REM Latency in Alzheimer’s Disease

D. L. Bliwise, J. Tinklenberg, J. A. Yesavage, H. Davies,
A. M. Pursley, D. E. Petta, L. Widrow, C. Guilleminault,
V. P. Zarcone, and W. C. Dement

Latency to the first episode of rapid eye movement sleep (REML) has been proposed as a potential biomarker for Alzheimer’s disease (AD). In this study, we compared REML values from 28 AD patients and 28 age- and sex-matched controls. We employed multiple definitions of REML and multiple cutoffs to classify patients and controls. Results indicated that the best REML definition and optimal cutoff criterion resulted in only 65% correct classifications. We discuss the longer REML in AD patients relative to controls in terms of both overall sleep disturbance and selective deterioration of the REM–cholinergic system. As REML may be relatively short in other forms of psychopathology (e.g., affective disorders), REML may still hold promise in the differential diagnosis of dementia and pseudodementia.

Introduction

The differential diagnosis of dementia and pseudodementia represents a formidable diagnostic problem in geriatric psychiatry (Wells 1979; Grunhaus et al. 1983). The value of existing biomarkers, such as the Dexamethasone Suppression Test (DST) for making this differential has been questioned (APA Task Force 1987). An alternative promising biomarker is the latency to the first episode of polysomnographically defined rapid eye movement sleep (REM latency, REML). Although some controversy exists as to the specificity of REML as a marker for psychopathology diagnosis (e.g., some depression subtypes, schizophrenia, schizoaffective disorder, panic disorder, borderline personality disorder, obsessive compulsive disorder [Reich et al. 1975; Insel et al. 1982; McNamara et al. 1984; Udde et al. 1984; Thase et al. 1986; Ganguli et al. 1987; Reynolds and Kupfer 1987; Zarcone et al. 1987]), the typical finding is that REML is decreased in patient groups relative to controls. That is, such patients enter REM sleep relatively early in the sleep period. Some have speculated that this could represent a cholinergic supersensitivity, particularly insofar as depressive illness is concerned (Gillin et al. 1982; Dube et al. 1985).

Comparatively less is known about REML in dementia. As REM sleep regulation may particularly depend on cholinergic systems (Gillin and Sitaram 1984) and cholinergic
deterioration is characteristic of Alzheimer's disease (AD), REM latency should be lengthened in AD (i.e., AD patients should enter REM relatively late in the sleep period). Existing data conflict on this issue, with some studies reporting longer REM latency in AD patients (Prinz et al. 1982b; Reynolds et al. 1983a), some studies reporting no differences from controls (Lowenstein, et al. 1982; Prinz et al. 1982a; Reynolds et al. 1985, 1986, 1987), and some studies reporting significant differences depending on the definition of REM latency used (Vitiello et al. 1984). One early study even reported a nonsignificant trend for shorter REM latency in chronic brain syndrome patients (Feinberg et al. 1967). More recent data suggested that REM latency may be prolonged in dementia relative to controls only after sleep deprivation (Reynolds et al. 1987).

In this study, we compared AD patients to nondemented aged control subjects. We employed multiple definitions of REM latency (Knowles et al. 1982; Reynolds et al. 1983b), as well as multiple cutoffs, to derive optimal sensitivities and specificities for classifying patients and controls (Reynolds et al. 1983b; Giles et al. 1986). As REM latency values decline with age (Kupfer et al. 1982a; Ulrich et al. 1980), even from decade to decade, we employed age-matched nondemented controls.

Methods

Subjects

AD patients were 28 ambulatory men (n = 18) and women (n = 10) mean age 67.6, SD 8.65 years who were living at home at the time of the study. All but one lived with a spouse or caregiver. Patients were recruited through the Alzheimer's disease support group at the Palo Alto Veterans Administration Hospital; however, two were originally seen as patients in the Sleep Disorders Center at Stanford Medical School. Following NINCDS criteria (McKann et al. 1984), 4 subjects carried a diagnosis of definite AD, 19 probable AD, and 5 possible AD. Mean duration of dementia was 5.0 ± 3.3 years. Mean Folstein Mini-Mental State Exam (Folstein et al. 1975) score (17.96 ± 6.87) and Hachinski rating scale (Rosen et al. 1980) score (1.58 ± 1.46) suggested mild dementia and relatively few signs of multiinfarct dementia, respectively. Patients with a history of chronic alcohol intake were excluded. Only 3 AD patients used psychotropic medication regularly (haloperidol, alprazolam, triazolam). The patients receiving the latter two medications were discontinued specifically for the study a week prior to the lab night. The patient receiving haloperidol had been discontinued by the family a month prior to the sleep lab study.

