Probabilistic sleep architecture models in patients with and without sleep apnea

MATT T. BIANCHI1,2, NATHANIEL A. EISEMAN1, SYDNEY S. CASH1, JOSEPH MIETUS3, CHUNG-KANG PENG3 and ROBERT J. THOMAS4

1Neurology Department, Massachusetts General Hospital, Boston, MA, USA, 2Division of Sleep Medicine, Harvard Medical School, Boston, MA, USA, 3Division of Interdisciplinary Medicine and Biotechnology, Beth Israel Deaconess Medical Center, Boston, MA, USA and 4Division of Pulmonary, Critical Care and Sleep Medicine, Department of Medicine, Beth Israel Deaconess Medical Center, Boston, MA, USA

Accepted in revised form 1 June 2011; received 25 October 2010

SUMMARY Sleep fragmentation of any cause is disruptive to the rejuvenating value of sleep. However, methods to quantify sleep architecture remain limited. We have previously shown that human sleep–wake stage distributions exhibit multi-exponential dynamics, which are fragmented by obstructive sleep apnea (OSA), suggesting that Markov models may be a useful method to quantify architecture in health and disease. Sleep stage data were obtained from two subsets of the Sleep Heart Health Study database: control subjects with no medications, no OSA, no medical co-morbidities and no sleepiness (n = 374); and subjects with severe OSA (n = 338). Sleep architecture was simplified into three stages: wake after sleep onset (WASO); non-rapid eye movement (NREM) sleep; and rapid eye movement (REM) sleep. The connectivity and transition rates among eight ‘generator’ states of a first-order continuous-time Markov model were inferred from the observed (‘phenotypic’) distributions: three exponentials each of NREM sleep and WASO; and two exponentials of REM sleep. Ultradian REM cycling was accomplished by imposing time-variation to REM state entry rates. Fragmentation in subjects with severe OSA involved faster transition probabilities as well as additional state transition paths within the model. The Markov models exhibit two important features of human sleep architecture: multi-exponential stage dynamics (accounting for observed bout distributions); and probabilistic transitions (an inherent source of variability). In addition, the model quantifies the fragmentation associated with severe OSA. Markov sleep models may prove important for quantifying sleep disruption to provide objective metrics to correlate with endpoints ranging from sleepiness to cardiovascular morbidity.

KEYWORDS adult, biological rhythms, computational model, human, sleep disorders

INTRODUCTION

Quantification of human sleep–wake transitions may involve multiple measurements and time scales of resolution (Thomas, 2006). For example, wrist actigraphy can measure dichotomous (sleep versus wake) staging over days or weeks (Ancoli-Israel et al., 2003). The gold standard for measuring sleep architecture is laboratory polysomnography (PSG), which reports the five sleep stages defined by the American Academy of Sleep Medicine (AASM): wake; rapid eye movement (REM) sleep; and three non-REM (NREM) substages (Silber et al., 2007). Certain stage transition patterns have long been recognized, such as the approximately 90-min ultradian REM cycle and the tendency of slow-wave sleep (stage N3) to occur early in the night (Fuller et al., 2006). However, these patterns may not be easily recognized in a single night of PSG monitoring, the typical clinical situation. For example, fragmentary transitions to light sleep or wake may be seen...
throughout the night, even without a sleep disorder such as sleep apnea. Also, the timing of REM sleep blocks may be variable, even in the absence of confounding medications. Ideally, a model of sleep architecture should capture the accepted ‘average’ patterns, while allowing night-to-night and/or inter-individual variation in architecture. Quantifying architecture may be of clinical value, as several groups have shown that routine summary metrics (such as sleep stage percentages and sleep efficiency) fail to identify fragmentation associated with apnea or other disorders (Bianchi et al., 2010; Klerman et al., 2004; Norman et al., 2006; Swihart et al., 2008).

