2023

Are parent-reported sleep logs essential? A comparison of three approaches to guide open source accelerometry-based nocturnal sleep processing in children

Christine W. St. Laurent

et. al.

Received: 7 July 2023 | Revised: 30 October 2023 | Accepted: 10 November 2023

DOI: 10.1111/jsr.14112

RESEARCH ARTICLE

Journal of Sleep Research · ESRS

Are parent-reported sleep logs essential? A comparison of three approaches to guide open source accelerometry-based nocturnal sleep processing in children

Sarah Burkart[1] | Michael W. Beets[1] | Christopher D. Pfledderer[2,3] | Lauren von Klinggraeff[1] | Xuanxuan Zhu[4] | Christine W. St. Laurent[5] | Vincent T. van Hees[6] | Bridget Armstrong[1] | R. Glenn Weaver[1] | Elizabeth L. Adams[1]

[1]Department of Exercise Science, Arnold School of Public Health, University of South Carolina, Columbia, South Carolina, USA

[2]University of Texas Health Science Center (UTHealth) at Houston, School of Public Health in Austin, Austin, Texas, USA

[3]Michael and Susan Dell Center for Healthy Living, UTHealth School of Public Health in Austin, Austin, Texas, USA

[4]Department of Epidemiology and Biostatistics, Arnold School of Public Health, University of South Carolina, Columbia, South Carolina, USA

[5]Department of Psychological and Brain Sciences, University of Massachusetts Amherst, Amherst, Massachusetts, USA

[6]Accelting, Almere, The Netherlands

Correspondence

Sarah Burkart, Department of Exercise Science, Arnold School of Public Health, University of South Carolina, Columbia, SC, US.

Email: sburkart@mailbox.sc.edu

Funding information

Eunice Kennedy Shriver National Institute of Child Health and Human Development, Grant/Award Number: F32HD105384; National Heart, Lung, and Blood Institute, Grant/Award Number: F31HL158016; National Institute of Diabetes and Digestive and Kidney Diseases, Grant/Award Number: R01DK116665; National Institute of General Medical Sciences, Grant/Award Number: P20GM130420

Summary

We examined the comparability of children's nocturnal sleep estimates using accelerometry data, processed with and without a sleep log. In a secondary analysis, we evaluated factors associated with disagreement between processing approaches. Children ($n = 722$, age 5–12 years) wore a wrist-based accelerometer for 14 days during Autumn 2020, Spring 2021, and/or Summer 2021. Outcomes included sleep period, duration, wake after sleep onset (WASO), and timing (onset, midpoint, waketime). Parents completed surveys including children's nightly bed/wake time. Data were processed with parent-reported bed/wake time (sleep log), the Heuristic algorithm looking at Distribution of Change in Z-Angle (HDCZA) algorithm (no log), and an 8 p.m.–8 a.m. window (generic log) using the R-package 'GGIR' (version 2.6-4). Mean/absolute bias and limits of agreement were calculated and visualised with Bland–Altman plots. Associations between child, home, and survey characteristics and disagreement were examined with tobit regression. Just over half of nights demonstrated no difference in sleep period between sleep log and no log approaches. Among all nights, the sleep log approach produced longer sleep periods (9.3 min; absolute mean bias [AMB] = 28.0 min), shorter duration (1.4 min; AMB = 14.0 min), greater WASO (11.0 min; AMB = 15.4 min), and earlier onset (13.4 min; AMB = 17.4 min), midpoint (8.8 min; AMB = 15.3 min), and waketime (3.9 min; AMB = 14.8 min) than no log. Factors associated with discrepancies included smartphone ownership, bedroom screens, nontraditional parent work schedule, and completion on weekend/summer nights (range = 0.4–10.2 min). The generic log resulted in greater AMB among sleep outcomes. Small mean differences were observed between nights with and without a sleep log. Discrepancies existed on weekends, in summer, and for children with smartphones and screens in the bedroom.

KEYWORDS
accelerometry, children, open-source processing, parent-report, sleep diary

1 | INTRODUCTION

Sleep is important for children's healthy development (Chaput et al., 2016), and various measurement techniques exist to evaluate children's sleep metrics. Accelerometry is a commonly used and well-accepted measure of multi-day free-living sleep in children (Meltzer, Montgomery-Downs et al., 2012; Sadeh, 2011; Spruyt et al., 2011). However, measuring sleep via accelerometry in children is not without limitations. Due to the single stream of data (i.e., movement), accelerometry processing algorithms struggle to differentiate absence of movement during quiet rest from sleep and movement during sleep periods as wake (Sadeh, 2015). To overcome this limitation, collecting participant-reported (e.g., parent or child) bed/wake times via a sleep log is recommended for studies utilising accelerometry to measure free-living sleep to augment device-based estimates (Meltzer, Montgomery-Downs et al., 2012).

A sleep log (or diary) is used in the processing of accelerometry data to help an algorithm more accurately classify a sleep period. Theoretically, the addition of more information (i.e., bed/wake times from a sleep log) in device-based processing should lead to more accurate sleep estimates. However, under real study conditions, there is a risk of missing sleep log data. When studying children's sleep, parents completing a daily survey of bed and wake times can be inherently burdensome, and often results in nights without a completed log (Tétreault et al., 2018). Parents may misreport bed and/or wake times or may not know what children are doing after they go to bed (e.g., engaging in screen time), especially as children get older and have access to screens in the bedroom. It is not uncommon for studies using sleep logs to insufficiently report compliance rates or remove nights without a sleep log, which influences the number of nights included in analyses and may introduce bias (Tétreault et al., 2018). It is estimated up to 28% of data may be lost due to participant non-compliance (e.g., device wear or log completion) (Acebo et al., 1999). Understanding factors that influence sleep log reporting can help inform when sleep logs are needed and when they may bias sleep estimates. Further, in instances where sleep logs are missing or to reduce participant burden completely, it may be useful to apply generic sleep information to help guide detection.

Two studies ($n = 38$ aged 8–12 years (Meltzer, Walsh et al., 2012) and $n = 60$ aged 6–10 years (Tétreault et al., 2018)) in children have compared estimates when processing device-based nocturnal sleep data with and without a sleep log. Small mean differences were observed between methods (e.g., <29 min for sleep duration) (Meltzer, Walsh et al., 2012; Tétreault et al., 2018). Both studies recommended a modest number of nights without a sleep log could be combined with nights with a sleep log, especially if researchers are interested in individual person-level differences (Meltzer, Walsh et al., 2012; Tétreault et al., 2018). However, these studies were conducted using algorithms with uncertain replicability, as comprehensive descriptions of algorithm methods do not exist. van Hees et al. compared their open-source algorithm for detecting the sleep period with and without a sleep log in healthy adults and sleep clinic patients undergoing polysomnography and found good agreement for detecting the sleep period (mean C-statistic = 0.83–0.86), although deviations were large (i.e., >2 h) for those with more fragmented sleep who had longer wake periods during the night (van Hees et al., 2015; van Hees et al., 2018). In 47 healthy adults' accelerometry processed with and without a sleep log, Plekhanova et al. (2020) identified poor reliability in sleep period and sleep duration detection between log approaches, so caution was advised when interpreting findings across studies with and without a sleep log. Given the current knowledge of how adult estimates differ with and without a sleep log when processing with an open-source algorithm, there is a need to understand if these estimates differ in children.

