Infant Behavior and Development 69 (2022) 101775

Contents lists available at ScienceDirect

Infant Behavior and Development

journal homepage: www.elsevier.com/locate/inbede

Novel and noninvasive methods for in-home sleep measurement and subsequent state coding in 12-month-old infants

Melissa N. Horger *

Graduate Center, City University of New York, USA

ARTICLE INFO

Keywords:
Sleep measurement
REM
NREM
Actigraphy
Cardiorespiratory
Infant

ABSTRACT

Research studying the role of sleep in development abounds but focuses on global aspects of sleep like quality or timing. Far fewer studies include the ultradian cycle, or patterns of REM and non-REM (NREM), because doing so is costly in time and resources. In a complete, lab-based sleep study, individuals are monitored by a technician overnight while wearing a host of sensors to capture brain activity, eye and limb movement, and cardiorespiratory rates. There is a need for creative, minimally disruptive solutions to study sleep that do not compromise the richness and accuracy of the measurements. The current pilot study parsed down the physiological measures to only movement and cardiorespiratory rates, creating a protocol simple enough for caregivers to incorporate into bedtime. Ten 12-month-old infants (+/- 3 weeks) wore an actigraph and wireless cardiorespiratory sensor for five nights of data collection. Of this, 92% were useable data. Actigraphy was analyzed with the Sadeh algorithm to delineate sleep from wake. Heart rate and respiration were then used to state score visually or via an algorithm; greater variability demarcated REM from NREM. Time spent in each state was compared between scoring methods as well as to published results from age matched infants who underwent polysomnography (PSG). Visually scored data, using a 1-hour viewing window, was in line with peer's PSG values. To automatically state score, epoch-by-epoch cardiorespiratory and actigraphy files were produced for each minute of data collection. Heart and respiratory rates were transformed into z-scores and iterations of scoring, using increasingly greater z score thresholds, were compared to determine which identified state proportions most similar to data collected with PSG. Based on these results, our novel method appears to be a feasible choice for studying the ultradian cycle. The combination of actigraphy and cardiorespiratory monitoring is uniquely advantageous because it is less resource intensive and more naturalistic, being put on by caregivers while still resulting in high rates of good data. Taken together, it is a quality option for infant researchers interested in incorporating sleep into their paradigms.

1. Introduction

The ultradian cycle is the repeated oscillation between sleep states over the course of the night and is comprised of rapid eye movement (REM) sleep, non-REM (NREM) sleep, and wake (Sheldon, 2012). Our knowledge about the developmental role of sleep states in human infants is relatively limited. Two recent reviews posit that both are integral to development and call for additional

* Corresponding address: 2800 Victory Blvd. 4S-108 Staten Island, NY 10314, USA.
E-mail address: mhorger@gradcenter.cuny.edu.

https://doi.org/10.1016/j.infbeh.2022.101775
Received 8 March 2022; Received in revised form 12 September 2022; Accepted 13 September 2022
Available online 18 September 2022

research to study this role through the second year when the decreasing proportion of REM sleep slows and stabilizes around adult levels (Cao et al., 2021; Knoop et al., 2020). The latter paper, by Knoop and colleagues (2020), specifically calls for new methods to study infants' ultradian cycles.

Presently, the gold standard for sleep measurement across development is polysomnography (PSG) (Carskadon & Dement, 2016). A full PSG records cardiorespiratory rates, brain activity, muscular activity, nasal air flow, sound, and movement (including eye as well as limb movements), making it prohibitively expensive to researchers and inconvenient to participants. Access to the resources required to design a paradigm including PSG is rare and commonly reserved for hospital settings; an observation bolstered by the literature (e.g., Knoop et al., 2020; Samson, 2021). Full sleep labs are equipped with an electroencephalogram (EEG), electromyogram (EMG), cardiorespiratory monitoring system, disposable electrodes, and audiovisual equipment, as well as a physical space large enough for all this and a bed. Additionally, a sleep technician or research assistant is required to be awake and present for the data collection and two are needed for coding.

Most of the physiological measures use one or more sensors that are physically attached to the person, amounting to approximately 6 points of connection in addition to an EEG cap (Werth et al., 2017). PSG is highly intrusive and more than one night of data collection is necessary because the disruption in routine or additional stress significantly alter sleep metrics (Grigg-Damberger et al., 2007). Indeed, four-month-old infants showed a decrease in REM sleep when they underwent a PSG in the lab as compared to at home (Bernstein et al., 1973) and this pattern was replicated in children and teens (Salgado et al., 2022; Scholle et al., 2003). As it is difficult for parents and infants to disrupt their schedules for consecutive days and nights, accounting for the first night effect is challenging (Grigg-Damberger et al., 2007).

State scoring, the process of delineating REM and NREM within a sleep period, was standardized by the American Academy of Sleep Medicine (AASM) which created the specific, age-based guidelines for analyzing PSG results (Scholle et al., 2011). While primarily related to the EEG output, additional criteria for REM include twitching, elevated and variable cardiorespiratory rates, and rapid eye movements (Grigg-Damberger et al., 2007; Knoop et al., 2020). These secondary characteristics and their success in automated state scoring are a relatively recent feature of the literature (Isler et al., 2016; Werth et al., 2017), though manual scoring has a longer history (e.g., Anders & Keener, 1975). Preterm infants are common participants due to efforts to preserve their sleep despite the bustling hospital environment. In the neonatal intensive care unit (NICU), cardiorespiratory monitoring is standard, and a recent review of studies using these metrics to state score concluded that heart rate and respiratory frequency were "valuable parameters for sleep classification in preterm infants" (de Groot et al., 2021, p8). To date, no research has extended this method to older infants or in-home settings.

1.1. *Current study*

The current study sought to pilot a new, naturalistic methodological technique to assess the ultradian cycle of one-year-old infants. The protocol was designed to be implemented in the home to mitigate first night effects, more readily permit multiple nights of recording, and increase participant comfort levels. Our method combined wireless cardiorespiratory monitoring with the most common technique for in-home, infant sleep measurement: actigraphy (Sadeh et al., 1991; Sadeh et al., 2015; Horger et al., 2021). Worn around an infant's ankle, an actigraph continuously records movement amplitude to delineate sleep from wake; however, it does not provide enough information to score the ultradian cycle (Sadeh et al., 2000). Instead, within the sleep periods, heart and respiratory rates were used to classify REM and NREM.

