CLINICAL REVIEW

Use of actigraphy for assessment in pediatric sleep research

Lisa J. Meltzer a,⁎, Hawley E. Montgomery-Downsb,⁎, Salvatore P. Insanac,f, Colleen M. Walsh d,g,h

⁎ National Jewish Health, 1400 Jackson Street, G311, Denver, CO 80206, USA
⁎⁎ West Virginia University, PO Box 6040, Morgantown, WV 26506, USA
© University of Pittsburgh Medical Center, 3811 O’Hara Street, E-1107, Pittsburgh, PA 15213, USA
⁎⁎⁎ Drexel University, 3141 Chestnut Street, Philadelphia, PA 19104, USA

A R T I C L E   I N F O

Article history:
Received 20 July 2011
Received in revised form 3 October 2011
Accepted 14 October 2011
Available online 15 March 2012

Keywords:
Actigraphy
Accelerometer
Adolescents
Children
Infants
Pediatric
Sleep
Wake

SUMMARY

The use of actigraphs, or ambulatory devices that estimate sleep–wake patterns from activity levels, has become common in pediatric research. Actigraphy provides a more objective measure than parent-report, and has gained popularity due to its ability to measure sleep–wake patterns for extended periods of time in the child’s natural environment. The purpose of this review is: 1) to provide comprehensive information on the historic and current uses of actigraphy in pediatric sleep research; 2) to review how actigraphy has been validated among pediatric populations; and 3) to offer recommendations for methodological areas that should be included in all studies that utilize actigraphy, including the definition and scoring of variables commonly reported. The poor specificity to detect wake after sleep onset was consistently noted across devices and age groups, thus raising concerns about what is an “acceptable” level of specificity for actigraphy. Other notable findings from this review include the lack of standard scoring rules or variable definitions. Suggestions for the use and reporting of actigraphy in pediatric research are provided.

© 2011 Elsevier Ltd. All rights reserved.

Introduction

Actigraphy is an objective, non-intrusive method for estimating sleep–wake patterns using activity-based monitoring. The use of actigraphy in research has gained significant popularity over the past 20 years, to the extent that the recent growth in published research studies that used actigraphy have outpaced studies that used polysomnography (PSG). 1 Actigraphy can be a particularly valuable methodology for use among pediatric populations, where the common reliance on parental report alone may limit the range and accuracy of information about children’s sleep. Consistent with the growth of actigraphy use across sleep research domains reported by Sadeh, 1 has there been also significant growth in the report of actigraphy specific to pediatric studies (Fig. 1), with the number of studies published in 2010 alone (n = 41) similar to the total number of studies published from 1991 to 2001 (n = 38).

The 2007 American Academy of Sleep Medicine practice parameters state that, “Actigraphy is indicated for delineating sleep patterns, and to document treatment responses in normal infants and children (in whom traditional sleep monitoring by polysomnography can be difficult to perform and/or interpret), and in special pediatric populations”. 2 Despite this recommendation and the growth in the number of studies that utilize actigraphy, there are currently no practice standards regarding recording parameters and scoring of the recorded actigraphy signals. Thus researchers and clinicians have little guidance in the use of this technology with little consistency across the body of literature reporting actigraphy results. This lack of uniformity has resulted in the inability to compare data values across studies. Further, despite the common use of these devices, no normative values are available for pediatric populations.

Validity of actigraphy in pediatric research

Another area of particular concern and utmost importance is the validity of actigraphy as a measure of sleep–wake patterns among children and adolescents of all ages. Many studies discuss the validity of actigraphy compared to PSG, but then cite studies validated...
on adult samples. Actigraphy output provides valuable information about activity levels that are ideal for visual analysis, which is useful for evaluating clinical treatment efficacy or to corroborate parental report of child sleep. However, actigraphically measured sleep estimates among pediatric samples (e.g., total sleep time or wake after sleep onset) should only be used when the recording device and particular sleep value are established as valid in comparison to a “gold standard” measure such as PSG or direct observation. Correlation statistics are often used to evaluate the validity of actigraphy when compared to a gold standard like PSG. However, correlations alone are not an appropriate way to validate these devices. A perfect correlation can be found between any two instruments, even if they have widely divergent measurement scales, as long as both measures increase at the same proportional rate. Some validation studies have relied only on correlation analyses, while others have overemphasized high sensitivity and underemphasized low specificity. Sensitivity and specificity bear further explanation because they are the most appropriate statistical method for validity assessment.

Sensitivity and specificity are most commonly used in the biomedical field for determining the quality of a novel diagnostic instrument. Sensitivity describes how accurately an instrument identifies people with a disorder (“true positive” cases); the more people who are inaccurately identified as not having the disorder (“false negative” cases), the less sensitive the instrument. On the other hand, specificity describes how accurately an instrument identifies people who do not have the disorder (“true negative” cases); the more people inaccurately identified as having the disorder (“false positive” cases), the less specific the instrument.