The main conceptual model of sleep transitions draws analogy to a ‘flip-flop’ switch controlling transitions between wake and sleep, and between REM and NREM sleep (Fuller et al., 2007; Lu et al., 2006). Several groups have taken more quantitative approaches based on survival analysis or considering stage transition probabilities (Chervin et al., 2009; Klerman et al., 2004; Yassouridis et al., 1999). Markov modeling of sleep dates back as early as 1973 (Kemp and Kamphuisen, 1986; Kim et al., 2009; Yang and Hursch, 1973). One of the limitations of the transition probability approach is that it does not consider the distribution of stage duration: if the distribution of observed bouts of, for example, REM sleep has multiple components, then simplifying REM sleep as a single state may overlook potentially useful information. Understanding that an observed sleep stage may have more than one ‘generator’ state is the foundation of a type of Markov modeling routinely performed in other fields that extracts model structure directly from the empirical distributions.

We previously reported that sleep architecture in subjects of the Sleep Hearth Health Study (SHHS) demonstrated multi-exponential wake, REM sleep and NREM sleep bout distributions (Bianchi et al., 2010), suggesting more complex dynamics than previously reported (Blumberg et al., 2005; Lo et al., 2004). Our goals here are to build a realistic model of sleep architecture that fulfills four criteria: probabilistic transitions; multi-exponential dynamics; a approximately 90-min REM cycle; and increased fragmentation with sleep apnea.

**MATERIALS AND METHODS**

**Database cohorts**

This study involved use of the SHHS, a large database of home-based PSG (Quan et al., 1997). The study was designed to include Approximately 6000 adults who were at least 40 years old, each of which underwent home PSG. The study recruited from feeder studies (in decreasing order of number of participants): Atherosclerosis Risk in Communities Study; Cardiovascular Health Study; Framingham Heart Study; New York Hypertension Cohorts; Tucson Epidemiologic Study of Airways Obstructive Diseases; the Health and Environment Study; and Strong Heart Study. Category IV Institutional Review Board (BIDMC) approved use of this data, which is anonymous and thus we do not require additional consent [all subjects provided written consent to participate in the SHHS, conforming with the Code of Ethics of the World Medical Association (Declaration of Helsinki)]. The pre-defined groups within this dataset included controls [defined as apnea–hypopnea index (AHI) < 5, ESS < 10], no medications and no cardiovascular co-morbidities, and subjects with severe obstructive sleep apnea (OSA; AHI > 30). We did not further control for co-morbidities, which were more common in the severe apnea group (Bianchi et al., 2010). In addition, the racial composition was mainly Caucasian, with 85.4% in the control group and 74.3% in the severe apnea group.

The bout durations were scored in the SHHS database in units of 30-s epochs according to standard clinical criteria. The scoring reliability in the SHHS has been previously described (Whitney et al., 1998). Based on a small sample of PSGs, inter- and intra-scorder values showed epoch by epoch kappa coefficients > 0.80. For the current analysis, NREM bouts were simplified such that we considered only a single NREM stage (by collapsing all NREM sub-stages to a single stage) for simplicity. This was done before bout durations were counted, in other words, a four-epoch sequence, such as N1, N2, N2, N3 would be scored as a single bout of NREM, of duration four epochs. Thus, the sleep architecture consisted only of three stages: wake after sleep onset (WASO); REM sleep; and NREM sleep.

The control and sleep apnea groups used in this study have been described in our previous manuscript (see Table 1 in Bianchi et al., 2010 for details). The sleep apnea group (n = 338) differed from the control group (n = 374) in several clinical respects: younger age (63.7 ± 10.5 versus 68.2 ± 6.3 years); more males (70.7 versus 35.6%); more co-morbidities (HTN 54.4%, diabetes 5.9%, heart attack history 7.1%, versus 0% in controls); and higher sleepiness scores. PSG sleep stage percentages showed little or no differences; PSG sleep stage percentages showed little or no differences; however, there was a greater number of arousals (36.4 versus 17.3 per hour) and greater AHI (47.4 versus 1.8 per hour).