2. Method

2.1. *Participants*

Ten 12-month-old infants ($M = 372.2$ days; $SD = 12.2$) and their families participated in five nights of data collection. All but one of the infants were full term. None had current heart or respiratory problems or dermatological ailments. Six infants were male and 4 were female. Seven caregivers described their infant's race and ethnicity as White, 2 as Hispanic, and 1 as Black, Asian/Pacific Island, and White.

Families were recruited on a rolling basis from the Child Development Lab database and via outreach at community events. They were compensated for their time and effort at a rate of $5 for each night of data collection and an infant onesie. The maximum amount that any family was eligible to receive was $25. The research was approved by the Institutional Review Board of the College of Staten Island and parents provided written informed consent for their child's participation at the beginning of the study.

2.2. *Apparatus*

2.2.1. *Actigraph*

Participants were given a MicroMini Motionlogger actigraph (Ambulatory Monitoring, Inc., Ardsley, NY) to collect actigraphy data. The actigraph was initialized before the initial meeting with the families. Once data collection was complete, the actigraphy data was downloaded via an interface and analyzed using the ActionW 2.7 software and the Sadeh algorithm (Sadeh et al., 1995). This program coded each minute as sleep or wake.

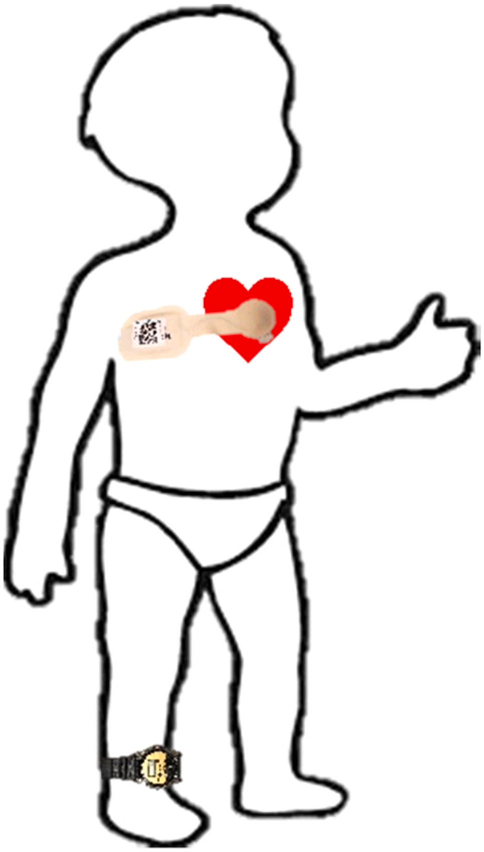

Fig. 1. Schematic of an infant wearing the Lifetouch sensors and actigraph.

2.2.2. *Heart rate and respiratory monitoring*

Lifetouch (Isansys Co., https://www.isansys.com/) is a wireless non-invasive body-worn device for collecting heart and respiratory rates. The infant size contains 2 female snaps that are attached to 2 electrodes with corresponding male snaps. After removing the paper backing, sensors were stuck on the infant's chest. The sensor emitted a soft blinking light to signal that it was on and collecting data. An android tablet, with the Isansys Gateway application installed, recorded measures from the Lifetouch sensor via Bluetooth and sent the data to the secure Lifeguard server installed within the University's IT network. Heart and respiration data could be seen and monitored by parents in real time on the tablet and once data was collected, it could be exported or viewed on a user interface, the Patient Status Engine.

2.2.3. *Research kit*

The Lifetouch sensors, tablet, actigraph, and a short book of instructions were housed in a bag given to parents labeled the Research Kit. In addition to these materials, it also contained extra female snap electrodes, cotton pads, adhesive remover, and petroleum jelly.

2.3. *Procedure*

Families were contacted via phone or email when their infant was 10–11 months old to explain the study and invite them to participate. If families agreed and their infant was the appropriate age (12 months +/- 3 weeks), the researcher scheduled the initial meeting at a time that was most convenient for them. Before COVID-19, during the initial meeting, the researcher explained the consent form and provided the Research Kit. After COVID-19, the researcher dropped off the equipment and explained the consent form virtually. The researcher provided her contact information (phone number and email) in case parents had any questions or concerns during their participation.

For five nights, parents turned on the Lifetouch sensor and put it on their infant before bed. They also placed the actigraph around their infant's left ankle (See Figure 1). Other than this, they were encouraged to have their nightly routine proceed as usual. The tablet continuously recorded data that was collected via the sensor, using Wi-Fi to transmit this data to a secure server at the lab's physical location. After five nights of data collection, the researcher scheduled a time to collect the materials, and families were thanked for their participation and given their monetary compensation.

Heart rate, respiration, and actigraphy data were labeled with the participant's number. Data from the actigraph was used to

