Waiting Time Distributions of Actigraphy Measured Sleep

J.E. Slaven*[,1], A. Mnatsakanova[1], S. Li[1,2], J.M. Violanti[3], C.M. Burchfiel[1], B.J. Vila[4] and M.E. Andrew[1]

[1]*Biostatistics and Epidemiology Branch, Health Effects Laboratory Division, National Institute for Occupational Safety and Health, Centers for Disease Control and Prevention, Morgantown, WV, USA*

[2]*Department of Statistics, West Virginia University, Morgantown, WV, USA*

[3]*School of Public Health and Health Professions, Department of Social and Preventive Medicine, State University of New York at Buffalo, NY, USA*

[4]*Criminal Justice Program and Sleep and Performance Research Center, Washington State University, Spokane, WA, USA*

Abstract: Sleep quality has a large impact on daily performance and general health. Among the different methods of objectively measuring sleep quality, actigraphy continues to be very popular. It can take continuous activity measurements over several days in order to determine sleep-wake cycles and calculate sleep variables, including the three standard sleep variables used in determining sleep quality: total sleep time, sleep efficiency, and wake-after-sleep onset, in which study analyses use the mean of these variables. In the case of wake-after-sleep onset, which calculates the amount of time between falling asleep and waking up, the average does not characterize wake-after-sleep times as it does not account for the total number of awakenings or the frequencies of wake-after-sleep times. Instead, we recommend using the entire distribution of wake-after-sleep onset times, which we will call waiting time distribution, which better characterizes wake-after-sleep onset than the average value. Sleep quality for each participant was determined by their total sleep time and sleep efficiency. Non-parametric statistics were utilized to determine differences in waiting time distributions between participants with good and poor quality of sleep. Discriminant analysis was performed to show that a distribution of waiting times discriminates better between qualities of sleep than the average wake-after-sleep onset time does. Waiting time distributions were also fit to standard probability distributions for utility and ease of understanding. Analyses show that the waiting time distribution categorizes sleep qualities better than the average wake-after-sleep onset variable, as well as giving more information and better characterizations.

Keywords: Actigraphy, sleep quality, waiting times, distribution analysis.

INTRODUCTION

Various mathematical and statistical models have been developed to describe sleep-wake cycles, from sinusoidal models of circadian rhythm to polysomnographic images of sleep stages [1-3]. Actigraphy also continues to be a popular method of analyzing, as well as determining, sleep-wake patterns. Not only has it been shown to correspond well with polysomnography [4] for sleep percentage and sleep latency, but also it allows researchers to collect data for many consecutive days, for both sleep and wake shifts.

Actigraphy generates a large set of sleep related variables, which can be used to categorize sleep quality. Clinicians typically look at three of the actigraphy derived variables: total sleep time, sleep efficiency (the amount of time spent sleeping while in bed), and wake-after-sleep onset [5-7].

Those three variables are generally reported with their averages, from data that are collected over many days. For total sleep time and sleep efficiency, the average is a suitable summery statistic, as both of these variables are usually normally distributed. However, wake-after-sleep onset is characterized by multiple awakenings each night, with varying lengths of sleep before awakening, making it very asymmetric. The skewness of the distribution also depends on the length of wake-to-sleep times, making a percentile, such as the median, a better summary statistic. This set of awakenings is more appropriately considered as a waiting time distribution (WTD).

Another benefit of using the entire distribution of wake-after-sleep onset times is that the distribution itself can be fit to standard probability distributions. This gives researchers the ability to use distributional parameters and probability theory, which can aid in predictive modeling, and even in survival analysis, as the calculations for these WTDs can also be viewed as a time to failure problem, where a time of unbroken sleep fails when the participant wakes up.

We plan to show that using the entire WTD, as opposed to the average wake-after-sleep onset time, can better differentiate between good and poor quality sleep. Parameters derived from the WTDs would then also aid in detecting differences between groups where sleep problems are common, such as shift workers and emergency response personnel.

