INTRODUCTION

THE CLINICAL DESCRIPTION OF RAPID EYE MOVEMENT BEHAVIOR DISORDER (REMBD)\(^1\) and the more recent report that a high proportion of patients with this syndrome develop idiopathic Parkinson's disease (PD)\(^2\) suggest functional connections between brain regions controlling waking movements (i.e., the basal ganglia) and brainstem centers well-recognized to control REM sleep atonia.\(^3\) Current concepts of the basal ganglia describe a family of segregated, yet parallel, cortical/subcortical pathways that allow integration of cortical and thalamic activity.\(^4,5\) We and others have shown that specific basal ganglia structures, such as the substantia nigra pars reticulata and the internal segment of the globus pallidus, have caudal connections (via inhibitory GABA pathways) to regions which modulate REM sleep atonia.\(^6,7\) The polysomnographic significance of these connections suggests that REM sleep identification in patients with neurodegenerative conditions primarily involving the motor system (such as idiopathic PD) may be difficult because of enhanced inhibition of neural systems necessary for the atonia of REM sleep. Alternatively, loss of REM sleep atonia might reflect degeneration of neural subgroups within the pedunculopontine (PPN) region and surround (e.g., locus coeruleus) and loss of their descending influences on spinal motor circuits.

As a part of our interest in understanding sleep mechanisms in such patients, we report here a study examining inter-rater reliability of sleep stages in PD. No studies have demonstrated that scorers can reliably agree upon REM sleep identification in such patients, but consensual agreement on what constitutes REM in PD or REMBD is essential to understand clinical features associated with that state. Additionally, reliable identification of arousal state may be particularly difficult in neurodegenerative diseases like PD because of factors such as frequent awakenings, poorly defined sleep landmarks (e.g., sleep spindles, K-complexes), and diffuse theta and delta frequencies dominating the electroencephalogram.\(^8\)

METHODS

Patients

We selected polysomnograms from 10 PD patients (X age=62.5, SD=9.5; range 45–79). All had longstanding PD (X disease duration=11.2 years). Hoehn-Yahr staging indicated moderate to advanced disease: stage 2 (n=2), stage 3 (n=6), 4 (n=2). One patient had additional evidence of periventricular small vessel disease by Magnetic Resonance Imaging. The 10 patients were selected from a database of nocturnal polysomnograms from 68 patients with Parkinson's disease. The selected recordings did not differ from the 58 remaining recordings on patient-related variables such as age, gender, disease duration, and Hoehn-Yahr staging. They also did not differ on polysomnographic measures such as Total Sleep Time, Sleep Efficiency, REM %, periodic leg movements (PLMS) index, PLMS with arousal index, and Respiratory

Abstract: Recently described functional connections between basal ganglia and brainstem circuits provide a neurobiologic basis for the absence of REM sleep atonia in Parkinson's disease (PD). However, identifying atypical REM sleep in PD may be problematic. Reliable sleep staging has never been demonstrated in such patients. In this study, 3 experienced scorers independently evaluated overnight polysomnograms from 10 (PD) patients. Results indicated good agreement for distinguishing REM from NREM sleep and waking. Reliable differentiation among NREM stages was more difficult to achieve. The results suggest that, despite suspension of REM sleep atonia accompanying PD, trained scorers can distinguish REM from wakefulness and NREM sleep.

Key words: Polysomnography; REM sleep; Atonia; Reliability; Parkinson's Disease

INTER-RATER RELIABILITY FOR IDENTIFICATION OF REM SLEEP IN PARKINSON'S DISEASE

Inter-rater Reliability for Identification of REM Sleep in Parkinson's Disease

Donald L. Bliwise PhD, Monique L. Williams RPSGT, Dainis Irbe MD, Farzaneh Pour Ansari MA, David B. Rye MD PhD

Sleep Disorders Center, Emory University Medical School, Wesley Woods Geriatric Hospital, Atlanta, Georgia 30329

SLEEP, Vol. 23, No. 5, 2000

Accepted for publication March 2000

Address correspondence to: Donald L. Bliwise PhD, Sleep Disorders Center, Emory University Medical School, Wesley Woods Geriatric Hospital, 1821 Clifton Road, N.E., Atlanta, Georgia 30329, Tel:404-728-4750, Fax:404-728-4756, E-mail:dbliwis@emory.edu.