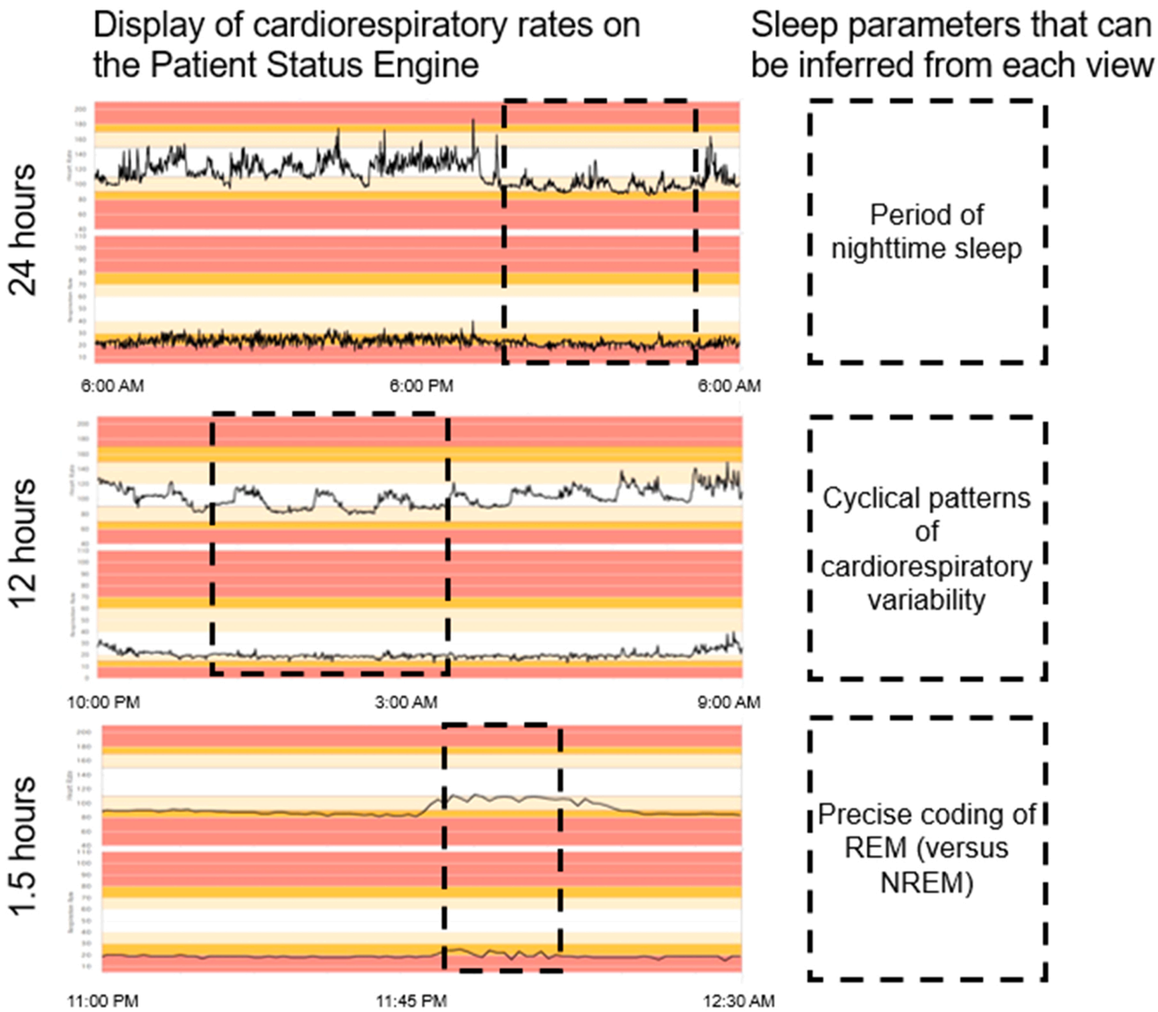

Fig. 2. Adjusting the viewer size on the Patient Status Engine produces data that can be visually coded into different outcomes.

distinguish sleep from wake. Heart rate and respiration were then used to state score visually or via an algorithm. Amount of time spent in REM, NREM, and indeterminant sleep were compared between scoring methods as well as to previously reported results from age matched infants who underwent polysomnography (Grigg-Damberger et al., 2007).

2.4. State coding

2.4.1. Manual visual coding

The start and end of each night were identified via actigraphy, using the Sadeh algorithm, and manually checked. Actigraphy was also used to demarcate wake episodes within the night. During periods of sleep, cardiorespiratory data was visually coded on the Patient Status Engine. Both heart and respiratory rates were visible simultaneously and the viewing window could be adjusted in size to see, for example, 2 h of data or 12 h of data (See Fig. 2). Previous work that included visual coding of cardiorespiratory parameters did not specify the scale at which data were coded (Isler et al., 2016). Additionally, it was unclear if the Patient Status Engine (a proprietary user interface) would be comparable to other data collection software. As such, the visual scale was initially set to 2 h, but this boundary was arbitrarily assigned. Iterative analysis was necessary to determine the most accurate scale for state scoring.

State coding was based on the variability as well as the time constraints. Per the former, REM was indicated by a noticeable increase in the variability of heart and respiratory rates, as displayed in the bottom panel of Fig. 2. REM periods needed to last at least 5 min while NREM periods had a threshold duration of at least 10 min (Grigg-Damberger et al., 2007). Each minute of sleep was coded as REM, NREM, or indeterminate. Indeterminate sleep were epochs that did not display a clear pattern or did not meet the timing constraints.

2.4.2. Automated coding algorithms

Epoch-by-epoch files were produced from the cardiorespiratory and actigraphy data for each minute of data collection. Again,

Table 1
Cumulative state durations (across all 5 nights) from the first and second visual passes.

Subject	Coding attempt	REM	NREM	IND
1	1	1138 (39%)	1369 (47%)	391 (13%)
	2	750 (26%)	1507 (52%)	639 (22%)
2	1	833 (34%)	1078 (45%)	485 (20%)
	2	736 (30%)	1172 (49%)	489 (20%)
3	1	1244 (39%)	1430 (45%)	493 (15%)
	2	1069 (33%)	1761 (55%)	337 (10%)
4	1	868 (37%)	1069 (46%)	372 (16%)
	2	641 (27%)	1432 (62%)	229 (9%)
5	1	1312 (46%)	1188 (41%)	354 (12%)
	2	952 (33%)	1597 (56%)	305 (10%)
6	1	1083 (40%)	1274 (47%)	344 (13%)
	2	722 (28%)	1573 (57%)	404 (15%)
7	1	985 (43%)	1051 (45%)	267 (11%)
	2	679 (29%)	1416 (61%)	208 (9%)
8	1	887 (40%)	893 (41%)	398 (18%)
	2	589 (27%)	1259 (58%)	318 (15%)
9	1	1154 (45%)	1001 (39%)	390 (15%)
	2	794 (31%)	1382 (54%)	386 (15%)
10	1	n/a	n/a	n/a
	2	563 (27%)	1223 (58%)	313 (15%)