Controls were 28 age- (mean age 67.6, SD 9.22) and sex- (18 men; 10 women) matched individuals selected from a previously described sample of 83 elderly research volunteers without insomnia (Bliwise et al. 1987). These volunteers were originally recruited from advertisements for good sleeping senior citizens. The advertisements were placed in local papers and on bulletin boards in university buildings, senior citizen centers, and churches. Their health status varied widely, but all were independently living, in self-reported good health, had no history of cancer, and were free from acute illness at the time of recording. None of these control subjects regularly used psychotropic medication. Their mean Mini-Mental State Exam score was 27.6 (SD 2.2).

Procedures

Both patients and controls underwent one night of polysomnography following standard procedures (Rechtschaffen and Kales 1968); sleep stages were scored in 30-sec epochs.
Although many sleep laboratory studies of nondemented elderly subjects employ at least one adaptation night (Kales et al. 1967; Roth et al. 1972; Webb and Campbell 1979), we studied AD patients for only one night because of our expectation that many would not tolerate the procedures for multiple nights. In addition, at least some evidence suggests that sleep disturbance is not amenable to laboratory adaptation effects in such patients (Reynolds et al. 1983a; Allen et al. 1987). As REML could be affected by the time an individual goes to bed, relative to his or her customary bed time (i.e., REML could be phase-dependent (Schulz and Lund 1985; van den Hoofdakker and Beersma 1985)), both patients and controls went to bed with lights off within 30 min of their usual bedtime, as reported by caregiver (patients) or self-report (controls). AD patients’ lights off time averaged over an hour earlier than those of controls (9:42 versus 10:53 PM, t = 4.59, p < 0.0001). Most AD patients were able to tolerate the laboratory procedures. Only two patients became so severely agitated that we were forced to remove electrodes and suspend the laboratory procedures. No polysomnographic data were collected from these individuals. These individuals were grossly demented (Mini-Mental State scores of 9 and 5). One subsequently died from pneumonia 15 months after the sleep lab study was attempted; an autopsy confirmed the diagnosis of AD in this case. The other individual remains alive, but now requires full-time custodial care. Another patient among the 28 became agitated during the night and was given triazolam, as was customary for this patient. Polysomnography results for this patient were based on data recorded prior to drug administration.

REML was defined in six different ways, representing all possible combinations of 3 sleep-onset definitions (Stage 1, Stage 2, and 8 min of sleep in 10 min of recording) and 2 intervening wakefulness definitions (included/excluded). These definitions followed those described by Knowles et al. (1982) and Reynolds et al. (1983b).

Data analyses employed the Kolmogorov-Smirnov single-sample test to test the distribution of each variable for normality. Group comparisons were then made using matched-pair t-tests for normally distributed data or the Wilcoxon matched-paired signed-rank test for nonnormally distributed data.

Results

Table 1 shows polysomnographic sleep data in AD patients and controls. Although the groups did not differ in the total amount of sleep, AD patients showed evidence of substantially poorer quality sleep, as indicated by a lower mean sleep efficiency, lower percentages of Stages 3 and 4, and a higher percentage of Stage 1.

Percentage of REM did not differentiate AD patients from controls; however, several of the REML measures distinguished the two groups. Specifically, those REML definitions including less stringent sleep-onset criteria (first epoch of Stage 1 or Stage 2) and those definitions including wakefulness distinguished the groups, whereas other definitions did not, though all means suggested that long REMLs were associated with AD. Figure 1 shows the frequency distribution of the Stage 1/wakefulness-included definition of REML in AD patients and controls.