The hypnograms in Fig. 1 were chosen from the control group, with the requirement that their total sleep time was at least 6 h, and their efficiency was > 70%. Sleep efficiency is the percentage of the recorded night after sleep onset that was scored as any sleep stage. Note that this is the only figure in which all five AASM -criteria staging are shown (that is, all three NREM stages are shown). However, in the remaining

| Table 1 Exponential fitting parameters: pair-wise fitting in control subjects |
|-----------------------------|-------------------------|-------------------------|-----------------------------|
| τ-values                  | From stage | To stage | τ-values |
| 0.6, 4                    | WASO       | NREM     | 1.6, 8.7, 65 |
| 0.2, 1                    | WASO       | REM      | 10          |
| 1.7, 11                   | NREM       | REM      | 3, 21       |
| 1.6, 8, 106               | NREM       | WASO     | 0.6, 3.8   |
| 5, 44                     | REM        | NREM     | 1.8, 12     |
| 16                        | REM        | WASO     | 0.8         |

NREM, non-rapid eye movement; REM, rapid eye movement; WASO, wake after sleep onset. Tau-values are in units of ‘epochs’.
Bout duration analysis and curve fitting

As described previously (Bianchi et al., 2010), we generated frequency–duration histograms from pooled bouts of each stage. Histograms were then subject to exponential fitting routines in Prism (GraphPad Software, LaJolla, CA, USA). Bin width was one epoch. The standard exponential decay function takes the general form $Y = Y_0 e^{-kt} + C$, where $k$ is the rate constant of decay (forced positive), $t$ is the time (in units of epochs), and $Y_0$ is the $Y$-intercept at time zero when the plateau term, $C$, is forced to zero, as in our fitting. Multiple exponential functions were sometimes required, and involved the linear sum of $i$ components defined by $\frac{1}{C_0} Y_0 e^{-kiX} / C_i$; the $Y_0$ of which corresponds to the intercept of each component at time zero. Note that, although the value of $Y_0$ gives some indication of the contribution of each exponential term, it does not reflect the number of bouts associated with that time constant or the area under the curve of that exponential component.

Deciding the optimal number of exponentials was accomplished as follows. Sequential one-, two-, three- or four-exponential fits were compared pair-wise with a non-linear sum-of-squares $F$-test for nested models (such as exponential functions with varying number of terms) using built-in Prism routines, until the additional component no longer significantly improved the fit by $F$-test criteria. In other words, for all cases, the distribution was first fit by comparing one- versus two-exponential functions, and only if the two-exponential function was significantly improved by the $F$-test standard was two- versus three-exponential comparison made, etc. For example, for the NREM bouts ending by a transition to REM sleep, the two- versus three-exponential fitting showed an $F$-value of 94.11, which was significant to $P < 0.0001$, in favor of the three-exponential fit (but the four-exponential fit did not converge, and thus the three-exponential was favored). For REM bouts that ended by a transition to NREM, the $F$-value from one- to two-exponential fit was 327.8 ($P < 0.0001$), but comparing two- versus three-exponential fit yielded a $F$-value of 2.3 ($P > 0.1$).

Simulations

The simulated transitions in Fig. 2 were generated using ion channel software called QUB (http://www.qub.buffalo.edu).
The models of control and OSA subjects from the SHHS database were simulated using Mathematica (Wolfram Research, Champaign, IL, USA). The continuous-time Markov model specifies rate constants such that the time spent in any given state is exponentially distributed, with respect to the reciprocal sum of the possible exit transition rates. The fitting parameters described above were used to define a transition matrix, such as that shown in Fig. 3. Transitions to REM states (from wake or NREM sleep states) were adjusted with a sinusoid-like function with period of 90 min and a mean value of 1. Specifically, the function was the sum of four Gaussians with means every 180 epochs and standard deviations of 40 epochs. This was normalized so as to guarantee an average value of 1. These values were chosen to approximate the classic 90-min NREM–REM cycle, with approximately 20% variance (which was arbitrarily chosen). In the dataset, the number of cycles was variable across subjects, as seen in Figs 1 and 4e, and thus for simplicity we approximated the number at four cycles (corresponding to approximately 6–7 h sleep). The duration of time spent in any given state was based on a random number selected from an exponential distribution with a parameter equal to the sum of the rate constants out of the given state. Simulated transitions were ‘scored’ into 30-s-long units of time, rounding down lengths of under 15 s. These values were chosen to follow the clinical standard of 30-s window duration, and states are scored in each epoch according to the features present in the majority (> 15 s) of each epoch’s window. Fitting of exponentials, and displaying the data, was performed with Prism as above.

