Estimating Sleep Patterns with and Adolescents: How Many N Reliable Measures?

Christine Acebo,[1] Avi Sadeh,[3] Ronald Seifer,[2] OrnaTzischinsky,[4] Am

(1) Sleep and Chronobiology Research Laboratory and (2) I Psychiatry and Human Behavior, E.P. Bradley Hospital/Brov Department of Psychology, Tel Aviv University, Tel Aviv, Isra Medicine, Technion-Israel Institute of Technology, Haifa, Isr Cross, Worcester, Mass

Study Objectives: This study provides estimates of reli commonly used actigraphic measures of sleep patterns, and stability of measures over several months.

Design and Setting: Data are from three studies that c graphs and associated validated algorithms [ASA]) on ch schedules.

Participants: Participants were 169 children aged 12-6

Measurements and Results: Up to 28% of weekly re because of illness, technical problems, and participant n should record for at least 1 full week. Reliability estimate for sleep start time, wake minutes, and sleep efficiency. M may require 7 or more nights for estimates of stable indiv for all measures. We found significant and high correlatio when weekend nights were included.

Conclusions: Five or more nights of usable recordings dren and adolescents.

Key words: Actigraphy; children; adolescents; sleep pa

SLEEP AND WAKEFULNESS change dramatically over the course of child development, yet our knowledge of the parameters of these phenomena and how they change remains sparse. The majority of previous objective studies,

Accepted for publication July, 1998

Address correspondence and requests for reprints to Christine Acebo, PhD, Sleep and Chronobiology Research Laboratory, E.P. Bradley Hospital/Brown University School of Medicine, 1011 Veterans Memorial Parkway, E. Providence, RI 02915, Email: Christine_Acebo@Brown.edu

Activity Monitoring in Children
ights Are Necessary for

y R. Wolfson,[5] Abigail Hafer,[1] and Mary A. Carskadon[1]

Developmental Psychopathology Laboratory, Department of
vn University School of Medicine, E. Providence, RI; (3)
el; (4) Sleep Laboratory, Baruch Rappaport Faculty of
ael; (5) Department of Psychology, The College of the Holy

ability for aggregated values from 1 to 7 recording nights for five
reliability as a function of night type (weeknight or weekend night),

btained 7 nights of actigraph data (using Mini Motionlogger acti-
ildren and adolescents living at home on self-selected sleep-wake

0 months, and 55 adolescents aged 11-16 years.
cordings may be unacceptable for analysis in young participants
oncompliance; studies aiming to collect 5 nights of actigraph data
es for values aggregated over any 5 nights were adequate (≥.70)
Measures of sleep minutes and sleep period were less reliable and
idual differences. Reliability for 1- or 2-night aggregates were poor
ns between summer and fall session measures for all five variables

are required to obtain reliable actigraph measures of sleep for chil-

tterns; reliability; individual differences

most of which focus either on the first year of life or on adolescence, have used techniques including direct behavioral observations,[1,2] laboratory studies,[3-8] video recordings,[9] and mattress or crib recordings.[10-11] The strengths and weaknesses of these techniques vary,[12] but none has the capacity for true objective ambulatory monitoring over extended periods. More recently, methods employing continuous measurement of activity by monitors with solid-state memory (actigraphs) have facilitated the assessment of naturalistic sleep/wake patterns. Actigraph monitors have evolved into small, lightweight, and unobtrusive devices capable of collecting time-based activity data over extended periods. In conjunction with validated algorithms for transforming activity data to sleep/wake estimates, this technology provides an opportunity for systematic documentation of natural sleep/wake behavior of infants, children, and adolescents.[13]

Current actigraph methodology, however, is not without shortcomings. For example, although leading proponents of sleep/wake assessment with actigraphy have cautioned that measurement characteristics vary in combination with type of device, type of algorithm, and type of participant,[14] relatively few studies have reported on the validity and reliability of actigraphic sleep/wake measures. We have demonstrated previously that the Mini Motionlogger actigraph and its associated sleep/wake scoring algorithms[15-16] provide valid measures of behavioral sleep/wake patterns for normal infants, children, and adults within documented nocturnal periods. The scoring algorithms are relatively insensitive to placement on the participant (eg, dominant vs nondominant wrist) and to sensitivity differences between individual devices.[15] Validity, however, is assessed in structured situations, such as during laboratory-based sleep studies where the kinds of difficulties faced in field-based ambulatory recording are not encountered. For example, laboratory bedtime, risetime, and total recording time are determined and documented by protocol, technical problems and artifacts are either resolved quickly or documented thoroughly, and participants who are sleep-deprived, ill, or use medications, caffeine, or nicotine are typically excluded. Even so, aggregation of data over multiple nights is routinely done for sleep measures obtained in the laboratory because of known or suspected instability of single-night measures,[17-20] particularly apparent during the first laboratory night.[21] Indeed, some older studies report data aggregated over as many as 4 and 5 nights[22] with the assumption that averaging values over multiple nights will ensure reproducible individual differences.

In classical test theory, observed score variance on a test is assumed to be the sum of true score variance and random error variance (with reliability defined as the ratio of true score variance to total variance). Furthermore, if a test is made longer by combining or aggregating parallel versions of it (ie, versions consisting of correlated items), the proportion of true-score variance will increase as the summed error variance approaches its expected value of zero.* Thus, lengthening a test with parallel forms should improve the accuracy of measurement if the true scores are not subject to systematic fluctuation across situations (eg, practice, developmental, psychosocial, or environmental

*Item here is used to refer to a single-night measurement.

effects). Aggregation of measures over multiple items is thus used to stabilize individual (true-score) differences, yielding measures that not only typify the individual but may also be expected to predict other variables. The best index of whether such aggregation is indeed useful is the reliability estimate of a single measurement item (ie, for 1 night of sleep) compared with the reliability of scores obtained by aggregating over multiple measurement items (ie, for multiple nights of sleep).

