The role of actigraphy in sleep medicine

Avi Sadeh and Christine Acebo

1Department of Psychology, Tel Aviv University, Tel Aviv, Israel, and 2E. P. Bradley Hospital/Brown University Medical School, Providence, RI, USA

KEYWORDS
actigraphy, actimetry, sleep, sleep disorders, monitoring, activity, motility

Summary During the last decade actigraphy (activity-based monitoring) has become an essential tool in sleep research and sleep medicine. The validity, reliability and limitations of actigraphy for documenting sleep–wake patterns have been addressed. Normative data on sleep–wake patterns across development have been collected. Multiple studies have documented the adequacy of actigraphy to distinguish between clinical groups and to identify certain sleep–wake disorders. Actigraphy has also been shown to be effective in documenting the effects of various behavioral and medical interventions on sleep–wake patterns. Actigraphy is less useful for documenting sleep–wake in individuals who have long motionless periods of wakefulness (e.g. insomnia patients) or who have disorders that involve altered motility patterns (e.g. sleep apnea). Potential users should be aware of a number of pitfalls of actigraphy: (1) validity has not been established for all scoring algorithms or devices, or for all clinical groups; (2) actigraphy is not sufficient for diagnosis of sleep disorders in individuals with motor disorders or high motility during sleep; (3) the use of computer scoring algorithms without controlling for potential artifacts can lead to inaccurate and misleading results. © 2002 Published by Elsevier Science Ltd

INTRODUCTION

In 1995 the Standards of Practice Committee of the American Sleep Disorders Association (ASDA) commissioned a task force to evaluate the role of actigraphy in sleep medicine. This effort led to a review paper on this topic [1] and was accompanied by ASDA's guidelines [2]. This recognition by ASDA was an important landmark in the acceptance of actigraphy by sleep researchers and sleep medicine clinicians. The use of actigraphy in sleep medicine and research continued to rise, as illustrated by the increasing number of publications over the years (Fig. 1).

The present review briefly addresses the work reviewed earlier [1] and describes significant newer contributions. This review is based on a literature search that included the Pubmed, Social and Science...
Citation Index and Psychlit databases. Keywords included sleep, actigraphy, actimetry, actimeter and actigraph. In addition, the reference lists of articles were scanned for additional citations. Papers that did not directly address sleep evaluation were excluded.

DEFINING ACTIGRAPHY

Activity monitoring has a long history both in medicine [3] and in sleep research [3, 4]. The term actigraphy refers to methods using miniaturized computerized wristwatch-like devices to monitor and collect data generated by movements. Most actigraphs contain an analogue system to detect movements. In some devices, a piezo-electric beam detects movement in two or three axes and the detected movements are translated to digital counts accumulated across pre-designed epoch intervals (e.g., 1 min) and stored in the internal memory. The actigraph can collect data continuously over an extended period (1 week or longer). Some devices are programmable and enable selection of specific modes of operation (e.g., variable movement frequency bandwidths, sensitivity levels or epoch intervals) whereas other devices have only one fixed mode. Data are downloaded to the computer using special interface units or other forms of communication channels. In this review we have chosen not to focus on the technical merits of the various instruments, modes of operations and specific brand names. New devices, scoring algorithms and operating procedures are constantly being developed and updated. Furthermore, no consensus has been reached among clinicians and researchers as to which device or algorithm is best. Thus, new users should read the scientific literature on each instrument carefully in order to reach an educated decision as to which combination of device, mode of operation and scoring algorithm is most suitable for a specific research protocol. The first stage in evaluating these questions should be based on the existing reliability and validity studies.

VALIDITY OF ACTIGRAPH MEASURES

As with any new methodology, a primary task is the establishment of psychometric data for measures so that researchers and clinicians may select valid and reliable measures appropriate for their needs and for the type of participants/patients they study. This task is complicated for actigraphic measures because there are multiple actigraph devices with differing mechanical properties, sensitivities and sampling capabilities. For example, a study comparing two of the most widely used actigraphs reported significant differences in the sensitivity of the two recording devices to movements. The authors concluded that the device with the lower sensitivity may fail to detect movements during sleep that indicate disruption or wakefulness [5]. Sampling rates also vary between devices and even between studies using the same instrument. Another study compared sleep–wake scoring of recordings made in different modes of operation (zero crossing, time above threshold, proportional integration) with two different devices and concluded that all modalities provided valid measures although the proportional integration method appeared to be most accurate [6].