RESULTS

Variability in human sleep architecture

Single-night PSG data often demonstrate greater variability and number of transitions than the classical approximately 90-min NREM–REM cycles. As shown in Fig. 1a, variable transitions among sleep stages, as well as between sleep and wakefulness, occur even in the absence of pathology such as OSA. These hypnograms are from seven control subjects of the SHHS database, scored according to AASM criteria. The variability in architecture is visually evident, as is the degree of fragmentation based on transitions among sleep stages as well as between sleep and wakefulness. Such variability between subjects may be due to heritable factors, but within- and between-subject variability may have roots in multiple factors, such as irregularities in sleep–wake routine, medications, daytime exposures (such as exercise, caffeine, alcohol), or even tolerance of the sleep recording system itself. There is also likely some degree of inherent variability in addition to these factors; that is, transitions appear probabilistic.

Within this seemingly variable architecture, the well-known approximately 90-min REM sleep cycling can be seen on average. Fig. 1b shows the counts of REM sleep epoch-by-epoch over the night, based on the seven hypnograms in Fig. 1a, which are aligned at sleep onset. The occurrence of waking epochs (WASO), in contrast, is more broadly distributed throughout the night (Fig. 1c).
We recently demonstrated that the distribution of bout durations for WASO, REM sleep and NREM sleep followed multi-exponential dynamics (Bianchi et al., 2010). Building a Markov model of sleep architecture requires mapping observed sleep stages to states within the model. In order to understand this mapping, it is useful to distinguish between the sleep stages as phenotypes (which are defined by visual scoring of PSG records), and the Markov model states as “generators” (which are inferred by statistical analysis of the observed stage duration distributions). Simple simulations show that an observed phenotypic stage may require multiple generator states in the Markov model.

Consider a two-state system with rate constants governing transitions between one sleep state (S1) and one wake state (W1). Repeated 8-h runs of this simple two-state system yield mono-exponential distributions of sleep and wake bout durations (Fig. 2a). The distribution of bout durations is assessed using frequency–duration histograms, which plot the relative number of events occurring for any given bout duration (collected in bins on the X-axis).

When one starts with a given Markov model, the number of states dictates the number of observed exponentials in the histograms of simulated bout durations. Experimentally, however, we start with observed sleep stages, and must infer the underlying model. In the case of human sleep, we pool data across subjects to generate bout duration histograms, given the need for many observations to ensure accurate fitting, as described previously (Bianchi et al., 2010). Observed bout durations in clinical PSG data have been reported to fit with one (Lo et al., 2004) or more (Bianchi et al., 2010) exponential functions in the literature. Each fitted exponential function of a given phenotypic stage implies a distinct generator state in the Markov model. Fig. 2b shows a three-state model, containing a second generator state, S2, which yields a phenotypic sleep stage identical to that yielded by the S1 state. In other words, to the observer, sleep–wake patterns produced by this Markov model have only two phenotypic stages (sleep and wake), similar to the model in Fig. 2a. However, there are now two exponential components to the sleep bout distribution, in contrast to the model in Fig. 2a.

In a Markov model, the mean time spent in any state (τ) is defined by the reciprocal of the sum of the exit rate constants. The wake bout durations for both models are mono-exponential, with τ = 10 min. For the two-state model, this 10-min value corresponds to the reciprocal of the single exit rate constant from state W1, which is 0.1 per minute. For the three-state model, this corresponds to the reciprocal of the sum of two exit rates (0.02 and 0.08), which is also 0.1 per minute. The distribution of sleep bout durations, in contrast, is quite different between the two- and three-state models. The second sleep state, S2, is much less stable than the original, with a fivefold greater exit rate constant compared with S1. Also, S2 is the more likely exit transition from the wake state, W1 (four times more likely than a transition from W1 to S1). Accordingly, the best fit of the sleep bout distribution histogram required a second exponential term (Fig. 2b). Thus, the observed phenotypic state, sleep, is dictated by two underlying generator states in this model.