Table 2 addresses the potential utility for the Stage 1/wakefulness-included definition for REML by examining the sensitivity, specificity, and percentage correct classification of patients by varying the criterion in 15-min increments. As would be expected, sensitivities and specificities vary inversely as the classification criterion increases in length. In terms of overall correct classification, a time of 105 min was optimal, meaning that AD patients were likely to take 105 min or longer to enter REM at the beginning of the
Table 1. Comparison of Polysomnographic Sleep Measures in Alzheimer's Patients and Controls (n = 56)

<table>
<thead>
<tr>
<th>Variable</th>
<th>AD mean (SE)</th>
<th>Controls mean (SE)</th>
<th>Comparison*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total sleep time (min)</td>
<td>363.9(103.7)</td>
<td>374.8(61.1)</td>
<td>NS</td>
</tr>
<tr>
<td>Sleep efficiency (%)</td>
<td>68.3 (15.7)</td>
<td>78.3(10.0)</td>
<td>t = 2.73, p &lt; 0.02</td>
</tr>
<tr>
<td>Stage 1 (%)</td>
<td>27.2 (15.0)</td>
<td>17.5(13.4)</td>
<td>t = 2.61, p &lt; 0.02</td>
</tr>
<tr>
<td>Stage 2 (%)</td>
<td>56.2 (11.9)</td>
<td>56.3(10.7)</td>
<td>NS</td>
</tr>
<tr>
<td>Stages 3 and 4 (%)</td>
<td>2.98(5.56)</td>
<td>8.50(6.61)</td>
<td>Z = 3.20, p &lt; 0.002</td>
</tr>
<tr>
<td>REM (%)</td>
<td>13.6 (6.1)</td>
<td>17.5(13.4)</td>
<td>NS</td>
</tr>
</tbody>
</table>

REM latency (REML) measures* (min)
- REML, Stage 1, W in 113.9 (63.1) 86.3(28.9) t = 2.15, p < 0.05
- REML, Stage 1, W out 85.8 (48.1) 73.8(21.9) NS
- REML, Stage 2, W in 105.8 (64.1) 82.2(28.5) t = 1.82, p < 0.09
- REML, Stage 2, W out 77.4 (41.1) 71.9(22.5) NS
- REML, 8/10, W in 90.4 (68.3) 78.6(29.7) NS
- REML, 8/10, W out 73.3 (41.8) 70.2(23.4) NS

*Paired comparisons used matched paired t-tests or Wilcoxon's signed-rank test (Z).
*REML definitions refer to point of onset (first epoch of Stage 1, first epoch of Stage 2, or first epoch of 8 min of sleep in 10 min of recording) and inclusion or exclusion of intervening wakefulness after sleep onset (W in, W out).

Despite the statistically significant kappa coefficient, the proportion of patients correctly classified with this criterion was less than 2 in 3.

We further examined relationships between dementia and REML by correlating all six REML measures with Mini-Mental State scores (Table 3). Only the correlation involving REML defined by excluding wakefulness and considering sleep onset by 8 min of sleep in 10 min of recording yielded a statistically significant relationship. In all cases, however, the relationships were positive, i.e., mild dementia was associated with longer REMLs and severe dementia was associated with shorter REMLs.

![Figure 1](image-url)  
**Figure 1.** Frequency distribution of REML (Stage 1 sleep onset/wakefulness-included definition) in 28 AD patients and matched controls.
Table 2. Sensitivities and Specificities for Different REML Values Using Best Definition in Table 1 (Sleep onset = Stage 1, Wake Time in)

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Percent correctly classified</th>
<th>Percent sensitivity (positive in AD)</th>
<th>Percent specificity (negative in controls)</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>≥ 60 min</td>
<td>57.2</td>
<td>92.9</td>
<td>21.4</td>
<td>0.4 NS</td>
</tr>
<tr>
<td>≥ 75 min</td>
<td>53.6</td>
<td>75.0</td>
<td>32.1</td>
<td>0.07 NS</td>
</tr>
<tr>
<td>≥ 90 min</td>
<td>60.7</td>
<td>57.1</td>
<td>64.3</td>
<td>0.21 NS</td>
</tr>
<tr>
<td>≥ 105 min</td>
<td>64.3</td>
<td>53.6</td>
<td>75.0</td>
<td>0.29, p &lt; 0.03</td>
</tr>
<tr>
<td>≥ 120 min</td>
<td>60.7</td>
<td>39.3</td>
<td>82.1</td>
<td>0.21 NS</td>
</tr>
</tbody>
</table>

Discussion

These data indicate that the specific definition of REML influences whether or not AD patients differ from controls on REML. In fact, only definitions including wakefulness significantly differentiated the two groups. This suggests that whatever the ultimate value REML may have as a biomarker, its utility may have more to do with the generally disturbed sleep experienced by AD patients than the specific deterioration of the REM–cholinergic system. Other reports using multivariate models have shown similarly that sleep variables reflecting poor sleep quality, such as lower sleep efficiency and lower percentages of Stages 3 and 4 distinguished AD patients from controls (Prinz et al. 1982a, b). Our results concur with those of Vitiello et al. (1984), who also reported that only a REML definition including wakefulness distinguished among different levels of dementia. We must point out that there are other protocols under which the integrity of the REM–cholinergic system in AD could be tested. For example, cholinergic agonists such as arecoline and RS-86 shorten REML in at least some depressed patients (Berger et al. 1985, Jones et al. 1985) and might be expected to lengthen REML in AD.