*Address correspondence to this author at the Biostatistics and Epidemiology Branch, Health Effects Laboratory Division, National Institute for Occupational Safety and Health, Centers for Disease Control and Prevention, 1095 Willowdale Rd, MS 4050, Morgantown, WV 26505, USA; Tel: 304-285-6299; Fax: 304-285-6112; E-mail: cto8@cdc.gov

1874-6209/08 2008 Bentham Science Publishers Ltd.

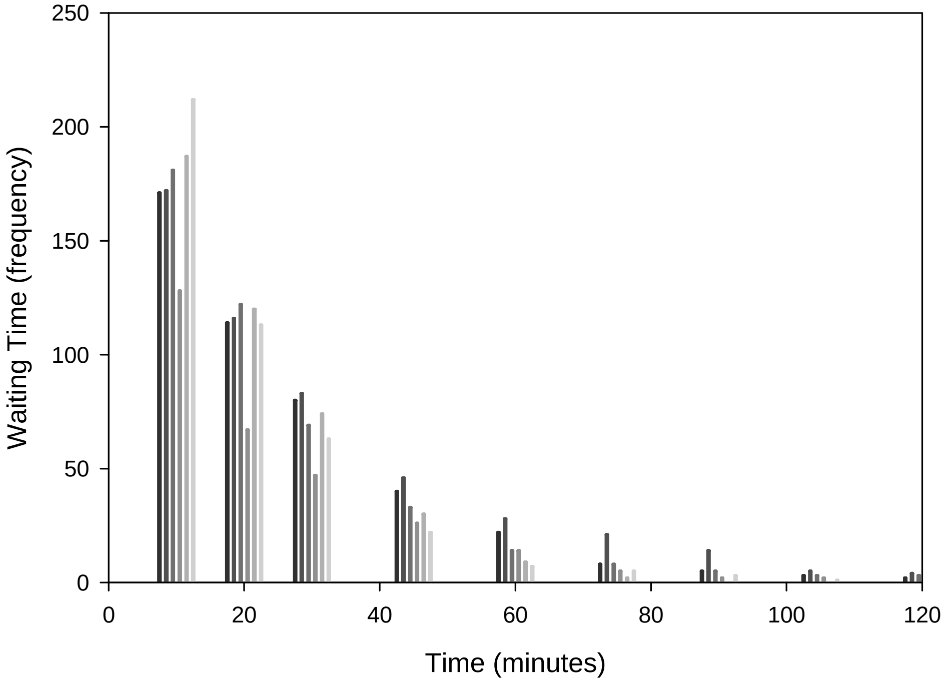

Fig. (1). Waiting time distributions for six (randomly chosen) participants, showing up to two hours of data, in time bins of 15, 20, 30, 45, 60, 75, 90, 105, and 120 minutes.

METHODS

In a study of health outcomes associated with stress among police officers one of the research questions under study is how stress affects sleep quality, as quality of sleep has been shown to affect health and overall performance; moreover, poor sleep can have negative impact on mental and physical characteristics as well as on social factors such as shift work [8]. To study this effect, officers were asked to wear an accelerometer in order to record their movement, which allows us to determine the quantity and quality of their sleep. This study uses data from the first 388 of an anticipated 500 officers from the Buffalo Police Department, NY. Data from the remaining 112 officers have not yet been collected. Subjects with corrupted data points or who were non-compliant were excluded from analyses [9], in order to keep the outcomes from being biased in either direction due to improper data, giving a final data set of 224 participants. A general description of the study design, methods, and participant characteristics for a pilot study of these officers has been reported [10].

The Motionlogger Actigraphs were worn for 15 days and removed only for short periods of time in order to protect them from water damage (e.g., while bathing or swimming). All phases, testing, and reports of the study were approved by the State University of New York at Buffalo Internal Review Board and the National Institute for Occupational Safety and Health Human Subjects Review Board.

The data from the accelerometers were transferred into computer files using Action4 software (Ambulatory Monitoring, Inc) and sleep was scored using the Primary Integration Mode. The files were then exported into Excel and finally transformed into SAS data files.