Note. Participant 10 was added later and as such, only coded once, using a one hour time window.

actigraphy was used to delineate sleep periods. All coding and computation were done in R (Version 4.0.0), using knitr, dplyr, ggplot2, lubridate, rmcorr, pwr, irr, BlandAltmanLeh, TOSTER, and magicfor packages, and stored on GitHub. Both heart and respiratory rates were included though, the literature on this is mixed as some research includes both metrics while others focused on respiration alone (for review, see de Groot et al., 2021). On an individual level, cardiorespiratory rates were averaged within each night, centered, and standardized around the mean to produce z scores. Then, they were compared to varying thresholds of greater than 0, 0.25, and 0.5 to ensure only minutes with the most variability would be scored as REM. Heart and respiratory z scores were compared to the thresholds separately using a series of loops and if then logic statements. If the value surpassed the threshold, the minute was scored as REM. If the minute was coded as REM by both heart and respiratory rates, it remained as REM; if not, it was changed to indeterminate.

2.5. Data analysis

2.5.1. Sample size analysis

Sample size was determined based on the method developed by Jan and Shieh (2018) to estimate the total number of necessary data points, or nights of collection. In the distribution specifications, the mean difference score was estimated to be zero with a standard deviation of 10 (indicating 10% of the epochs coded). The standard deviation was estimated based on the results of research comparing sleep state coding using polysomnographic, behavioral, and respiratory variability (Isler et al., 2016). The resulting recommended sample size was N = 52. While this did not account for repeated measures, an ANOVA was used to test the hypothesis that there were no subject differences in accuracy of sleep state coding.

2.5.2. Methodological comparisons

Sleep state durations and proportions, as determined by the automated algorithms, were compared to the published PSG results of age matched peers (from Scholle et al., 2011) to identify the most accurate threshold. Then, results from only this algorithm were compared to the results of visual coding in a series of Bland-Altman plots. Bland-Altman plots detect the magnitude and direction of difference between two methods. They also establish a cut off, the limits of agreement, between which all difference scores should fall. When establishing the boundaries, standard deviations were calculated using the components of variance methods to account for repeated measures (Bland & Altman, 1999). Lastly, paired samples equivalence tests were run as a more stringent comparison. The equivalence interval was defined based on concordance rates (90–100%) reported for visual respiratory state scoring and set to 0–10% (Isler et al., 2016).

3. Results

3.1. Preliminary analysis

Data could not be reviewed or analyzed until the equipment was returned after the participation interval. As such, researchers were not notified (unless by parents) if the equipment fell off in the night or malfunctioned in some way. Even so, 46 of the 50 scheduled nights (92%) were useable data, recording for the entire nighttime interval. Within these 46 nights, each epoch was screened for impossible values (resulting from data collection errors), resulting in 28,989 good epochs (96.8%) that were included in the subsequent analyses. A one-way ANOVA was used to check for a main effect of subject and compared durations derived from visual state