By its nature, actigraphic assessment of sleep/wake patterns outside the laboratory loses control over the myriad of factors that we know can impact on the stability of sleep measures. Few studies, however, have addressed reliability of measures and, by extension, the number of recording nights required to obtain aggregated measures that both characterize individuals and have predictive validity.[13,14,16,23,24] If actigraphic sleep/wake estimates are to be useful as accurate measures of individual differences, then we must know how many nights to record.

The purpose of this study was to estimate the reliability of aggregate values based on 1 to 7 nights of recording for five commonly used actigraphic measures of sleep patterns. We also assessed reliability as a function of type of night (weeknight or weekend night). Finally, we assessed the stability of aggregated measures over 2 months for a subset of participants. Such evaluations are critical to assure the most appropriate and efficient use of actigraphy for the study of sleep patterns.

METHODS

Data for this report were drawn from three studies that obtained at least 7 nights of actigraph data on children and adolescents living at home on self- or parent-selected sleep-wake schedules.[25-28] All procedures were approved by the E.P. Bradley Hospital Institutional Review Board for the Protection of Human Subjects (IRB), and parents provided informed consent for participants. A brief description of each study is provided below.

Study 1: Infant/Childhood Activity Study (child groups).—Boys and girls aged 12-60 months took part in a project aimed at describing normative behavioral sleep-wake patterns.[25,26] Potential participants were located from a variety of sources, screened, and recruited into groups of 12-, 18-, 24-, 30-, 36-, 48-, and 60-month-olds (with 2-month windows around each age). Successful completion of the study required at least 4 nights of actigraph data when (1) the child was not sick or taking medications; (2) the actigraph was on the child and collecting data; (3) the data corresponded to the mother's diary; and (4) the diary was completed appropriately. Of the 169 children (84 boys and 85 girls) who completed the study, 154 were Caucasian, 3 African-American, 1 Asian, 1 Hispanic, and 10 of mixed racial descent. Fifty additional participants aged 12-60 months were enrolled but dropped from the study for a variety of reasons, including illness (n=20), child's refusal to wear actigraph (n=11), technical problems (n=10), mother's lack of compliance with procedures (n=3), and discovery of exclusion criteria (see below) after assessments were complete (n=6). Data from 15 participants came from a second or third week of recording due to technical problems or illness during the first week.

Inclusion criteria were that children were full-term infants with no medical complications at birth, lived full-time with their biological mothers, and had mothers who could read and write English. Children were excluded for major genetic, medical, psychological/behavioral problems, or serious sleep problems (defined by parent). Also excluded were children whose parents reported mental health problems or sleep disorders in the child's first-degree relatives, and children who were taking medications that might affect sleep or alertness. Because the actigraph records motion, children who routinely co-slept with parents or siblings were not included in the study, although some periods of co-sleeping did occur and were documented by parents.

Participants who passed a telephone screening interview were scheduled for a home visit. At this time, mothers were interviewed to obtain demographic and health information, the actigraph was put on the child and the mother was instructed in its care and use, instructions were provided for diary completion, and a booklet of additional questionnaires was left for parents to complete during the week. This project used Mini Motionlogger actigraphs (Advanced Model, Ambulatory Monitoring Inc., Ardsley, NY), weighing 57 g, set for 1-minute recording bins and zero crossing mode (a measure of movement frequency). The device accumulates one count each time movement causes the sensor signal to cross a fixed reference signal. Amplifier setting 18 was used for this device, which yields a sensitivity of .05 g in a frequency range 2 to 3 Hz. Parents of children under 36 months of age placed the actigraph on the child's left ankle; for children 36 months or older, the actigraph was placed on the nondominant wrist (as reported by parents). Mothers kept a diary of the child's bedtimes and risetimes, as well as times when the actigraph was off or exposed to external motion (such as car rides). Research staff telephoned mothers in the middle of the week to ask about any problems, illnesses, or special events, and to remind them about the second home visit. The second home visit was made at the end of the week to check the dairy and questionnaires for completeness and to collect the actigraph and forms. The actigraph was downloaded into the computer immediately upon return to the laboratory, the record was printed and checked against the mother's diary, and a follow-up phone call was made to resolve any discrepancies or to ask about missing data.

Study 2: Circadian Rhythms Project (adolescent group A).—Adolescent Group A included 19 participants (10 boys, ages 11.2 to 14. 1, mean 12.7±1.0 years; 9 girls ages 12.2 to 14.4, mean 13.1±0.7) participating in a study investigating sleep and circadian rhythm parameters with home and laboratory assessments.[27] Participants were recruited through a laboratory open house for parents and children at which procedures were explained and demonstrated. Exclusion criteria included self-reported sleep schedules that varied by greater than 3 hours across a week or that indicated a pattern of chronic insufficient sleep; current sleep problem or sleep disorder in self or parent; major genetic, medical, psychological or behavioral problems; mental retardation or learning disabilities; current use of psychoactive agents or other drugs that may affect sleep and alertness; or reported family history (first-degree relative) of diagnosed major depressive disorder, bipolar depression, or schizoaffective disorder. Screening was performed by interview, questionnaire, and child and parental report. Eighteen of the group A adolescents were Caucasian and one was of mixed African-American/Native-American descent.