The aim of this study was to compare nocturnal sleep outcomes generated with R package 'GGIR' software (Migueles et al., 2019) when using a parent-reported sleep log to guide parameter detection compared to the default algorithm without a sleep log in a large sample of elementary school-aged children's free-living sleep. We also examined if using a generic 8 p.m.-8 a.m. sleep log to guide sleep period detection resulted in differences compared to the default algorithm. As no criterion measure of children's sleep was used in this study, we instead focus on the agreement between methods rather than accuracy. Finally, to understand characteristics associated with disagreement between processing methods, we examined if child, home, or survey completion characteristics were associated with these differences.

2 | METHODS

2.1 | Participants

Children were recruited through two neighbouring school districts in the southeastern United States to participate in a multi-year longitudinal cohort study between September 2019 and January 2021 (recruitment paused and re-started due to COVID-19 school closures). All children in kindergarten through fifth grade at participating schools were invited to participate. Families were informed of the study by paper copies sent home, a link to online information distributed via email, text, and school newsletter, and researcher presence at school events. Parents completed an informed consent packet either via hard copy, which was signed and returned to the child's school or via online survey platform. No exclusion criteria were used prior to recruitment. Data collection began only after schools re-opened.

2.2 | Procedure

Children were asked to wear an Actigraph GT9X accelerometer (Actigraph LLC, Pensacola, FL, USA) on their non-dominant wrist 24 h/day for 14 days during Autumn 2020, Spring 2021, and Summer 2021. At the beginning of each wave, parents completed a survey describing their child, family, and home characteristics. The accelerometer protocol was identical across all three waves of data collection. Prior to the measurement period, participants were mailed an envelope that contained an accelerometer mounted in an adjustable wristband, an instruction sheet with frequently asked questions, and a prepaid return envelope. Participants were instructed to wear the accelerometer snug on the non-dominant wrist at all times, only to remove the device during water-based activities (e.g., bathing, swimming) or during organised sports (e.g., if rules banning jewellery during competition exist). Parents received a brief survey texted to their smartphone nightly during each 14-day period in which they provided information about their child's day and previous night's sleep (e.g., reasons for non-wear, bed/wake times). Upon return, accelerometers were downloaded, and data were prepared for processing. Parents were compensated $25 at each measurement period for returning the accelerometer and completing at least 10 of the nightly diaries. Children received their choice of new sports equipment (e.g., basketball, soccer ball, football, jump rope) upon return of the accelerometer. This study was approved by the lead author's Institutional Review Board (Pro00080382).

2.3 | Accelerometry

The Actigraph GT9X accelerometer is a triaxial research-grade accelerometer frequently used in studies measuring children's free-living 24-h behaviours (i.e., physical activity, sedentary behaviour, sleep). Actigraph GT9X accelerometers were initialised and downloaded using Actilife software (version 6.13.4, Actigraph LLC). Accelerometers were initialised to record data at a frequency of 30 Hz and begin data collection at 7:00 a.m. on the day preceding device delivery. Stop time was not used. Idle sleep mode was enabled to preserve battery life and the display was turned off to limit distractions for children while attending school. Data were downloaded and saved in raw format as .gt3x files and converted to .csv files for processing.

2.4 | Sleep log

Parents were prompted to complete a brief (<5 min) nightly survey sent to their smartphone via text message at 7:30 p.m. during each 14-day measurement period in which they were asked to report the time at which their child went to bed on the previous night and woke up that morning. Parents also reported reasons for accelerometer non-wear (e.g., device lost/misplaced, wristband broken, uncomfortable). Nightly surveys expired 24 h after they were sent to limit potential recall bias. The text message prompt approach aimed to increase ease of completion for participants and provide information to the research team as to when the log was completed. Throughout this manuscript, the term 'sleep log' will be used to describe the parent-reported sleep log. Data from nightly sleep logs were exported from Qualtrics (Qualtrics, Provo, UT, USA) and reformatted to fit the advanced sleep log criteria for processing in GGIR.

2.5 | Accelerometry processing

Data were processed with GGIR in three ways: (i) with a parent-reported sleep log (van Hees et al., 2018), (ii) without a sleep log using GGIR's default Heuristic algorithm looking at Distribution of Change in Z-Angle (HDCZA) algorithm (van Hees et al., 2015), and (iii) with a generic sleep log, using the GGIR package (Version 2.6–4) in R (Version 4.0.3) (Migueles et al., 2019). Briefly, the sleep detection in GGIR works as follows. First, it classifies sustained inactivity bouts in which the wrist z-angle does not change by >5° for ≥5 min (van Hees et al., 2015). Next, the sleep period time window is defined from the start of the first and the end of the last sustained inactivity bout that overlap with a guider window. The guider is either the sleep log, the estimate of the HDCZA algorithm that does not rely on a sleep log, or a generic sleep log. The HDCZA algorithm identifies all periods of >60 min in the day with lack of arm angle changes, ignores gaps between those periods <30 min, and then identifies the largest of those merged periods as what is referred to as the guider window.

The GGIR script used can be found in File S1. The following sleep variables produced in GGIR's part 4 full output were examined: sleep period duration (sleep onset to wake time including any time spent awake during the night; 'sptduration'), sleep onset (time at which the sleep period began; 'sleeponset_ts'), waketime (time at which sleep period ended, waking time; 'wakeup_ts'), sleep duration (accumulated nocturnal sustained inactivity bouts within the sleep period; 'sleepdurationinspt'), and wake after sleep onset (WASO; sleep duration subtracted from sleep period). Midpoint was calculated as halfway between onset and waketime.

In the no log approach, raw accelerometry data were processed using the default HDCZA algorithm, which is described in detail elsewhere (van Hees et al., 2018). The 'loglocation' argument was not specified, a new output folder was created, and the 'def.noc.sleep' argument was set to a single integer (i.e., def.noc.sleep = c(1)), which detects the sleep period using the HDCZA algorithm (van Hees et al., 2018). In the sleep log approach, parent-reported bed and wake times were entered into a .csv file in GGIR's advanced sleep log format. If a parent-reported log was incomplete or missing for a given night, this night was excluded from analysis. The sleep log was denoted by directing GGIR to the location of the appropriate .csv with the 'loglocation' argument. In the generic log approach, research staff created an alternative log where the parent-reported sleep log entries were altered to a generic 8:00 p.m. bedtime and 8:00 a.m. waketime. For example, if a parent reported a 10:00 p.m. bedtime and a 6:00 a.m. waketime for their child on a given night, these values were replaced with an 8:00 p.m. bedtime and 8:00 a.m. waketime. Changes were only made to the log on nights when parents completed the log. If the parent-reported log was incomplete or missing for a given night, no manipulation occurred to ensure equal sample sizes of nights within a child across all three approaches.