The R statistical software package (The R Core Development team, the R Project for Statistical Computing, http://www.r-project.org) was used to generate the WTDs for each participant. SAS version 9.1 (SAS Institute, Cary, NC) was then used for all statistical analyses.

Distributions were calculated by finding the frequency of wake-after-sleep events, where each event had its length of time between falling asleep and waking up recorded. Time periods analyzed ranged from fifteen minutes to eight hours, for every fifteen minute interval. Fig. (**1**) shows an example of waiting time distributions for six participants, showing up to two hours of the distribution.

To compare how sleep quality can be classified between the WTD and the average wake-after-sleep onset variable, sleep quality was rated as good or poor according to each participant's total sleep time and sleep efficiency. Using the meta-analysis values from a published study [5], we used the authors' criteria of categorizing sleep quality; we considered an average sleep time per night of at least 6.5 hours to and an average nightly sleep efficiency of at least 85% to be indicative of good sleep quality. If at least one of these scores was rated good, then the participants had their overall sleep quality rated as good. If neither of those two variables was rated good, then the participant was rated as having poor overall sleep quality.

After calculating WTD and sleep quality for each participant, the 50th, 75th, and 90th percentiles of their time bins were determined. Since distributions were non-symmetric and heavily skewed to the right, indicating that the median and other percentile markers are better parameters to test than the mean time, non-parametric analyses were performed. Proc Npar1way was used to test for differences between sleep quality groups (good quality versus poor quality) at each percentile using the Wilcoxon and Kolmogorov-Smirnov tests.

As the WTD can be considered as a time-to-failure problem, we fit the participants' data to several of the standard time-to-failure distributions. The beta, exponential, gamma, and Weibull distributions were compared to the empirical distributions to determine how closely they matched, using Proc Capability. The Anderson-Darling and Kolmogorov-Smirnov goodness-of-fit tests were performed to determine if the empirical distribution truly fit the theoretical distribution.

Parameters for each tested distribution were then averaged together by sleep quality. Analysis of variance using Proc Mixed was performed to determine if the parameters were significantly different from each other across sleep quality groups.

To ensure that the entire WTD can classify sleep quality at least as well as the average wake-after-sleep onset time, discriminant analysis was performed to generate error rates in sleep quality classification, using the previous quality classifications derived from total sleep time and sleep efficiency.

To help researchers look for differences between groups with regard to the different lengths of wake-after-sleep onset times subjects have, survival functions, a function which finds the probability that an individual from a given group is still asleep at x minutes, were calculated to determine if there were differences between sleep qualities with respect to the probability of an awakening at different time periods, where the Weibull distribution's survival function was calculated with equation 1.

$$S(x) = \exp(-(\chi^{\gamma})),\ x \geq 0; \gamma > 0, \tag{1}$$

with shape parameter γ. S(x) itself falls in the interval (0, 1), where S(x) = 1 when x = 0 and S(x) $\rightarrow$ 0 as x $\rightarrow \infty$.

RESULTS

After analyzing each participant's WTD, it was noticed that there were very few occurrences at high time amounts, indicating that even participants with good sleep quality, would rarely sleep more than a few hours before some kind of awakening event. Because of this, the decision was made to only use time bins for up to two hours, as the frequency of occurrences where participants slept longer than that before awakening was much less than 1%.

Nonparametric analysis showed that there was no difference between the two groups at the 50th or 75th percentile, but there was a significant difference at the 90th percentile, up near the right-hand tail of the distributions. This indicates that, regardless of sleep quality, most participants have similar waiting times of sleep to wake, except that those with good quality sleep can have at least some longer sleep periods than those with poor sleep quality, indicating that the median, and thus the mean, wake-after-sleep onset time cannot accurately differentiate between sleep qualities. The mean, minimum, and maximum values for all three percentiles by sleep quality and the results of the non-parametric tests are given in Table **1**.