Computer algorithms have been written to allow the automatic scoring of actigraph records, primarily to infer sleep–wake measures, although some programs also provide measures of circadian parameters. These scoring programs, although useful, must be examined carefully. A number of scoring programs have been developed, each with different algorithms for discriminating sleep from wake [7–12]. Some of these algorithms have been validated against polysomnographic measures but generally only for scoring data from a specific brand of actigraph with a specific mode of operation. In addition, scoring programs usually yield more measures than simple estimates of total sleep or wake; few of these additional measures have been subjected to validity or reliability analyses. The number of scoring programs and the number of variables available from these programs continues to increase as the technique becomes more widely used, resulting in the proliferation of variables that have no demonstrable validity or reliability yet are automatically generated when the scoring program is run.

Studies reporting the validity of actigraph measures of behavioral sleep and wake up through the early 1990s were extensively described and summarized in the 1995 ASDA review [1]. The conclusion of those authors was that the epoch-by-epoch agreement for actigraphically estimated sleep during laboratory studies is reasonably high (greater than 0.85 agreement) for normal individuals
and some patient samples at a variety of ages regardless of the algorithm employed. Measures of wake have lower validity. Polysomnography (PSG)—actigraph correlations between whole night measures of sleep and sleep efficiency are also reasonable (generally above 0.80) for normal individuals and some types of patients. More recent studies have not appreciably altered these conclusions.

It is important to note that most validation studies have been conducted in the laboratory where artifacts are controlled and recording conditions are standardized. Most research and clinical studies, on the other hand, are conducted in the home where control is lost. Individuals in the home spend hours on the couch watching television (very quietly), forget to put the actigraph on again after a shower, drink lots of coffee or alcohol before bedtime and use medications that may affect movement and/or sleep. Detailed documentation in daily logs of sleep–wake periods and artifact-related information is critical as are scoring procedures that restrict algorithm scoring to documented periods. These peripheral measures provide the only way to take back some of the control lost outside the laboratory.

Increased wakefulness during the night generally decreases the accuracy of actigraph measures [13]. Decreased accuracy has been most well documented for insomnia patients: quiet wakefulness is often miscoded as sleep [9, 14, 15]. The impact of this inaccuracy is largely on measures of wake during the night and sleep onset measures [1, 16], although discrepancies between PSG and actigraph measures of total night-time sleep have been greater than 2 h in some insomnia patients [15]. Jean-Louis et al. [12] have suggested that using different scoring algorithms for insomnia patients may improve accuracy. Multiple algorithms, however, hamper the use of the methodology for between-group studies and limit cross-study comparisons.

Actigraph measures are also less accurate in other patients with disturbed sleep, such as those with movement disorders [15] and shift workers [17]. Validity has not been established for many other clinical groups such as psychiatric patients on or off medications. One exception is a study that assessed actigraphic sleep–wake algorithms in patients with major depression [18]. In this study, actigraph measures from a standard algorithm and a special algorithm optimized for the depressed sample were compared with PSG scores. The results indicated that the standard algorithm was less accurate for the depressed sample than for normal adults. For instance, using the standard algorithm the correlation between PSG- and actigraph-derived measures of total sleep time was 0.85; for sleep efficiency the correlation was 0.65. The authors concluded that actigraphy should not be used as a surrogate for PSG but rather as a complementary method.

Validity of measures in older adults appears to depend on how otherwise healthy and unimpaired they are [19, 20]. However, even in demented elderly people reasonable validity has been demonstrated in a comparison of actigraph measures (using the Actillume) with PSG sleep–wake scorings [20]. Some suggestions have been made that the amount of nocturnal activity changes with increasing age [21, 22], although the direction of change is not clear nor to what extent activity change reflects change in nocturnal sleep and wake. Recent studies indicate that algorithms are not necessarily equally accurate for adults at different ages. For example, an algorithm optimized for healthy menopausal women was less accurate when used to score records from healthy young adults [23].