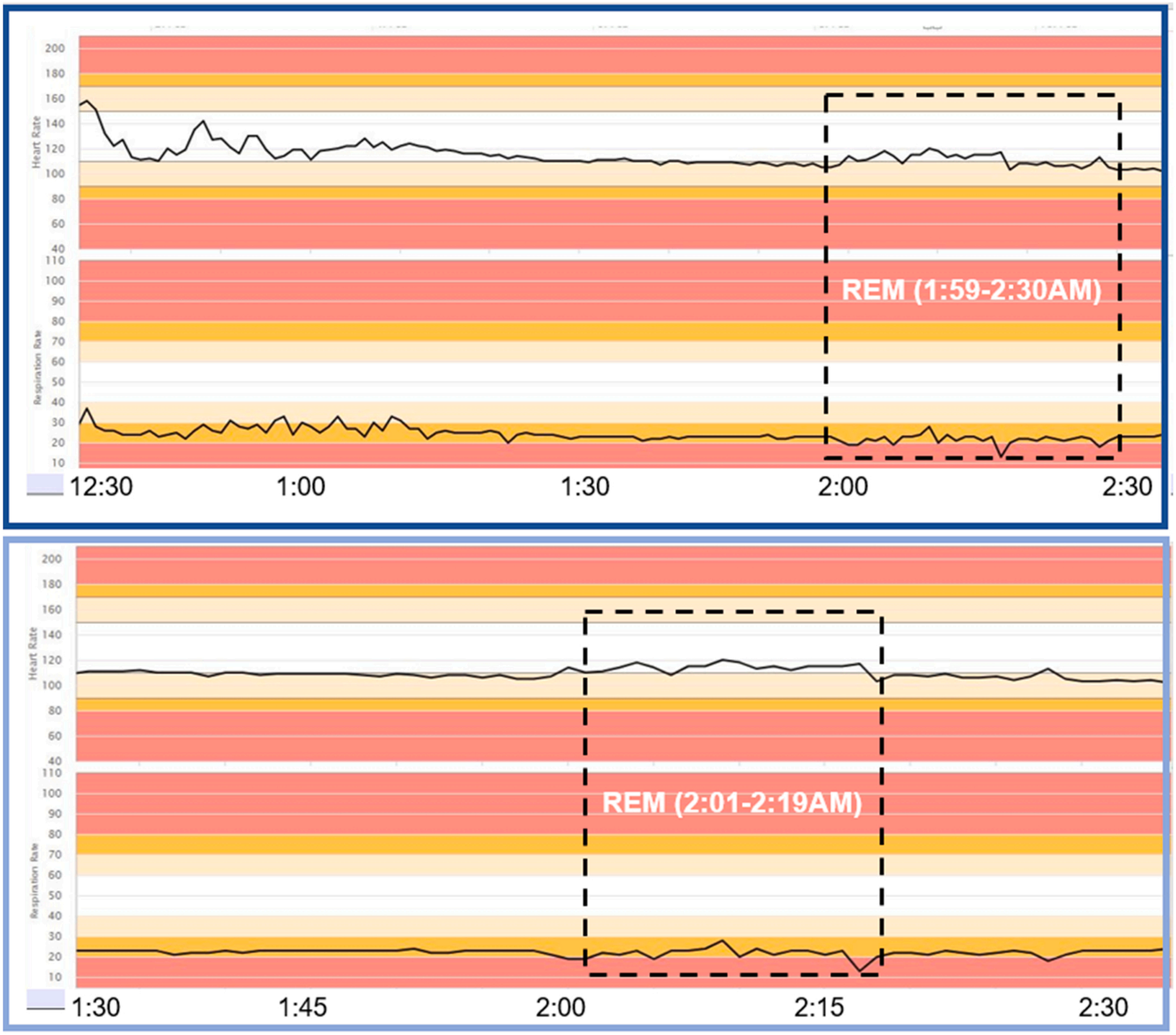

Fig. 3. State coding example using the 2-hour window and the 1-hour-window.

scoring, as we would not expect individual differences to persist after averaging across nights. It was nonsignificant across all states [REM: $F(9, 18) = 1.28$, $p = 0.36$, NREM: $F(9,18) = 0.99$, $p = 0.5$, IND: $F(9,18) = 2.05$, $p = 0.15$].

3.2. Visual coding

Before visual coding, actigraphy data was used to delineate sleep from wakefulness. The first iteration of visual coding is displayed in the top rows in Table 1 The Patient Status Engine viewing window was set to display 2 h at a time. When the cursor was moved over the data, the timestamp and current value appeared. Each minute, after being confirmed by actigraphy as sleep, was scored as REM, NREM, or indeterminate. Total duration for each state was summed for each participant across the entire data collection period.

The resulting data displayed a bias toward overestimating REM. Polysomnographic results stipulated that one-year-olds spend approximately 21–30% of their time in REM (Scholle et al., 2011). To address this bias, a second attempt at visual coding modified the procedure by shrinking the viewing window and focusing more heavily on respiration than heart rate. Fig. 3 contains an example of how this modification changed the coding. The top panel (outlined in dark blue) shows 2 h of data while the lower panel (outlined in light blue) includes only one hour. The dashed black boxes highlight the same REM episode, but greater zoom allows more specificity in determining the onset and offset. During the first pass, the REM was coded as 1:59–2:30AM and during the second, it was demarcated as 2:01–2:19AM. Resulting data after re-scoring were much more in line with Scholle and colleague's (2011) review, as seen in the bottom rows of Table 1.

Table 2
Frequency of interrater reliability outcomes.

		Coder 2		
		REM	*NREM*	*IND*
Coder 1	***REM***	2813	72	178
	NREM	70	1162	228
	IND	156	184	333

Note. 6027 epochs were included. IND = indeterminate sleep.

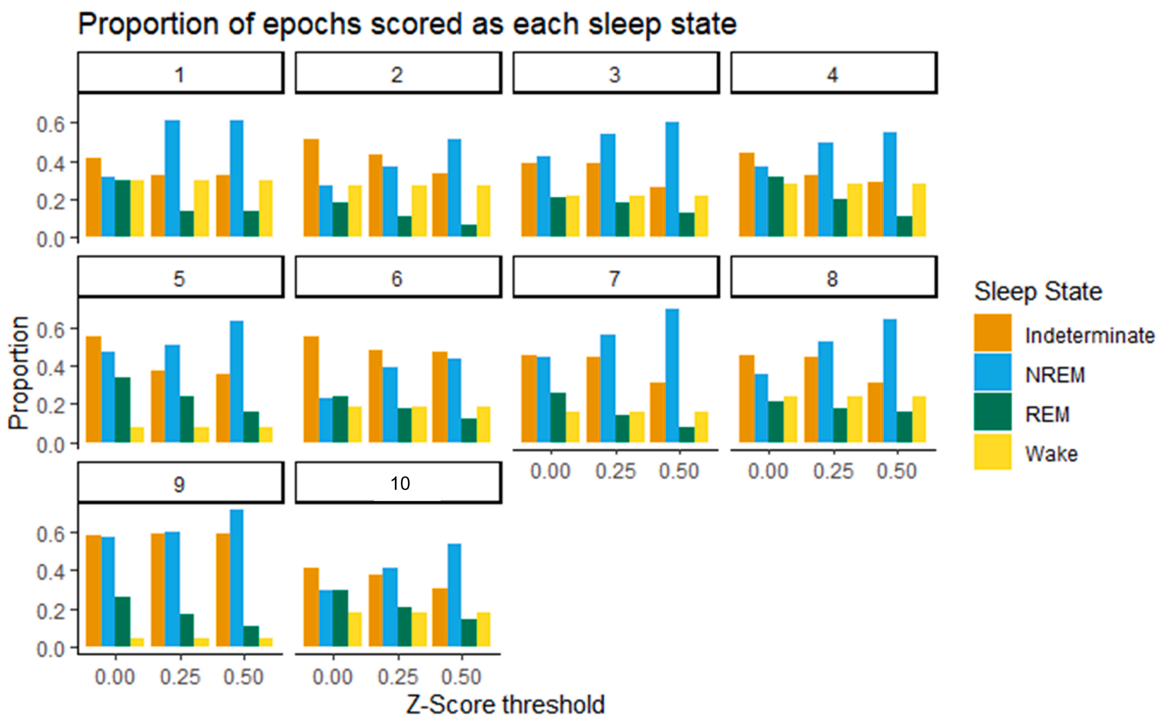

Fig. 4. Proportion of time spent in each state across the different z score thresholds for each infant.

3.2.1. *Interrater reliability*

One randomly chosen night from each infant was viewed by a secondary coder with the 1-hour window, resulting in 6027 minutes, or epochs, of reliability coding. Total epochs in which coders agreed ranged from 77% to 91% ($M = 85\%$) across participants and Cohen's Kappa ($\kappa = 0.77$) indicated substantial agreement (McHugh, 2012). Table 2 presents the frequency of each possible instance of match or mismatch. Matches in epochs defining REM and NREM were most common while mismatches typically related to indeterminate sleep.