Anderson-Darling and Kolmogorov-Smirnov Goodness-of-Fit tests show that the Weibull distribution fit the empirical data best, with 98% of the participants' data fitting it. The other distributions that were tested, the gamma, exponential, and beta, performed much worse, with fitting rates of 54%, 3%, and 0% respectively. The Weibull distribution is frequently used as a way of determining time to failure. Waiting time distributions of sleep are also time to failure problems, as we are determining the length of time a participant sleeps before failing to sleep longer by waking up. When the empirical WTDs and the respective Weibull parameters are averaged by sleep quality group (good and poor), the empirical distributions still fit the Weibull. Both the shape and scale Weibull parameters are also significantly different when compared by sleep quality with the scale p-value = 0.0061 and the shape p-value < 0.001. Fig. (**2**) shows an example of the empirical distribution overlaid with the Weibull distribution for one of the (poor sleep quality) participants.

Discriminant analysis results indicate that the waiting time distribution's Weibull parameters of shape and scale can identify sleep quality based on total sleep time and sleep efficiency better than the average wake-after-sleep onset value. The WTD's Weibull parameters correctly identified over 70% of participants' sleep quality, with an error rate of 29.56%. The average wake-after-sleep onset value correctly identified approximately 65% of participants' sleep quality, with an error rate of 34.85%.

As the shape and scale Weibull parameters obtained from WTD analysis better classified sleep quality than the wake-after-sleep onset average, correlation analysis was performed to see if there was any correlation between the two parameters and the average onset variable. The Pearson correlation coefficient for the shape parameter is -0.26 with a p-value of < 0.0001, indicating it is significantly correlated with the average wake-after-sleep onset. The scale parameter has a correlation coefficient of -0.06 and is not significant with a p-value of 0.33.

Table 1. The Mean, Minimum, and Maximum Time Values for the 50th, 75th, and 90th Percentiles by Sleep Quality and p-Values from the Wilcoxon and the Kolmogorov-Smirnov (KS) Nonparametric Tests, Comparing Percentiles between Sleep Qualities

	Good Quality			Poor Quality			p- value	
Percentile	**Mean**	**Min**	**Max**	**Mean**	**Min**	**Max**	**Wilcoxon**	**KS**
50th	15	15	15	15	15	15	0.9999	0.9999
75th	15.36	15	30	15	15	15	0.2332	0.9999
90th	25.21	15	60	16.53	15	30	0.0001	0.0001