Despite the fact that some definitions worked better than others, the proportion of AD patients and controls correctly classified with the best of the REML measures was mediocre. Whether or not other cross-validations will yield higher classifications remains to be seen. We must emphasize that we did not have a depressed group in this study. As depression should be associated with relatively shorter REML values and AD with relatively longer REML values (with controls occupying an intermediary position), our results do not vitiate the potential value of REML for symptomatic patients with a mixed clinical picture. One preliminary longitudinal study has examined patients with such a mixed presentation and found that REML did not significantly differentiate patients whose

Table 3. Rank Order Correlations between Six REML Measures and Mini-Mental State Scores for AD Patients (n = 28)

<table>
<thead>
<tr>
<th>Definition*</th>
<th>rho</th>
</tr>
</thead>
<tbody>
<tr>
<td>REML, Stage 1, W in</td>
<td>0.22</td>
</tr>
<tr>
<td>REML, Stage 1, W out</td>
<td>0.26</td>
</tr>
<tr>
<td>REML, Stage 2, W in</td>
<td>0.21</td>
</tr>
<tr>
<td>REML, Stage 2, W out</td>
<td>0.29</td>
</tr>
<tr>
<td>REML, 8/10, W in</td>
<td>0.28</td>
</tr>
<tr>
<td>REML, 8/10, W out</td>
<td>0.38 (p &lt; 0.05)</td>
</tr>
</tbody>
</table>

*See Table 1 for specification of definitions.
clinical status improved (presumably depressed) from those whose clinical status deteriorated (presumably demented) (Reynolds et al. 1986). A more recently published follow-up of these patients (Reynolds et al. 1988), however, found that one definition of REM latency (together with percentage of REM sleep, sleep maintenance, and indeterminate NREM sleep) provided statistically significant classification of patients with mixed syndromes.

Another limitation in our study was our reliance on single night sleep lab data. As mentioned earlier, some (Reynolds et al. 1983a; Allen et al. 1987) sleep lab studies of demented patients do not show adaptation effects. Nonetheless, it is possible that our reliance on a single night of data may have led to a chance difference between AD patients and controls if, in fact, the former were more likely to have disturbed sleep. In addition, in the case of affective disorders, Anseusse et al. (1985a,b) have shown that internight variation in REM latency may have diagnostic utility above and beyond the values generated in a single night. Although the single night is a drawback in our work, we would contend that if REM latency is to hold promise as a biomarker, it must be cost effective. Even a single night of polysomnography may be considerably more expensive than the cost of an overnight DST, and the cost of multiple nights of recording clearly would be prohibitive. Although some work on automated, ambulatory systems for recording REM sleep has occurred (McPartland et al. 1974; Kayed et al. 1979; Kupfer et al. 1982b; Helfund et al. 1986), these are not yet consensually agreed upon, may be expensive, and are subject to methodological criticism (Boukadoum and Kionas 1986). For these reasons, we believe that a single night sleep study is representative of how a potential sleep-related biomarker would be utilized in practice at this time.

The positive correlation between REM latency (using the 8/10, waking-excluded definition) and Mini-Mental State score contrast with the data of Vitiello and coworkers (1984). Why shorter REM latency should be associated with more severe dementia in our study is unknown. At the very least, the relationship between REM sleep and AD may be more complicated than heretofore recognized. It is of interest in this regard that in the largest series of polysomnographically evaluated geriatric patients studied to date, REM latency was shorter in affective disorder than in AD, but also apparently shorter in AD than in controls (Reynolds et al. 1988).

Finally, we must point out that our classification results are relevant only for REM latency with the definitions employed here. Other definitions of REM latency, and indeed other aspects of REM sleep (density, activity), may still provide both higher sensitivity and specificity than the frankly modest classification accuracy we have reported. We have focused on REM latency because its units (minutes) are intuitively appealing and its measurement does not require elaborate signal processing. At the present time, however, it would appear that measures of overall sleep disturbance are likely to provide discriminations of AD patients at least as good as those measures specifically related to REM sleep.

References