An ideal test accurately identifies 100% of both positive cases (is highly sensitive) and negative cases (is highly specific). When it comes to actigraphy, researchers have established a convention of considering sensitivity to be the proportion of epochs scored as sleep using polysomnography that are accurately identified as sleep by actigraphy. Specificity, on the other hand, is the proportion of polysomnography-scored wake epochs accurately identified as wake by actigraphy. An actigraph, or algorithm, that incorrectly scores sleep as wake has low sensitivity, and an actigraph, or algorithm, that incorrectly scores wake as sleep has low specificity. For example, if an actigraph scores an entire sleep period as sleep, it would be 100% sensitive at the expense of poor specificity in identifying wake during the sleep period. Both sensitivity and specificity are inherently important since the incorrect scoring of sleep or wake can result in under or over-estimates of reported sleep variables (e.g., total sleep time, wake after sleep onset).

Thus, examinations of sensitivity and specificity, and the relation between the two (represented by the “likelihood ratio”), are necessary for establishing the extent of actigraphy’s validity. Another assessment technique commonly used to examine instrument validity is the Bland–Altman concordance technique. This approach provides a visual representation of agreement by plotting the new method against the gold standard. The difference between the measures for each participant are fitted to lines that represent the ideal (no difference), plus either standard deviations or time discrepancies to show each participant’s deviation from the ideal.

For a more detailed description of validation and reliability issues pertaining to use of actigraphy, the reader is referred to several comprehensive reviews.

### Study purpose

The purpose of this paper is to provide a comprehensive review and synthesis of the research literature that used actigraphy as a measure of sleep–wake patterns or circadian rhythms among pediatric populations (ages 0–18 years, inclusive). This review highlights areas that should be considered in the design, execution, and report of sleep–wake variables in pediatric research studies that utilize actigraphy. In addition, a thorough description of the existing research that has examined the validity of these devices is provided. It is important to note that the purpose of this paper is not to provide recommendations about which brand to use or specific ways to improve existing technologies. Further, the clinical utility of actigraphy is beyond the scope of this review. Rather we objectively report on the use of existing technologies, as well as provide evidence-based recommendations for the proper use and reporting of actigraphy in pediatric research.

### Methods

A literature search was conducted using the Cumulative Index to Nursing and Allied Health (CINAHL), PsychINFO, and Pubmed. Search terms were (actigraph or accelerometer or actimeter or actiwatch) and (newborn or infant or toddler or preschool or child or children or adolescent or adolescence or youth or pediatric or premature infant) and (sleep or rest or circadian). In addition to the electronic search, reference lists from the identified articles and other review papers were reviewed for additional studies. Inclusion criteria for this review were: 1) a primarily pediatric study sample (ages 0–18 years); 2) an actigraphic device worn on the wrist or ankle/calf (for infants and toddlers) to measure some aspect of sleep or circadian rhythms; and 3) the study was
published in English. Papers were excluded if 1) the study sample was primarily adults (with no discreet pediatric data presented); 2) the outcome variable was daytime activity, with no report of actigraphic-derived sleep variable; 3) the report was a single subject case study, 4) the paper was not published in a peer-review journal (e.g., dissertation).

As this review focused on actigraphy methodological issues (as opposed to clinical outcomes), meta-analytic techniques were not used. Instead, results are presented descriptively and are divided into the following sections: 1) Study sample: papers reviewed and demographic information, 2) Actigraphy use: devices, software, and placement, 3) Validation: methods and statistics, and 4) Scoring and interpretation: algorithms and scoring rules. Where appropriate, results are further broken down by age group to highlight developmental issues.

Results

Study sample

Using the previously described search terms, 491 abstracts for papers published through December 2010 (including electronic publications ahead of print) were identified. An additional 15 articles were identified through a review of reference lists of eligible articles or review papers. Based on a review of these 506 abstracts, 269 papers met inclusion criteria for review. An additional 41 papers were excluded during full paper review because they included primarily adults (n = 10), the outcome variable was daytime activity rather than sleep or circadian rhythms (n = 8), the actigraph was not placed on the wrist or ankle (e.g., belt or upper arm, n = 12), wrist actigraphy was not used (e.g., piezo-electric pads, n = 8), or it was a case study/report (n = 3). Thus the analyses are based on a final review of 228 papers.

During the review, multiple studies were identified (e.g., same authors, same grant funding) as potentially including the same or overlapping samples and methodologies. To confirm the overlap, primary or corresponding authors were contacted via email. These articles are identified with a dagger (†) in the References section, highlighting that the study methodologies may not be independent. To prevent overinflating our results (e.g., the frequency with which certain devices were used), analyses were rerun using only one representative paper per sample. As seen in Table 1, no significant differences were found between the full sample and the limited sample in terms of the proportion of reported actigraph brands, placement of actigraphs, algorithms/sensitivities, or epoch lengths. Thus, the following results include all 228 papers reviewed.

In total, actigraphy was reported for 16,788 subjects (mean subjects per study = 73.6, standard deviation ± 74.9, median = 48.0, range = 3–387 subjects). Age was divided into developmental groups as follows: infants (0–11.9 months), toddlers (12–35.9 months), preschoolers (3–5 years), school-aged (6–12 years), and adolescents (13–18 years). Eighty-two studies (36.0%) reviewed subjects from two age groups, 13 studies (5.7%) included subjects from three age groups, and six studies (2.6%) included subjects from four age groups. School-aged children were most commonly studied (n = 136 studies, 59.6%), followed by adolescents (n = 89, 39.0%), preschoolers (n = 55, 24.1%), infants, (n = 41, 18.0%), and toddlers (n = 33, 14.5%).