**Markov model of sleep architecture: control subjects**

The number of required states to model sleep–wake distributions in control subjects was inferred from the number of exponential components required to fit the distribution of pooled NREM and WASO bouts (three each) and REM bouts (two), drawn from our prior analysis (Bianchi et al., 2010). In other words, we defined three generator states for the observed NREM sleep and WASO stages, and two generator states for REM sleep (Fig. 3a), for a total of eight

![Figure 3. Eight-state Markov model of sleep–wake activity in control subjects. The connectivity was inferred by adjacent-stage analysis (see text; see Table 1 for fitting data). Note that all transitions are reversible except W2 to R2. (a) Matrix of rate constants for each transition. Gray shading indicates transitions that were either not considered (transitions within a phenotypic class, such as NR1 to NR2), or not observed in the adjacent-stage analysis (such as NR3 to R1).](image)
states. Note that the three NREM generators do not correspond to AASM-defined substages N1–N3 (each of which, when analysed separately, is actually multi-exponential; Bianchi et al., 2010). The following fitted exponential time constants (in units of epochs) were used to constrain the sum of the exit rate constants from each Markov generator state: WASO (0.6, 3.1, 16); NREM (1.7, 7.8, 44); and REM (3.8, 19). In other words, each time constant requires a state, and determines the mean time spent in that state (and thus restricts the sum of the possible exit rate constants).

The connectivity of these eight states was determined by exponential fitting of subsets of the pooled bouts, conditioned on the adjacent stage. There are six possible transition pairs.
(e.g., an observed NREM bout can terminate with a transition to WASO or to REM sleep). In each pair, we fitted the distribution of the ‘from’ stage and the ‘to’ stage (Table 1). For example, the subset of REM bouts that terminated by a transition to NREM was fitted with two exponentials (as were all pooled REM bouts), while the subset of NREM bouts that followed the REM bouts was fitted with two exponentials (one fewer than the three exponentials required for the complete NREM pooled data). From this fitting strategy, we infer that both R1 and R2 are connected with N1 and N2 (but not with N3), as the two NREM exponentials were similar to the fast and intermediate values when all NREM bouts were considered. This strategy was used for all six possible stage pairs to define the model connectivity of Fig. 3a. Note that there was only one asymmetry uncovered in this analysis: W2 can transition to R2, but the reverse transition could not be justified by the pair-wise fitting. Also, note that the third (longest) WASO component was never resolved during the fitting routines of this pair-wise analysis; we presume this was related to the small contribution of this component in the larger, pooled dataset, and thus the smaller, pair-wise analysis was underpowered to detect it. For simplicity, we placed this state, W3, in tandem with W2.

The next step involved assigning rate constants. As stated above, the mean duration of each state is defined by the reciprocal sum of the exit rate constants. Each τ-value obtained from fitting the pooled stage duration histograms was used to constrain the sum of the exit rate constants for its generator state (Bianchi et al., 2010). For example, state NR2 corresponds to the intermediate τ = 7.8 (from the three-exponential fit of pooled NREM stage durations), and the four exit rate constants must add to its reciprocal, 0.13. The constraints used for each state are listed in the figure legend of Fig. 3. Within the constraints set by the sum of the exit rates, the relative division of this ‘exit opportunity’ among the exit paths was assigned manually to approximate the relative proportion of WASO, NREM and REM sleep stages in the SHHS data (optimizing so many free parameters is beyond the scope of the current modeling). The final set of parameters used in the Markov model is shown in Fig. 3b.

Because transitions in this type of Markov model are time independent and memory-less, there is no constraint on the timing of REM blocks, allowing REM to spread throughout a simulated night of sleep (data not shown). To account for the approximately 90-min ultradian rhythm of REM sleep, we imposed a time-variation adjustment factor on the rate constants for entry into R1 and R2 (see Materials and methods). The mean rate adjustment factor was 1, such that the overall bout durations would not be compromised by this imposed time dependence.

Figure 4 shows that this model produces realistic hypnograms that exhibit several important features of typical human sleep architecture. First, there is variability in the sleep architecture between simulated nights, which is an inherent property of the Markov model's probabilistic state transitions (Fig. 4a). Second, the ultradian REM rhythm can be seen by counting REM epochs over each minute of the night (Fig. 4b). Third, fragmentary wake bouts are spread fairly evenly throughout the night (Fig. 4c), consistent with the SHHS data (compare with Fig. 1). Finally, the distribution of bout durations for each stage follows the predicted patterns of three exponentials for NREM sleep and WASO, and two exponentials for REM sleep (Fig. 5). This is expected because the parameters of the model were derived from these exponential values.