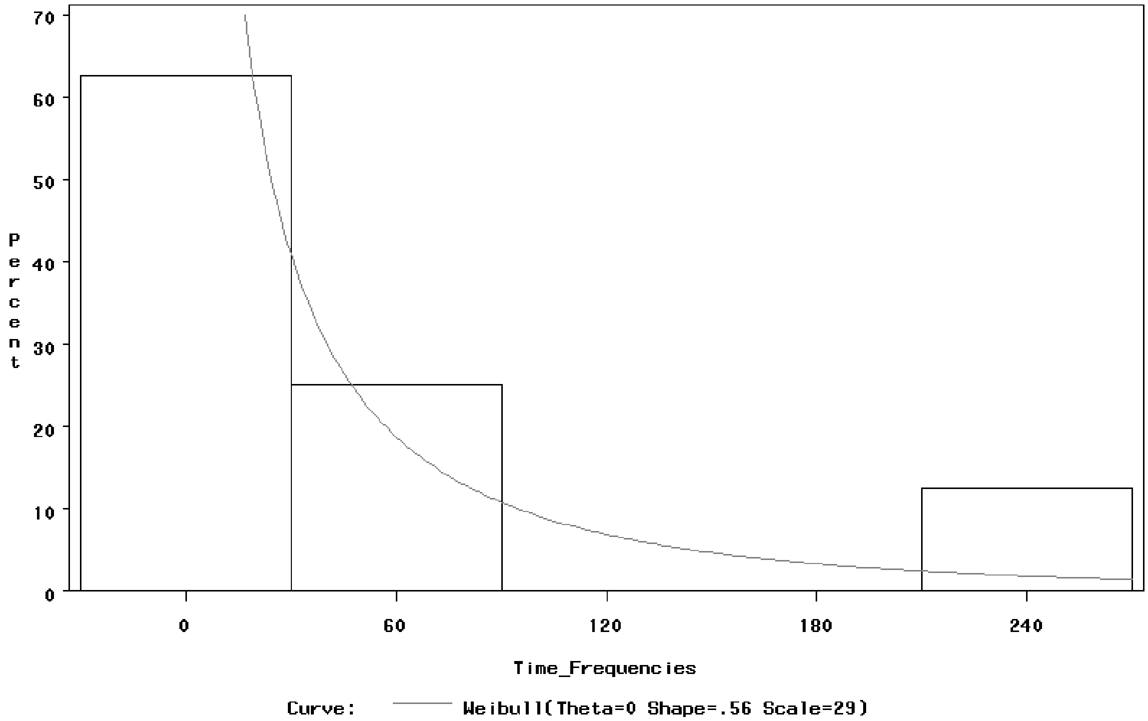

Fig. (2). Example of a Weibull distribution being fitted over a participant's empirical distribution, with the number of occurring frequencies of time bins.

Survival analysis shows that the probability of a participant's 'surviving,' that is, not waking up, is higher for those with good sleep quality. Participants who have poor sleep quality tend to awaken more often and/or sleep less time before they awake, regardless of the length of time of the awake period. Fig. (**3**) shows the survival curves for good and poor quality sleep, using the average Weibull parameters for each sleep quality. The figure shows that participants with poor quality sleep have a greater probability of waking up as more time passes.

DISCUSSION

Waiting time distributions offer an additional method of analyzing sleep data that have the ability to provide even more information on participant sleep patterns. WTD can be used in place of the average wake-after-sleep onset variable,

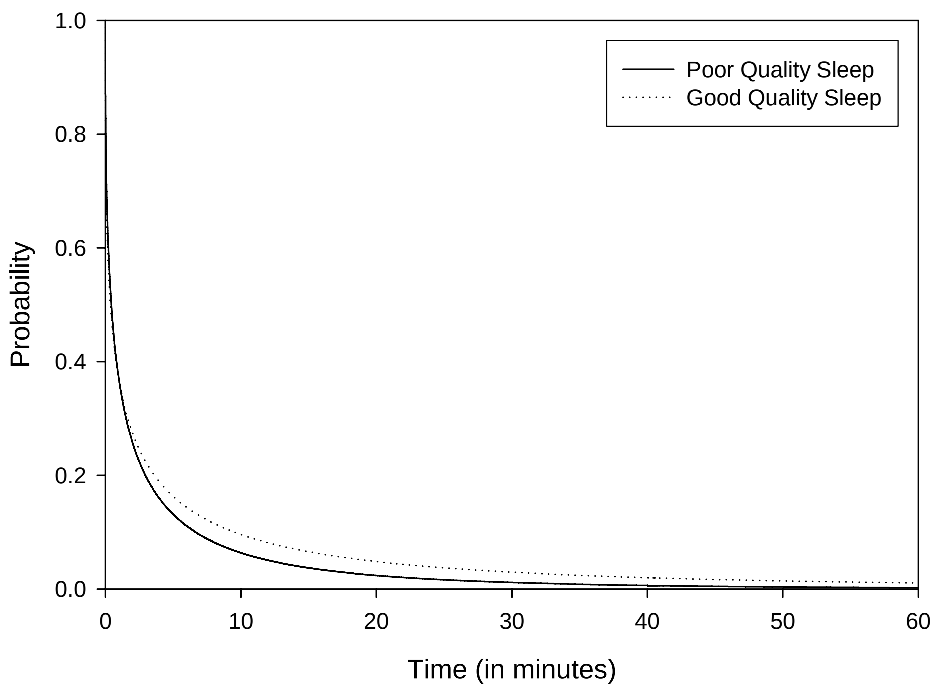

Fig. (3). Survival curve calculated with the Weibull survival function for good and poor sleep qualities, using the average Weibull parameters for each sleep quality group. The y-axis shows the probability of a participant staying asleep, at the time given on the x-axis, before an awakening event.

as the 90th percentile of the WTD is more capable of differentiating sleep quality than the average is. Although both of the error rates were fairly high in regards to classification, it must be remembered that sleep quality is typically categorized using three sleep variables, as opposed to the two used here, since the third variable is one of the variables involved in the comparison.