Fifty-three studies (23.2%) included children with developmental disorders (attention-deficit hyperactivity disorder = 23, autism spectrum disorder = 19, intellectual disability = 11). Forty-two studies (18.4%) included children with chronic illnesses, including cancer, asthma, chronic pain, diabetes, obesity, among many others. Insomnia was the primary sleep disorder studied

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comparison between full sample of studies (All papers) and subgroup of studies with overlapping data eliminated (Limited sample).</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Brand of actigraphy</td>
</tr>
<tr>
<td>Ambulatory-Monitoring Inc.</td>
</tr>
<tr>
<td>Mini-Mitter</td>
</tr>
<tr>
<td>Cambridge Actiwatch</td>
</tr>
<tr>
<td>Other brands</td>
</tr>
<tr>
<td>Not reported</td>
</tr>
<tr>
<td>Actigraph placement&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Non-dominant wrist</td>
</tr>
<tr>
<td>Dominant wrist</td>
</tr>
<tr>
<td>Wrist unspecified</td>
</tr>
<tr>
<td>Not reported</td>
</tr>
<tr>
<td>AMI algorithms</td>
</tr>
<tr>
<td>Sadeh</td>
</tr>
<tr>
<td>Cole–Kripke</td>
</tr>
<tr>
<td>UCSD</td>
</tr>
<tr>
<td>Not reported</td>
</tr>
<tr>
<td>AMI epoch length&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>30-s</td>
</tr>
<tr>
<td>1-min</td>
</tr>
<tr>
<td>Not reported</td>
</tr>
<tr>
<td>Mini-Mitter sensitivity&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Medium</td>
</tr>
<tr>
<td>Automatic</td>
</tr>
<tr>
<td>Multiple</td>
</tr>
<tr>
<td>Study specific</td>
</tr>
<tr>
<td>Not reported</td>
</tr>
<tr>
<td>Mini-Mitter epoch length&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>15-s</td>
</tr>
<tr>
<td>30-s</td>
</tr>
<tr>
<td>1-min</td>
</tr>
<tr>
<td>Not reported</td>
</tr>
</tbody>
</table>

UCSD: University of California, San Diego.  
<sup>a</sup> Limited sample includes only one manuscript when multiple studies were published using overlapping data.  
<sup>b</sup> One study used both wrists.  
<sup>c</sup> One study used the Low sensitivity and one study used the High sensitivity.  
<sup>d</sup> Two studies used a 15-min epoch, one study a 15-s or 30-s epoch, and one study a 10-s epoch.

(n = 32, 14.0%); seven studies (3.1%) focused on youth with obstructive sleep apnea.

Actigraphy use

The majority of studies (n = 130, 56.8%) used Ambulatory Monitoring Inc. actigraphs (Ardsley, New York). Mini-Mitter actigraphs (now owned by Phillips-Respironics, Bend, Oregon) were the second most commonly used (n = 52, 22.7%), Cambridge Actiwatch<sup>c</sup> actigraphs (Cambridge, UK) were the third most commonly used (n = 29, 12.6%). Eight studies (3.4%) reported using a brand other than these three (i.e., Gaehwiler (n = 3), American Military Inc., Healthdyne, Individual Monitoring Systems, Somnitor, Somnowatch). Ten studies (4.4%) did not specify the type of actigraph used.

Computer-based software is interfaced with devices to convert actigraphically measured activity counts into sleep and wake epochs; some software programs provide automatic scoring options for certain variables (e.g., sleep onset time). Although most investigators used the software that was provided by the actigraph’s manufacturer, 60 studies (26.3%) did not report on the type of scoring software used and 10 studies (4.4%) used a software program created specifically for that particular study or laboratory.  

Actigraph placement for pediatric populations is most commonly recommended on the non-dominant wrist for older children and the ankle/calft for infants and toddlers.<sup>8</sup>
review, wrist placement was most common (68.5%); 102 studies (44.7%) used the non-dominant wrist, nine studies (3.9%) used the dominant wrist, one study (0.4%) used both wrists, and 46 studies (20.2%) did not report which wrist was used. Fifty studies (22%) placed the actigraph on an ankle (or calf for some infants); all but one of these studies focused on young children (under five years). Actigraphy placement was not reported in 19 studies (n = 8.3%).

Validation

Our review revealed 15 published validation-type studies that used pediatric samples, all but three of which were published in the last decade. Among these 15 papers, one was for the purpose of algorithm development rather than assessment, and two were comparisons of actigraphy placement only. Among the remaining 12 studies, two either only used correlation statistics or did not report sensitivity and specificity. The remaining 10 validation-type studies reported sensitivity and specificity; of these, seven were direct comparisons between actigraphy and polysomnography; two compared actigraphy to direct observation; and one compared actigraphy to videosomnography.

Together these 10 validation-type studies reported results from a cumulative 931 children (Table 2). Five studies (total n = 449 subjects) focused specifically on infants, one study (n = 11 subjects) focused on toddlers, one study (n = 58 subjects) focused on preschool-age children, one study (n = 16 subjects) focused on adolescents, and two studies (total n = 175 subjects) included a range of ages that encompassed all of these groups. Among these, only one study included children with developmental disorders and two studies (with overlapping samples) included children with sleep disordered breathing.