Markov model of sleep architecture: severe OSA

We next extended the model to account for the architecture fragmentation observed in subjects with severe OSA, again constraining parameters based on our prior analysis of SHHS subjects (Bianchi et al., 2010). The number of generator states required to fit the NREM and REM sleep bout duration histograms was unchanged in the severe OSA group (eight); however, the τ-values were faster. Although the third, longest WASO component was not observed in pooled analysis of the OSA group, we kept this state as defined in the control model, because interestingly it could be identified in the fitting of pair-wise bout durations. The fitting values obtained in our prior study were used to constrain the exit rate constants from each state (in units of epochs): WASO (0.5, 3.7); NREM (1.0, 4.8, 33); and REM (1.9, 16). Pair-wise fitting was again carried out, which defined five additional points of connectivity: R1 and R2 to NR3 (and vice versa); and R2 to W2 (see Table 2). Thus, the OSA model can exhibit fragmentation based on two quantitative differences compared with the control model: additional transitions are available, and several of the constraining exit rate constant values are smaller. The final transition rate constants are shown in Fig. 6b. As with the control model, we again applied the time-varying feature to the R1 and R2 entry rate constants to impose a approximately 90-min REM cycle (in the case of OSA subjects, this involved the two additional entry rate constants from N3).

This model of OSA sleep architecture yields hypnograms that recapitulate important features of the SHHS data. Fig. 7 shows simulated hypnograms from the OSA model, along with the counts of REM and wake epochs across multiple nights. For comparison, hypnograms are shown from several subjects in the SHHS with severe OSA.

DISCUSSION

This study demonstrates that a continuous-time Markov model can capture several salient features of human sleep architecture: (i) probabilistic variability in bout duration; (ii) multi-exponential bout length distributions; and (iii) REM ultradian cycling. Furthermore, the fragmentation associated with severe OSA was quantified by two differences in the model: existing transition rates were faster; and additional
state transitions occurred that were not evident in the control cohorts. The results suggest that Markov models can be used to quantify sleep fragmentation in health and disrupted sleep. Improved characterization of disrupted sleep architecture may aid in the clinical prediction of endpoints ranging from sleepiness to cardiovascular co-morbidities.

**Quantifying sleep fragmentation**

Standard PSG metrics such as sleep efficiency and stage percentages are insensitive to sleep fragmentation, which can be better characterized by bout duration statistics (Bianchi et al., 2010; Norman et al., 2006; Swihart et al., 2008). For example, a sleep efficiency of, for example, 80% could be attributed to prolonged sleep latency, a block of wake in the middle of the night, early awakening, or many brief awakenings throughout the night – and each of these may have distinct etiologies and clinical impact. Other metrics that have been proposed to reflect fragmentation include survival analysis (a measure of sleep bout duration; Norman et al., 2006; Punjabi et al., 1999; Swihart et al., 2008), autonomic arousals (Lombardi et al., 2008), and proportion in stage N1 (Wesenberg et al., 1999). The arousal index is also a measure of fragmentation at a finer time-scale than standard sleep stage scoring (Krystal et al., 2002; Thomas, 2003). It may be that arousals of various duration or ‘intensity’ have graded effects on fragmenting sleep. Thus, summary metrics such as the routinely used arousal index (arousals per hour of sleep) or some composite metric that weights a spectrum of arousals may provide an important metric of fragmentation that can be performed on small groups or even individuals, as an alternative to more complex state transition modeling.

---

**Figure 5.** Distribution of bout durations from the Markov model. Frequency–duration histograms (1-epoch bins) were constructed for bouts of wake after sleep onset (WASO) (a), non-rapid eye movement (NREM) sleep (b) and rapid eye movement (REM) sleep (c). The best mono-exponential fit (gray dashed line) is overlaid on the simulated data to illustrate the need for multi-exponential fitting of these distributions.

