



A comparative study of methods for automatic detection of rapid eye movement abnormal muscular activity in narcolepsy

Olesen, Alexander Neergaard; Cesari, Matteo; Christensen, Julie Anja Engelhard; Sørensen, Helge Bjarup Dissing; Mignot, Emmanuel; Jennum, Poul

Published in:
Sleep Medicine

Link to article, DOI:
[10.1016/j.sleep.2017.11.1141](https://doi.org/10.1016/j.sleep.2017.11.1141)

Publication date:
2018

Document Version
Peer reviewed version

[Link back to DTU Orbit](#)

Citation (APA):

Olesen, A. N., Cesari, M., Christensen, J. A. E., Sørensen, H. B. D., Mignot, E., & Jennum, P. (2018). A comparative study of methods for automatic detection of rapid eye movement abnormal muscular activity in narcolepsy. *Sleep Medicine*, 44, 97-105. DOI: 10.1016/j.sleep.2017.11.1141

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Accepted Manuscript

A comparative study of methods for automatic detection of rapid eye movement abnormal muscular activity in narcolepsy

Alexander Neergaard Olesen, Matteo Cesari, Julie Anja Engelhard Christensen, Helge Bjarup Dissing Sorensen, Emmanuel Mignot, Poul Jennum



PII: S1389-9457(17)31592-7

DOI: [10.1016/j.sleep.2017.11.1141](https://doi.org/10.1016/j.sleep.2017.11.1141)

Reference: SLEEP 3590

To appear in: *Sleep Medicine*

Received Date: 24 June 2017

Revised Date: 3 October 2017

Accepted Date: 13 November 2017

Please cite this article as: Olesen AN, Cesari M, Christensen JAE, Sorensen HBD, Mignot E, Jennum P, A comparative study of methods for automatic detection of rapid eye movement abnormal muscular activity in narcolepsy, *Sleep Medicine* (2018), doi: 10.1016/j.sleep.2017.11.1141.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

A comparative study of methods for automatic detection of rapid eye movement abnormal muscular activity in narcolepsy

Alexander Neergaard Olesen^{a,*}, Matteo Cesari^a, Julie Anja Engelhard Christensen^b, Helge Bjarup Dissing Sorensen^a, Emmanuel Mignot^c, Poul Jennum^b

^aDepartment of Electrical Engineering, Technical University of Denmark, Kgs. Lyngby, Denmark

^bDanish Center for Sleep Medicine, Department of Clinical Neurophysiology, Rigshospitalet, Glostrup, Denmark

^cStanford Center for Sleep Sciences and Medicine, Stanford University, Palo Alto, CA, USA

***Corresponding author.** Noeddebogade 4, 5.tv., DK-2200 Copenhagen, Denmark. Tel.: +45 29840968 (mobile).

E-mail address: aneol@elektro.dtu.dk (Alexander Neergaard Olesen)

Abstract

Objective: To evaluate rapid eye movement (REM) muscular activity in narcolepsy by applying five algorithms to electromyogram (EMG) recordings, and to investigate its value for narcolepsy diagnosis.

Patients/methods: A modified version of phasic EMG metric (mPEM), muscle activity index (MAI), REM atonia index (RAI), supra-threshold REM EMG activity metric (STREAM), and Frandsen method (FR) were calculated from polysomnography recordings of 20 healthy controls, 18 clinic controls (subjects suspected with narcolepsy but finally diagnosed without any sleep abnormality), 16 narcolepsy type 1 without REM sleep behavior disorder (RBD), 9 narcolepsy type 1 with RBD, and 18 narcolepsy type 2. Diagnostic value of metrics in differentiating between groups was quantified by area under the receiver operating characteristic curve (AUC). Correlations among the metrics and cerebrospinal fluid hypocretin-1 (CSF-hcrt-1) values were calculated using linear models.

Results: All metrics excluding STREAM found significantly higher muscular activity in narcolepsy 1 cases versus controls ($p < 0.05$). Moreover, RAI showed high sensitivity in the detection of RBD. The mPEM achieved the highest AUC in differentiating healthy controls from narcoleptic subjects. The RAI best differentiated between narcolepsy 1 and 2. Lower CSF-hcrt-1 values correlated with high muscular activity quantified by mPEM, sMAI, IMAI, PEM and FR ($p < 0.05$).

Conclusions: This automatic analysis showed higher number of muscle activations in narcolepsy 1 compared to controls. This finding might play a supportive role in diagnosing narcolepsy and in discriminating narcolepsy subtypes. Moreover, the negative correlation between CSF-hcrt-1 level and REM muscular activity supported a role for hypocretin in the control of motor tone during REM sleep.

Keywords:

Computer-assisted analysis

Electromyogram (EMG)

Narcolepsy

Polysomnography

Rapid eye movement (REM) sleep behavior disorder

REM sleep without atonia

Introduction

Narcolepsy is a neurological sleep disorder characterized by excessive daytime sleepiness (EDS), sudden onsets of rapid eye movement (REM) sleep periods during wakefulness, and a general disturbed pattern of nocturnal sleep. The disorder is currently recognized as either narcolepsy type 1 (NT1) or narcolepsy type 2 (NT2) [1]. NT1 patients are associated with cataplexy, which for the majority of patients is caused by a loss of hypothalamic hypocretinergic cells leading to lower cerebrospinal fluid (CSF) hypocretin-1 (hcr1) levels than in the general population. The loss of cells is likely due to an autoimmune-mediated destruction of hypocretinergic brain cells [2–5]. The mechanism for NT2 is unknown but is most likely heterogeneous [6].

The current diagnosis of NT1 is based on medical history of patients exhibiting EDS for at least 3 months and the presence of one or both of the following: 1) cataplexy combined with a positive multiple sleep latency test (MSLT) with a mean sleep latency <8 minutes and the presence of at least two sleep onset REM periods (REM sleep latency ≤ 15 minutes) during the polysomnogram (PSG) and 4-5 daytime nap opportunities; and 2) CSF-hcr1 levels ≤ 110 pg/mL, or less than one third of the mean values obtained in normal subjects with the same standardized assay. The NT2 present normal CSF-hcr1 levels [1].

Recent studies have also shown that a positive MSLT in NT2 is often not a repeatable clinical feature, raising questions regarding the diagnostic entity [7]. Indeed, false positive MSLTs are associated with shift work and sleep deprivation [8]. Due to variation in interpretation and day-to-day variations in PSG/MSLT [9,10], other diagnostic procedures have been proposed for diagnosing narcolepsy, including reports of sleep stages [11-16], and autonomic disturbance findings during the PSG [17].

A promising supportive diagnostic procedure for narcolepsy might be the analysis of abnormal muscular activity during REM sleep. In fact, in NT1, REM sleep without atonia (RSWA) and clinical signs indicative of REM sleep behavior disorder (RBD) have been frequently reported [18-20]. Current hypotheses claim that an association between RBD and NT1 might be caused by a common dysregulation of the dopamine system [21]. Therefore, abnormal REM muscular activity could also be considered a diagnostic feature. Unfortunately, there is a lack of standardization in the identification of abnormal motor activity during sleep, which has to be overcome in order to make this feature useful for diagnostic purposes. Recently, there has been an increased focus on automatizing the quantification of muscle activity using either the original definitions of phasic and tonic activations [22] or novel computer detection approaches derived thereof [23,24].