For one phase of this study, participants wore wrist actigraphs identical to those described above on the non-dominant wrist[29] for 7 days while living at home on their self-selected sleep-wake schedules. No participants had summer school or similar obligations during this week. In addition to wearing the actigraph, participants completed daily diaries and called the laboratory each morning to report bedtime, risetime, and whether the actigraph was working. Follow-up telephone calls were made by research staff to all participants at midweek and whenever a participant failed to make a morning call to the lab. The only constraints on these youngsters' sleep patterns were that they were to sleep at home each night and to avoid staying awake all night. Participants came to the lab at the end of the week, at which time the actigraph was downloaded, the actigraph record was printed, and discrepancies between record and diary were resolved before subjects left the laboratory or by phone call the next day.

Study 3: School Transition Project (adolescent group B).—The third group were adolescents taking part in a longitudinal study investigating sleep/wake pattern changes across the 9th to 10th grade transition.[28] These participants were recruited from Rhode Island junior high schools, and we have included data from 36 (13 boys, ages 14.0 to 15.7, mean 14.9±.4 years; 23 girls ages 14.4 to 16.2, mean 15.0±.5) who completed the summer and/or fall sessions. These data allow comparisons of sleep parameters from actigraphy between times with and without school schedules. Thirty adolescents were Caucasian, two were African-American, two were Hispanic, one was Bermudan, and one was of mixed ethnicity. Data from both sessions

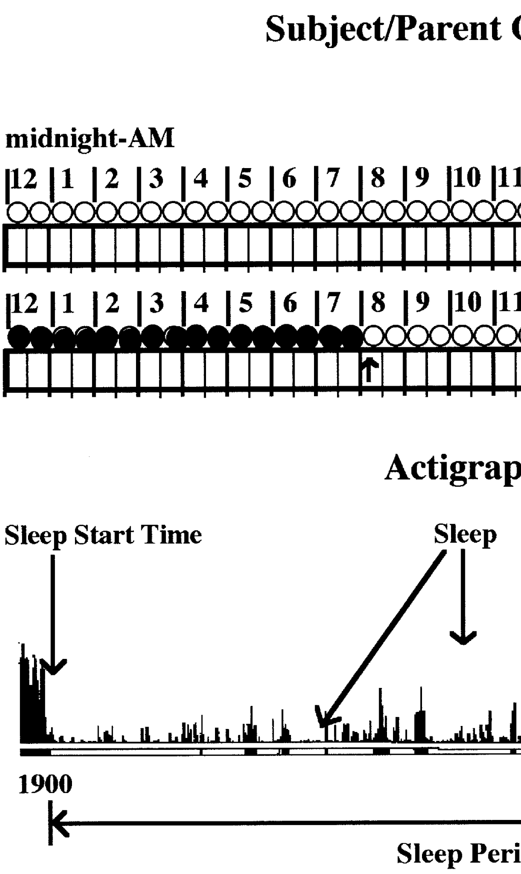
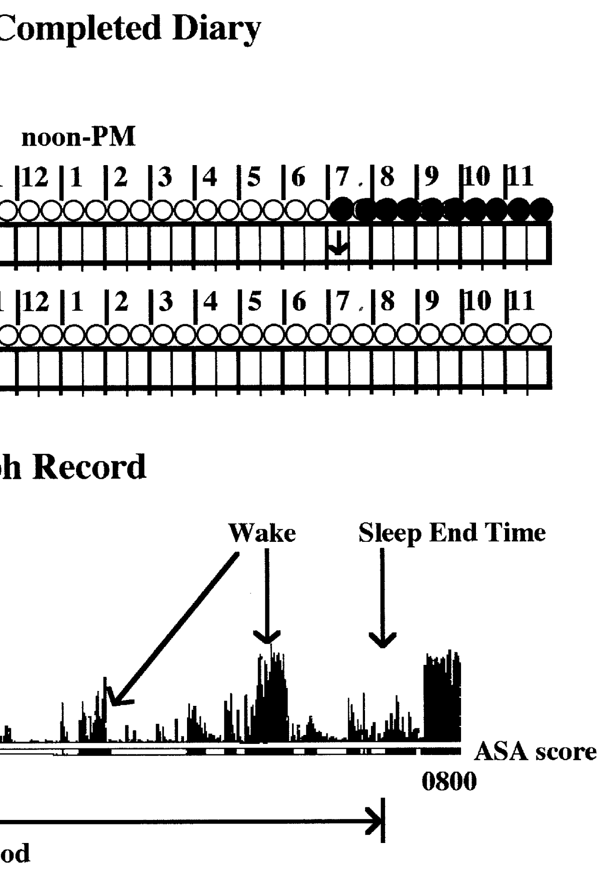


Figure 1.—Illustration of actigraph scoring procedures. An example of 1 nig
the top of the figure. Bedtime was indicated by a down arrow, risetime by a
were darkened. An example of the corresponding night of actigraphy data is
applied to each nightly record to obtain minutes of sleep and wake during p
and 30 minutes after the reported rising time (the scoring interval). In this ex
ened bar under the actigraph record indicates minutes scored as wake by th
Sleep start time was defined as the first minute of at least 3 consecutive min
as the last minute of at least 5 consecutive minutes of scored sleep just prio
of elapsed minutes from sleep start time to sleep offset time; sleep minutes a
as the number of minutes scored as wake during sleep period; and sleep eff

were available for 29 of the 36 participants. Actigraph data were not available for two participants during the fall session because of technical problems in one and because the other withdrew from the study. Data are unavailable for five participants during the summer assessment due to technical problems for two, and as a result of irreconcilable differences between actigraph record and diary report for three. Exclusion criteria included parent- or child-reported current sleep disorder, chronic medical illness, use of psychoactive drugs, or evidence of learning disabilities/psychopathology.

