IQR 0–17 ^a	14, IQR 3–31 ^a

^a $p < 0.05$ for the pairwise comparison (using paired t tests to compare pWLR separately with cWLR, Cole-Kripke, and Sadeh, respectively). Please refer to an expanded Table S2 reporting specific p values, included in the supplement.