Although these validation-type studies utilized different devices, recording parameters, and scoring algorithms, they uniformly reported high sensitivity and low specificity (Table 2). The infant studies reported ranges of sensitivities = 83.4–99.3 and specificities = 17.0–97.8; the toddler study reported a sensitivity = 87.7 and a specificity = 76.9; the preschooler study reported a sensitivity = 97.0 and a specificity = 24.0; and the adolescent study reported a sensitivity = 95.0 and a specificity = 74.5. Studies that included multiple age groups reported

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Actigraphy validation studies compared to polysomnography (n = 6), direct observation (n = 2), or videosomnography (n = 1) that report sensitivity and specificity.</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>Age(s)</td>
</tr>
<tr>
<td>Hyde et al.</td>
<td>45</td>
</tr>
<tr>
<td>Insana et al.</td>
<td>22</td>
</tr>
<tr>
<td>O’Driscoll et al.</td>
<td>130</td>
</tr>
<tr>
<td>Sadeh et al.</td>
<td>11</td>
</tr>
<tr>
<td>Sadeh et al.</td>
<td>16</td>
</tr>
<tr>
<td>Sadeh et al.</td>
<td>10</td>
</tr>
<tr>
<td>Sadeh et al.</td>
<td>11</td>
</tr>
<tr>
<td>Sadeh et al.</td>
<td>10</td>
</tr>
<tr>
<td>Sadeh et al.</td>
<td>10</td>
</tr>
<tr>
<td>Sadeh et al.</td>
<td>58</td>
</tr>
<tr>
<td>So et al.</td>
<td>8</td>
</tr>
<tr>
<td>Sitnick et al.</td>
<td>13</td>
</tr>
<tr>
<td>Sung et al.</td>
<td>8</td>
</tr>
<tr>
<td>Sung et al.</td>
<td>20</td>
</tr>
<tr>
<td>Tilmann et al.</td>
<td>354</td>
</tr>
</tbody>
</table>


a These were the default settings. Tilmann et al. also evaluated optimized and extended performances, as well as novel algorithms.

b Novel algorithm calibrated on the 3-month infant data.
a sensitivity range = 82.2−90.1 and a specificity range = 50.9−72.8. In other words, across devices, placement sites, and algorithms, actigraphy was consistently good at accurately identifying sleep periods, but less accurate in identifying wake after sleep onset among pediatric populations. In fact, 55% of the specificities reported in Table 2 were less than 60.0.

Two examples provide a meaningful and practical perspective. First, the default wake threshold values for actigraphy from 14-month old infants underestimated total sleep time by more than 1 h among 55% of infants when compared to polysomnography; furthermore the range of differences between actigraphy and polysomnography for total sleep time ranged from 6 to 276 min.20 Second, videosomnography on preschool-age children with and without developmental disorders underestimated sleep time and overestimated wake after sleep onset.24 More concerning, these authors reported instances when the child was clearly awake and moving on the video recording, yet the actigraph reported that the child was asleep. Although variability in actigraphy placement (ankle versus wrist) and sensitivity thresholds may have contributed to these findings, more research and development are clearly needed to improve the validity, in particular the specificity, of actigraphy for use among all pediatric populations.

Scoring and interpretation

Scoring algorithms or wake threshold sensitivity levels were not reported in 72 studies (31.4%). The most common epoch length studied was 1-min (n = 139, 60.4%); only 16 studies (7.0%) used 30-s epochs. Epoch length was not reported in 63 studies (27.4%). Since algorithms/sensitivities and epoch lengths are somewhat dependent on the device used, the following sections present this information for the three most commonly used actigraphy brands.

AMI devices (n = 130)

The most commonly used scoring algorithm for the AMI devices was the Sadah algorithm (n = 85 studies, 65.4%). Seven studies (5.4%) used the Cole–Kripke algorithm, and two studies (1.5%) used the University of California, San Diego (UCSD) algorithm. No scoring algorithm was reported for 36 studies (27.7%) that used AMI devices. One-min epochs were the most commonly reported epoch length in 85 studies (65.4%). Only two studies that used AMI devices used a 30-s epoch length (1.5%). Forty-three studies (33.1%) did not report the epoch length used. AMI devices offer three different data collection modes; 43 studies (33.0%) used the zero crossing mode (ZCM), seven studies (5.4%) used the time above threshold mode (TAT), and one study (0.8%) used the proportional integration measure (PIM). Seventy-nine (60.8%) studies did not report the data collection mode.

Mini-Mitter devices (n = 52)

The most commonly used sensitivity (wake threshold) setting for the Mini-Mitter devices was the medium sensitivity (n = 23 studies, 44.2%). Four studies (7.7%) used multiple thresholds, while other studies reported using the low sensitivity (n = 1, 1.9%), high sensitivity (n = 1, 1.9%), or automatic (n = 2, 3.9%) thresholds. Two studies (3.9%) used a sensitivity threshold designed for that study, while 19 studies did not report a sensitivity/threshold setting (36.5%). One-min epochs were the most commonly reported epoch length in 31 studies (59.6%). Six studies (11.5%) used a 30-s epoch length and four studies (7.7%) used a 15-s epoch length. Nine studies (17.3%) did not report the epoch length that was used.