The current study attempted to objectively quantify EMG activity in narcolepsy patients, focusing on five different automatic detection approaches using quantitative measures derived from the chin EMG: 1) a modified version of the phasic electromyographic metric (mPEM) [25]; 2) the muscle activity index (MAI) [26]; 3) the REM atonia index (RAI) [27]; 4) the suprathreshold REM EMG activity metric (STREAM) [28]; and 5) the automated method proposed by Frandsen et al. (FR) [29]. The current study aimed to: 1) investigate the previously mentioned computer detection approaches with respect to detection of EMG changes in narcoleptic sleep; 2) investigate whether the automated analysis of EMG might be used as a supportive diagnostic tool for narcolepsy using a single overnight PSG; and 3) investigate the relationship between CSF-hcrt-1 values and abnormal muscular activity in REM sleep.

Methods

Subjects and recordings

A total of 81 subjects were included in this study, which consisted of one healthy control group and four patient groups. The healthy control (C) group comprised 20 healthy subjects with no previously diagnosed sleep disorders and no suspicion of narcolepsy. Patients were admitted to the sleep clinic under suspicion of narcolepsy and underwent standard PSG-MSLT procedures [30], including CSF-hcrt-1 measurements [31]. After clinical tests, patients were divided into four groups: 1) 18 patients with normal CSF-hcrt-1 levels and no diagnosis of narcolepsy or other primary sleep disorders (sleep clinic controls, SCC); 2) 16 patients diagnosed with NT1 (NT1-RBD); 3) 9 patients diagnosed with NT1 and RBD (NT1+RBD); and 4) 18 subjects diagnosed with NT2. The RBD diagnosis was based on anamnesis, questionnaires and video-PSGs, which were scored by technicians according to American Academy of Sleep Medicine (AASM) rules. Subject demographics are shown in Table 1, including gender, age, CSF-hcrt-1 level, BMI, sleep stage distributions, and known medication. All patients were advised to discontinue medications known to affect normal sleep behavior, such as centrally acting stimulants and medications for cataplexy, at least 2 weeks prior to the PSG recording.

All PSGs were recorded at the Danish Center for Sleep Medicine in accordance with AASM standards [30], with a sampling frequency of 256 Hz. Only the chin EMG channel was used in the current study. Ten PSGs (two C, three SCC, two NT1-RBD, and three NT2 patients) were dismissed from further analysis due to unacceptable noise. The EMG signals were recorded with different amplifier systems, in which the lowest cut-off frequency for the built-in anti-aliasing filter was 70 Hz. For the sake of uniformity, all EMG signals were low-pass filtered with a fourth-order Butterworth zero-phase low-pass filter with a cut-off at 70 Hz. Moreover, recordings were digitally filtered with a 10 Hz Butterworth high-pass filter and notch filtered at 50 Hz using a zero-phase filtering procedure. All algorithms were implemented in MATLAB (R2016a, The MathWorks, Natick, MA, USA).

The study was accepted by the Danish Health Authority and the Data Protection Agency. All data were anonymized. The work was carried out in accordance with the Declaration of Helsinki.

Table 1. Demographic data for the five studied groups.

Parameter	C	SCC	NT1-RBD	NT1+RBD	NT2	<i>p</i>
Total count	20	18	16	9	18	-
Fraction of men	0.45	0.39	0.63	0.67	0.44	0.509
Age, years, $\mu \pm \sigma$	34.1 \pm 7.8	37.7 \pm 11.4	34.7 \pm 18.7	26.7 \pm 21.1	29.2 \pm 11.7	0.241
BMI, kg/m ² , $\mu \pm \sigma$	25.6 \pm 6.7	23.9 \pm 4.2	27.5 \pm 3.8	22.8 \pm 4.6	23.1 \pm 3.5	0.036
CSF hypocretin 1 level	-	401.9 \pm 62.2	33.8 \pm 29.8	27..3 \pm 21.5	332.9 \pm 98.7	-
Sleep efficiency, %, $\mu \pm \sigma$	91.1 \pm 6.8	86.4 \pm 15.1	85.4 \pm 10.9	87.7 \pm 9.3	93.4 \pm 3.1	0.133
Time in bed, minutes, $\mu \pm \sigma$	471.9 \pm 40.9	478.9 \pm 33.8	472.5 \pm 54.4	455.1 \pm 68.0	511.5 \pm 142.2	0.863
REML, minutes, $\mu \pm \sigma$	100.4 \pm 52.5	97.1 \pm 33.2	88.31 \pm 98.4	34.1 \pm 44.0	109.3 \pm 114.6	0.011
W, %, $\mu \pm \sigma$	8.9 \pm 6.8	13.6 \pm 15.1	14.5 \pm 11.0	12.3 \pm 9.3	6.6 \pm 3.1	0.149
REM, %, $\mu \pm \sigma$	20.4 \pm 4.9	18.2 \pm 7.4	17.4 \pm 7.2	22.7 \pm 5.2	20.2 \pm 6.9	0.163
N1, %, $\mu \pm \sigma$	4.8 \pm 2.2	5.7 \pm 3.3	11.8 \pm 7.8	10.4 \pm 7.6	5.3 \pm 3.1	0.008
N2, %, $\mu \pm \sigma$	45.0 \pm 8.3	43.1 \pm 10.7	42.4 \pm 12.1	35.7 \pm 9.7	48.1 \pm 11.6	0.135
N3, %, $\mu \pm \sigma$	20.8 \pm 8.1	19.3 \pm 7.8	13.9 \pm 6.9	18.8 \pm 13.8	19.8 \pm 8.5	0.071
Fraction of subjects taking stimulant medication	0	0.06	0.25	0.33	0.17	0.059
Fraction of subjects taking anti-cataplexy medication	0	0	0	0.33	0.22	0.003
Fraction of subjects taking both stimulant and anti-cataplexy medication	0	0	0	0.33	0	<0.001

C, healthy controls; SCC, sleep clinic controls; NT1 \pm RBD, type 1 narcolepsy with/without RBD; NT2, type 2 narcolepsy; BMI, body mass index; CSF, cerebrospinal fluid; REML, REM sleep latency; REM, rapid eye movement;

p-values were obtained with Kruskal-Wallis tests and Chi-squared tests for fractions.

Modified phasic electromyographic metric (mPEM)

An automatic way of calculating the PEM was developed in accordance with the original visual scoring method [25]. However, because of the absence of a well-defined pre-sleep period, the definition of the baseline level was changed as described in the following. A phasic activation must fulfill three criteria: 1) activation above four times the baseline; 2) activation between 0.1-2.5 seconds in duration; and 3) clear return to baseline after each activation. In order to meet these criteria, 1) the baseline was defined in 30-second epochs as the median of the absolute value of the EMG in these epochs. Then, 2) 2.5-second mini-epochs were considered and for each of them activations were found when the EMG was above four times the baseline previously defined, and the activations lasting <0.1 second were removed. Finally, 3) activations within the last 10 samples of the 2.5-second mini-epochs were removed from further analysis, as these would have no clear return to baseline. The fraction of mini-epochs in REM sleep meeting the requirements was reported as the PEM. It is worth noticing that the method proposed in [25] was applied to EMG signals filtered between 10-100 Hz.

Muscle activity index (MAI)

To detect abnormal muscular activity, the amplitude EMG signal was calculated from the original EMG signal as the difference between an upper and lower envelope calculated by linear interpolation between the upper and lower waveform peaks, respectively. The amplitude signal was then smoothed over 0.025 seconds using a moving average filter. A baseline was extracted by smoothing the amplitude signal over 200 seconds and the threshold was defined as four times the baseline. In the original paper, the threshold was defined as twice the baseline level [26]; however, in the current study it was found that this threshold level was too susceptible to noise and it was decided to set the threshold at four times the baseline instead (an example supporting this choice is shown in Appendix Fig. A1). This is in accordance with the original specifications used by Lapierre and Montplaisir [22], and has been extensively used in other studies [18,25,29,32]. The EMG activity was detected when the amplitude signal exceeded the threshold curve, and clusters of activity <1 -second distance were defined as one unique event. The long-lasting and short-lasting muscle activity indexes (IMAI and sMAI) were defined as the number of activity events/hour of REM sleep above and below 0.5 seconds in length, respectively. It should be noticed that the method proposed in [26] was applied to EMG filtered in the range 10-100 Hz.