3.3. *Automated algorithms*

The three z score thresholds were used to score each epoch of the nightly sleep data separately. Total state proportions are displayed for each infant in Fig. 4. Durations of each state were compared to results using PSG and, while none were a perfect match, the threshold of greater than 0.25 was the most accurate when considering all states (see Fig. 5).

3.3.1. *Bland-Altman plots*

Bland-Altman plots comparing visual scoring to the automated algorithm ($z > 0.25$) were produced in Fig. 6. While only 3 points exceeded the limits of agreement, the algorithm consistently underestimated the amount of time spent in REM and overestimated the time spent in indeterminate. The measures agreed more in estimating the time spent in NREM as evidenced by difference scores approaching zero. To address the bias, the algorithm was adjusted to only include respiratory rate. Prior researchers relied on this metric over heart rate (Isler et al., 2016). Based on the excess of indeterminate sleep, the final step of the algorithm – comparing state determinations from heart and respiratory rates – was the likely source of the bias.

Figure 7 displays Bland-Altman plots comparing the algorithm using only respiration (and not heart rate) to visual coding. Only REM and NREM were scored because indeterminate (IND) was previously indicated by a disagreement between heart and respiration.

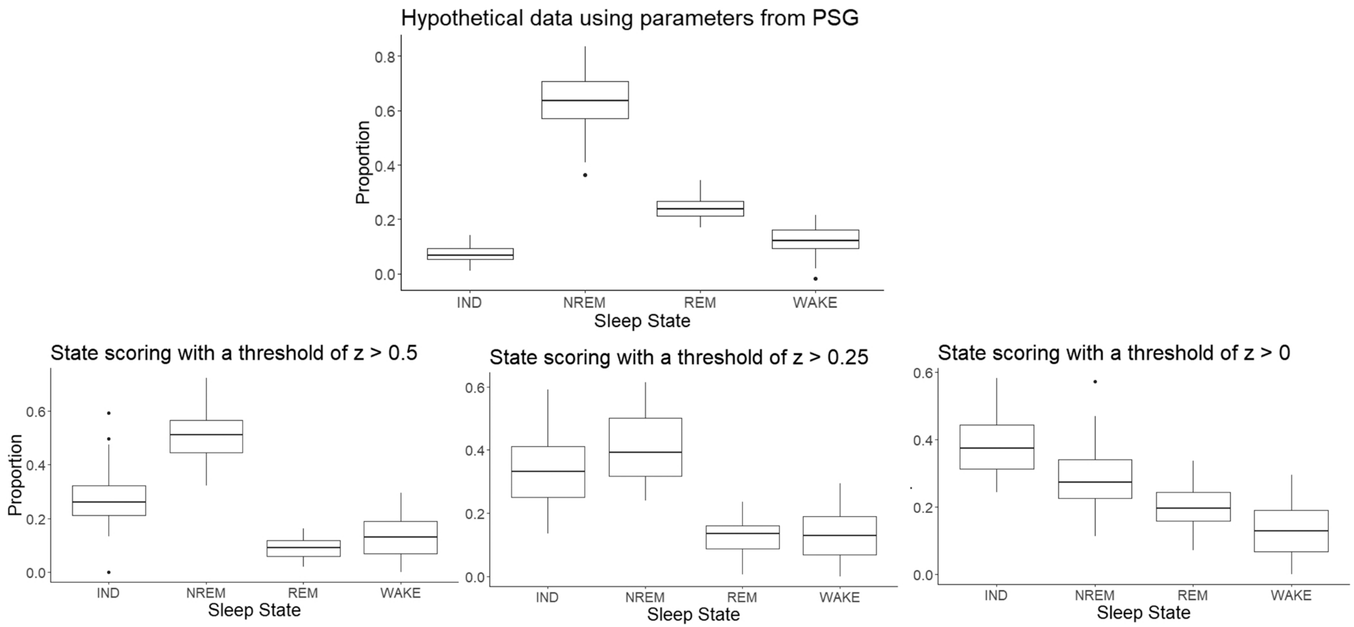

Fig. 5. Box and whisker plots of proportion of the night spent in each sleep state across coding thresholds. *Note.* The top plot was generated using rnorm and the means and standard deviations listed by Scholle et al. (2011) for 1 year old infants. Time awake was not explicitly listed but was calculated by subtracting total sleep time from total sleep period. N1 was used as a substitute for indeterminate sleep.

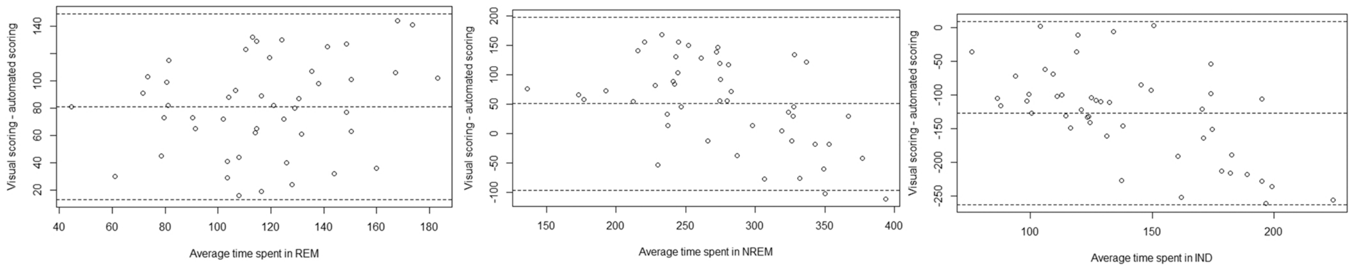

Fig. 6. Bland-Altman plots with limits of agreement comparing visual scoring to the automated algorithm (z < 0.25) across each sleep state.