Cambridge devices (n = 29)

The most commonly used sensitivity (wake threshold) setting for the Cambridge devices was the medium sensitivity (n = 14 studies, 48.3%). Two studies used the automatic (6.9%) threshold. One study (3.4%) used a sensitivity threshold designed for that study, while 12 studies did not report a sensitivity/threshold setting (41.4%). One-min epochs were the most commonly reported epoch length in 17 studies (58.6%). Four studies (13.8%) used a 30-s epoch length, three studies (10.3%) used a 2-min epoch length, and one study (3.4%) used a 15-s epoch length. Four studies (13.8%) did not report the epoch length that was used.

As recommended by Aco & et al.23 five to seven nights was the most common data collection length across devices (n = 106, 46.5%). Three studies (1.3%) did not report a data collection length. Less than five nights of actigraphy was reported in 27.6% of studies (n = 63), 11 of these (4.8%) only included one night of recording (notably 9 of the 11 studies were validation studies that compared a single night of actigraphy to concurrent PSG recordings). Fifty-six studies (24%) included more than seven nights of actigraphy data collection, 14 studies (6.1%) included more than 2 weeks of actigraphy. (Note: studies that included multiple non-consecutive data collection times, e.g., three nights each on three separate occasions, were coded as three nights since each measurement time was averaged across three nights, not nine nights). We attempted to determine how many weekdays and weekend days were included across studies, but over 70% of studies did not break down how many of each were included, and we avoided inference based solely on the number of days studied.

Daily sleep diaries are necessary to help score sleep onset and sleep offset times, as well as identify artifact, including times when the device is removed. Only about half of the reviewed studies reported using some form (e.g., paper, electronic) of parent-reported daily sleep diary (52.6%, n = 120), while 70 studies (30.7%) used a child-reported diary. Fourteen studies (6.1%) included both a parent- and child-reported sleep diary. In addition to paper or electronic diaries, seven studies (3.1%) used a daily telephone call to or from parents, and five studies used a daily telephone call to or from children. Twenty-three studies (10.1%) used an event marker to indicate sleep onset and sleep offset (16 out of 23 studies also included a parent or child sleep diary).

Scoring rules

The use of specific actigraphically measured sleep variables and their corresponding reported definitions are indicated in Table 3. The most frequently reported actigraphically measured sleep variables across studies were “sleep duration” (n = 156, 68.4%), “sleep efficiency” (n = 131, 57.5%), and “bedtime/sleep onset” (n = 109, 47.8%). One-hundred and fifty-nine studies (69.9%) provided a definition for the reported actigraphy measures.

Many actigraphy variables, including “bedtime/sleep onset”, “wake time/sleep offset”, and “nocturnal wake” were calculated using specific time-based scoring rules. There is a lack of consistency in these rules. The most common scoring definition for “bedtime/sleep onset” was the number of consecutive minutes of decreased activity below a specific threshold. The three most frequently-used rules for “bedtime/sleep onset” were 3 (n = 18), 10 (n = 9), and 15 (n = 4) consecutive minutes scored as sleep, although studies also used 1, 5, and 20-min rules. The primary scoring definition for “wake time” was the last minute of a predetermined number of minutes scored as sleep prior to wake. The three most frequently utilized time rules for “wake time” were 5 (n = 18), 10 (n = 9), and 1 (n = 3) min scored as sleep before wake, although some studies used 3 and 15 min rules.

The primary scoring definition for “nocturnal wake frequency” was a consecutive number of minutes scored as wake during a sleep interval. The number of minutes required to score “nocturnal wake
frequency” was ≥5 min of wake in all studies that reported this particular scoring definition (n = 30).

Although Table 3 captures many of the commonly reported actigraphy variables, several nontraditional actigraphy measures were also found. Examples of variables that lack consistent definition across studies include: active sleep, quiet sleep, motionless sleep, activity mean, standard deviation of activity mean, variability in sleep schedule, coefficient of variation, motor activity level, actigraph pattern rating, activity per four hour interval, wake-onset instability, motionless sleep percentage, sleep quality, extremely long sleep, very short sleep index, insufficient sleep, prolonged sleep onset latency, activity level sleep transition rates, and variability in sleep schedule.

### Discussion

#### Use of actigraphy in pediatric sleep research

The purpose of this review was to describe methodological issues with actigraphy in pediatric sleep research. Although actigraphy is by no means a “gold-standard” measure of sleep, the rapid growth in its use over the past 20 years in pediatric research and the sheer volume of studies published that utilize actigraphy demonstrate the importance of these devices as an estimate of sleep—wake patterns. Despite the growth in actigraphy use, this field of research among infants). A more recent validation study demonstrated that different devices and scoring algorithms do not perform equally across age groups.25

### Areas for further validation

As previously described, statistics used to evaluate the validity of actigraphy devices (e.g., correlations versus sensitivity and specificity) varied across studies, calling into further question the validity of actigraphy to estimate sleep—wake patterns among pediatric populations. Since sleep changes rapidly across