Procedures for this group were the same as for Adolescent group A described above. Data were taken from the first 7 nights of the summer and fall studies when participants lived at home on self-selected sleep-wake schedules, wore actigraphs on nondominant wrists, completed



ht of a daily diary used by mothers of 1- to 5-year-old children is shown on
n up arrow, and half-hour interval circles when the child was judged asleep
shown on the bottom of the figure. The sleep scoring algorithm (ASA) was
ortions of the record encompassing 30 minutes before the reported bedtime
ample, the algorithm was applied between 0630 and 0830 hours. The dark-
e ASA algorithm. Open portions of the bar indicate minutes scored as sleep.
utes of scored sleep within the scoring interval; sleep end time was defined
r to the end of the scoring interval; sleep period was defined as the number
s the number of minutes during sleep period scored as sleep; wake minutes
iciency as (sleep minutes/ sleep period) x 100.

daily diaries, and called the laboratory each morning to report bedtime, risetime, and whether the actigraph was working.

Measures

Sleep/wake measures were estimated from actigraphic data using the validated ASA algorithm.[15,16] Each nightly record was scored for the portion indicated as nighttime sleep by parent (children 12-60 months) or adolescent (group A, group B) report. Specifically, the algorithm was applied during portions of the record encompassing 30 minutes before reported bedtime through 30 minutes after reported rising time (the scoring interval). The variables assessed for this report are defined below and illustrated in Fig. 1. All final scoring required reasonable correspondence between parent's/child's completed diary and the

Table 1.—Distribution of the number of subjects with at most 4, 5, 6, or 7 scored nights

Age Group	4	5	6	7	Total n
12 mo.		4	11	9	24
18 mo.	2	5	4	18	29
24 mo.	3	3	4	12	22
30 mo.	3	2	9	7	21
36 mo.	3	3	5	10	21
48 mo.	1	3	5	15	24
60 mo.	3	6	7	12	28
Grp A:SU	1	2	1	15	19
Grp B:SU			7	24	31
Grp B:FA	2	6	7	19	34
Total	18	34	60	141	253

Grp A:SU=Adolescent group A studied during summer.
Grp B:SU=Adolescent group B studied during summer.
Grp B:FA=Adolescent group B studied during fall.

actigraph record. For example, if an adolescent recorded a diary bedtime of 2300 but continuous sleep could be scored beginning at 2130, the adolescent was questioned about the discrepancy. If no reasonable explanation was offered, the night was considered unscorable.

Sleep start time: the first minute of at least 3 consecutive minutes of sleep as scored by the ASA algorithm
Sleep end time: the last minute of at least 5 consecutive minutes of sleep just prior to the end of the scoring interval as scored by the ASA algorithm (This variable was not analyzed for this report.)
Sleep period: elapsed minutes from sleep start time to sleep end time
Sleep minutes: minutes of sleep period that are scored as sleep by the ASA algorithm (indicated by an open bar on the ASA score bar in Fig. 1)
Wake minutes: the number of minutes awake during sleep period (indicated by a solid bar on the ASA score bar in Fig. 1)
Sleep efficiency: percentage of the sleep period spent asleep as scored by the ASA algorithm ([sleep minutes/ sleep period] x 100)

Nights were not scored as usable data if the child was sick or taking medications, the child slept with another person for the entire night, the actigraph was off or not working for all or part of the night, or the diary indicated unusual external motion, such as sleeping in a car or train. Table 1 shows the distribution of participants in each age group with

Table 2.—Sleep start

	Number of Nights						
Age Group	1	2	3	4	5	6	7
12 mo.	.42	.59	.69	**.74**	**.78**	**.81**	**.84**
18 mo.	.44	.61	**.70**	**.76**	**.79**	**.82**	**.84**
24 mo.	.38	.55	.65	**.71**	**.76**	**.79**	**.81**
30 mo.	.48	.65	**.74**	**.79**	**.82**	**.85**	**.87**
36 mo.	.63	**.77**	**.84**	**.87**	**.89**	**.91**	**.92**
48 mo.	.50	.67	**.75**	**.80**	**.84**	**.86**	**.88**
60 mo.	.42	.59	.68	**.74**	**.78**	**.81**	**.83**
Grp A:SU	.57	**.72**	**.80**	**.84**	**.87**	**.89**	**.90**
Grp B:SU	.46	.63	**.72**	**.78**	**.81**	**.84**	**.86**
Grp B:FA	.17	.29	.38	.45	.51	.56	.59

Reliability estimates from aggregated mean values over successive numbers of recording nights for each group of subjects. Each column shows the estimated reliability for a value based on a given number of nights (Winer, 1971). Values ≥ .70 are highlighted in bold and indicate acceptable reliability
Grp A:SU=Adolescent Group A studied during summer
Grp B:SU=Adolescent Group B studied during summer
Grp B:FA=Adolescent Group B studied during fall

respect to number of nights scored.

Data Analysis

We calculated the intraclass correlation estimating the reliability of measures assessed on a single night of recording within each age group for each measure, using parameters derived from a random effects analysis of variance on all available data.[30] By applying the Spearman-Brown prophecy formula,[30] we then were able to estimate the effects that changes in the length of the test (number of recorded nights) would have on reliability. Such reliability coefficients indicate the proportion of true-score variance in aggregated test scores, with coefficients above .70 indicating adequate reliability according to criteria that have been proposed for an individual difference measure.[31] Missing data were left missing, and thus reliability estimates may include some increased variability.