Additionally, the empirical waiting time distributions can be modeled with the Weibull distribution. Just as the 90th percentile, the estimated Weibull shape and scale parameters also categorize sleep quality better than the average wake-after-sleep onset value. This type of model can also be used for its distributional properties, which include survival analysis, moment generating functions, and probability calculations. The Weibull parameters may perform better in discriminating sleep quality than the average wake-after-sleep onset because of the additional parameter. The shape parameter for the Weibull distribution gives some of the same information as the average wake-after-sleep onset variable, but with the addition of the scale parameter, giving a more detailed model to help differentiate the sleep quality classifications.

Although analyzing sleep patterns with WTD takes more time than using the average wake-after-sleep onset value, it can add valuable, additional information for the researchers, as it not only categorizes sleep quality at least as well as the average wake-after-sleep onset time. Also, as the distribution of wake-after-sleep times is asymmetric, the sleep-to–wake onset average is an improper parameter to use; higher percentiles differentiate between sleep qualities much better, as the number of longer wake-after-sleep onset times is much greater in people with good sleep quality as opposed to those with poor sleep quality. Waiting time distributions also provide all of the benefits that the theory of probability distributions has to offer, including parameter estimation and survival time analysis. Compared to the average wake-after-sleep onset time, waiting time distributions give many more methods of analyzing participant characteristics, comparing sleep quality between groups, and modeling sleep.

DISCLAIMER

The findings and conclusions in this report are those of the author(s) and do not necessarily represent the views of the National Institute for Occupational Safety and Health.

REFERENCES

[1] Achermann P, Borbely A. Mathematical models of sleep regulation. Front Biosci 2003; 8: 683-693.

[2] Jean-Louis G, von Gizycki H, Zizi F, *et al*. Determination of sleep and wakefulness with actigraph data analysis software (ADAS). Sleep 1996; 19: 739-743.

[3] Fanfulla F, Patruno V, Bruschi C, Rampulla C. Obstructive sleep apnoea syndrome: is the "half-night polysomnography" an adequate method for evaluating sleep profile and respiratory events. Eur Respir J 1997; 10: 1725-1729.

[4] Cole R, Kripke D, Gruen W, Mullaney D, Gillin J. Automatic sleep/wake identification from wrist activity. Sleep 1992; 15: 461-469.

[5] Ohayon M, Carskadon M, Guilleminault C, Vitiello M. Meta-analysis of quantitative sleep parameters from childhood to old age in healthy individuals: developing normative sleep values across the human lifespan. Sleep 2004; 27: 1255-1273.

[6] Ancoli-Israel S, Cole R, Alessi C, Chambers M, Moorcroft W, Pollak C. The role of actigraphy in the study of sleep and circadian rhythms. Sleep 2005; 26: 342-392.

[7] Sadeh A, Acebo C. The role of actigraphy in sleep medicine. Sleep Med Rev 2002; 6: 113-124.

[8] Vila B. Impact of long work hours on police officers and the communities they serve. Am J Ind Med 2006; 49: 972-80.

[9] Slaven J, Andrew M, Violanti J, Burchfiel C, Vila B. A statistical test to determine quality of accelerometer data. Physiol Meas 2006; 27: 413-423.

[10] Violanti J, Burchfiel C, Miller D, *et al.* The Buffalo Cardio-Metabolic Occupational Police Stress (BCOPS) pilot study: methods and participant characteristics. Ann Epidemiol 2006; 16: 148-56.

Received: December 18, 2007 Revised: January 03, 2008 Accepted: January 10, 2008