<table>
<thead>
<tr>
<th>Variable</th>
<th>Reported (%)</th>
<th>Definition (%)</th>
<th>Primary scoring definition</th>
<th>Secondary scoring definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bedtime/Sleep onset</td>
<td>109 (47.8)</td>
<td>81 (74.3)</td>
<td>Time of first (predetermined) consecutive minute with decreased activity below a threshold: (n = 38, 34.9)</td>
<td>Identified by diary: (n = 13, 11.9)</td>
</tr>
<tr>
<td>Sleep onset latency</td>
<td>68 (29.8)</td>
<td>46 (66.2)</td>
<td>From bedtime to sleep onset: (n = 20, 29.4)</td>
<td>From lights out to sleep onset: (n = 17, 25.0)</td>
</tr>
<tr>
<td>Nocturnal wake frequency</td>
<td>75 (32.9)</td>
<td>51 (88.0)</td>
<td>Consecutive minutes of wake during a sleep interval: (n = 20, 40.0)</td>
<td>Number of transitions from sleep to wake: (n = 17, 22.7)</td>
</tr>
<tr>
<td>Nocturnal wake duration</td>
<td>30 (11.2)</td>
<td>17 (56.7)</td>
<td>Minutes awake between sleep onset and offset: (n = 13, 44.3)</td>
<td>Number of awakenings: (n = 3, 10.0)</td>
</tr>
<tr>
<td>Wake after sleep onset</td>
<td>44 (19.3)</td>
<td>28 (63.6)</td>
<td>Minutes scored as wake during sleep duration period: (n = 20, 45.5)</td>
<td>Percentage of time spent awake after sleep onset: (n = 4, 9.1)</td>
</tr>
<tr>
<td>Midpoint</td>
<td>3 (1.3)</td>
<td>3 (100.0)</td>
<td>Half of the sleep duration: (n = 3, 100.0)</td>
<td>—</td>
</tr>
<tr>
<td>Wake time/Sleep offset</td>
<td>78 (34.2)</td>
<td>55 (70.5)</td>
<td>Time of last (predetermined) consecutive minute of sleep: (n = 33, 42.3)</td>
<td>Identified as rise/end/sleep offset time: (n = 7, 9.0)</td>
</tr>
<tr>
<td>Naps</td>
<td>22 (9.4)</td>
<td>14 (63.6)</td>
<td>Duration of daytime napping: (n = 9, 40.9)</td>
<td>Percent of sleep during the daytime period: (n = 3, 13.6)</td>
</tr>
<tr>
<td>Sleep period</td>
<td>61 (26.8)</td>
<td>53 (86.9)</td>
<td>Time from sleep onset to awakening: (n = 40, 65.6)</td>
<td>Time spent asleep: (n = 6, 9.8)</td>
</tr>
<tr>
<td>Sleep duration</td>
<td>156 (68.4)</td>
<td>111 (71.2)</td>
<td>Time spent sleeping during a determined interval: (n = 69, 44.2)</td>
<td>Time from bedtime/sleep onset to sleep end: (n = 28, 18.0)</td>
</tr>
<tr>
<td>Sleep efficiency</td>
<td>131 (57.5)</td>
<td>86 (65.7)</td>
<td>Percentage of time spent asleep during a defined interval: (n = 51, 38.9)</td>
<td>Total sleep time divided by time in bed: (n = 18, 13.7)</td>
</tr>
<tr>
<td>Fragmentation</td>
<td>14 (6.1)</td>
<td>12 (85.7)</td>
<td>Index of activity during sleep: (n = 4, 28.6)</td>
<td>Measure of movement or wake during sleep: (n = 3, 21.4)</td>
</tr>
<tr>
<td>Longest sleep</td>
<td>25 (11.0)</td>
<td>13 (52.0)</td>
<td>Longest continuous episode scored as sleep: (n = 12, 48.0)</td>
<td>Longest duration of no activity: (n = 1, 4.0)</td>
</tr>
</tbody>
</table>
development, additional research is needed to demonstrate the validity of different devices and scoring thresholds for each developmental age group. Because of known differences in pediatriac sleep there are separate standards for scoring PSG for children and adults. \textsuperscript{27}–\textsuperscript{30} Similarly, developmental differences need to be considered in the validation of actigraphy.

Beyond further development of actigraphic devices and algorithms/sensitivities, additional validation studies are needed to address the appropriate placement of the devices. This review only included studies that used a wrist or ankle/calf (for infants) placement because the validity of these placements has been assessed in more than one study. Further research is needed to validate the use of actigraphs placed on the upper arm, belt, or shirt pocket. Additional studies that compare two brands of devices side-by-side (and ideally against PSG) are also needed. In this review, only the Weiss et al.\textsuperscript{14} study compared two brands of actigraphic devices, reporting strong correlations for TST ($r = 0.88$) but less strong correlations for sleep efficiency ($r = 0.65$), a likely result of the low sensitivity of these devices to detect wake after sleep onset.

Further research is also needed to examine the validity of algorithms and wake sensitivity thresholds. Algorithms and sensitivity thresholds are likely to be sensitive to developmental state and/or sleep disorders (e.g., sleep disordered breathing). Users should select the appropriate algorithm/sensitivity threshold based on existing validation studies, and should be cautious about simply using default settings and/or autoscoring features (e.g., AMI auto set down interval or Mini-Mitter automatic set rest interval).