**Figure 6.** Markov model of SHHS subjects with severe OSA. (a) Eight generator states defined the sleep apnea Markov model, the connectivity was inferred by adjacent-state analysis (see text; see Table 2 for fitting data). The gray shading indicates transitions observed in the OSA group but not in the control group (compare with Fig. 3). (b) The matrix contains rate constants for each transition in the model. Gray shading indicates transitions that were either not considered (transitions within a class, such as NR1 to NR2), or not found in the adjacent-state analysis (such as NR3 to W3).
When correlations between clinical metrics and subjective or objective endpoints are small or non-significant, one potential reason for false negatives is heterogeneity permitted by coarse metrics such as sleep efficiency. Quantitative assessment of sleep architecture may thus prove useful for improving the currently controversial correlations between subjective sleepiness and PSG/MLST metrics (Chervin and Aldrich, 1999; Gottlieb et al., 1999). Sleep bout distributions as a measure of architecture have proved useful in characterizing subjects with pain and/or fatigue (Togo et al., 2008), and novel metrics of sleep stability based on cardiopulmonary coupling, which also capture fragmentation, are gaining increasing utility (Thomas et al., 2010; Yang et al., 2010). Given the diversity of causes of non-refreshing sleep (Ohayon, 2008), quantifying sleep architecture phenotypes may be of clinical benefit in domains of treatment and prognosis.

Models based on state transitions represent a useful class of tools in part because they can accommodate diverse definitions of state. For example, multi-exponential bout distributions are found in NREM sleep whether one considers each substage separately, or collapses them into a single composite stage (Bianchi et al., 2010). Rodent sleep bout distributions also

Figure 7. Simulated hypnograms from the Markov model of obstructive sleep apnea (OSA). (a) Randomly chosen simulated hypnograms from the severe OSA model, with wake (W), REM (R) and NREM (N) sleep stages. The time legend applies to (d) as well. (b) The number of observed REM sleep bouts plotted below the hypnograms shows the imposed probabilistic REM sleep ultradian rhythm. Counts of REM sleep are shown for n = 7 (solid line) and n = 30 (dashed line) simulated hypnograms. (c) Counts of wake across the night, for n = 7 (solid line) and n = 30 (dashed line) simulated patients were more evenly distributed. (d) Single-night hypnogram data are aligned from seven SHHS subjects with severe OSA. Staging, as in prior figures, is simplified to WASO, NREM and REM sleep. The time legend in (a) applies. (e) Counts of REM sleep are plotted, as in (b), for n = 7 (solid line) and n = 30 (dashed line) subjects. (f) Counts of WASO are plotted, as in (c), for n = 7 (solid line) and n = 30 (dashed line) subjects. REM, rapid eye movement; WASO, wake after sleep onset.
exhibit exponential behavior (Jojo et al., 2006; Lo et al., 2004), raising the possibility that experimental manipulations (such as lesions or mutations) could link states and/or transition probabilities to circuits or signaling systems.

One question raised by prior work involves the use of power law fitting to sleep–wake distributions, especially wake bouts (Blumberg et al., 2005; Diniz Behn et al., 2008; Lo et al., 2004). The underlying physiological implications of power law models (self-organizing, scale-free patterns) may be different than that of exponential (stochastic) models. We recently addressed the complexity of model fitting in this regard, by demonstrating that simulated multi-exponential distributions could be well-fit with a power law model, and vice versa (Chu-Shore et al., 2010). This work demonstrated that, even when using strict fitting criteria, power law and multi-exponential distributions are statistically similar. Furthermore, the parameters over which a multi-exponential process is most likely to mimic a power law are very near the time constants and proportions we reported for human sleep bouts. Given this confound in model fit, the related question of model choice is thus difficult, especially for non-nested models (such as power law and multi-exponential). Here, our modeling is taken to be a parsimonious interpretation of the data, in which Markov connectivity and transition probabilities can be constrained directly by empiric data. In contrast, any number of complex interacting systems of equations can yield power law dynamics. It is in this sense that we suggest a Markov model is more parsimonious, despite the increase in degrees of freedom (that is, a multi-exponential model, compared with the single power law exponent parameter).

Assumptions and limitations of the Markov sleep model

The Markov sleep model incorporated several simplifying assumptions. First, we collapsed the NREM sleep substages into a single stage, such that the model yields only three observed stages (wake, NREM sleep and REM sleep). Second, the time spent in any given state was dependent only upon its exit rates, and not directly influenced by any prior history of state transitions. Sleep architecture is not however a ‘memory-less’ process, as it is known that prior history (such as wake duration prior to sleep or REM sleep deprivation) can influence the occurrence of certain stages. Although the single-night PSG dataset used to generate and constrain the model lacked information relevant to assess circadian and homeostatic influences, data from forced desynchrony and deprivation studies may inform improvements in these domains, respectively. Third, we manually imposed a 90 ± 20-min cycle of time-variation in the rate constants governing entry into the two REM states. The model required this constraint to allow for the known approximately 90-min REM sleep rhythm.