Additional analyses with data from adolescent groups assessed the effects of night of the week on reliability of measures, because adolescents are known to manifest different sleep patterns between weekday (or school) nights and weekend nights.[32-35] Reliability coefficients were calculated within each adolescent group for 5-night mean values that included only 5 weeknights, and 5-night means that included 3 weeknights and 2 weekend nights. In this way we aimed to keep the size of the aggregate the same

Table 3.—Sleep period

	Number of Nights						
Age Group	1	2	3	4	5	6	7
12 mo.	.37	.54	.64	**.70**	**.74**	**.78**	**.80**
18 mo.	.26	.42	.52	.59	.64	.68	.72
24 mo.	.15	.27	.35	.42	.48	.52	.56
30 mo.	.26	.41	.51	.58	.64	.68	**.71**
36 mo.	.30	.46	.56	.63	.68	**.72**	**.75**
48 mo.	.13	.22	.30	.37	.42	.46	.50
60 mo.	.28	.44	.54	.61	.67	**.71**	**.74**
Grp A:SU	.21	.34	.44	.51	.56	.61	.64
Grp B:SU	.16	.28	.37	.43	.49	.54	.57
Grp B:FA	.10	.18	.25	.31	.36	.40	.44

Reliability estimates from aggregated mean values over successive numbers of recording nights for each group of subjects. Each column shows the estimated reliability for a value based on a given number of nights (Winer, 1971). Values ≥ .70 are highlighted in bold and indicate acceptable reliability
Grp A:SU=Adolescent Group A studied during summer
Grp B:SU=Adolescent Group B studied during summer
Grp B:FA=Adolescent Group B studied during fall

Table 5.—Wake minutes

	Number of Nights						
Age group	1	2	3	4	5	6	7
12 mo.	.47	.64	.73	.78	.82	.84	.86
18 mo.	.33	.50	.60	.66	.71	.75	.78
24 mo.	.31	.48	.58	.65	.70	.73	.76
30 mo.	.46	.63	**.72**	**.77**	.81	.84	.86
36 mo.	.06	.11	.16	.20	.24	.27	.30
48 mo.	.36	.53	.63	.69	.74	.77	.80
60 mo.	.34	.50	.60	.67	.72	.75	.78
Grp A:SU	.60	.75	.82	.85	.88	.90	.91
Grp B:SU	.63	.77	.84	.87	.90	.91	.92
Grp B:FA	.53	.69	.77	.82	.85	.87	.89

Reliability estimates from aggregated mean values over successive numbers of recording nights for each group of subjects. Each column shows the estimated reliability for a value based on a given number of nights (Winer, 1971). Values ≥.70 are highlighted in bold and indicate acceptable reliability
Grp A:SU=Adolescent Group A studied during summer
Grp B:SU=Adolescent Group B studied during summer
Grp B:FA=Adolescent Group B studied during fall

and contrast presumably homogeneous weeknights with heterogeneous weeknight/weekend night combinations (the latter perhaps having lower internal consistency). Finally, Pearson product-moment correlations were calcu-

Table 4.—Sleep minutes

	Number of Nights						
Age group	1	2	3	4	5	6	7
12 mo.	.44	.61	**.70**	**.76**	**.80**	**.82**	**.85**
18 mo.	.22	.36	.46	.53	.59	.63	.67
24 mo.	.32	.48	.58	.65	**.70**	**.74**	**.77**
30 mo.	.31	.47	.57	.64	.69	**.73**	**.75**
36 mo.	.21	.35	.44	.52	.57	.62	.65
48 mo.	.32	.49	.59	.66	**.71**	**.74**	**.77**
60 mo.	.51	.68	**.76**	**.81**	**.84**	**.86**	**.88**
Grp A:SU	.37	.54	.64	**.70**	**.75**	**.78**	**.81**
Grp B:SU	.30	.46	.56	.63	.68	**.72**	**.75**
Grp B:FA	.25	.40	.50	.57	.62	.66	**.70**

Reliability estimates from aggregated mean values over successive numbers of recording nights for each group of subjects. Each column shows the estimated reliability for a value based on a given number of nights (Winer, 1971). Values $\geq$.70 are highlighted in bold and indicate acceptable reliability
Grp A:SU=Adolescent Group A studied during summer
Grp B:SU=Adolescent Group B studied during summer
Grp B:FA=Adolescent Group B studied during fall

Table 6.—Sleep efficiency

	Number of Nights						
Age Group	1	2	3	4	5	6	7
12 mo.	.53	.69	**.77**	**.82**	**.85**	**.87**	**.89**
18 mo.	.33	.49	.59	.66	**.71**	**.74**	**.77**
24 mo.	.36	.53	.63	.69	**.74**	**.77**	**.80**
30 mo.	.45	.62	**.71**	**.76**	**.80**	**.83**	**.85**
36 mo.	.05	.10	.15	.19	.22	.26	.29
48 mo.	.33	.50	.60	.67	**.72**	**.75**	**.78**
60 mo.	.40	.58	.67	**.73**	**.77**	**.80**	**.83**
Grp A:SU	.63	**.77**	**.84**	**.87**	**.89**	**.91**	**.92**
Grp B:SU	.63	**.78**	**.84**	**.87**	**.90**	**.91**	**.92**
Grp B:FA	.59	**.74**	**.81**	**.85**	**.88**	**.89**	**.91**

Reliability estimates from aggregated mean values over successive numbers of recording nights for each group of subjects. Each column shows the estimated reliability for a value based on a given number of nights (Winer, 1971). Values $\geq$ .70 are highlighted in bold and indicate acceptable reliability
Grp A:SU=Adolescent Group A studied during summer
Grp B:SU=Adolescent Group B studied during summer
Grp B:FA=Adolescent Group B studied during fall

lated between summer and fall session mean values from various combinations and numbers of nights for group B adolescents to describe stability over longer periods of time.