### Reporting of actigraphy data

Significant inconsistencies were noted in this review for how variables were defined, with some definitions deviating from what the variable may be traditionally understood to represent. For example, some authors described “bedtime” as the first minute of actigraphically identified sleep, whereas others used this as a definition of “sleep onset”. As another example, some authors reported “wake time” as the number of minutes the child was awake after sleep onset, whereas others used the same term to describe the parent-reported time the child awakened in the morning. Differences in variable terminology and meaning might appear subtle or fluid, or a function of semantics, but it is essential that sleep variables and their meanings be used consistently for effective communication and interpretation. Due to the variety of ways that a variable can be calculated and/or defined, standardized definitions are clearly needed. Table 4 provides a recommended list of common variable names and definitions that should be considered by researchers for inclusion in reports of studies that use actigraphy.

Due to the lack of reported data across studies, one variable not examined in this review was the amount of data that were lost due to technical failure or participant non-adherence to wearing the device. Investigators who have used actigraphy understand that these two issues are quite common. Furthermore, data loss can be a result of a lack of diary data, which is needed to identify artifact and set sleep intervals. Acebo et al.\textsuperscript{22} found that up to 28% of weekly recording may be unusable due to child illness, non-adherence, or device failure. Similarly, a recent study of healthy adults found that 27% of actigraphy nights were unusable due to missing data (compared to only 1.5% of sleep diary nights).\textsuperscript{31}

Finally, there has been little validation of many of the variables automatically calculated by manufacturer scoring programs (e.g., sleep bouts, wake bouts, motionless sleep or immobile time, circadian parameters).\textsuperscript{9} Additional work is needed to validate these variables.

### Table 4

<table>
<thead>
<tr>
<th>Proposed variable definitions for actigraphy when used in pediatric research.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reported variables</strong></td>
</tr>
<tr>
<td><strong>Bedtime</strong></td>
</tr>
<tr>
<td><strong>Wake time</strong></td>
</tr>
<tr>
<td><strong>Sleep opportunity</strong> (Time in bed)</td>
</tr>
<tr>
<td><strong>Actigraphy variables</strong></td>
</tr>
<tr>
<td><strong>Sleep onset</strong></td>
</tr>
<tr>
<td><strong>Sleep offset</strong></td>
</tr>
<tr>
<td><strong>Total sleep time (TST)</strong></td>
</tr>
<tr>
<td><strong>Sleep onset latency</strong></td>
</tr>
<tr>
<td><strong>Wake after sleep onset</strong></td>
</tr>
<tr>
<td><strong>Sleep efficiency</strong></td>
</tr>
<tr>
<td><strong>Night waking</strong></td>
</tr>
<tr>
<td><strong>Night waking frequency</strong></td>
</tr>
<tr>
<td><strong>Night waking duration</strong></td>
</tr>
<tr>
<td><strong>24 h sleep duration</strong></td>
</tr>
<tr>
<td><strong>Midpoint</strong></td>
</tr>
<tr>
<td><strong>Naps (Frequency and/or duration)</strong></td>
</tr>
</tbody>
</table>

| **Less commonly reported**                    |
| **Clock time halfway between sleep onset and sleep offset** |
| **Scoring rules similar to “night waking” required to define “nap” in terms of the predetermined number of minutes asleep preceded and followed by predetermined number of minutes of wake. Then can define nap frequency and nap duration** |

### Considerations for the use of actigraphy in pediatric sleep research

A comprehensive review of the methodological challenges and considerations for the general use of actigraphy can be found in a recent review of actigraphy use with adult populations by Berger et al.\textsuperscript{32} The following provides a highlight of these considerations specific to pediatric sleep research.

### Choosing the device

The selection of an actigraph for use in a research study should be based on a number of factors. First, researchers should consider the cost of an actigraphy system. This includes the device, which can vary in cost depending on the features (e.g., event marker, light sensor, off-wrist detection). Actigraphy interface and software are additional expenses. Some devices require new batteries to be installed every 45 days (or preferably changed for each new subject), which over the long-term can result in increased cost. Other devices have rechargeable batteries. For software, some programs may require only one license to be installed on multiple computers, while other programs may require a separate license be purchased for every computer, resulting in increased expense. Finally, it is
important to consider warranties, maintenance, and technical support (for the device and software); for example, one company suggests that device maintenance occurs every 12 months at a cost of more than $100 per device.

Another area of consideration is the population of youth for the study. Actigraph devices vary by size and weight, which may impact the comfort level of participants, in particular, young children. Consideration should also be given to the durability and water resistance of the devices. Most importantly, actigraphy systems should be validated among the population of interest. If the selected device or settings have not yet been validated and reported in the literature, researchers should perform a validation study prior to use in their primary study.

### Study design

Once the actigraphy system has been chosen, researchers need to consider the use of the device in the specific study. For example, how will the actigraph be delivered to and collected from study participants? What is the length of time for data collection? Based on the existing literature, to obtain five nights of actigraphic data, a minimum of seven recording nights is usually necessary. If the difference in weekday and weekend sleep is of interest, then 2 weeks of data should be collected. Other data collection considerations include the epoch length and the mode of data collection (e.g., ZCM, TAT, PIM, or TRI).