REM atonia index (RAI)

The EMG signal was rectified and the amplitude in each 1-second mini-epoch was calculated as the average amplitude in that mini-epoch minus the minimum amplitude in a surrounding moving window of 30 seconds, in order to reduce the noise influence. The obtained amplitudes (*amp*) of each 1-second mini-epoch were then divided into 20 stage-specific categories according to their values (ie, $amp \leq 1 \mu V$, $1 \mu V \leq amp \leq 2 \mu V$, ..., $18 \mu V \leq amp \leq 19 \mu V$ and $amp \geq 20 \mu V$). By considering only the 1-second mini-epochs in REM stage, the RAI was calculated as the percentage of $amp \leq 1 \mu V$ divided by the sum of the number of mini-epochs with $amp \leq 1 \mu V$ and those with amplitudes $> 2 \mu V$ (ie, excluding those with $1 \mu V < amp \leq 2 \mu V$) [27,33]. It has to be noted that the method proposed in [27,33] was applied to EMG signals filtered between 10-100 Hz.

Supra-threshold REM activity metric (STREAM)

Each recording was divided into 3-second mini-epochs and the sample variance for each mini-epoch calculated. The STREAM was computed as the percentage of REM mini-epochs with sample variance larger than a threshold defined as four times the fifth percentile of the observed sample variance during all NREM mini-epochs [28].

Frandsen method (FR)

Frandsen et al. proposed an optimized method for quantifying submental motor activity during REM sleep. After the filtering, an amplitude curve (AC) was generated by calculating the difference between the highest and lowest EMG values within an ~200 ms moving window. Then, the AC was divided in 30-second segments and the median amplitude (MA) calculated for each of them. For each of these segments, the baseline was identified as the lowest MA in the period 30 minutes before and 30 minutes after the current segment. Muscular activity was then identified when the AC was above four times the baseline and with a minimum duration of 0.3 seconds. Moreover, clusters of muscular activity at a distance < 0.5 seconds were considered as a unique muscular activation. Finally, the REM sleep mini-epoch index (FR) was calculated as the percentage of 3-second mini epochs in REM sleep with muscle activity events in $> 50\%$ of the time. The implementation of the different algorithms was validated on a visual basis.

Statistical analysis

Between-group comparisons were performed using a Kruskal-Wallis test. Post-hoc comparisons were performed with a Mann-Whitney U-test with Tukey-Kramer correction at a significance level of $p=0.05$, and the effect size for each comparison was calculated according to [34]. Diagnostic value

and performance were quantified by sensitivity, specificity and the area under the receiver operating characteristic (ROC) curve (AUC). In particular, performance was measured using AUC values and the largest sensitivity at which there was a minimum of 97% specificity. Linear regression models were used to assess relationships between metrics and CSF-hcrt-1 levels, with correlation being evaluated using Pearson's correlation coefficient and Student's *t*-test for significance. For all the metrics (RAI excluded) and CSF-hcrt-1 values, log-transformation was used, while for the RAI metric the $\log(100\text{-RAI})$ was applied to obtain positive correlations. All model residuals were checked for normality.

Results

Fig. 1 a)-f) show mPEM, STREAM, sMAI, IMAI, RAI and FR metrics, respectively, visualized by boxplots with group means designated by crosses, medians designated by horizontal lines, and the 25th and 75th percentiles designated by the bottom and top end of the blue box. Average and standard error of the mean values of the different metrics are shown in Table 2 together with the *p*-value of the between-groups comparison. Table 3 shows the *p*-values of the post-hoc comparison tests and the effect size of each comparison.

A comparison of the diagnostic value and performance of these different metrics is shown Table 4. The choice of the different combinations shown for AUC analysis is explained in the discussion section. Finally, the results of correlation analysis are shown in Fig. 2.

Table 2. Electromyogram analyses results across subject groups.

	C	SCC	NT1-RBD	NT1+RBD	NT2	<i>p</i>
sMAI, hours REM	87.1 (3.0)	91.5 (3.4)	161.8 (5.0)	127.4 (5.3)	137.8 (5.4)	0.0056
lMAI, hours REM	46.8 (2.0)	35.0 (0.9)	104.9 (4.3)	72.2 (5.2)	54.8 (2.3)	0.0015
mPEM, %	10.8 (0.3)	13.6 (0.4)	25.7 (0.6)	19.2 (1.0)	15.8 (0.7)	0.0004
RAI, %	96.9 (0.2)	96.1 (0.3)	93.0 (0.3)	89.8 (1.4)	96.4 (0.2)	0.0008
STREAM, %	7.9 (0.2)	14.1 (0.9)	15.2 (0.9)	16.6 (1.7)	19.4 (1.2)	0.0529
FR, %	11.2 (0.5)	13.3 (0.8)	20.8 (0.8)	19.5 (1.2)	12.9 (0.6)	0.0050

Results are displayed as mean (standard error of the mean), as well as the *p*-values for the Kruskal-Wallis test for each feature.

C, healthy controls; SCC, sleep clinic controls; NT1±RBD, type 1 narcolepsy with/without RBD; NT2, narcolepsy type 2; sMAI, short muscle activity index; lMAI, long muscle activity index; mPEM, modified phasic electromyographic metric; RAI, REM atonia index; STREAM, supra-threshold REM EMG activity metric; FR, Frandsen method; REM, rapid eye movement

Table 3. *P*-values obtained for the post-hoc comparisons by using a Mann-Whitney U-test with Tukey-Kramer correction highlighted by significance ($p < 0.05$) and relative effect size in parentheses.

	sMAI	IMAI	mPEM	RAI	FR
C vs SCC	0.9948 (-0.03)	0.9757 (0.07)	0.8822 (-0.17)	0.9944 (0.06)	0.9940 (-0.11)
C vs NT1-RBD	0.0153 (-0.50)	0.0121 (-0.49)	0.0002 (-0.66)	0.0167 (0.48)	0.0154 (-0.46)
C vs NT1+RBD	0.1668 (-0.46)	0.2735 (-0.42)	0.1100 (-0.50)	0.0079 (0.61)	0.0584 (-0.45)
C vs NT2	0.1392 (-0.38)	0.9101 (-0.11)	0.4083 (-0.25)	0.9481 (0.14)	0.7535 (-0.21)
SCC vs NT1-RBD	0.0660 (-0.45)	0.0026 (-0.60)	0.0122 (-0.54)	0.0721 (0.44)	0.0687 (-0.43)
SCC vs NT1+RBD	0.3527 (-0.37)	0.1111 (-0.53)	0.5198 (-0.31)	0.0322 (0.53)	0.1634 (-0.47)
SCC vs NT2	0.3505 (-0.24)	0.6292 (-0.22)	0.9394 (-0.14)	0.9978 (0.06)	0.9441 (-0.15)
NT1-RBD vs NT1+RBD	0.9901 (0.13)	0.9389 (0.23)	0.7320 (0.32)	0.9763 (0.06)	1 (0.06)
NT1-RBD vs NT2	0.9257 (0.11)	0.1600 (0.39)	0.1055 (0.40)	0.1522 (-0.38)	0.3336 (0.36)
NT1+RBD vs NT2	0.9992 (0.00)	0.7484 (0.16)	0.8974 (0.15)	0.0690 (-0.55)	0.5016 (0.33)