Table 7.—Reliability of measures aggregated over weeknights vs weeknights plus weekend nights

		Adolescent Group		
		Group A	Group B	Group B
		Summer	Summer	Fall
Sleep Start Time	5 Weeknights	**.88**	**.77**	**.79**
	3 Weeknights + 2 Weekend Nights	**.87**	**.83**	.35
Sleep Period Time	5 Weeknights	.65	.49	**.72**
	3 Weeknights + 2 Weekend Nights	.54	.45	.34
Sleep Minutes	5 Weeknights	**.74**	.68	.65
	3 Weeknights + 2 Weekend Nights	**.80**	.65	**.77**
Wake Minutes	5 Weeknights	**.89**	**.89**	**.87**
	3 Weeknights + 2 Weekend Nights	**.88**	**.89**	**.84**
Sleep Efficiency	5 Weeknights	**.89**	**.89**	**.87**
	3 Weeknights + 2 Weekend Nights	**.91**	**.90**	**.87**

Reliability estimates from mean values over either 5 weeknights (or school nights in the Fall) or 3 weeknights and 2 weekend nights for Adolescent groups. Values $\geq$ .70 are highlighted in bold and indicate acceptable reliability.

RESULTS

Reliability estimates for aggregated values over successive numbers of recording nights (1 to 7) for each measure and each group are presented in Tables 2 through 6. Columns show reliability estimates for mean values based on a given number of nights. As shown in the tables, increasing aggregation was associated with greater reliability (eg, 5-night mean values were more reliable than 4-night mean values). Adequate reliability (≥70 indicated in bold on the table), however, was infrequently reached for most measures with fewer than 5 nights. For certain variables, even 7 nights of recording failed to provide this level of reliability. Marked differences were clear among measures; adequate reliability was evident for sleep start (Table 2) for most groups with 3 nights of recording, but only five of 10 groups had adequate reliability for sleep period (Table 3) even with 7 nights of recording. Six or 7 nights produced reliable estimates of sleep minutes (Table 4) in most subject groups. Wake minutes (Table 5) and sleep efficiency (Table 6) estimates were adequate with 5 nights. Reliability estimates were markedly low for 36-month-old children on wake minutes and sleep efficiency, and none of the adolescent samples showed adequate reliability for sleep period. Examination of means, standard deviations, and coefficients of variation indicated that restriction of range was not a reason for low reliability values.

Table 8.—Correlations between summer and fall session mean values aggregated over various combinations of nights (n= 14 to 26, adolescent group B)

	WTh	FSa	FSaSuM	SuMTW	All Nights
Sleep Start Time	.58 **	.68 ***	.72 ***	.66 **	.75 **
Sleep Period Time	.01	.69 ***	.53 **	.19	.45
Sleep Minutes	.28	.66 ***	.59 **	.32	.59*
Wake Minutes	.73 ***	.63 ***	.69 ***	.84 ***	.88 ***
Sleep Efficiency	.69 ***	.63 ***	.69 **	.86 ***	.88 ***

* $p < .05$; ** $p < .01$; *** $p < .001$

Each column shows Pearson correlations between summer and Fall sessions for mean values of specific combinations of nights. WTh (Wednesday and Thursday), FSa (Firday and Saturday), FSaSuM (Friday,Saturday, Sunday, and Monday), SuMTW (Sunday, Monday, Tuesday, and Wednesday), All Nights (all scored nights regardless of day of the week).

Note: True error-corrected significance of $p<.002$ (Bonferroni method) is indicated for correlations marked by ***.

Reliability estimates for 5-night mean values aggregated over weeknights only or over a combination of weeknights and weekend nights are presented in Table 7. Few meaningful differences were apparent between the two sets of estimates, except for a loss of adequate reliability for sleep start and sleep period during the fall session (when students were going to school), when weeknights and weekend nights were combined.

To evaluate the influence of type of recording night on assessed stability over time, correlations between summer and fall sessions for adolescent group B were computed for mean actigraphy values aggregated over various combinations and numbers of nights (see Table 8). The variables sleep start, wake minutes, and sleep efficiency showed significant stability over the 2-month interval for all combinations of nights, indicating excellent stability over several months for these measures. Sleep period and sleep minutes were significantly correlated only when weekend nights were included in the average.

DISCUSSION

Actigraphy has become a widely used methodology because it appears at first glance to be easy, affordable, and user-friendly. Our goal was to evaluate some of the psychometric properties of this technique by scoring actigraph records with algorithms validated for the specific device used (Mini-Act actigraph) and for the kinds of participants studied (children and adolescents). The first question we addressed was how much acceptable data can be expected from a 7-day recording period. Our experience with the children and adolescents in these three studies indicates that a loss of up to 28% of weekly recordings may be expected; the main reasons were illness, technical problems, and participant noncompliance (including failure to wear the actigraph or complete the diary regularly and accurately). Further, from 7-day recordings, 44% of participants with acceptable data sets had fewer than 7 scorable nights, with 93% yielding at least 5 nights (not necessarily consecutive). Thus, from a practical point of view, studies aiming to collect 5 nights of actigraph data should likely record for at least a week.