It is worth noting that the large number of transitions required for histogram fitting currently limits the application of this method of modeling to pooled population data. We also limited our investigation to subjects with AHI < 5 versus AHI > 30, categories that are quite different in sleep physiology. OSA encompasses a spectrum of disease severity, and because averaging over many subjects was required to obtain differences in architecture dynamics (Bianchi et al., 2010), characterizing intermediate severity groups in terms of Markov modeling is rather limited. Despite this limitation, patients with severe sleep apnea are most likely to demonstrate medical and symptomatic consequences (cardiovascular risk, sleepiness, motor vehicle accidents), and thus represent a source of important clinico-pathological correlation. Given the nearly exclusive practice of single-night diagnostic PSG data available in routine clinical practice, individualized modeling must await advances in reliably recording sleep architecture longitudinally, likely in the home setting. The modeling performed here is based on pooled data from hundreds of patients; presumably within-patient measurement will exhibit less variability, and durations of weeks (rather than months) may suffice. Perhaps advancements in home monitoring will allow discovery of heterogeneities in Markov model parameters within groups that otherwise appear clinically homogenous. It is also likely that architecture analysis in populations chosen to be homogeneous across many clinical features will allow phenotyping to be performed with less need for pooling. However, to bring this type of modeling to the level of an individual, multiple nights of data are required simply because the number of transitions required for histogram fitting is far greater than that observed on a single night.

One of the strengths of this model resides in the extraction of its structure and parameters from empiric human data. The number of generator states was extracted from exponential fitting of bout durations, and the connectivity among these states was extracted by pair-wise stage analysis. This process, however, assumes that the time constants of the bout duration histograms can be mapped to model rate constants. When models contain ‘tandem’ generator states (transitions occurring between phenotypically identical states), the observed bout distributions are prolonged. This extraction of parameters from empiric data also allows the model to account for multi-exponential behavior, which is not modeled by traditional discrete time Markov (probability-containing matrix) approaches to sleep architecture (Kemp and Kamphuisen, 1986; Kim et al., 2009; Yang and Hursch, 1973).

CONCLUSION

In summary, our results extend the utility of Markov state transition modeling to the multi-exponential characteristics of human sleep–wake stage durations. An important benefit of this approach is that it captures both the probabilistic nature of transitions and the spectrum of stabilities (reflected in exponential time constants) observed in empiric PSG data. Further work is necessary to properly incorporate other fundamental aspects of sleep, such as the homeostatic sleep drive and circadian influences, which may preferentially affect certain transition probabilities. The ultimate goal of using Markov modeling to improve phenotyping of patients with sleep disorders will require repeated testing in individuals, most likely with advancements in home sleep monitors.

© 2011 European Sleep Research Society, J. Sleep Res.
CONFLICT OF INTEREST
Dr Thomas has consulted for Total Sleep Holdings; has a patent for CO2 adjunctive therapy for complex sleep apnea, ECG-based method to assess sleep stability and phenotype sleep apnea. Dr Thomas, Dr Peng and Mr Mietus are part co-inventors of the sleep spectrogram method (licensed by the BIDMC to Embla), and share patent rights and royalties. Mr Mietus has financial interests in DynaDx Corp. Dr Peng has financial interests in DynaDx Corp. Dr Bianchi has a patent pending on a novel home sleep monitoring device. Drs Cash and Westover have no conflicts to report. Mr Eiseman has no conflicts to report.

ACKNOWLEDGEMENTS
The authors thank Scott McKinney, and Drs Andrew Phillips, Elizabeth Klerman and Catherine Chu-Shore for valuable discussions and comments on the manuscript. Dr Bianchi receives funding from the Department of Neurology, Massachusetts General Hospital, and the Clinical Investigator Training Program: Harvard/MIT Health Sciences and Technology – Beth Israel Deaconess Medical Center, in collaboration with Pfizer, Inc. and Merck & Co. This funding source had no role in the design, interpretation or publication of this study. This paper represents the work of the authors, and not the Sleep Heart Health Study.

REFERENCES

© 2011 European Sleep Research Society, J. Sleep Res.