C, healthy controls; SCC, sleep clinic controls; NT1±RBD, type 1 narcolepsy with/without RBD; NT2, narcolepsy type 2; sMAI, short muscle activity index; IMAI, long muscle activity index; mPEM, modified phasic electromyographic metric; RAI, REM atonia index; FR, Frandsen method

Table 4. Comparison of the diagnostic value and performance of the five different metrics, given as area under the curve (AUC) (sensitivity/specificity). Sensitivity was calculated at the point of a minimum of 97% specificity.

	sMAI	IMAI	mPEM	RAI	STREAM	FR
C vs NT1+NT2	0.784 (0.158/1)	0.708 (0.132/1)	0.794 (0.395/1)	0.741 (0.222/0.974)	0.737 (0.290/1)	0.731 (0.132/1)
SCC vs NT1+NT2	0.725 (0.079/1)	0.779 (0.368/1)	0.712 (0.132/1)	0.700 (0.133/0.974)	0.597 (0.026/1)	0.714 (0/1)
NT1 vs NT2	0.542 (0.044/1)	0.693 (0.261/1)	0.696 (0/1)	0.788 (0.267/1)	0.533 (0/1)	0.725 (0.044/1)
NT1-RBD vs NT1+RBD	0.587 (0.214/1)	0.651 (0.071/1)	0.706 (0.071/1)	0.540 (0.286/1)	0.587 (0/1)	0.540 (0.143/1)

C, healthy controls; SCC, sleep clinic controls; NT1±RBD, type 1 narcolepsy with/without RBD; NT2, narcolepsy type 2; sMAI, short muscle activity index; IMAI, long muscle activity index; mPEM, modified phasic electromyographic metric; RAI, RAI atonia index; STREAM, supra-threshold REM EMG activity metric; FR, Frandsen method

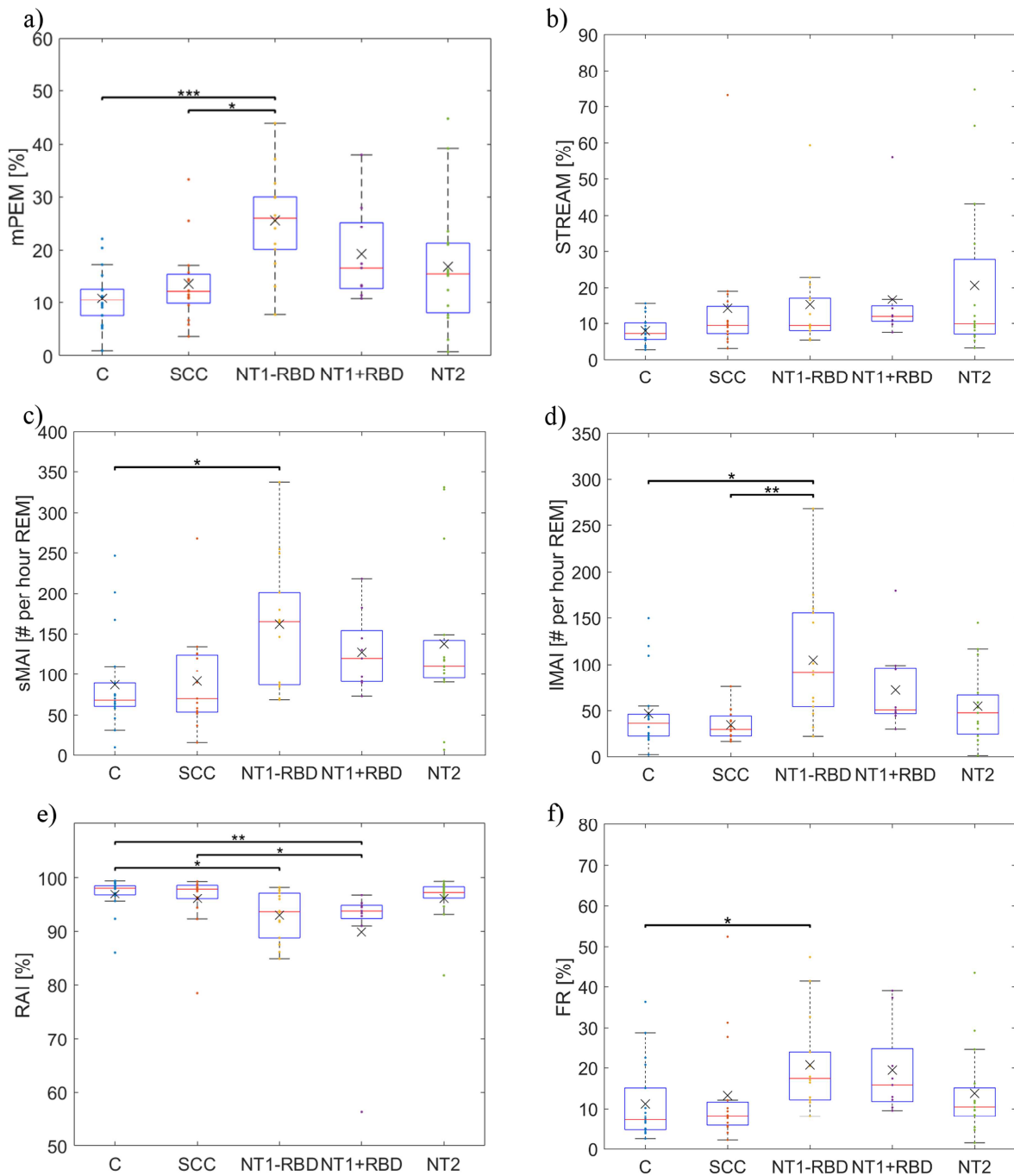


Fig. 1. a) Modified phasic electromyogram metric (mPEM); b) supra-threshold rapid eye movement electromyogram activity metric (STREAM); c) short muscle activity index (sMAI); d) long muscle activity index (lMAI); e) rapid eye movement atonia index (RAI); and e) Frandsen metric (FR) across subject groups shown as box plots with crosses denoting group means.

C, healthy controls; SCC, sleep clinic controls; NT1±RBD, type 1 narcolepsy with/without RBD; NT2, narcolepsy type 2