Inter-rater Reliability—Bliwise et al
Disturbance Index. There was a suggestion that the individuals selected represented somewhat better sleepers when compared to individuals not selected for the reliability study; sleep latencies to 10 consecutive minutes of sleep were significantly shorter in those subjects selected (29.4 vs. 62.3 mins, t = 2.66, p < .02).

All 10 patients were treated with cabidopa/L-dopa (mean daily L-dopa dose = 860 mg) and all received other medications including: various serotonin re-uptake inhibitors (sertraline, trazodone, venlafaxine, paroxetine) (n=4), monoamine oxidase inhibitors (selegiline) (n=7), other dopamine agonists (bromocriptine, pergolide) (n=5), benzodiazepines (clonazepam, diazepam) (n=3), amantadine (n=3), and risperidone or clozapine (n=2). Sleep apnea was minimal in this group. Mean Respiratory Disturbance Index was 6.2 (SD=7.4); only one patient had an RDI in excess of 15 events per hour (24.3). Mean low oxygen saturation (%) was 88.0 (SD=5.1). Mean (SD) PLM Index was 20.8 (25.0); mean (SD) PLMS with arousal Index was 3.4 (6.2).

Polysomnographic Recordings

All recordings were made on Grass Model 78 polysomnographs using paper recordings run at 10 mm/sec using a standard montage for sleep disorders evaluations. The following channels were employed: electroencephalogram (C3-A2, Oz-A1), left and right monopolar electrooculograms, mentalis electromyogram, respiratory airflow (nasal/oral thermistors), respiratory effort (thoracic/abdominal piezoelectric bands), single channel electrocardiogram (modified lead II), and pulse oximetry. Additionally four channels of surface electromyogram, with electrodes placed over the left and right anterior tibialis and brachoradialis, were recorded. Sleep stages were scored in 30-second epochs.

Because polysomnograms in PD patients might be expected to contain REM sleep without atonia (i.e., muscle tone was not suspended) and because all three scorers were experienced in reviewing polysomnograms, we asked scorers to determine REM sleep using whatever information was available to them on the polysomnogram. The identical paper recording was used by all three scorers; no notes, underlining or remarks were allowed on any of the recordings.

Scorers

Three different individuals independently scored each recording manually by entering scores on a paper data entry form. All three scorers were experienced in technical aspects of polysomnography in dementia patients in general and in PD patients in particular. Scorers were told that the records scored were all derived from PD patients. Scorer A was an American Academy of Sleep Medicine certified sleep disorders specialist; Scorer B was an RPSGT with four years experience; Scorer C was a PGY III neurology resident who formerly worked as a polysomnographic technologist and had seven years experience. Scorers did not discuss stage scoring during the period in which the reliability study was performed. Scorers were instructed to score the polysomnograms by reference to the Rechtschaffen and Kales 9 scoring criteria but taking into account the findings of Schenck et al.1,2 that allow for the possibility of REM sleep with atonia in PD. Scorers were not told how to ascertain REM sleep in the absence of atonia, except to use the presence of rapid eye movements and desynchronized EEG as guides. Scorers were encouraged to use other channels (i.e., leg or arm EMG, heart rate or respiration channels) to aid in the determination of REM.

Data Analyses

Data were analyzed as percentage agreements among all possible scorer pairs (AB, BC, AC). Additionally, we computed Kappa coefficients to describe the strength of associations. Data were tabulated and analyzed by an individual not involved in the scoring process. The total number of epochs across all 10 cases was 8,395; epoch totals for each individual case varied between 675 and 1147.