Several studies have reported sex differences in actigraph estimates of total nocturnal sleep with females manifesting more sleep than males [21, 24–27]. The increased sleep in older female adults is consistent with other subjective and objective reports, but laboratory PSG studies typically have found few sex differences in sleep–wake measures in infants, children and young adults. Further research is necessary to clarify whether males and females are differentially impacted by environmental and/or psychosocial influences in the home environment or whether sex differences in activity level during the night [28] lead to differential accuracy of sleep–wake scoring for males and females.

Monk and colleagues have suggested [29] that activity counts during sleep vary as a function of sleep stage. This observation is intriguing but also points to the fact that even normal sleep may have levels of activity that challenge algorithm accuracy.

The results of the studies mentioned above indicate that activity amounts and patterns during sleep are more or less variable as a function of age, sex and physical and mental health status. This variability may limit the usefulness of actigraphy for studies aimed at comparing sleep measures between developmental stages or between patient groups. This issue requires further study.
INDIVIDUAL DIFFERENCES: RELIABILITY AND STABILITY OF MEASURES

Aggregation of measures over multiple nights stabilizes individual differences and thus yields measures that may also be expected to predict other variables. The best index of whether aggregation is useful is the reliability estimate of a single measurement item compared with the reliability of scores obtained by aggregating scores over multiple nights. Only a few studies have addressed the stability of actigraph measures across nights for individuals [1, 26]. Acebo et al. [30] reported reliability estimates for primary actigraph variables based on mean values aggregated over one to seven nights of recording in pre-school children and adolescents. Results indicated that reliability for values aggregated over five nights was adequate (<0.70) for sleep start time, wake minutes during the night and sleep efficiency. Measures of sleep minutes during the night and sleep period (sleep onset to sleep offset) were less reliable and required aggregation over seven or more nights to acquire measures that reflected stable individual differences. Reliability for one- or two-night aggregated values was poor for all measures. In addition, despite weeknight and weekend night differences in adolescent subjects, measures that included both types of night showed more stability across sessions (summer and fall recording sessions) than aggregates of either type of night separately.

Studies addressing these reliability and stability issues in samples of older adults and in patient samples are needed. In addition to providing information about stability and instability of sleep–wake measures over the lifespan, the night-to-night variability of sleep–wake measures may provide insight into the dysregulation of sleep–wake patterning seen at older ages and also may be an important individual characteristic associated with psychopathology [31].

PREDICTIVE VALIDITY

A second kind of validity that is important to assess for any measure is predictive validity, that is how well the measure correlates with or predicts other characteristics of individuals that should be related within a theoretical framework. Actigraph measures are beginning to fare relatively well with respect to this issue. For example, in a number of studies poor sleep quality or insufficient sleep documented by actigraphy has been associated with increased daytime sleepiness [24, 26, 32]. Other studies have demonstrated associations between actigraphic sleep parameters and temperament or other individual characteristics [26, 31, 33–37]. Many of the studies described below give evidence for the predictive validity of some actigraph measures in some samples. We must remember, however, that a measure may be extremely inaccurate in measuring what it is purported to measure while still correlating well with another variable. In this case, positive results may lead to incorrect conclusions and future hypotheses. Alternatively, an inaccurate measure may fail to relate to other characteristics as it theoretically should. Without good assessments of measure validity and reliability, it may be impossible to determine whether failure to obtain predicted results is a function of invalid or unreliable measures or whether the hypothesized theoretical connection is in error.

ARTIFACTS AND DATA LOSS

A major strength of actigraphy, the ability to monitor behavioral activity and to infer sleep–wake patterns over long periods of time in the home, is also a source of some weaknesses. Acebo et al. [30] reported that up to 28% of weekly recordings of children and adolescents were insufficient for analysis. The main reasons for data loss included participant non-compliance (inability to complete the diary or log and failure to wear the actigraph), illness and technical problems. As noted above, detailed participant/patient logs are essential for accurate scoring of records. Times when the actigraph is off the wrist will otherwise be scored as sleep. Showers (with the actigraph off) just before bedtime or after risetime are remarkably easy to mistake for sleep. Conversely, activity of co-sleeping bed partners or sleep during car rides may be scored as waking. A child with a common cold taking cough medicine may have markedly disturbed sleep, yet that night’s sleep may not be characteristic of his/her sleep in general. For these reasons, the log should contain information about bedtimes and risetimes, times when the actigraph is not on the wrist and times of external motion or unusual events. When the actigraph data are retrieved, participants/
patients should be queried about times when the log and actigraph records are not in agreement. Actigraphs are mechanical devices and, as such, subject to breakage. Children and adolescents are remarkably capable of bending metal parts, dislodging event buttons and otherwise damaging the instruments. Data loss may also occur when curious wearers of any age remove the battery cover to “see what’s inside”. Finally, instruments may lose calibration and fail in other ways. Unlike laboratory studies, where technical problems and artifacts are recognized quickly and either resolved or thoroughly documented, problems that occur during a week of home recording often lead to complete loss of data. For these reasons, it is important to devise and follow procedures for testing batteries and maintaining equipment.