2.6 | Associations with disagreement

Associations between child, home, and survey characteristics and disagreement between the sleep log and no log approaches were examined. Parents reported their child's date of birth (from which age was calculated at the start of each wave, continuous), gender (male/female), and smartphone ownership (yes/no). They also provided information about the home including number of screens accessible for the child in the bedroom/home (continuous – range 0 to 10+ screens, 10+ screens was analysed as 10), number of children living in the home (continuous), if the child shared a bedroom (yes/no), and which shifts the parent worked only if they reported working full- or part-time (e.g., early morning, day, evening, night, multiple shifts). We were also able to extract data from the nightly survey completions (i.e., school night [Sunday–Thursday] versus weekend night [Friday–Saturday], summer versus school year, time to completion from when the survey text message was sent) to examine if when a survey was completed was associated with disagreement.

2.7 | Data analysis

For inclusion in these analyses, children were required to have at least 1 night of valid sleep data (defined as ≥16 h of wear time between noon and noon, or between 6:00 p.m. and 6:00 p.m. if the parent reported a waketime after noon) (van Hees et al., 2015) and a corresponding parent-reported sleep log entry for that night. Nights were excluded if: (i) parent-reported bedtime or waketime was not entered (i.e., log not completed), (ii) the device-detected sleep period (onset to waketime) was <160 min, (iii) the device-detected waketime was before 5:00 a.m. or after 1:00 p.m., or (iv) the device-detected onset was before 7:00 p.m. or after 6:00 a.m. (Barreira et al., 2015).These parameters were used to focus analyses solely on nocturnal sleep.

Descriptive statistics were calculated for sample characteristics and sleep outcomes of interest across each wave of data collection. As sleep outcomes were assessed on multiple days across three waves within the same children, linear mixed models were used to estimate means at each wave to account for clustering of nights within children. We compared output from GGIR between approaches (i.e., with and without a log). Similar to Galland et al. (2016), we discuss agreement rather than accuracy as none of these approaches are considered the 'gold standard' of sleep assessment in children. We assessed agreement of sleep outcomes between processing approaches (i.e., sleep log versus no log and generic log versus no log) by calculating mean bias and limits of agreement (LoA) separately for each sleep outcome of interest (onset, midpoint, waketime, sleep period duration, sleep duration, and WASO) (Bland & Altman, 1986). Comparisons across approaches were visually represented with Bland–Altman plots with the no log approach on the *x*-axis and difference between sleep log or generic log and no log on the *y*-axis (e.g., sleep log – no log or generic log – no log). Positive bias values indicate an overestimation of the sleep parameter of interest in the type of log used (i.e., sleep log or generic log) compared to the no log approach. Negative bias values indicate an underestimation in the type of log used compared to the no log approach. Multilevel models with the difference between approaches as the independent variable and the no log approach as the dependent variable were used to determine whether statistically significant trends in bias were present.

Disagreement was calculated as the sleep period ('sptduration') identified in the no log approach subtracted from the sleep log approach (i.e., sleep log – no log). No difference in 'sptduration' (to three decimal places as produced by GGIR) between approaches indicated perfect agreement. The absolute value of differences between approaches was calculated and used in subsequent analyses. The percentage of nights that demonstrated agreement, as well as those that fell within 10-min intervals ranging from 0 to 60+ min (i.e., >0 to ≤10 min, >10 to ≤20 min, etc.) were described. We conducted multilevel tobit regression analyses with nights nested within children to examine if child, home, and/or survey completion characteristics were associated with disagreement between the sleep log and no log approaches. Tobit regression was selected as the absolute value of the difference between approaches (dependent variable) was bounded at zero. All analyses were conducted in Stata (version 16; StataCorp LLC, College Station, TX, USA).

3 | RESULTS

After removal for non-wear, device malfunction, and lost devices, analyses included 697 children in Autumn, 522 children in Spring, and 410 children in Summer (Table 1). A total of 722 unique children participated in at least one wave of data collection, which resulted in 13,466 nights of sleep data across all three waves. Children had a minimum of one and a maximum of 16 valid 24-h periods in any wave of data collection paired with a parent-reported sleep log. Notably, 6429 nights with valid accelerometer data (~32%) were unusable due to the absence of a completed sleep log. Descriptive data for sleep outcomes are presented in Table 2 by approach (i.e., sleep log, no log, generic log).

3.1 | Sleep log vs. No log

The mean sleep period was 9.1 h in the sleep log approach and 9.0 h in the no log approach (Table 2). As shown in Figure 1a and Table 3,

TABLE 1 Sample characteristics by each wave of data collection.

Characteristic	Autumn 2020	Spring 2021	Summer 2021
Number of children	697	522	410
Number of nights	5814	4355	3297
Mean (SD):			
Nights per child	8.4 (3.9)	8.3 (3.9)	8.1 (4.0)
Time to complete nightly survey (h)[a]	2.5 (4.1)	2.7 (4.2)	3.1 (4.6)
Age (years)	8.3 (1.7)	8.8 (1.7)	9.1 (1.7)
Number of screens in home[b]	7.6 (2.5)	7.6 (2.5)	7.6 (2.6)
Number of screens in bedroom[b]	1.2 (1.6)	1.2 (1.6)	1.1 (1.7)
Number of children in home	2.5 (1.0)	2.5 (1.0)	2.5 (1.0)
N (%):			
Gender			
Female	355 (50.9)	250 (47.9)	196 (47.8)
Male	342 (49.1)	272 (52.1)	214 (52.2)
Race			
Black	210 (30.1)	131 (25.1)	91 (22.2)
White	408 (58.5)	334 (64.0)	276 (67.3)
Other[c]	79 (11.4)	57 (10.9)	43 (10.5)
Household income			
<$10,000– $29,999/year	89 (12.8)	54 (10.4)	39 (9.5)
$30,000–$59,999/year	166 (23.8)	113 (21.6)	76 (18.5)
$60,000–$99,999/year	202 (29.9)	153 (29.3)	126 (30.7)
>$100,000/year	225 (32.3)	193 (37.0)	159 (38.8)
Not reported	15 (2.2)	9 (1.7)	10 (2.4)
Smartphone ownership	170 (24.4)	140 (26.8)	101 (24.6)
Share bedroom	264 (37.9)	183 (35.1)	145 (35.4)
Parent work schedule			
Day (8:00 a.m.–6:00 p.m.)	441 (63.3)	302 (57.9)	252 (61.5)
Early morning, night, multiple	153 (22.0)	104 (19.9)	61 (14.9)
Not reported/do not work	103 (14.8)	116 (22.2)	97 (23.7)

[a]Time to complete represents the amount of time from when a parent was sent the nightly survey (7:30 p.m., reporting previous night's bedtime and that morning's wake time) to when it was submitted.
[b]Number of screens in the home/bedroom ranged from 0 to ≥10 screens. Those selecting ≥10 screens in the home/bedroom were analysed as 10 screens.
[c]Other encompasses American Indian, Asian, Hawaiian or Pacific Islander, More than One Race, Not Reported.

sleep period identification was overestimated by 9.3 min in the sleep log approach compared to the no log approach (absolute mean bias [AMB] = 28.0 min) while sleep duration (Figure 1b, Table 3) was underestimated by 1.4 min (AMB = 14.0 min). WASO (Figure 1c, Table 3) was overestimated by 11.0 min (AMB = 15.4 min). Sleep timing was underestimated by 13.4, 8.8, and 3.9 min for onset (Figure 1d, Table 3, AMB = 17.4 min), midpoint (Figure 1e, Table 3, AMB = 15.3 min), and waketime (Figure 1f, Table 3, AMB = 12.8 min), respectively. Use of the parent sleep log resulted in absolute differences ranging from 14 to 28 min among sleep outcomes of interest.