* $p < 0.05$

** $p < 0.01$

*** $p < 0.001$

ACCEPTED MANUSCRIPT

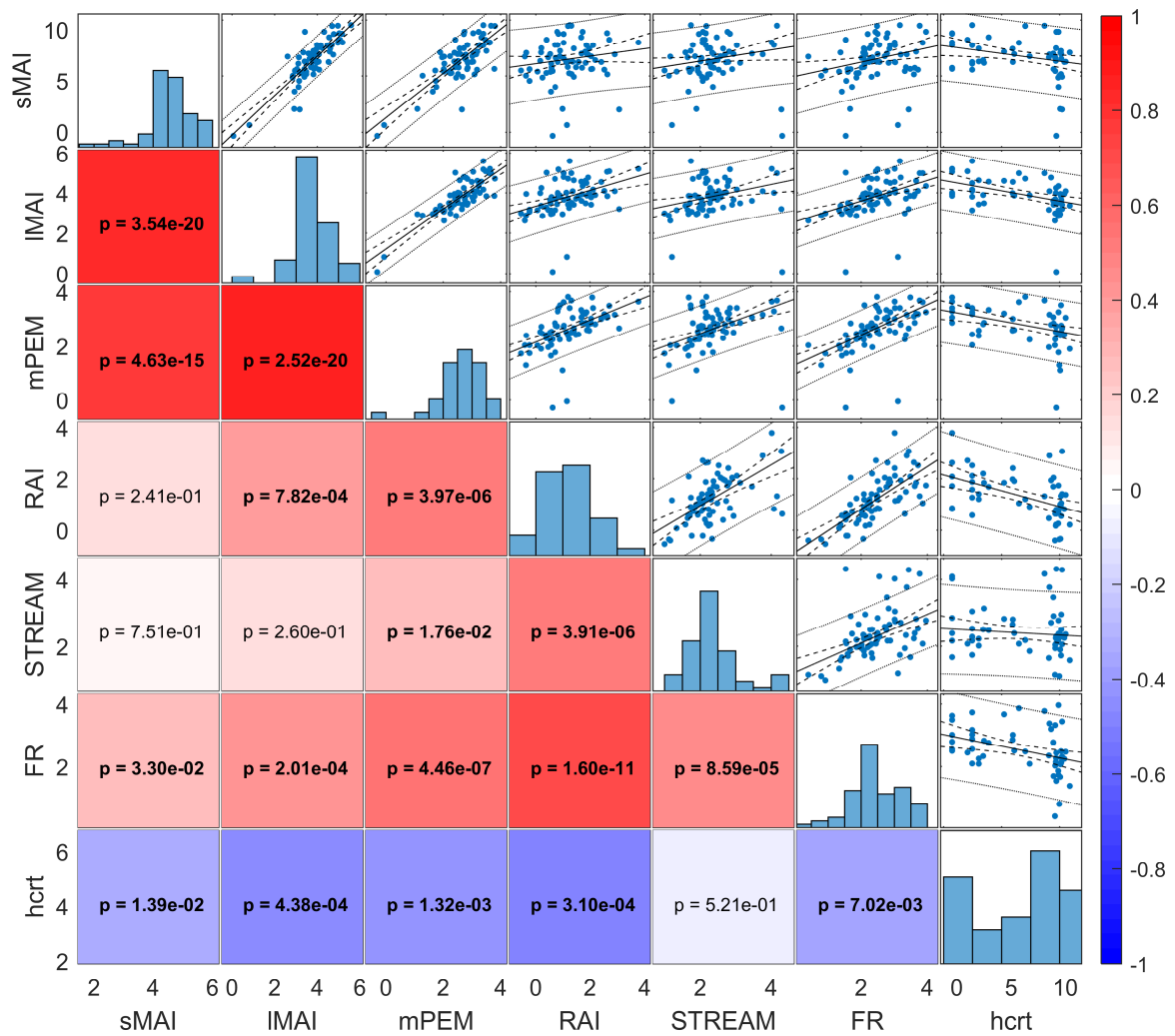


Fig. 2. Correlation plots and analysis between metrics. All the metrics (RAI excluded) and CSF-1-hcrt levels are log-transformed, and $\log(100-RAI)$ was used. Colors in the left boxes represent the correlation values, and significant p -values for the correlation are highlighted. In the diagonal, the distributions of the metrics are shown.

sMAI, short muscle activity index; IMAI, long muscle activity index; mPEM, modified phasic electromyographic metric; RAI, REM atonia index; REM, rapid eye movement; STREAM, supra-threshold REM electromyography activity metric; FR, Frandsen method; hcrt, hypocretin

Discussion

The results in Fig. 1a-f) and Tables 2 and 3 show that all the metrics (STREAM excluded) found elevated REM EMG activity in NT1-RBD compared to C, which is in agreement with the muscular abnormalities in NT1 observed in previous studies [18-20]. Increase in REM EMG activity was also observed when SCC were compared to NT1-RBD, but statistical difference was obtained only by IMAI and mPEM, while sMAI, RAI and FR were close to significance level. All the metrics presented a medium effect size when C/SCC were compared to NT1-RBD, with mPEM showing the highest value. These results further confirm that NT1 subjects are characterized by abnormal REM muscular activity. Compared to the other metrics, RAI was more sensitive to detect RBD as a comorbidity of NT1, since it was the only one showing statistical difference between NT1+RBD and C/SCC with a medium effect size. It is interesting to observe that RAI presented significant or close-to-significance *p*-values when NT1±RBD patients were compared to other groups. This means that this metric has the potential to quantify the abnormal muscular movements during REM that have previously been observed [18–20]. However, it must be noticed that the authors of RAI proposed that healthy subjects are characterized by values >90% [27]. If this threshold was applied in this study, the majority of subjects would have been considered healthy. Low sample size in the current study and inter-clinic variability due to differences in recording equipment may have influenced these results. It may be that introducing subject-specific amplitude criteria based on NREM value features, such as in the STREAM calculations, would improve applicability to other datasets.

The AUC analysis performed in Table 4 was performed in order to understand whether automatically detected REM EMG activity could be used as a tool to diagnose and differentiate narcolepsy with high specificity, according to the scheme shown in Fig. 3. When a subject is referred to the clinic, it is first desirable to understand whether he/she has narcolepsy, then which type it is, and in case of NT1, whether RBD is present or not. Fig. 3 illustrates the decision hierarchy on the role of RSWA in the diagnosis of narcolepsy in the top tier, and RBD as a comorbidity of NT1 in the lowest tier.

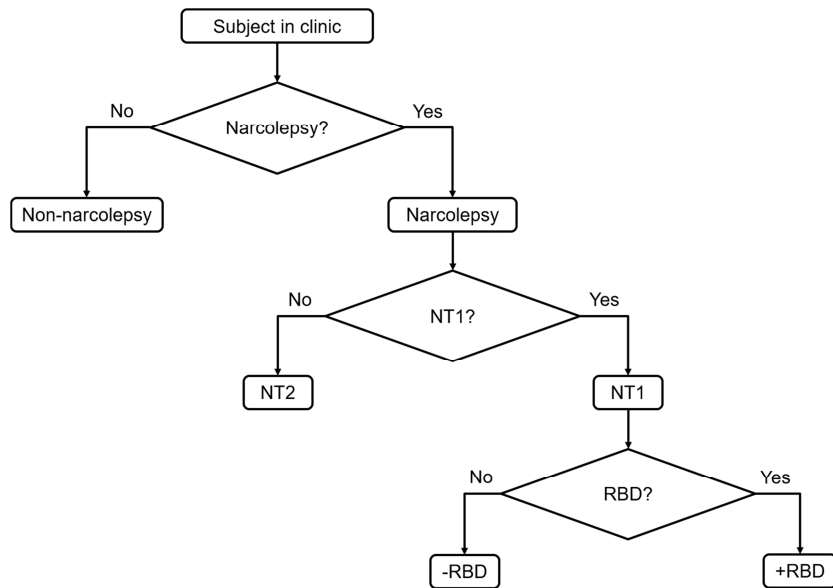


Fig. 3. Schematic representation of the diagnostic decisions made with the proposed automatic analysis of REM-EMG activity. The presence of narcolepsy is first analyzed, then which type of narcolepsy, and finally in case of NT1 the presence of comorbid RBD is noted.

The results in Table 4 show that the applied metrics showed promising diagnostic performances in discriminating healthy subjects (C/SCC) from the narcoleptic ones, with mPEM having the highest AUC (0.79) when C were considered, and IMAI having the highest one (0.78) when SCC were taken into account. Testing NT1 against NT2 yielded an AUC of 0.79 using the RAI metric, thus showing that it has potential in distinguishing between narcolepsy types. Finally, mPEM had the highest AUC value (0.71) in distinguishing NT1+RBD and NT1-RBD, while the other metrics did not show satisfactory results. These results are generally promising because they show that automated EMG analysis during single-night PSG has the potential of discriminating between healthy controls and narcolepsy and between narcolepsy subtypes. The current study chose to focus on highly specific features for narcolepsy detection (at least 97% specificity) instead of seeking optimal diagnostic points balancing specificity and sensitivity in the ROC curve. Previous studies have likewise described highly specific features obtained from other sources, such as power spectral information in the EEG [11] and sleep stage transitions [12], and found sensitivities that were comparable to those found in the current study. However, since the obtained sensitivities were relatively low, this type of analysis might be considered as a supportive tool in the diagnosis of narcolepsy in combination with other tools, although extension to a larger sample will be needed to confirm this. In order to provide a more complete overview of the AUC analysis, the sensitivity and specificity obtained in correspondence of the optimal point of the ROC curve are presented in appendix Table A. 1.