**ACTIGRAPH PLACEMENT**

In most studies the actigraph has been placed on the non-dominant wrist. Assessments of placement differences have indicated that some scoring algorithms are relatively insensitive to wrist placement (dominant vs non-dominant) in spite of significant differences in activity levels [38]. On the other hand, Violani and colleagues have reported shifts of motor activity dominance across the night between wrists [39]. Studies of activity recorded from multiple body sites have found significant differences between sites [40, 41]. Standardization of placement would benefit cross-study comparisons.

**Practice Points**

Actigraph methodologies should always include:

1. Algorithms and instruments that are appropriate for the task.
2. Measures with documented validity, aggregated over multiple nights.
3. Assessment and elimination of potential artifacts, and standardized placement.
4. Documentation of sleep–wake and artifact-related behavior with daily logs.

**DEVELOPMENTAL STUDIES AND NORMATIVE DATA**

The goal of obtaining normative data across the life-span on the basis of large-scale actigraphic studies has almost been reached over the last few years. Studies describing normal samples of individuals ranging from newborns to elderly individuals have been completed or nearly completed.

**Infants and children**

Sleep–wake patterns of 220 full-term babies have been described during the first 48 h in the hospital nursery [42]. This study provided data on sleep schedule, sleep duration and sleep quality of the newborns. Perhaps the most striking finding was the very high variability between infants vis-à-vis their sleep duration and sleep quality under relatively standard conditions, before being exposed to variations in familial psychosocial and physical environments.

Acebo and colleagues have completed an actigraphic study of sleep–wake patterns in 169 infants and young children (1–5 years of age) [43]. The results of this study indicated that nocturnal sleep does not change dramatically over this age range; rather, the major developmental changes are associated with a decrease in daytime sleep.

Sleep in school-aged children has been described in two recent studies [26, 33]. Aronen and colleagues described sleep and behavior problems in 49 school-age children and found an association between low true sleep time and externalizing behavior problems reported by teachers [33]. Sadegh et al. described 140 school-age children (second, fourth and sixth grades) and found a strong delay in sleep onset and reduction of sleep duration from second to sixth grade. In addition, a high prevalence of sleep fragmentation (in 18% of the children) was indicated by the actigraphic measures [26]. Significant sex differences were found, with girls spending more time in sleep in general and in motionless sleep in particular.

Carskadon and colleagues collected repeated measurements of behavioral sleep in adolescents using actigraphy as well as other measures of sleep, sleepiness and circadian rhythms [24]. Students were assessed twice (ninth and tenth grades) to determine the effects of a change to an earlier school start time. Actigraphy measures reflected earlier morning rise time and shortening of total sleep time associated with the transition from ninth to tenth grade. A significant sex difference was found, with girls sleeping more than boys.
Adults and elderly people
Surprisingly, studies providing normative data on young adults are still missing, although older adults are included in some of the aging studies. Investigations of sleep in non-clinical samples of elderly people have provided normative data on sleep and related phenomena in the elderly [19, 21, 28, 44–49]. Other studies have focused on describing sleep in clinical disorders associated with aging such as Alzheimer’s [50–54], Parkinson’s disease [55], and dementia [20, 56–59]. For example, studies have documented that the level of dementia in demented patients is associated with weakened organization of the sleep–wake system [52, 56, 59]. Actigraph sleep of very demented patients has been characterized as highly fragmented with increased activity during the night and increased amounts of sleep during the daytime. Overall these studies have documented the phenomenon of sleep fragmentation associated with aging and its disorders, although, as noted previously, increased fragmentation also decreases the accuracy of actigraph measures. Other studies have assessed the outcomes of various interventions for improving sleep in clinical and non-clinical samples of elderly people. These studies have documented positive effects of bright-light treatment [54, 57], melatonin [60–63] and physical exercise [64, 65] on sleep in the elderly.