There were statistically significant negative trends in bias such that as each sleep outcome increased or shifted later, the sleep log approach underestimated sleep period (coefficient ± standard error −0.64 ± 0.01, 95% confidence interval [CI] = −0.66, −0.62), sleep duration (B = −0.76 ± 0.02, 95% CI = −0.79, −0.72), WASO (B = −0.21 ± 0.01, 95% CI = −0.23, −0.19), onset (B = −0.86 ± 0.01, 95% CI = −0.89, −0.83), midpoint (B = −1.17 ± 0.02, 95% CI = −1.21, −1.14), and waketime (B = −1.04 ± 0.01, 95% CI = −1.07, −1.02).

Among all nights, 53.1% of nights (*n* = 7150) demonstrated perfect agreement in identifying the sleep period between the sleep log and no log approach (Figure 2). An additional 21.0% of nights fell within 30 min of the identified sleep period (total 74.1% of nights) and 33.4% of nights fell within 60 min (total 86.5% of nights) of the identified sleep period. Finally, 13.5% of nights demonstrated disagreement of >60 min. Additional reporting of agreement within 10-min windows is presented in Figure 2 and Table S1.

We conducted multilevel tobit regressions to examine the relationship between the absolute value of differences in sleep period identification and child, home, and survey completion characteristics. Regression coefficients are presented in Table 5. Child ownership of a smartphone was associated with a 9.1 min increase in absolute difference between approaches (95% CI = 5.66, 12.54). Each additional screen in the bedroom (range = 0–10) was associated with a 3.2 min increase in absolute difference between approaches (95% CI = 2.34, 4.13). When parents completed a nightly survey on a weekend night or summer night, this was associated with a 7.9 and 8.0 min increase in absolute difference between approaches, respectively. Each additional screen in the home was associated with a 0.8 min decrease in absolute difference between approaches. Compared to working during the day, working other times (i.e., early morning, night, multiple shifts) was associated with a 3.2 min increase in absolute difference between approaches among parents who reported working full- or part-time. Each additional hour it took for parents to complete the nightly survey (range 0–24 h) was associated with a 0.4 min absolute difference. As parents were asked to recall bedtime from the previous night, even an immediate response would require a parent to recall bedtime from up to 24 h ago. Thus, if a parent took nearly 24 h to complete the nightly survey from when it was sent, a parent may be recalling ~48 h ago, which would translate to an ~9.2 min absolute difference between approaches.

3.2 | Generic log versus No log

The sleep period was 9.5 h in the generic log approach and 9.0 h in the no log approach (Table 2). As shown in Figure 1a and Table 4, sleep period identification was overestimated by 30.1 min in the generic log approach compared to the no log approach (AMB = 46.9 min) while sleep duration (Figure 1b, Table 4) was underestimated by 0.03 min (AMB = 18.6 min). WASO (Figure 1c, Table 4) was overestimated by 30.3 min (AMB = 33.5 min). Sleep timing was underestimated by 36.1, 21.1, and 5.8 min for onset (Figure 1d, Table 4, AMB = 36.7 min), midpoint (Figure 1e, Table 4, AMB = 27.0 min), and waketime (Figure 1f, Table 4, AMB = 21.6 min), respectively. Use of the generic log resulted in absolute differences ranging from 18 to 47 min among sleep outcomes of interest.

There were statistically significant negative trends in bias such that as each sleep outcome increased or shifted later, the generic log approach underestimated sleep period ($B = -0.48 \pm 0.01$, 95% CI = −0.49, −0.46), sleep duration ($B = -0.73 \pm 0.02$, 95% CI = −0.76, −0.70), WASO ($B = -0.12 \pm 0.01$, 95% CI = −0.13, −0.11), onset ($B = -0.59 \pm 0.01$, 95% CI = −0.61, −0.57), midpoint ($B = -1.03 \pm 0.01$, 95% CI = −1.05, −1.01), and waketime ($B = -1.29 \pm 0.01$, 95% CI = −1.30, −1.27).

TABLE 2 Sleep characteristics of the sample using each of the three processing methods.

Sleep outcomes, mean ± SE	Autumn 2020			Spring 2021			Summer 2021		
	Log	No log	Generic log	Log	No log	Generic log	Log	No log	Generic log
Onset	11:00 p.m. ± 2.28	11:12 p.m. ± 2.60	9:40 p.m. ± 1.95	10:52 p.m. ± 2.43	11:03 p.m. ± 2.90	9:30 p.m. ± 2.08	10:40 p.m.± 3.68	11:00 p.m.± 4.47	10:05 p.m.± 3.18
Midpoint	2:32 a.m. ± 1.65	2:40 a.m. ± 1.98	2:23 a.m. ± 1.35	2:22 a.m. ± 1.83	2:28 a.m. ± 2.17	2:13 a.m. ± 1.50	3:16 a.m. ± 3.08	3:30 a.m. ± 3.75	2:50 a.m. ±2.12
Waketime	7:05 a.m. ± 1.33	7:10 a.m. ±1.60	7:06 a.m. ± 1.03	6:53 a.m. ± 1.62	6:55 a.m. ± 1.82	6:57 a.m. ± 1.32	7:53 a.m. ± 2.88	8:00 a.m. ± 3.38	7:37 a.m. ±1.52
Sleep period duration, h	9.11 ± 0.03	8.99 ± 0.03	9.46 ± 0.03	9.06 ± 0.03	8.89 ± 0.03	9.43 ± 0.03	9.21 ± 0.04	9.03 ± 0.04	9.53 ± 0.04
Sleep duration, h	7.63 ± 0.03	7.68 ± 0.03	7.69 ± 0.03	7.62 ± 0.04	7.62 ± 0.04	7.69 ± 0.04	7.65 ± 0.04	7.72 ± 0.04	7.53 ± 0.05
WASO, min	1.49 ± 0.03	1.33 ± 0.04	1.77 ± 0.03	1.44 ± 0.04	1.28 ± 0.04	1.73 ± 0.03	1.56 ± 0.04	1.32 ± 0.04	2.00 ± 0.04

Abbreviations: SE, standard error; WASO, wake after sleep onset.