RESULTS

As would be expected for the sleep of advanced PD patients, the amount of recorded sleep was relatively low in the 10 patients with a mean of 304 minutes of total sleep and 55 minutes of REM recorded (means of each patient's

<table>
<thead>
<tr>
<th>Scorer Pair</th>
<th>Overall Agreement (Wake, NREM, REM)</th>
<th>Error Category (%)</th>
<th>Overall Agreement (W/W, 1/1, 2/2, 3/3, 4/4, R/R)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AB</td>
<td>88.92 % Kappa .78 95% CI .77-.80</td>
<td>W/N 7.52</td>
<td>W/W .068 1/1 .288 2/2 .113 3/3 .644 4/4 1.50 R/R 4.08</td>
</tr>
<tr>
<td>BC</td>
<td>85.48 % Kappa .72 95% CI .70-.73</td>
<td>R/N 9.75</td>
<td>81.66 % Kappa .76 95% CI .75-.77</td>
</tr>
<tr>
<td>AC</td>
<td>86.72 % Kappa .76 95% CI .74-.77</td>
<td></td>
<td>67.62 % Kappa -.04 95% CI -.05-.02</td>
</tr>
</tbody>
</table>

Note: W = waking; N = NREM sleep; R = REM sleep

SLEEP, Vol. 23, No. 5, 2000

Inter-rater Reliability—Bliwise et al
Figure 1—Example of disagreement among scorers in differentiating REM-sleep from NREM-sleep. Scorers B and C scored epochs 174, 175 and 176 as NREM-sleep due to intrusion of slow frequency EEG signal during periods of phasic rapid eye movements. Out of phase eye movements in epoch 175 (asterisks) lead scorer A to score each epoch as REM-sleep.

Figure 2—Example of disagreement among scorers in differentiating REM-sleep from wake and NREM-sleep. Scorers A and B scored epochs 906-908 as REM-sleep based on identification of: a) consistently “slow” alpha frequencies; b) rapid, out of phase eye movements in epoch 906; and c) bursts of phasic EMG activity in the limbs typical of REM-sleep. Scorer C scored epoch 906 as NREM-sleep and 907 as wake reflecting a greater weight placed on a quantitative assessment of EMG activity and prolonged tonic elevation of EMG activity in the left arm in epoch 907. Note the presence of the 4-5 Hz tremor characteristic for parkinsonism in the left arm and leg at the beginning of epoch 908 (diamond lines). Conventional teaching that tremor disappears in sleep lead to the scoring of epoch 908 as wake by scorer C.
Figures 3 & 4—Six consecutive epochs demonstrating relative scoring agreement/disagreement in transition from consensually agreed upon waking to REM-sleep. Epoch scorers for scorers A, B, and C, respectively, were 528 (W, W, W); 529 (W, N, W); 530 (W, N, W); 531 (N, N, W); 532 (R, N, W) and 533 (R, R, W). Reliability in scoring REM-sleep onset is compromised by an elevation of chin EMG tone that persists through these six epochs.
and one patient in the latter pair received an SSRI. and two with stage 3 disease. One patient in the former pair stage 2 (including the patient with white matter disease) likely due to four patients, including two with Hoehn-Yahr

Disagreement among NREM stages appeared to be primarily due to four patients, including two with Hoehn-Yahr stages 2 and 3 and 1 and 4 virtually non-existent. The stage categories of disagreement showing the largest scorer pair discrepancy were stage 2 vs. 3 (A/B [2.72], B/C [9.66], A/C [9.77]) and stage 3 vs. 4 (A/B [0.0], B/C [2.20], A/C [3.31]).

Figures 1 to 4 provide examples of scored epochs showing typical agreements and disagreements among scorers across 12 different epochs. The continued presence of slow EEG activity throughout the recording presented formidable challenge in discerning REM sleep, and Figure 1 shows an example of REM sleep accompanied by such slow EEG activity. Although scorers relied upon EEG, EOG, and chin EMG in attempting to derive state, the presence of phasic muscle activity in all limbs (Figure 2) may have contributed to identification of REM sleep in some epochs in some patients. In other cases, the unmistakable progression of stages from wakefulness to REM sleep was made considerably more difficult by persistent, tonic EMG elevations associated with REM (Figures 3 and 4). Within these Figures (c.f. especially Figure 2), intermittent recurrent tremor may be detected and may have predisposed scorers to consider these epochs as wakefulness.

