

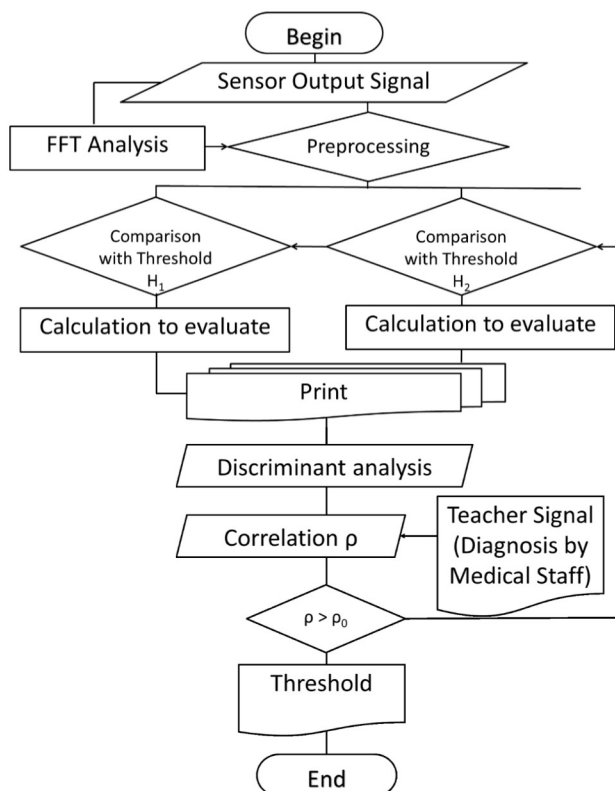


Fig. 1. An example of the flow for an automatic algorithm imitating assessment in accordance of the empirical rule not clearly written.

ing overnight [12–14].

As essential diagnostic criteria of RBD in the ICSD, in addition to abnormal behaviors in daily life or during video monitoring using polysomnography (PSG), the presence of RWA, i.e., increased tonic or phasic EMG activity despite REM sleep suggested by eye movements and brain waves on the PSG records, was adopted [15]. Since single night PSG often shows no abnormal behavior, the confirmation of RWA is an important issue.

3. Materials and Methods

PSG was performed on 12 males aged 60–80 years (mean $\pm$ standard deviation: 72.7 ± 1.7 years) with suspected RBD who visited the Good Sleep Center, Nagoya City University Hospital and 13 healthy males aged 20–60 years (40.5 ± 13.2 years). Surface EMG activity of the chin muscles was recorded using Alic 5 (Philips Respironics GK, Tokyo) overnight at a sampling frequency of 200 Hz. PSG recording was initiated after devices such as a sensor were applied at 19:00. The lights out time was 22:00 and the lights on time was 5:30, at which point recording was completed.

In the present study, the EMG activity of the chin muscles that appeared during REM sleep was extracted. Since muscle activity may be classified into tonic and phasic components, RWA was assessed in each component, and the percentage of the RWA duration in the total REM sleep duration was compared between automatic and visual assessments.

4. Assessment Methods and Development of Automatic Algorithms

In the present study, algorithms were developed for 3 visual assessment methods to evaluate RWA: the AASM scoring manual [16], the method described by Montplaisir *et al.* [17], and the SINBAR method [18]. Tonic muscle activity was extracted from EMG records, its duration (T_{tonic}) was calculated, and its ratio to the entire REM sleep time (T_{tonic}/T) was outputted. Phasic muscle activity was also extracted, its duration (T_{phasic}) was calculated, and its ratio (T_{phasic}/T) was outputted. A plane K spanned by these ratios T_{tonic}/T , T_{phasic}/T is herein defined as an evaluation space. According to each visual assessment method, the 2 calculated percentages were plotted on (T_{tonic}/T)-(T_{phasic}/T) plane K. Visual assessment methods to evaluate RWA may also be compared using (T_{tonic}/T) and (T_{phasic}/T) values on this plot.

In the statistical analysis, $p < 0.05$ was regarded as significant.

5. Results and Discussion

Various visual assessment methods are used to assess RWA. Each method is complex, and there is no internationally standardized method. The diagnosis of RWA is based on a physician's visual assessment, and an RWA assessment method that may contribute to an evaluation of the severity of the disorder or treatment is awaited. In this study, algorithms for the automatic assessment of RWA were developed.

Based on the 3 visual assessment methods described above, respective automatic algorithms for the assessment of RWA were constructed. PSG was performed on 12 subjects with suspected RBD and 13 healthy subjects. In the former group, chin EMG records were analyzed using our automatic algorithms. The percentage of the RWA duration in the total REM sleep duration was assessed separately for the tonic and phasic components, and compared with that calculated using each visual assessment method by physicians and medical technologists. As a result, we identified a method that provided results that did not significantly differ from those obtained by visual assessments.

In addition, chin EMG records in 12 subjects with suspected RBD and 13 healthy subjects were analyzed using the automatic algorithms developed in this study. In each evaluation method, a scatter plot on plane K was produced (Fig. 2). The group with suspected RBD was discriminated from the healthy group in accordance with the linear discriminant analysis (LDA). At first, F-test was conducted as a statistical test to examine whether the centers of gravity (CoG) overlap for each group where the significant level was set to be 0.05. F-value was estimated from the Wilks' Lambda Λ and compared with the critical value $F_0(2, 22) = 3.4434$ because the number of the explanation variable equals to 2 and the sum of the samples equals to 25. In the LDA, Wilk's lambda tests how well each level of independent variable contributes to the model. The scale ranges from 0 to 1, where 0 means total discrimination, and 1 means no discrimination. Significant difference of the CoG could be observed from the F-test if the F-value is greater than the critical value. Based on this statistical evidence, it