Journal of Sleep Research ESRS

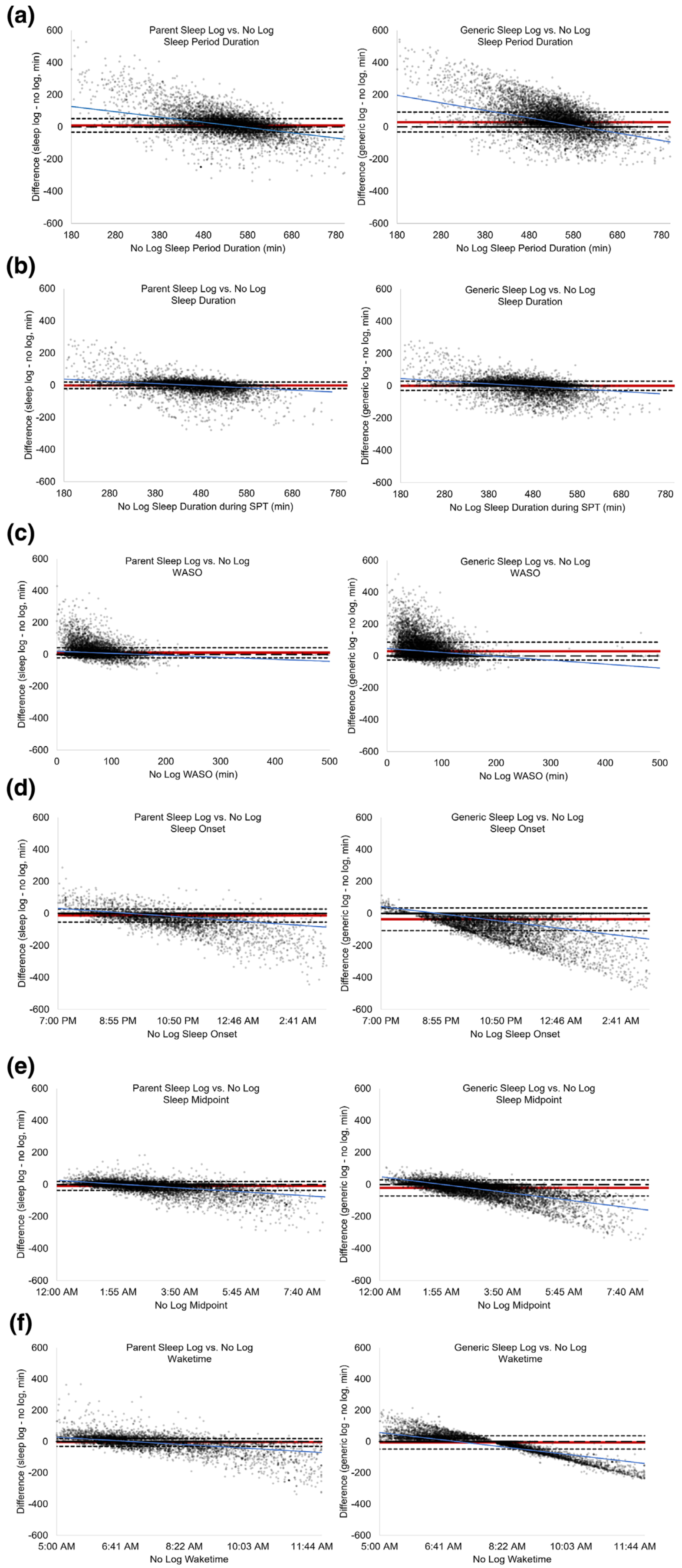

FIGURE 1 Bland–Altman plot of agreement for sleep period time identification (a), sleep duration (b), wake after sleep onset (WASO) (c), sleep onset (d), sleep midpoint (e), and waketime (f) comparing the sleep log (left side) and generic log (right side) to the no log approach. The red line indicates the mean bias while the dashed lines indicate the 95% limits of agreement. The solid blue line indicates the regression line where the difference between the log approach and no log approach is the dependent variable, and the no log approach sleep outcome of interest is the independent variable.

TABLE 3 Mean and absolute agreement comparing the sleep log and no log approaches.

	Agreement, mean bias (min)			Agreement, absolute bias (min)		
	Bias	SD	95% limits of agreement	Bias	SD	95% limits of agreement
Sleep period	9.34	21.34	(−32.48, 51.15)	28.03	23.18	(−17.41, 73.47)
Sleep duration	−1.36	10.39	(−21.73, 19.01)	14.03	11.04	(−7.61, 35.67)
Onset	−13.39	20.82	(−54.19, 27.41)	17.44	20.64	(−23.00, 57.89)
Midpoint	−8.83	14.14	(−36.53, 18.88)	15.25	13.86	(−11.90, 42.41)
Waketime	−3.87	13.67	(−30.67, 22.92)	14.78	12.75	(−10.21, 39.78)
WASO	11.01	16.12	(−20.59, 42.61)	15.38	16.39	(−16.75, 47.51)

Note: Positive bias values indicate an overestimation of the sleep parameter of interest in the type of log used (i.e., sleep log or generic log) compared to the no log approach. Negative bias values indicate an underestimation in the type of log used compared to the no log approach.
Abbreviations: SD, standard deviation; WASO, wake after sleep onset.