The next question was how much night-to-night consistency was evident for children and adolescents across the age ranges studied. For all measures, estimates over 1 or 2 nights were poor. In contrast, reliability estimates for mean values aggregated over any 5 nights of recording were generally adequate for sleep start time, wake minutes, and sleep efficiency. Measures of sleep minutes and sleep period were less reliable, and may require 7 nights or more to obtain adequate estimates of stable individual differences. Extrapolating from these data combined with the data for the number of acceptable records, a recording time of at least a week should be used to obtain commonly used actigraph measures to reliably characterize individuals. This suggestion is consistent with the data presented in other studies that assessed stability across nights for other devices and other scoring algorithms.[24,36] Adequate reliability is especially important for clinical applications, where the unit of analysis is the individual and unreliable data may result in incorrect diagnostic classification.[23]

The poor reliability estimates for sleep start time in the adolescent group studied during the school year led to further analyses. Reliability estimates of 5-night aggregates, some including only weeknights and others including a combination of weeknights and weekend nights, indicated that sleep period stability was poor for adolescents examined during summer sessions (regardless of the mix of nights) and for the weeknight/weekend night combination for adolescents examined during the fall session. Sleep start time reliability for the combined weeknight/weekend night measure was also poor for the adolescents studied during the fall. Together, these findings highlight the often-noted instability of adolescent sleep scheduling.[32-35] Nevertheless, we found significant and high correlations between summer and fall session measures for all five actigraph variables when weekend nights were included. This finding supports the usefulness of these measures as indicators of stable individual differences in sleep-wake patterning in adolescents if weekend nights are included in the aggregated measures.

We do not understand why 36-month-old infants showed such low reliability on wake minutes and sleep efficiency. Simply put, individual 36-month old children varied widely from night to night on both wake minutes and sleep efficiency. We do not know whether this is a characteristic of the children in our sample or a function of a particular developmental instability at this age.

Because of the limitations of any sleep-scoring algorithm, measurement error may not be equivalent across all records. For example, measurement error for wake minutes may be higher for records with large amounts of activity than for records with small amounts of activity during sleep. Reliability estimates may thus have some error as a function of our sleep-scoring algorithm.

This study contributes to a relatively new literature, assessing some ages of children with one specific instrument and algorithm for estimating sleep. The generalizability of our conclusions will increase as more information becomes available on the degree of similarity across instruments and similarity of phenomena across age groups and clinical populations.

From these three studies, we conclude that:

1. Actigraphy is a methodology that requires care in its implementation and yields reliable data for externally documented bedtime periods.
2. Individual differences on some sleep parameters that we expect to be reproducible (eg, sleep start, wake minutes, and sleep efficiency) likely require aggregation over 5 or more nights of usable recording. We suspect this is a function of intraindividual variability resulting from situational and other influences that can affect any sleep/wake measure when sleep/wake schedule is not artificially controlled, rather than error due to the recording methodology. There is a wealth of prior evidence that sleep is influenced by such factors as amount of prior sleep,[37] presence of bed partner,[38-40] fixed or subject-selected scheduling,[18-20,41] and so forth.
3. Individual differences in sleep period and sleep minutes likely require more than 7 nights of recording. Again, we suspect this reflects the joint contribution of both reliable individual differences and situational factors. In fact, instability of sleep scheduling in teenagers is a subject of considerable current interest.[27,42,43]
4. The social context in which recordings are made may influence the psychometric properties of the data obtained (eg, summer vs school-year schedules for adolescents).

In sum, our findings support continued use of actigraphy for studies of child and adolescent sleep/wake patterns, but only when appropriate attention is given to pragmatic issues around the data collection and understanding of the psychometric properties of shorter vs longer data collection periods.

ACKNOWLEDGMENTS

This work was supported by NIH grants MH146757, MH145945, and MH52415. Dr Carskadon is supported by MH01358. We thank Elizabeth Yoder, Stephanie Shimada, Rachel Westerman, Semra Aytur, Camille Brown, Jennifer

Lane, Deborah Gutman, Dana Kovalchick, Katherine Minard, Jennifer Taylor, PhD, Clayton Bennett Jr., Catherine Darley, Jenifer Wicks, Katherine Sharkey, and Liza Kelly for their assistance with this project.

REFERENCES

1. Thoman EB, Korner AF, Kraemer HC. Individual consistency in behavioral states in neonates. Dev Psychobiol 1976;9:271-283

2. Becker PT, Thoman EB. Organization of sleeping and waking states in infants: consistency across contexts. Physiol Behav 1983;31:405-410.

3. Kahn A, Fisher C, Edwards A et al. Twenty-four hour sleep patterns: comparison between 2- to 3-year old and 4- to 6-year-old children. Arch Gen Psychiatry 1973;29:380-385.

4. Williams R, Karaçan I, Hursch C. EEG of Human Sleep. New York: J Wiley & Sons, 1974.

5. Hoppenbrouwers T, Hodgman JE, Harper RM, Sterman MB. Temporal distribution of sleep states, somatic activity and autonomic activity during the first half year of life. Sleep 1982;5:131-144.

6. Fagioli 1, Salzarulo P. Sleep states development in the first year of life assessed through 24-hr recordings. Early Hum Dev 1982;68:215-228.

7. Coble P, Kupfer D, Reynolds C, Houck R EEG sleep of healthy children 6 to 12 years of age. In: Guilleminault C, ed. Sleep and Its Disorders in Children. New York: Raven Press, 1987: 29-41.

8. Carskadon MA, Keenen S, Dement WC. Nighttime sleep and daytime sleep tendency in preadolescents. In: Guilleminault C, ed. Sleep and Its Disorders in Children. New York: Raven Press, 1987: 43-52.

9. Anders TF & Keener M Developmental course of night-time sleep-wake patterns in full-term and premature infants during the first year of life. Sleep 1985;8:173-192.

10. Sander LW, Stechler G, Burns P, Julia H. Early mother-infant interaction and 24-hour patterns of activity and sleep. J Am Acad Child Psychiatry 1970;9(l):103-123.

11. Thoman EB & Whitney MP. Sleep states of infants monitored in the home: individual differences, developmental trends, and origins of diurnal cyclicity. Inf Behav Dev 1989; 12:59-75

12. Thoman, EB & Acebo, C. Monitoring of Sleep in Neonates and Young Children. In: Ferber R, Kryger M, eds. Principles and Practices of Sleep Medicine in the Child. Philadelphia, PA. W.B. Saunders, 1995:55-68.