Fig. 2 shows the correlation analysis between metrics and CSF-hcrt-1 levels. The linear relationships between variables were evaluated using Pearson's correlation coefficient. However, since this is a parametric method sensitive to outliers in data, the correlation analysis was also conducted using Spearman's correlation coefficient included in Appendix Fig. A. 2. These results were similar to the ones reported in Fig. 2. The analysis in Fig. 2 shows that the different metrics were generally characterized by significant correlations among them. These results show that the metrics investigated in this study generally agreed on the quantification of abnormal muscular activity. It is particularly interesting to observe a significant negative correlation between log-transformed CSF-hcrt-1 values and metrics. This means that more muscular activity was observed in patients having lower CSF-hcrt-1 levels, thus supporting the hypothesis that CSF-hcrt-1 plays an important role in regulating muscle tone in narcolepsy during REM sleep.

This is the first study to compare multiple EMG quantification methods on a single dataset containing NT1 and NT2 patients with the aim of investigating the role of automatically detected RSWA as a supportive tool in the diagnosis and differentiation of narcolepsy. Previous studies on automatic EMG quantification have focused primarily on RBD, iRBD in connection with neurodegenerative diseases such as Parkinson's, or NT1 alone [25-29,35-37]. However, the currently implemented automatic methods present important limitations with regard to artifact removals. Specifically, none of the algorithms embed automatic detection of events that might affect the results in the calculation of abnormal muscular activity, such as sleep apneas, arousals, and noise. It is believed that adding the detection of such features should be taken into account when developing future automatic/data-driven EMG analysis methods.

This study was limited by the following factors: 1) although a total of 81 subjects (71 after artifact rejection) were enrolled, the number of subjects in each group was very low. This limited the extent of the conclusions drawn based on the statistical analyses, and further studies with more data including idiopathic RBD and/or Parkinson groups should confirm the findings of this study. The small sample size might also explain the fact that NT1+RBD subjects presented a very similar muscular activity level as NT1-RBD, which was a relatively unexpected result. 2) The EMG analysis was performed on submental recordings only, and could be extended to include traces of legs and arms, since this has been shown to better differentiate RBD patients from controls in some studies [38]. However, the submental was not affected by periodic leg movements to the same degree as lower limb recordings [25]. 3) Many other detection approaches have been proposed in the literature, which could also have been investigated in this study. However, it chose to focus on methods which could be adapted to an automatic setting using chin EMG only, excluding those using multiple traces, such as the SINBAR method [38,39]. 4) Because of different hardware systems were used for

PSG recording, all EMG signals at 70 Hz had to be digitally low-pass filtered post hoc. This was a critical point, since some of the proposed algorithms (ie, mPEM, RAI and MAI) have originally been applied to EMG signals low-pass filtered at 100 Hz. It is unclear at this point whether the lack of frequency components in the 70-100 Hz range in the analyzed signals would have a significant effect on the final results, since at least one study used the same low-pass cut-off point at 70 Hz for analyzing RWSA in patients with RBD and comorbid OSA [40]. Further studies are needed to understand whether the different filtering might affect the results in a significant way.

ACCEPTED MANUSCRIPT

Conclusions

In summary, five automated methods evaluating nocturnal EMG activity were implemented in order to evaluate whether automatically detected RSWA could be used as a supportive tool in the diagnosis of narcolepsy and differentiation of narcolepsy subtypes. The results show that abnormal muscular activity is detected by most methods in NT1 subjects and that RAI is more sensitive than the other methods in identifying RBD as a comorbidity of NT1. The analysis of the diagnostic performances of the automatic methods show that mPEM yields promising results in the diagnosis of narcolepsy and identification of RBD in NT1, while RAI has the highest AUC in differentiating between NT1 and NT2. Moreover, it was found that CSF-hcrt-1 are associated with RSWA, underlying that hypocretin plays a role in the regulation of muscle activity during sleep.

Sources of support: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interests: None.

Author contribution: ANO and MC: main authors, implementation of the automated methods and statistical analysis; JAEC: study design, manuscript revision; HBDS: creation, initiation and management of the project; EM: creation, initiation and management of the project; PJ: creation, initiation and management of the project, main supervisor.

ACCEPTED MANUSCRIPT

References

- [1] American Academy of Sleep Medicine. International classification of sleep disorders. 3rd ed. Darien, IL: American Academy of Sleep Medicine; 2014.
- [2] Partinen M, Kornum BR, Plazzi G, Jennum P, Julkunen I, Vaarala O. Narcolepsy as an autoimmune disease: the role of H1N1 infection and vaccination. *Lancet Neurol* 2014;13:600–13.
- [3] Partinen M, Kornum BR, Plazzi G, Jennum P, Julkunen I, Vaarala O. Does autoreactivity have a role in narcolepsy? *Lancet Neurol* 2014;13:1072–3.
- [4] Peyron C, Faraco J, Rogers W, Ripley B, Overeem S, Charnay Y, et al. A mutation in a case of early onset narcolepsy and a generalized absence of hypocretin peptides in human narcoleptic brains. *Nat Med* 2000;6:991–7.
- [5] Hallmayer J, Faraco J, Lin L, Hesselson S, Winkelmann J, Kawashima M, et al. Narcolepsy is strongly associated with the T-cell receptor alpha locus. *Nat Genet* 2009;41:708–11.
- [6] Kornum BR, Knudsen S, Ollila HM, Pizza F, Jennum PJ, Dauvilliers Y, et al. Narcolepsy. *Nat Rev Dis Prim* 2017;3:16100.
- [7] Trotti LM, Staab BA, Rye DB. Test-retest reliability of the multiple sleep latency test in narcolepsy without cataplexy and idiopathic hypersomnia. *J Clin Sleep Med* 2013;9:789–95.
- [8] Goldbart A, Peppard P, Finn L, Ruoff CM, Barnet J, Young T, et al. Narcolepsy and predictors of positive MSLTs in the Wisconsin Sleep Cohort. *Sleep* 2014;37:1043–51.
- [9] Chen L, Ho CKW, Lam VKH, Fong SYY, Lam SP, Wing YK, et al. Interrater and intrarater reliability in multiple sleep latency test. *J Clin Neurophysiol* 2008;25:218–21.
- [10] Pataka A, Yoon C-H, Poddar A, Riha RL. Assessment of multiple sleep latency testing in adults in Europe. *Sleep Med* 2013;14:136–9.
- [11] Christensen JAE, Munk EGS, Peppard PE, Young T, Mignot E, Sorensen HBD, et al. The diagnostic value of power spectra analysis of the sleep electroencephalography in narcoleptic patients. *Sleep Med* 2015;16:1516–27.
- [12] Christensen JAE, Carrillo O, Leary EB, Peppard PE, Young T, Sorensen HBD, et al. Sleep stage transitions during polysomnographic recordings as diagnostic features of type 1 narcolepsy. *Sleep Med* 2015;16:1558–66.