Examination of disagreement by case showed that scorer discordance was consistent across patients. The two patients showing the largest wake vs. REM scoring disagreements (average scorer pair disagreements of 2.91% and 4.34%) and the two patients showing the largest NREM versus REM scoring (average scorer pair disagreements of 5.44% and 6.08%) received either tricyclics or SSRIs. Although these data suggest that such medications may have contributed to unreliability of REM identification, it should be emphasized that these levels of disagreement approximated the levels of disagreement seen for the patients as a group (c.f. columns 5 and 6 of data in Table 1). Disagreement among NREM stages appeared to be primarily due to four patients, including two with Hoehn-Yahr stage 2 (including the patient with white matter disease) and two with stage 3 disease. One patient in the former pair and one patient in the latter pair received an SSRI.

DISCUSSION

The drop in muscle tone near or at the beginning of REM sleep has been recognized for many decades as a hallmark of REM sleep. Subsequent studies in normal subjects confirmed that the mentalis EMG falls, on average, about five minutes before the start of tonic REM sleep and remains low following the end of REM in about 50% of all REM periods. Although drops in tonic EMG levels have been considered critical in normals to identify REM, it may well be that phasic EMG activity, typically ignored in the scoring of REM, may become particularly important in identifying REM sleep in PD. Similarly, in the case of REMBD, a condition considered to represent disassociated components of arousal, and a potential prodrome of PD, a diagnosis is made typically on the basis of a patient’s history, their tendency for dream enactment behavior, and on the laboratory night, specific episodes of the latter, including dream recall. Because REMBD episodes may not always occur on the laboratory night, identification of this condition on the basis of characteristic polysomnographically assessed motor activation could be of critical importance not only for clinical purposes but in other contexts as well (e.g., medical/legal).

Altered definitions of REM sleep lead to critical re-evaluation of the neuroanatomic and neurochemical basis for phasic vs. tonic components of this state. The neural generators for both the tonic and phasic events of REM sleep are located in the PPN region situated in the dorsolateral pons. Pathological involvement of these neurons in neurodegenerative conditions may be modulated by the ventral forebrain, including basal ganglia output pathways, and could result in transient elevations in EMG tone and/or excessive phasic EMG activity. While some animal data suggest that cholinergic PPN neurons are responsible for REM-sleep-atonia and admixed glutamatergic neurons modulate phasic events, the precise cellular elements and their chemical signatures are not known. Therefore, it remains unclear whether the REM-sleep specific EMG elevations seen here are due to: a) increased bombardment of motor neurons with excitatory drive; b) the same excitatory drive impacting upon less hyperpolarized motor neurons (reflective of insufficient atonia generating pathways); or c) some combination of these two.

Somewhat surprising in our data was the less reliable identification of NREM sleep stages, particularly insofar as discrimination of stages 2, 3, and 4 are concerned. While to some extent this could reflect the arbitrary differentiation of these stages by percentage of epoch occupied by delta activity (and therefore the scorer’s ability to individually count delta waves), a number of other studies have suggested that trained human scorers achieve relatively good discrimination among these NREM stages. Therefore, a likely explanation of the disagreements among the three scorers here may be the diffuse slow activity apparent in
the EEGs of many of these patients, a problem common in many neurodegenerative diseases. To what extent “normal” delta and “pathological” delta can be discriminated by visual recognition or by higher order signal processing (coherence, non-linear dynamics) remains an open question. On the other hand, without a clear functional rationale for distinguishing these types of delta activity within sleep the value of achieving such a differentiation is unclear.

The data presented here suggest that sleep with patients in advanced neurodegenerative disease (specifically, PD) is best scored with a simple three-fold classification of waking, REM, and NREM, thereby collapsing all of the latter into a category of “indeterminate” NREM sleep, as was proposed originally by Reynolds and others. These data also suggest that further examination of specific components of sleep in PD patients (e.g., breathing disruptions, muscle activity) is probably best accomplished by limiting comparisons to REM vs. NREM sleep. Additionally, we would recommend that analysis of phasic muscle activity, recorded not only from the mentalis/submentalis, but also limb leads, be employed routinely in PD to assist in REM sleep identification.

ACKNOWLEDGEMENTS

Supported By: NS-35345, AG-10643

REFERENCES