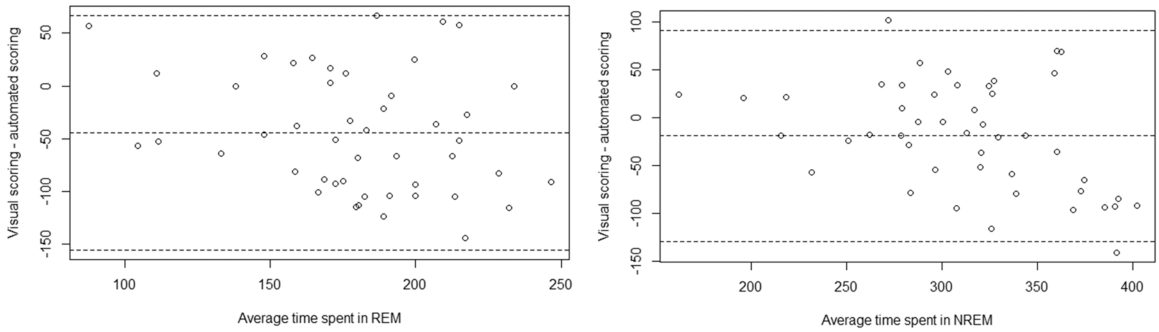

Fig. 7. Bland-Altman plots with limits of agreement comparing visual scoring to the automated algorithm (z < 0.25) using only respiration across each sleep state.

Difference scores were closer to zero, with a lesser, but consistent bias; the algorithm estimated more time spent in both states than visual coding. Again, 3 points exceeded the limits of agreement.

3.3.2. Equivalence tests

Infants slept an average of 624 min per night. As such, the equivalence interval was defined as − 62–62 min, accounting for a 10% mismatch. Visual state scoring was compared to durations of REM and NREM as defined by the respiration only algorithm. Methods were not equivalent in identifying REM sleep durations, $t(9) = 0.975$, $p = 0.177$, but were in marking NREM, $t(9) = 2.399$, $p < 0.05$.

4. Discussion

The current pilot project assessed the feasibility of a novel sleep measurement technique – wireless cardiorespiratory sensors and actigraphy – to study the ultradian cycle in one-year-old infants. Overwhelmingly, parents were able to implement the procedure with fidelity, resulting in less than 10% data loss. A procedure was developed for visually scoring the data in the Isansys Patient Status Engine with the appropriate viewing scale of one hour. The resulting data and proportion of time spent in each state was comparable to the results of age-matched peers who underwent polysomnography (Scholle et al., 2011). Interrater reliability was relatively high and in line with kappa values ($M_k = 0.76$) reported by a recent review of PSG scoring (Lee et al., 2022).

Several z score thresholds for the automated algorithm were compared and 0.25 emerged as most similar to age-matched PSG data. Scoring by the automated algorithm was not significantly different (exceeding the limits of agreement) than visual scoring, but it was systematically different. This bias was captured by the dotted line on Figs. 6 and 7 – the average difference score between the two methods. Adjusting the algorithm to only use respiration and dropping the final comparisons of this and heart rate was able to mitigate some bias though it was not eliminated entirely. Earlier iterations were skewed by an over classification of indeterminate sleep, or instances when heart and respiratory judgements disagreed. It may be that respiration is a more reliable index of sleep states, but it may also be an inadequate coding scheme for heart rate data, which may instead require more finesse. When assessed with a test of equivalence, the two methods were only designated as such for scoring NREM sleep.

The iterative approach to setting the z score threshold and viewing window was exploratory and will require replication for broader use. Differences between the current technique and those in the literature precluded an a priori approach. Reported PSG comparisons were built on simulated data using population parameters from age-matched peers. The lack of a concurrent gold standard comparison is the primary limitation of the work as an assessment of the accuracy of our method is contingent upon a ground truth comparison.

Future work is necessary to replicate the protocol. The final sample fell short of the total nights suggested by the power analysis and it is yet unknown how the method will cope with data from older or younger participants. Most research on ultradian cycle development is focused on newborns and early infancy because this period contains the most drastic changes. There is also a growing body of literature on premature infants' sleep in the NICU as new automated scoring techniques using cardiorespiratory parameters are developed to capitalize on standard monitoring procedures (de Groot et al., 2021). However, moving beyond these ages is worthwhile as the role of sleep states is suggested to continue to change from the second to third year of life (Cao et al., 2020).

5. Conclusions

While researchers have been studying sleep for decades, our knowledge is remarkably limited. Sleep creates a unique methodological conundrum that is further compounded when studied in infant populations. Techniques that provide the most detailed information about sleep are also the most invasive; the combined change in routine and application of unfamiliar equipment have been known to alter typical sleep patterns (Grigg-Damberger et al., 2007). The need for creative, minimally disruptive solutions to study sleep that do not compromise the richness and accuracy of the measurements is paramount. The implementation of this technology has the potential to fill this void.

Based on our preliminary results, the novel method appears to be a feasible choice for researchers interested in studying the ultradian cycle. The combination of actigraphy and cardiorespiratory monitoring is uniquely advantageous because it is less resource intensive and more naturalistic, being put on by caregivers while still resulting in high rates of good data. Such flexibility also proved invaluable in response to the waves of COVID-19 lockdowns, further underscoring the benefit of in-home data collection. Additionally, the use of an open-source platform (R version 4.0.0) for creating algorithms promotes transparency and collaboration. In its current form, the R code can be used to estimate total amounts of REM and NREM in a night. However, it can be built upon and specified in novel ways that will allow researchers to tailor it to their own research questions. Taken together, these are small steps in making sleep research more accessible to both scientists and the general public.

Author statement

The author does not have competing interests to declare. Research presented here was conceptualized, carried out, analyzed, and written up by the single author.

Data availability

I have shared my data and code in the Appendix.