- [13] Kempfner J, Jennum P, Sorensen HBD, Christensen JAE, Nikolic M. Automatic SLEEP staging: From young adults to elderly patients using multi-class support vector machine. 35th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., Osaka, Japan: 2013: 5777–80.
- [14] Sorensen GL, Knudsen S, Jennum P. Sleep Transitions in Hypocretin-Deficient Narcolepsy. *Sleep* 2013;36:1173–7.
- [15] Jensen JB, Sorensen HBD, Kempfner J, Sorensen GL, Knudsen S, Jennum P. Sleep-wake transition in narcolepsy and healthy controls using a support vector machine. *J Clin Neurophysiol* 2014;31:397–401.
- [16] Luca G, Haba-Rubio J, Dauvilliers Y, Lammers GJ, Overeem S, Donjacour CE, et al. Clinical, polysomnographic and genome-wide association analyses of narcolepsy with cataplexy: a European Narcolepsy Network study. *J Sleep Res* 2013;22:482–95.
- [17] Sorensen GL, Knudsen S, Petersen ER, Kempfner J, Gammeltoft S, Sorensen HBD, et al. Attenuated heart rate response is associated with hypocretin deficiency in patients with narcolepsy. *Sleep* 2013;36:91–8.
- [18] Dauvilliers Y, Rompré S, Gagnon J-F, Vendette M, Petit D, Montplaisir J. REM sleep characteristics in narcolepsy and REM sleep behavior disorder. *Sleep* 2007;30:844–9.
- [19] Knudsen S, Gammeltoft S, Jennum PJ. Rapid eye movement sleep behaviour disorder in patients with narcolepsy is associated with hypocretin-1 deficiency. *Brain* 2010;133:568–79.
- [20] Dauvilliers Y, Jennum P, Plazzi G. Rapid eye movement sleep behavior disorder and rapid eye movement sleep without atonia in narcolepsy. *Sleep Med* 2013;14:775–81.
- [21] Fraigne JJ, Torontali ZA, Snow MB, Peever JH. REM sleep at its core - circuits, neurotransmitters, and pathophysiology. *Front Neurol* 2015;6.
- [22] Lapierre O, Montplaisir J. Polysomnographic features of REM sleep behavior disorder: development of a scoring method. *Neurology* 1992;42:1371–4.
- [23] Fulda S, Plazzi G, Ferri R. Scoring atonia during normal and pathological rapid eye movement sleep: Visual and automatic quantification methods. *Sleep Biol Rhythms* 2013;11:40–51.
- [24] Frauscher B, Ehrmann L, Högl B. Defining muscle activities for assessment of rapid eye movement sleep behavior disorder: From a qualitative to a quantitative diagnostic level. *Sleep Med* 2013;14:729–33.

- [25] Bliwise DL, He L, Ansari FP, Rye DB. Quantification of electromyographic activity during sleep: a phasic electromyographic metric. *J Clin Neurophysiol* 2006;23:59–67.
- [26] Mayer G, Kesper K, Ploch T, Canisius S, Penzel T, Oertel W, et al. Quantification of tonic and phasic muscle activity in REM sleep behavior disorder. *J Clin Neurophysiol* 2008;25:48–55.
- [27] Ferri R, Manconi M, Plazzi G, Bruni O, Vandi S, Montagna P, et al. A quantitative statistical analysis of the submental muscle EMG amplitude during sleep in normal controls and patients with REM sleep behavior disorder. *J Sleep Res* 2008;17:89–100.
- [28] Burns JW, Consens FB, Little RJ, Angell KJ, Gilman S, Chervin RD. EMG variance during polysomnography as an assessment for REM sleep behavior disorder. *Sleep* 2007;30:1771–8.
- [29] Frandsen R, Nikolic M, Zoetmulder M, Kempfner L, Jennum P. Analysis of automated quantification of motor activity in REM sleep behaviour disorder. *J Sleep Res* 2015;24:583–90.
- [30] Berry RB, Brooks R, Gamaldo CE, Harding SM, Lloyd RM, Marcus CL, et al. *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications, Version 2.3*. Darien, Illinois: The American Academy of Sleep Medicine; 2016.
- [31] Knudsen S, Jennum PJ, Alving J, Sheikh SP, Gammeltoft S. Validation of the ICSD-2 criteria for CSF hypocretin-1 measurements in the diagnosis of narcolepsy in the Danish population. *Sleep* 2010;33:169–76.
- [32] Fairley JA, Georgoulas G, Mehta NA, Gray AG, Bliwise DL. Computer detection approaches for identification of phasic electromyographic (EMG) activity during human sleep. *Biomed Signal Process Control* 2012;7:606–15.
- [33] Ferri R, Rundo F, Manconi M, Plazzi G, Bruni O, Oldani A, et al. Improved computation of the atonia index in normal controls and patients with REM sleep behavior disorder. *Sleep Med* 2010;11:947–9.
- [34] Tomczak M, Tomczak E. The need to report effect size estimates revisited. An overview of some recommended measures of effect size 2014;1:19–25.
- [35] Ferri R, Bruni O, Fulda S, Zucconi M, Plazzi G. A quantitative analysis of the submental muscle electromyographic amplitude during rapid eye movement sleep across the lifespan. *J*

Sleep Res 2012;21:257–63.

- [36] Ferri R, Franceschini C, Zucconi M, Vandi S, Poli F, Bruni O, et al. Searching for a marker of REM sleep behavior disorder: submental muscle EMG amplitude analysis during sleep in patients with narcolepsy/cataplexy. *Sleep* 2008;31:1409–17.
- [37] Figorilli M, Ferri R, Zibetti M, Beudin P, Puligheddu M, Lopiano L, et al. Comparison between Automatic and Visual Scorings of REM Sleep without Atonia for the Diagnosis of REM Sleep Behavior Disorder in Parkinson Disease. *Sleep* 2016:13–6.
- [38] Frauscher B, Iranzo A, Gaig C, Gschliesser V, Guaita M, Raffelseder V, et al. Normative EMG Values during REM Sleep for the Diagnosis of REM Sleep Behavior Disorder. *Sleep* 2012;35:835–47.
- [39] Frauscher B, Gabelia D, Biermayr M, Stefani A, Hackner H, Mitterling T, et al. Validation of an Integrated Software for the Detection of Rapid Eye Movement Sleep Behavior Disorder. *Sleep* 2014;37:1663–71.
- [40] McCarter SJ, St. Louis EK, Duwell EJ, Timm PC, Sandness DJ, Boeve BF, et al. Diagnostic Thresholds for Quantitative REM Sleep Phasic Burst Duration, Phasic and Tonic Muscle Activity, and REM Atonia Index in REM Sleep Behavior Disorder with and without Comorbid Obstructive Sleep Apnea. *Sleep* 2014;37:1649–62.