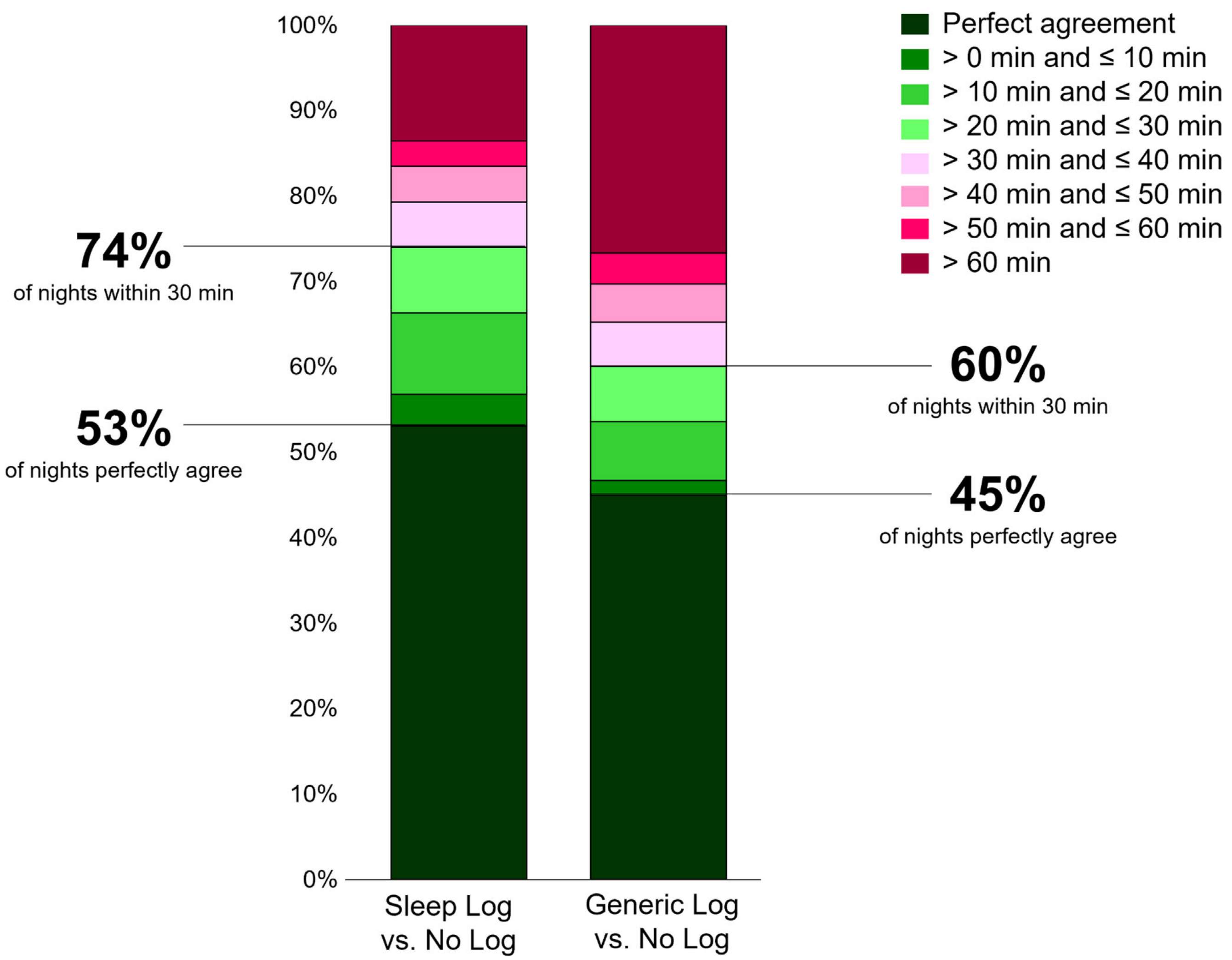

FIGURE 2 Magnitude of disagreement in sleep period identification between the sleep log and no log approaches and between the generic log and no log approaches.

Among all nights, 44.9% of nights ($n = 6051$) demonstrated perfect agreement in identifying the sleep period between the sleep log and no log approach (Figure 2). An additional 15.2% of nights fell within 30 min of the identified sleep period (total 60.1% of nights) and 28.4% fell within 60 min (total 73.3% of nights) of the identified sleep period. Finally, 26.7% of nights demonstrated disagreement >60 min. Additional reporting of agreement within 10-min windows is presented in Figure 2 and Table S2.

TABLE 4 Mean and absolute agreement comparing the generic log and no log approaches.

	Agreement, mean bias (min)			Agreement, absolute bias (min)		
	Bias	SD	95% limits of agreement	Bias	SD	95% limits of agreement
Sleep period	30.05	31.60	(−31.88, 91.98)	46.88	33.10	(−18.00, 111.76)
Sleep duration	−0.03	14.86	(−29.15, 29.09)	18.59	14.65	(−10.13, 47.30)
Onset	−36.09	36.52	(−107.70, 35.48)	36.67	35.63	(−33.16, 106.49)
Midpoint	−21.07	25.44	(−70.93, 28.79)	27.01	22.91	(−17.90, 71.82)
Waketime	−5.75	21.82	(−48.50, 37.01)	21.60	18.07	(−13.82, 57.02)
WASO	30.32	28.48	(−25.50, 86.14)	33.49	28.41	(−22.20, 89.18)

Note: Positive bias values indicate an overestimation of the sleep parameter of interest in the type of log used (i.e., sleep log or generic log) compared to the no log approach. Negative bias values indicate an underestimation in the type of log used compared to the no log approach.
Abbreviations: SD, standard deviation; WASO, wake after sleep onset.

TABLE 5 Multilevel tobit regression analyses examining child, home environment, and survey completion characteristics as predictors of disagreement (absolute value of the difference between approaches).

	Simple regression				Multiple regression			
	Coef.	SE	*p*	95% CI	Coef.	SE	*p*	95% CI
Child								
Age	1.57	0.51	0.002	0.57, 2.58	0.78	0.53	0.140	−0.25, 1.81
Sex (female)	1.17	1.73	0.498	−2.21, 4.55	−0.44	1.68	0.794	−3.72, 2.85
Smartphone ownership	12.88	1.58	0.000	9.79, 15.98	9.10	1.76	0.000	5.66, 12.54
Home environment								
Number of screens in bedroom	3.34	0.41	0.000	2.53, 4.15	3.24	0.46	0.000	2.34, 4.13
Number of screens in home	−0.74	0.26	0.004	−1.24, −0.24	−1.18	0.28	0.000	−1.72, −0.64
Number of children in home	0.25	0.81	0.754	−1.33, 1.83	−0.73	0.92	0.430	−2.54, 1.08
Shared bedroom	3.93	1.49	0.008	1.01, 6.85	3.18	1.67	0.057	−0.10, 6.45
Parent work schedule[a]								
Early morning	5.41	4.42	0.222	−3.27, 14.09	9.46	4.23	0.025	1.16, 17.75
Evening	9.98	4.4	0.023	1.36, 18.6	10.17	4.23	0.016	1.88, 18.46
Night	−9.82	4.12	0.017	−17.9, −1.74	−5.78	3.96	0.166	−13.25, 2.28
Multiple shifts	4.84	1.81	0.007	1.3, 8.38	5.38	1.73	0.002	1.99, 8.78
Survey completion								
Weekend[b]	7.76	0.87	0.000	6.05, 9.46	7.88	0.98	0.000	5.96, 9.80
Summer[c]	8.52	0.95	0.000	6.65, 10.39	8.01	1.12	0.000	5.82, 10.19
Time to complete (h)	0.52	0.11	0.000	0.31, 0.73	0.36	0.16	0.002	0.14, 0.59

[a]Referent group for parent work schedule = 9 a.m.–5 p.m.
[b]Referent group for weekend (Friday–Saturday night) = weeknights (Sunday–Thursday night).
[c]Referent group for summer = school year (i.e., combined autumn and spring measures).
Abbreviations: CI, confidence interval; Coef., coefficient; SE, standard error.

4 | DISCUSSION

This study aimed to compare sleep outcomes generated by the GGIR package using a parent-reported sleep log to guide sleep period detection compared to the default HDCZA algorithm with no log in a large sample of elementary school-aged children. We also compared the use of a generic 8 p.m.–8 a.m. sleep log to the default algorithm. In a secondary analysis, we aimed to examine associations between child, home, and survey completion characteristics and disagreement between processing approaches.

When comparing the sleep log to the no log approach, the sleep log approach demonstrated longer sleep periods, similar sleep duration, greater WASO, and earlier estimates of onset, midpoint, and waketime. Use of the generic log resulted in drastically earlier onset and greater sleep period and WASO compared with the no log approach. Due to the larger mean and absolute differences produced by the generic log, it appears that a generic log should not be used in place of incomplete or absence of parent-reported sleep information. High WASO values across approaches are likely a function of accelerometry's inherent challenges differentiating absence of movement during quiet rest from sleep and movement during sleep periods as wake, which results in overestimated wake during sleep periods (Meltzer, Montgomery-Downs et al., 2012). Children also may move or change positions during sleep more often than adults (De Koninck et al., 1992), which may lead to detection of wake during the sleep period when the child is likely asleep.