Acknowledgements

This work was supported by two Doctoral Student Research Grants (Graduate Center, City University of New York) in 2018 and 2020, a PsiChi Graduate Research Grant, and a Provost's Dissertation Fellowship (Graduate Center, City University of New York) to M. N.H. and by the National Science Foundation, Division of Behavioral and Cognitive Science [1551703] to Sarah E. Berger. Portions of this research were presented at the 54th Annual Meeting of the International Society of Developmental Psychobiology, Chicago, IL, November, 2021; and the International Pediatric Sleep Association Congress Online, February, 2021. The manuscript was modified from Chapter 5: *Novel methods for sleep state coding* of M.N.H's dissertation. The author is currently a postdoctoral fellow at Temple University. I gratefully acknowledge the assistance of my advisor, Dr. Sarah E. Berger, and the Child Development Lab and thank the parents and infants who participated.

Appendix. Raw data files and R Studio code.

All participants' minute-by-minute cardiorespiratory and actigraphy data can be found on the Github repository: https://github.com/mhorger/StateScoring-HRRR-acti. The template for the automated scoring algorithm and completed exemplar code are uploaded as well.

References

Bernstein, P., Emde, R., & Campos, J. (1973). REM sleep in four-month infants under home and laboratory conditions. *Psychosomatic Medicine, 35*(4), 322–329.
Bland, J. M., & Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistical Methods in Medical Research, 8*(2), 135–160.
Cao, J., Herman, A. B., West, G. B., Poe, G., & Savage, V. M. (2020). Unraveling why we sleep: Quantitative analysis reveals abrupt transition from neural reorganization to repair in early development. *Science Advances, 6*(38), eaba0398.
Carskadon, M. A., & Dement, W. C. (2016). Normal human sleep. In H. Kryger, T. Roth, & W. C. Dement (Eds.), *Principles and Practice of Sleep Medicine (15-24)*. New York, NY: Elsevier.
de Groot, E. R., Knoop, M. S., van den Hoogen, A., Wang, X., Long, X., Pillen, S., Benders, M., & Dudink, J. (2021). The value of cardiorespiratory parameters for sleep state classification in preterm infants: A systematic review. *Sleeping Medicine Reviews, 58*, Article 101462.
Grigg-Damberger, M., Gozal, D., Marcus, C. L., Quan, S. F., Rosen, C. L., Chervin, R. D., Wise, M., Picchietti, D. L., Sheldon, S. H., & Iber, C. (2007). The visual scoring of sleep and arousal in infants and children. *Journal of Clinical Sleeping Medicine, 3*(2), 43.
Horger, M. N., Marsiliani, R., DeMasi, A., Allia, A. M., & Berger, S. E. (2021). Researcher choices for infant sleep assessment: Parent report, actigraphy, and a novel video system. *The Journal of Genetic Psychology, 182*(4), 218–235. In press.
Isler, J. R., Thai, T., Myers, M. M., & Fifer, W. P. (2016). An automated method for coding sleep states in human infants based on respiratory rate variability. *Developmental Psychobiology, 58*(8), 1108–1115. https://doi.org/10.1002/dev.21482
Jan, S. L., & Shieh, G. (2018). The Bland-Altman range of agreement: Exact interval procedure and sample size determination. *Computers in Biology and Medicine, 100*, 247–252.
Knoop, M. S., de-Groot, E. R., & Dudink, J. (2020). Current ideas about the roles of rapid eye movement and non–rapid eye movement sleep in brain development. *Acta Paediatrica, 110*(1), 36–44. https://doi.org/10.1111/apa.15485
Lee, Y. J., Lee, J. Y., Cho, J. H., & Choi, J. H. (2022). Interrater reliability of sleep stage scoring: a meta-analysis. *Journal of Clinical Sleeping Medicine, 18*(1), 193–202.
McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica, 22*(3), 276–282.
Sadeh, A., Acebo, C., Seifer, R., Aytur, S., & Carskadon, M. A. (1995). Activity-based assessment of sleep-wake patterns during the 1st year of life. *Infant Behavior and Development, 18*(3), 329–337.
Sadeh, A., De Marcas, G., Guri, Y., Berger, A., Tikotzky, L., & Bar-Haim, Y. (2015). Infant sleep predicts attention regulation and behavior problems at 3-4 years of age. *Developmental Neuropsychology, 40*(3), 122–123.
Sadeh, A., Lavie, P., Scher, A., Tirosh, E., & Epstein, R. (1991). Actigraphic home-monitoring sleep-disturbed and control infants and young children: a new method for pediatric assessment of sleep-wake patterns. *Pediatrics, 87*(4), 494–499.
Sadeh, A., Raviv, A., & Gruber, R. (2000). Sleep patterns and sleep disruptions in school-age children. *Developmental Psychology, 36*(3), 291.
Salgado, E. L. M., Guerrero, F. A., Santos, A. D. J. R., & del Olmo Alcántara, B. E. (2022). Evaluation of the first-night effect of polysomnographic recording on cardiac autonomic activity in children with autism spectrum disorder. *Salud Mental, 44*(6), 307–314.
Samson, D. R. (2021). Taking the sleep lab to the field: Biometric techniques for quantifying sleep and circadian rhythms in humans. *American Journal of Human Biology, 33*(6). https://doi.org/10.1002/ajhb.23541

Scholle, S., Scholle, H. C., Kemper, A., Glaser, S., Rieger, B., Kemper, G., & Zwacka, G. (2003). First night effect in children and adolescents undergoing polysomnography for sleep-disordered breathing. *Clinical Neurophysiology, 114*(11), 2138–2145.

Scholle, S., Wiater, A., & Scholle, H. C. (2011). Normative values of polysomnographic parameters in childhood and adolescence: Cardiorespiratory parameters. *Sleeping Medicine, 12*(10), 988–996. https://doi.org/10.1016/j.sleep.2011.05.006

Sheldon, S. H. (2012). Development of sleep in infants and children. *Principles and Practice of Pediatric Sleep Medicine: Second Edition* (pp. 17–23). Elsevier,. https://doi.org/10.1016/B978-1-4557-0318-0.00003-6

Werth, J., Atallah, L., Andriessen, P., Long, X., Zwartkruis-Pelgrim, E., & Aarts, R. M. (2017). Unobtrusive sleep state measurements in preterm infants – A review. *Sleeping Medicine Reviews, 32*, 109–122. https://doi.org/10.1016/j.smrv.2016.03.005