Practice Points
1. Some normative data exist for most age groups from infancy to old age.
2. Developmental trends in actigraph-estimated sleep–wake patterns are generally consistent with those derived from other sleep assessment methods.
3. Neurological and other medical conditions associated with aging probably decrease the accuracy of actigraphy for sleep–wake identification.

CLINICAL SLEEP DISORDERS
Insomnia
The validity of actigraphic assessment of sleep in insomnia has been discussed in a number of publications [1, 14, 15, 66–69]. The basic fact that insomnia patients can remain inactive for prolonged periods in their attempts to fall asleep challenges the accuracy of actigraphic sleep–wake identification at the individual level. More appropriate uses of actigraphy may be for assessment during extended clinical trials or in situations where analysis is at the group level [1, 14, 15, 66–68].

Another approach has been to investigate some of the cognitive components of insomnia. Wicklow studied the role of intrusive thoughts in sleep-onset insomnia and found that the number of intrusive thoughts prior to sleep was a strong predictor of actigraphically assessed sleep latency [70]. In another study, subjective and actigraphic sleep measures provided diverging information on sleep in insomnia associated with chronic pain [71]. Reported pain level was associated with reported measures of sleep quality whereas no such correlation was found with actigraphic sleep measures.

Breathing-related disorders
Although actigraph measures appear to reflect variations in sleep due to breathing-related disturbances, it has become clear that actigraphy cannot be used for identification of breathing-related sleep disturbances [1, 2]. Recent studies have provided additional information indicating that fragmented sleep or more activity during sleep is associated with breathing disorders [32, 72]. Middelkoop and colleagues have concluded, however, that actigraphy is insufficient to identify reliably individuals who suffer from sleep apnea [72]. On the other hand, comparisons between groups of subjects may provide meaningful information. Sadeh and coworkers documented increased activity levels and reduced motionless sleep percent in a group of non-symptomatic asthmatic children compared with control children [32]. Furthermore, actigraphic sleep fragmentation was associated with reduced pulmonary function as measured by peak flow meter.

Periodic leg movements
The idea of actigraphic identification of periodic leg movements (PLMs) is very appealing. A full examination of this approach provided strong support that special devices placed on the foot and tailored scoring algorithm can provide reliable measures of PLMs [73]. A high correlation (0.91) was found between actigraphically scored and PSG scored PLMs. The authors concluded, however, that
a general index of restlessness is not sufficient to obtain either a reliable estimate of PLMs or a useful algorithm. In subsequent studies this method has been applied to demonstrate positive therapeutic effects of medication for restless leg syndrome [74, 75]. A recent study that evaluated the validity of actigraph measures of leg movements indicated that actigraphy underestimated leg electromyographic activity although a relatively high correlation (0.78) was found between the two methods [76]. On the basis of night-to-night error estimates the authors concluded that despite the discrepancy actigraphy could be used to obtain follow-up measures in intervention studies.

**Sleep-schedule disorders**

Although actigraph procedures for assessment and treatment of sleep-schedule disorders have not been systematically established, actigraphy has been accepted as a useful method for assessing sleep in individuals suffering from such disorders [1, 2]. In recent years, actigraphy has been used in intervention studies assessing phase-shifting effects of melatonin administration in normal individuals [77] and in individuals suffering from sleep-phase delay syndrome [78, 79].

**Transient disturbances associated with jet lag and shift work**

Recent studies have demonstrated the usefulness of actigraphy in assessing jet lag and shift work related phenomena [17, 80–87]. For example, in one study, zopiclone administration after a westward flight led to increased sleep duration and sleep consolidation in comparison with placebo [80]. Another study assessed the validity of actigraphic measures versus PSG in a shift-work situation [86]. Other shift-work studies have shown differences in sleep time and sleep quality on different shift schedules [84] and demonstrated that sleep quality can be improved following light treatment and melatonin administration [82].

**Practice Points**

1. Actigraphy should not be used as the sole method for diagnosing sleep disorders.
2. Actigraphy is useful as a complementary assessment method for insomnia, sleep–wake schedule disorders and restless leg syndrome.
3. Actigraphy data reflect disrupted sleep and increased motility in breathing-related disorders; however, the lack of specificity of these activity patterns precludes their use for diagnosis.
4. Actigraphy appears to be a valid tool for assessing changes in sleep schedule and sleep quality induced by jet lag and shift work and related interventions.