When comparing the sleep log approach to the no log approach, just over half of all nights demonstrated no difference in estimates of sleep period. Nearly three-quarters of all nights fell within 30 min of the detected sleep period. For researchers weighing the benefits and burden of including a sleep log, this provides important information in that decision-making process. There were low levels of mean bias for sleep outcomes of interest (i.e., <14 min), although the 95% LoA were wide. While 95% of the differences between approaches fall within the LoA range, individual night differences can be upwards of 1 h depending on the outcome of interest. To date, there are no other studies examining the influence of a parent-reported sleep log on sleep estimates produced by GGIR in children. In a similar comparison between the presence and absence of a sleep log in 47 adults over 4–7 nights of accelerometer wear, (Plekhanova et al. (2020) reported low levels of mean bias (<17 min) for onset and waketime, but greater mean bias for sleep period and duration (25–28 min). While the mean bias in sleep period and duration were smaller for children in the present study (<10 min), similar wide LoA were observed. Evidence in adults suggests poor reliability in sleep period and sleep duration detection with and without a sleep log (Plekhanova et al., 2020). With our results indicating an AMB of <30 min, combining nights with and without a sleep log in the same dataset may produce biased estimates of sleep period duration. As average effects of interventions designed to increase sleep duration in children range from 8 to 45 min (Busch et al., 2017; Fangupo et al., 2021), comparing nights with and without sleep logs may appear to indicate an intervention effect, when it could result from bias introduced by combining nights using both methods in the dataset. We caution future work to consider these differences before combining nights with and without a sleep log. This contrasts with the Meltzer, Montgomery-Downs, et al. (2012) and Tétreault et al. (2018) recommendations that some nights can be scored without a log and combined, although their findings are based on different devices and algorithms.

Child smartphone ownership, number of screens in the bedroom, parent work schedule, and type of night (weekend, summer) were statistically significantly associated with differences in sleep period identification when comparing the sleep log and no log approaches. Ownership of a smartphone and number of screens in the bedroom may help explain discrepancies between approaches. Parents could report the time their child went to bed but might be unaware of children accessing smartphones or other screens in the bedroom after that time. Screen use can delay falling asleep through a variety of mechanisms including time displacement, increased arousal due to screen content, social interaction, and the effect of light from screens (Cain & Gradisar, 2010; Gooley et al., 2011; Hale & Guan, 2015; Higuchi et al., 2005). It is not surprising that discrepancies existed between approaches on weekends and during summer as these are typically unstructured times for children (Brazendale et al., 2017). Weekends and summer periods typically result in later bedtimes and waketimes as children do not attend school. As these types of nights are less structured and unlikely to follow a typical routine, it may be more difficult for parents to recall accurate bed and wake times. Finally, it was hypothesised that the longer parents took to complete the survey would result in greater discrepancies between approaches as it would become more difficult to recall an accurate bedtime. While statistically significant, the magnitude of absolute difference was small, suggesting time to completion is not a strong driver of differences between approaches. Researchers may want to consider these factors and their potential influence on parent-report sleep log data when using logs to guide sleep processing.

This study is strengthened by the volume of free-living data collected in a large sample of elementary school-aged children. Over 13,000 valid nights of device-based sleep data were collected concurrently with parent-reported sleep logs over 1 year during three waves of data collection. By simultaneously collecting device-based and parent-report measures, we were also able to examine variables associated with disagreement between approaches to identify factors that may be important to consider when deciding to include a sleep log in future studies. Further, data were processed with GGIR, an open-source package in R, which facilitates transparency and reproducibility of raw accelerometry processing. Results should be interpreted within the context of this study's limitations. First, we did not compare our findings against polysomnography, the 'gold standard' for sleep assessment. Future validation studies are needed to understand how well GGIR estimates sleep in children with research-grade wearable devices. Next, only nights with a completed parent sleep log were included in analyses. It is likely that nights without a sleep log are systematically different than those with a log, and thus could influence findings. We also did not collect data pertaining to other parent/caregiver work schedules or whether the child split nights between different caregivers/households. These factors may influence sleep log completion and should be examined in future work. As this was a secondary analysis, we did not collect information about diagnosed sleep disorders or other conditions/medications that may influence children's sleep. Because the original validation of the default HDCZA algorithm showed better accuracy in adult healthy sleepers compared to sleep clinic patients (van Hees et al., 2018), future studies should examine if this relationship extends to younger populations. Finally, our sample consisted of children between the ages of 5 and 12 years. Additional work is needed to examine the performance of open-source processing methods in younger children and adolescents.

5 | CONCLUSION

Overall, findings suggest small differences between nights processed with and without a sleep log to guide sleep period detection in children. Although mean bias was <15 min between approaches, absolute bias neared a 30-min difference with wide LoA, which indicates larger differences on individual nights. When a parent sleep log was used, the absolute differences in sleep outcomes of interest (i.e., duration, WASO, sleep timing) ranged from 14 to 28 min. As there are no standards for paediatric sleep measured by accelerometry (Galland et al., 2016), we are unable to infer the clinical meaningfulness of these discrepancies between approaches. When a parent sleep log was used, the discrepancies were associated with child smartphone ownership, screens in the bedroom, parent work schedule, and weekend and summer nights. These factors may influence accurate parent reporting of bed and wake times, and lead to discrepancies between approaches. Based on these findings, we encourage researchers to consider these biases prior to combining nights processed with and without a sleep log in the same dataset. If sleep logs are inconsistently completed, processing all nights the same way (i.e., without a log) may reduce bias. Researchers are urged to transparently report if and how sleep logs are used in processing, the number of nights with and without a log, and how this may influence estimates of interest.

AUTHOR CONTRIBUTIONS

Sarah Burkart: Conceptualization; investigation; writing – original draft; methodology; writing – review and editing; formal analysis; visualization; software; data curation; project administration. **Michael W. Beets:** Conceptualization; funding acquisition; visualization; writing – review and editing; methodology. **Christopher D. Pfledderer:** Writing – review and editing; methodology. **Lauren von Klinggraeff:** Writing – original draft; writing – review and editing. **Xuanxuan Zhu:** Writing – review and editing; software; data curation. **Christine W. St. Laurent:** Writing – review and editing. **Vincent T. van Hees:** Writing – review and editing; methodology; software. **Bridget Armstrong:** Conceptualization; writing – review and editing; methodology; formal analysis. **R. Glenn Weaver:** Conceptualization; writing – review and editing; methodology. **Elizabeth L. Adams:** Writing – review and editing.