APPENDIX A

Table A.1. Comparison of the diagnostic value and performance of the five different metrics, given as optimal values of sensitivity/specificity.

	sMAI	IMAI	mPEM	RAI	STREAM	FR
C vs NT1+NT2	0.868/0.722	0.868/0.444	0.763/0.778	0.444/0.895	0.842/0.556	0.921/0.556
SCC vs NT1+NT2	0.947/0.467	0.868/0.533	0.790/0.600	0.533/0.842	0.974/0.133	0.921/0.467
NT1 vs NT2	1.000/0.133	0.870/0.467	0.957/0.333	0.800/0.739	0.913/0.267	0.783/0.733
NT1-RBD vs NT1+RBD	0.643/0.778	0.786/0.667	0.857/0.667	1.000/0.111	0.889/0.571	0.643/0.556

C, healthy controls; SCC, sleep clinic controls; NT1RBD, type 1 narcolepsy with/without RBD; NT2, narcolepsy type 2; sMAI, short muscle activity index; IMAI, long muscle activity index; mPEM, modified phasic electromyographic metric; RAI, RAI atonia index; STREAM, supra-threshold REM EMG activity metric; FR, Frandsen method

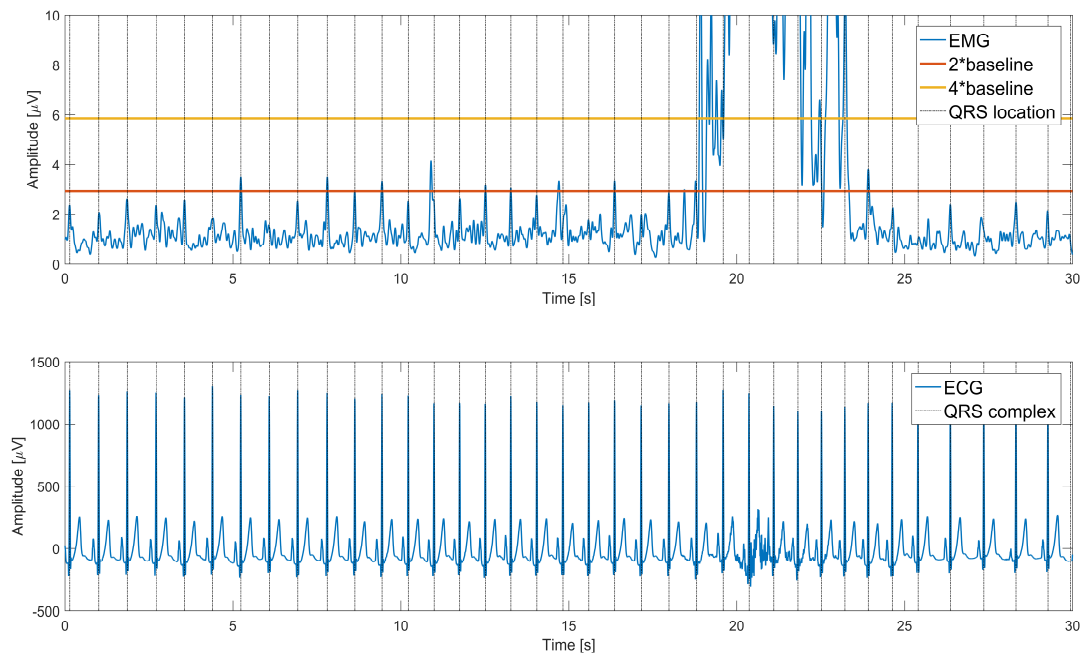


Fig. A. 1. Example of threshold chosen for identifying abnormal muscular activity with the muscle activity index method. It can be seen that if the threshold is set to twice the background level, the ECG artifacts in the EMG signal are wrongly considered as movements, while if the threshold is set to four times the baseline level, ECG artifacts are not seen as movements, while increased EMG activity is still accurately classified as movement. ECG, electrocardiogram; EMG, electromyogram

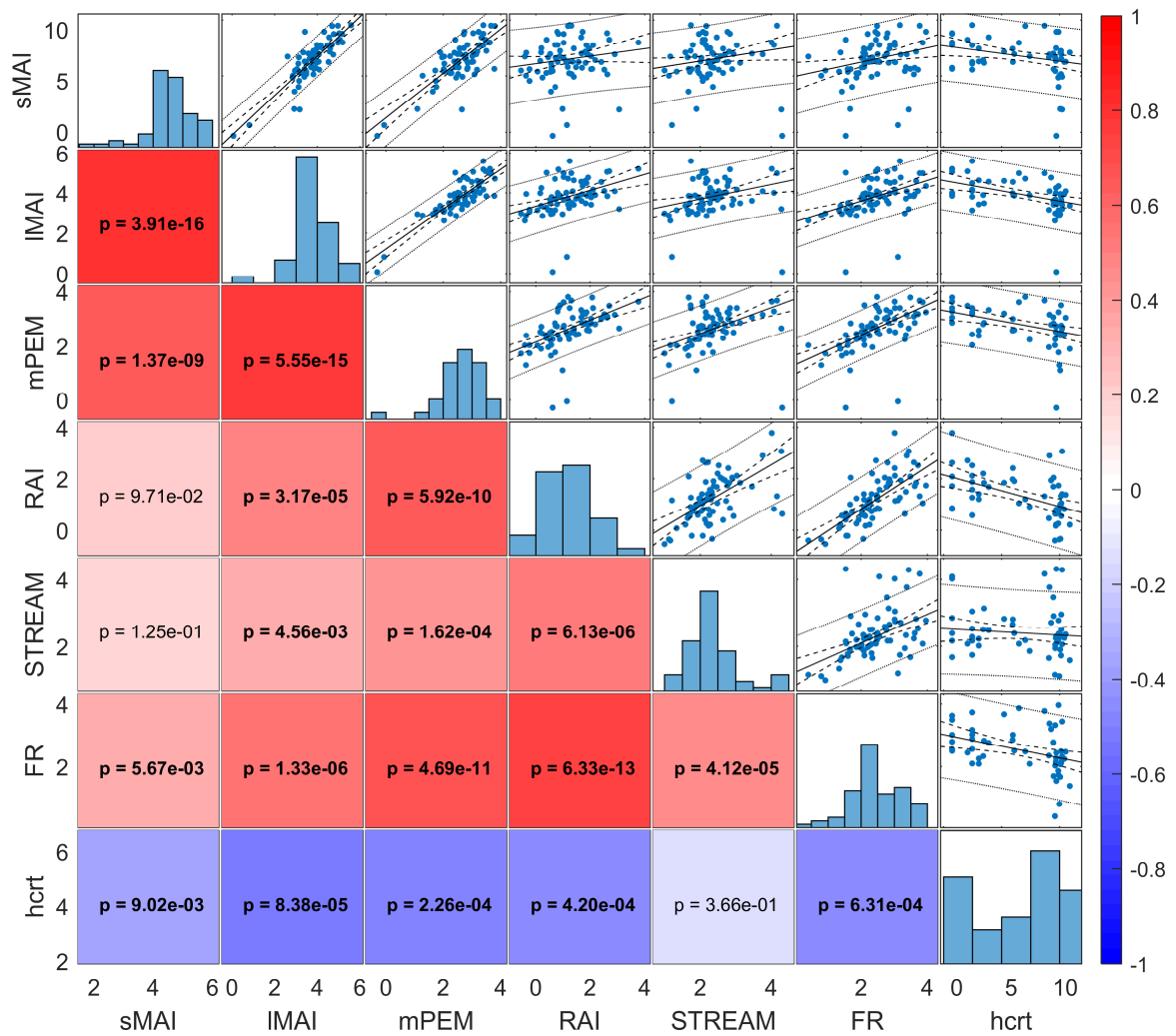


Fig. A. 2. Spearman correlation plots and analysis between metrics. All the metrics (RAI excluded) and CSF-hcrt-1 levels are log-transformed, and $\log(100-RAI)$ was used. Colors in the left boxes represent the correlation values, and significant p -values for the correlation are highlighted. In the diagonal, the distributions of the metrics are shown.

sMAI, short muscle activity index; IMAI, long muscle activity index; mPEM, modified phasic electromyographic metric; RAI, REM atonia index; STREAM, supra-threshold REM EMG activity metric; FR, Frandsen method

HIGHLIGHTS

- Comparison of five methods for automatic analysis of electromyogram (EMG) activity in narcolepsy.
- Narcolepsy type 1 is characterized by increased rapid eye movement (REM) EMG activity.
- Automatic EMG analysis in REM was a possible supportive tool in narcolepsy diagnosis.
- Lower CSF hypocretin 1 level correlated with increased REM EMG activity.