13. Sadeh A, Lavie P, Scher A, Tirosh E, Epstein R. Actigraphic home-monitoring sleep-disturbed and control infants and young children: a new method for pediatric assessment of sleep-wake patterns. Pediatrics 1991;87:494-9.

14. Sadeh A, Hauri PJ, Kripke DF, Lavie P The role of actigraphy in the evaluation of sleep disorders. Sleep 1995;18:288-302.

15. Sadeh A, Sharkey K, Carskadon MA. Activity-based sleep-wake identification: an empirical test of methodological issues. Sleep 1994; 17:201-7.

16. Sadeh A, Acebo C, Seifer R, Aytur S, Carskadon MA. Activity-based assessment of sleep-wake patterns during the first year of life. Infant Behav Develop 1995;18:329-337.

17. Webb W'B & Agnew HW. Measurement and characteristics of nocturnal sleep. In: Abt LE, Reiss BF, eds. Progress in Clinical Psychology. New York: Grune and Stratton, 1969:2-27.

18. Moses J, Lubin A, Naitoh P, Johnson LC. Reliability of sleep measures. Psychophysiol 1972;9:78-82.

19. Feinberg I. Some observations of the reliability of REM variables. Psychophysiol 1974;11:68-72

20. Bixler EO, Kales A, Jacoby JA, Soldatos CR, Vela-Bueno A. Nocturnal sleep and wakefulness: effects of age and sex in normal sleepers. Intern J Neuroscience 1984;23:3342.

21. Agnew HW, Webb WB, Williams RL. The first night effect: An EEG study of sleep. Psychophysiol 1966;2:263-266.

22. Feinberg 1, Koresko RL, Heller N. EEG sleep patterns as a function of normal and pathological aging in man. J Psychiat Res 1967, Vol 5, 107-144.
23. Sadeh A. Evaluating night wakings in sleep-disturbed infants: a methodological study of parental reports and actigraphy. Sleep 1996; 19(10):757-762.
24. Jean-Louis G, von Gizycki H, Zizi F, Spielman AJ, Fookson J, Taub H. Sleep/wake activity in normal individuals: a longitudinal actigraphic assessment. Sleep Res 1995;24A:475.
25. Acebo C, Sadeh A, Seifer R, Tzischinsky 0, Dickstein S, Aytur S, Brown C, Shimada S, Yoder E, Hafer A, Carskadon MA. Mothers' assessment of sleep behaviors in young children: scale reliability and validation versus actigraphy. Sleep Res 1994;23:96.
26. Tzischinsky 0, Aytur S, Lane J, Westerman R, Hafer A, Seifer R, Acebo C, Sadeh A, Carskadon MA. Temperament and actigraphically monitored night sleep in healthy toddlers. Sleep Res 1994;23:107.
27. Carskadon MA, Acebo C, Richardson GS, Tate BA, Seifer R. An approach to studying circadian rhythms of adolescent humans. J of Biol Rhythms 1997;12(3):278-289.
28. Carskadon MA, Acebo C, Wolfson AR, Tzischinsky 0, Darley C. Rem sleep on MSLTs in high school students is related to circadian phase. Sleep Research 1997;26:705.
29. Oldfield RC. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 1971;9:97-113.
30. Winer B. Statistical Principles in Experimental Design. (2nd Edition). New York: McGrawHill, 1971.
31. Rosenthal R, Rosnow RL. Essentials of Behavioral Research: Methods and Data Analysis. 2nd Edition. New York: McGraw-Hill, 1991.
32. Billiard M, Alperovitch A, Perot C, James A. Excessive daytime somnolence in young men, prevalence and contributing factors. Sleep 1987; 10:297-305.
33. Strauch I, Meier B. Sleep need in adolescents: A longitudinal approach. Sleep 1988;11:378-386.
34. Carskadon MA. Adolescent sleepiness: Increased risk in a high-risk population. Alcohol, Drugs, and Driving 1990;5/6:317-328.
35. Szymczak JT, Jasinska M, Pawlak E, Swierzykowska M. Annual and weekly changes in the sleep-wake rhythm of schoolchildren. Sleep 1993; 16(5):433-435.
36. van Hilten JJ, Braat EAM, van der Velde EA, Middelkoop HAM, Kerkhof GA, Kamphuisen HAC. Ambulatory activity monitoring during sleep: an evaluation of internight and intrasubject variability in healthy persons ages 50-98 years. Sleep 1993;16(2):146-150.
37. Aserinsky E. The maximal capacity for sleep: Rapid eye movement density as an index of sleep satiety. Biological Psychiatry 1969; 1: 147-159.
38. Mosko S, Richard C, McKenna J. Infant sleep and arousals during bedsharing. Ped Pulmonol 1995;20:340.
39. Mosko S, Richard C, McKenna J, Drummond S. Infant sleep architecture during bedsharing and possible implications for SIDS. Sleep 1996; 19:677-84
40. Monroe LJ Transient changes in EEG sleep patterns of married good sleepers: The effects of altering sleeping arrangement. Psychophysiol 1969;6:330-337.
41. Bunnell DE, Horvath SM. Reliability of the sleep cycle as a scoring unit. Psychophysiol 1984;21(4):482-485.
42. Carskadon MA, Vieira C, Acebo C. Association between puberty and delayed phase preference. Sleep 1993;16(3):258-262.
43. Acebo C, Carskadon, MA. Irregular sleep/wake patterns in adolescents. In Carskadon MA, ed. Adolescent Sleep Patterns: Biological, Social, and Psychological Influences. New York: Cambridge University Press, in press.