ACKNOWLEDGEMENTS

Research reported in this publication was supported in part by the National Institute of General Medical Sciences (NIGMS) under Award Number P20GM130420 (Burkart, Beets, Armstrong, Weaver, Adams), the National Heart, Lung, and Blood Institute (NHLBI) under Award Number F31HL158016 (von Klinggraeff), the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) under Award Number F32HD105384 (St. Laurent), and by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) under Award Number R01DK116665 (Beets). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

CONFLICT OF INTEREST STATEMENT

The authors have no conflict of interest to report.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Sarah Burkart https://orcid.org/0000-0002-0840-1680
Christopher D. Pfledderer https://orcid.org/0000-0001-7503-8554
Lauren von Klinggraeff https://orcid.org/0000-0002-4417-0701
Christine W. St. Laurent https://orcid.org/0000-0003-1932-3355
Vincent T. van Hees https://orcid.org/0000-0003-0182-9008
Bridget Armstrong https://orcid.org/0000-0002-1909-8170
R. Glenn Weaver https://orcid.org/0000-0001-5889-974X
Elizabeth L. Adams https://orcid.org/0000-0002-5602-8470

REFERENCES

Acebo, C., Sadeh, A., Seifer, R., Tzischinsky, O., Wolfson, A. R., Hafer, A., & Carskadon, M. A. (1999). Estimating sleep patterns with activity monitoring in children and adolescents: How many nights are necessary for reliable measures? *Sleep*, *22*(1), 95–103.

Barreira, T. V., Schuna, J. J., Mire, E. F., et al. (2015). Identifying children's nocturnal sleep using 24-h waist accelerometry. *Medicine and Science in Sports and Exercise*, *47*(5), 937–943.

Bland, J. M., & Altman, D. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*, *327*(8476), 307–310.

Brazendale, K., Beets, M. W., Weaver, R. G., Pate, R. R., Turner-McGrievy, G. M., Kaczynski, A. T., Chandler, J. L., Bohnert, A., & von Hippel, P. T. (2017). Understanding differences between summer vs. school obesogenic behaviors of children: The structured days hypothesis. *International Journal of Behavioral Nutrition and Physical Activity*, *14*(1), 100.

Busch, V., Altenburg, T. M., Harmsen, I. A., & Chinapaw, M. J. (2017). Interventions that stimulate healthy sleep in school-aged children: A systematic literature review. *The European Journal of Public Health*, *27*(1), 53–65.

Cain, N., & Gradisar, M. (2010). Electronic media use and sleep in school-aged children and adolescents: A review. *Sleep Medicine*, *11*(8), 735–742.

Chaput, J.-P., Gray, C. E., Poitras, V. J., Carson, V., Gruber, R., Olds, T., Weiss, S. K., Connor Gorber, S., Kho, M. E., Sampson, M., Belanger, K., Eryuzlu, S., Callender, L., & Tremblay, M. S. (2016). Systematic review of the relationships between sleep duration and health indicators in school-aged children and youth. *Applied Physiology, Nutrition, and Metabolism*, *41*(6), S266–S282.

De Koninck, J., Lorrain, D., & Gagnon, P. (1992). Sleep positions and position shifts in five age groups: An ontogenetic picture. *Sleep*, *15*(2), 143–149.

Fangupo, L. J., Haszard, J. J., Reynolds, A. N., Lucas, A. W., McIntosh, D. R., Richards, R., Camp, J., Galland, B. C., Smith, C., & Taylor, R. W. (2021). Do sleep interventions change sleep duration in children aged 0–5 years? A systematic review and meta-analysis of randomised controlled trials. *Sleep Medicine Reviews*, *59*, 101498.

Galland, B., Meredith-Jones, K., Gray, A., Sayers, R., Lawrence, J., Taylor, B., & Taylor, R. (2016). Criteria for nap identification in infants and young children using 24-h actigraphy and agreement with parental diary. *Sleep Medicine*, *19*, 85–92.

Gooley, J. J., Chamberlain, K., Smith, K. A., Khalsa, S. B. S., Rajaratnam, S. M. W., van Reen, E., Zeitzer, J. M., Czeisler, C. A., & Lockley, S. W. (2011). Exposure to room light before bedtime suppresses melatonin onset and shortens melatonin duration in humans. *The Journal of Clinical Endocrinology & Metabolism*, *96*(3), E463–E472.

Hale, L., & Guan, S. (2015). Screen time and sleep among school-aged children and adolescents: A systematic literature review. *Sleep Medicine Reviews*, *21*, 50–58.

Higuchi, S., Motohashi, Y., Liu, Y., & Maeda, A. (2005). Effects of playing a computer game using a bright display on presleep physiological variables, sleep latency, slow wave sleep and REM sleep. *Journal of Sleep Research*, *14*(3), 267–273.

Meltzer, L. J., Montgomery-Downs, H. E., Insana, S. P., & Walsh, C. M. (2012a). Use of actigraphy for assessment in pediatric sleep research. *Sleep Medicine Reviews*, *16*(5), 463–475.

Meltzer, L. J., Walsh, C. M., Traylor, J., & Westin, A. M. (2012b). Direct comparison of two new actigraphs and polysomnography in children and adolescents. *Sleep*, *35*(1), 159–166.

Migueles, J. H., Rowlands, A. V., Huber, F., Sabia, S., & van Hees, V. T. (2019). GGIR: A research community–driven open source R package for generating physical activity and sleep outcomes from multi-day raw accelerometer data. *Journal for the Measurement of Physical Behaviour*, *2*(3), 188–196.

Plekhanova, T., Rowlands, A. V., Yates, T., Hall, A., Brady, E. M., Davies, M., Khunti, K., & Edwardson, C. L. (2020). Equivalency of sleep estimates: Comparison of three research-grade accelerometers. *Journal for the Measurement of Physical Behaviour*, *3*(4), 294–303.

Sadeh, A. (2011). The role and validity of actigraphy in sleep medicine: An update. *Sleep Medicine Reviews*, *15*(4), 259–267.

Sadeh, A., III. (2015). Sleep assessment methods. *Monographs of the Society for Research in Child Development*, *80*(1), 33–48.

Spruyt, K., Gozal, D., Dayyat, E., Roman, A., & Molfese, D. L. (2011). Sleep assessments in healthy school-aged children using actigraphy: Concordance with polysomnography. *Journal of Sleep Research*, *20*(1pt2), 223–232.

Tétreault, É., Bélanger, M.-È., Bernier, A., & Carrier, J. (2018). Actigraphy data in pediatric research: The role of sleep diaries. *Sleep Medicine*, *47*, 86–92.

van Hees, V. T., Sabia, S., Anderson, K. N., Denton, S. J., Oliver, J., Catt, M., Abell, J. G., Kivimäki, M., Trenell, M. I., & Singh-Manoux, A. (2015). A novel, open access method to assess sleep duration using a wrist-worn accelerometer. *PLoS One*, *10*(11), e0142533.

van Hees, V. T., Sabia, S., Jones, S. E., Wood, A. R., Anderson, K. N., Kivimäki, M., Frayling, T. M., Pack, A. I., Bucan, M., Trenell, M. I., Mazzotti, D. R., Gehrman, P. R., Singh-Manoux, B. A., & Weedon, M. N. (2018). Estimating sleep parameters using an accelerometer without sleep diary. *Scientific Reports*, *8*(1), 1–11.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Burkart, S., Beets, M. W., Pfledderer, C. D., von Klinggraeff, L., Zhu, X., St. Laurent, C. W., van Hees, V. T., Armstrong, B., Weaver, R. G., & Adams, E. L. (2023). Are parent-reported sleep logs essential? A comparison of three approaches to guide open source accelerometry-based nocturnal sleep processing in children. *Journal of Sleep Research*, e14112. https://doi.org/10.1111/jsr.14112