**CLINICAL INTERVENTION STUDIES**

Because actigraphy allows objective assessment of sleep for extended periods with minimal inconvenience it has become a very useful tool for intervention studies. Recent actigraphic studies have documented improvement of sleep following administration of drugs [74, 75, 88, 89], melatonin [61–63, 78, 79, 82, 90–92], light exposure [57, 82, 93, 94] and behavioral interventions [10, 65, 95]. A few studies have failed to document changes in sleep following various interventions [53, 60, 96, 97]. One study demonstrated the impact of caffeine consumption on actigraphic sleep measures [98].

The advantages of actigraphy for intervention research are quite evident. The ability to monitor sleep for extended periods provides researchers with high statistical power for analysis of effects, and a growing number of studies support the sensitivity of actigraph measures used repeatedly during intervention studies. However, those designing studies of patients with specific neurological conditions or disorders that are associated with disturbances in the motor system need to remember that the ultimate measure from actigraphy is body movement. Evident change in actigraph measures may be a function of changes in motor behavior rather than sleep or wakefulness.

**Practice Points**

1. Actigraphy is a useful instrument for intervention studies.
2. Actigraphy is most sensitive when a within-subject design is used. In such designs the same device should be used for all trials to minimize interdevice variability.
3. Actigraphy may yield more accurate measures during very long assessment periods than daily logs.
4. Actigraphy may be useful with young children or participants who are unable to complete daily sleep logs, although without logs artifacts may lead to a marked decrease in the accuracy of scores.

CONCLUSIONS AND RECOMMENDATIONS

Actigraph recording instruments along with accompanying scoring software should be carefully evaluated with the research or clinical purpose clearly in mind. We include the following recommendations for researchers and clinicians interested in adding actigraphy to protocols:

- If measures of sleep quality (sleep efficiency, wake after sleep onset) are important, the device chosen must be sensitive to small movements during sleep and the algorithm must show documented validity for such measures.
- Actigraphy is not the best method if interest is in the precise timing of sleep and wake epochs or precise durations to sleep onsets. The largest discrepancies between actigraph and PSG measures are typically around sleep–wake and wake–sleep transitions.
- Caution must be used in evaluating actigraph measures from research participants or patients who are likely to have large amounts of movement during sleep or disrupted sleep. In these situations, within-subject designs and analyses are likely to be more accurate and revealing than between-subject comparisons [14].
- Actigraph measures are unlikely to provide sufficiently accurate information to assign diagnoses of sleep disorders. On the other hand, they may have utility for screening and assessment of treatment.
- Interpretation of results from actigraph studies should always be done with the knowledge that the actigraph records movement and that the scoring algorithm infers sleep and wakefulness. Anything that distorts, exaggerates or dampens movement is likely to result in incorrect translation to sleep–wake measures.
- Record for at least five nights and preferably a week to obtain aggregated measures that reliably characterize individuals and thus are more likely to predict other characteristics. Be aware of potential weekday/weekend differences and derive variables accordingly.
- Devise and follow procedures for testing batteries, testing and maintaining actigraphs and for obtaining participant/subject documentation of bedtimes, risetimes, actigraph off times, external motion and other unusual events.
- Maintain a good relationship with the manufacturer of the devices so that broken units may be repaired quickly and software updates received in a timely manner.

Research Agenda

1. Further comparison of actigraph devices and algorithms.
2. Collection of normative data across the life span with special attention to assessment of age- and sex-related differences in motor behavior during sleep.
3. Further exploration of the actigraphic methods for clinical syndromes such as sleep apnea syndrome or PLM syndrome.
4. Further comparison between actigraphy and subjective reports or PSG in assessing sleep for extended periods in special populations or under special circumstances (e.g. hospitalization).

REFERENCES


* The most important references are denoted by an asterisk.
ACTIGRAPHY IN SLEEP MEDICINE


35. Glod CA, Teicher MH, Hartman CR, et al. Increased...


92. Etzioni A, Luboshitzky R, Tiosano D, et al. Melatonin replacement corrects sleep disturbances in a child...