Variables to be assessed should be selected a priori and be well-defined. Consideration should be given to the research questions, as well as the validity of the variables of interest. For example, without accurate sleep diaries or event marker use, it is not possible to determine sleep onset latency. With this in mind, decisions need to be made about who will complete the daily sleep diary (i.e., parent, child); the frequency of diary completion (e.g., daily versus once in the morning and/or once in the evening); and the format of the diaries (e.g., paper, electronic, web-based, and/or phone call).

### How to handle data

Prior to data collection (and download) decisions need to be made regarding how the data will be processed and analyzed. First, diary data should be used to reduce artifact, including the removal of devices (e.g., for bathing, swimming), prolonged periods of quiet activity (e.g., going to the movies, watching television in bed), periods of activity that should be scored as sleep (e.g., child sleeping in a moving vehicle), movement or night awakenings that may be due to bed sharing (i.e., parent movement that may be detected by the child’s actigraph recording during sleep), and any reason for an atypical night of sleep (e.g., child illness, sleepover).

Second, the scoring algorithm or wake threshold sensitivity level also needs to be selected. Again, the decision about how to choose the appropriate specifications for a study should be based on the study population and previously published validation studies. Rules for how to handle missing data and/or artifact should be well-defined. Further, scoring rules should be clearly identified a priori (e.g., is sleep onset first of 3, 5, or 10 consecutive minutes of sleep?). Third, decisions also need to be made about whom, or how many people, will be involved with scoring the data and how reliability will be supported. Further, if multiple personnel clean data and/or score files, inter-rater reliability should be examined. Finally, for studies involving multiple groups (e.g., children with attention deficit/hyperactivity disorder (ADHD) versus typically developing children) or treatments (e.g., sleep restriction versus sleep extension) personnel who are scoring and interpreting actigraphy data should be blinded to the group/treatment. Without this blinding the ability to grade papers in systematic reviews is limited.

### Reporting actigraphy in research papers

While there are currently no standards for what should be reported in research papers, the following suggestions are offered for authors, reviewers, and consumers of the literature (Table 5).

#### Device/system information

A basic amount of device information should be included in all reports. This includes the name of the device, the specific model, and the name (and location) of the manufacturer. In addition, detailed information should be provided on the placement of the device (non-dominant or dominant wrist, left or right ankle, etc.). Researchers should also include the epoch length measured, the mode of data collection, and the algorithm or wake sensitivity threshold. Finally, the type and version of software used should be reported.

#### Data preparation and reporting

Multiple aspects of the data preparation are important to include in the research report. First, authors should provide detailed information about how they handled missing data or potential artifact. Second, the amount of data lost due to technical failure, artifact, or participant non-adherence to the study protocol needs to be included. Third, scoring rules should be clearly defined.

In terms of actual data reporting, a clear definition of variables (See Table 4 for suggestions) should be included; this permits cross-
study comparisons of results. Finally, the number of days included (including weekdays/weekends) should be reported.

Concluding remarks

Actigraphy has become a central part of pediatric sleep research in the past 20 years. While there are a number of benefits to actigraphy, researchers and consumers of the literature also need to be aware of the limitations and threats to validity, most notably poor specificity to detect wake after sleep onset, how artifact influences results, and the lack of consistency in terms of scoring rules and reported variables. Although we have provided some suggestions based on this review, the field of pediatric sleep would likely benefit from a task force that provides standard recommendations for the use and reporting of actigraphy in pediatric sleep research, as well as to define what is “acceptable” in terms of sensitivity and specificity.

Conflict of Interest Statement

The authors have no conflicts of interest to disclose.

Practice points

1. Actigraphy provides a non-intrusive, cost-effective way to objectivity estimate pediatric sleep—wake patterns for an extended period of time within the child’s natural environment.

2. The use of actigraphy also requires concurrent sleep diaries in order to identify artifact and to corroborate or guide identification of sleep periods on actigraphy output.

3. The selection of devices, scoring algorithms or sensitivities, and device placement should be based on existing validation studies done within the population of interest, considering developmental stage, developmental disorders, and/or sleep disorders.

4. The benefits of using actigraphy should be considered alongside the limitations of these devices, in particular the poor specificity of actigraphs to detect wake after sleep onset.

Research agenda

1. Pediatric sleep researchers should work collaboratively with device-producing companies to ensure that the need for highly specific and sensitive devices is met.

2. Novel devices and algorithms should undergo validation assessment among youth of all ages, as well as among youth with and without sleep disorders (e.g., sleep disordered breathing).

3. Further development of device utilization for disordered populations (e.g., restless legs syndrome or periodic limb movement disorder) is needed.

4. A task force should be convened to provide standard recommendations for the use and reporting of actigraphy in pediatric sleep research, as well as to determine what is “acceptable” in terms of sensitivity and specificity of these devices.

Acknowledgments

We thank Devon Ambler for her assistance with the references. Support for this manuscript was provided by K23 MH066772.

References


* The most important references are denoted by an asterisk.

1 Overlapping studies.

2 Studies that include children with autism spectrum disorders.

3 Studies that include children with intellectual disorders.

4 Studies that include children with chronic medical illnesses.

5 Studies that include children with ADHD.


Oyane NM, Bjorvatn B. Sleep disturbances in adolescents and young adults. Sleep 2006 May; 29(5):739–45.


