Original Article

Demographic and nap-related variance of the MSLT: results from 2,498 suspected hypersomnia patients
Clinical MSLT variance

Alyssa Cairns [a, *, 1], Lynn Marie Trotti [d, 1], Richard Bogan [a, b, c, 1]

[a] SleepMed, Inc. Columbia, SC, USA
[b] The University of South Carolina Medical School, Columbia, SC, USA
[c] Bogan Sleep Consultants, LLC, Columbia, SC, USA
[d] Department of Neurology, Emory University School of Medicine, Atlanta, GA, USA

ABSTRACT

Objective: The purpose of this study was to enhance our understanding of clinical trends in sleep and rapid eye movement (REM) propensity on the multiple sleep latency test (MSLT). Demographic variables of interest included early childhood/advanced age, gender, race, and REM-suppressant use.
Methods: Nocturnal sleep studies and 5-nap MSLTs were retrieved from a large repository of deidentified studies from various US sleep clinics between 2007 and 2015. Studies were signal processed, human-edited, and underwent rigorous quality assurance for inclusion.
Results: The final sample consisted of N = 2498 MSLTs (24.2% Black; 34.2% Men; Age 4–89). In adults (age ≥ 21), sleep propensity modestly decreased across nap (90% at nap 1 to 80% at nap 5; p < 0.001). Children ≤12 years were least likely to fall asleep on any nap (~55% at nap 5). REM propensity troughed at nap 4 (13%) and varied with age. Advanced age (≥60 years; OR: 0.28, p < 0.001), REM-suppressant use (OR:0.52, p < 0.001), and female sex (men OR: 1.48, p = 0.012) was associated with a decreased proportion of ≥2 REMs in adjusted logistic models. Children often demonstrated only 1 REM and generally had long sleep latencies, yielding a low proportion of MSLTs consistent with narcolepsy (11.0% vs. 19.2% and 16.8% in those between 13–20 and 21–59, respectively; p = 0.003).
Conclusions: MSLT outcomes vary greatly across age, gender, and use of psychotropic medication. Demographic variance should be considered when interpreting MSLT results. Robust age effect question the appropriateness of the MSLT as currently designed and implemented for children and older adults.

© 2018 Elsevier B.V. All rights reserved.

1. Introduction

Central nervous system hypersomnias are characterized by profound sleepiness in the absence of other explanatory factors [1]. Included in this general category are narcolepsy and idiopathic hypersomnia (IH). Narcolepsy is diagnostically partitioned into type 1 (N1) and 2 (N2), with N1 believed to be caused by auto-immune mediated destruction of neuropeptide hypocretin neurons [2]. Due to the salience and prevalence of the REM-intrusion phenomenon characteristic of N1, especially cataplexy, N1 often presents as a distinct clinical entity from N2, which often shares more symptom overlap with IH [3].

The multiple sleep latency test (MSLT) is the gold-standard for the assessment of hypersomnolence and is considered confirmatory for narcolepsy when the patient has a mean sleep latency (MSL) of ≤8 min and ≥2 REM periods (sleep onset REM periods, SOREMPs) between the Polysomnogram (PSG) and MSLT. A diagnosis of IH is assigned when the patient demonstrates high sleep pressure (MSL ≤8 min) but has <2 SOREMPs [1]. The MSLT has demonstrated good sensitivity, specificity, and reliability for N1 [4] (although between 14% and 29% have negative MSLTs) [5,6], however the test may be less robust for the differentiation of N2 versus IH [7–9].

Recent research has highlighted the poor reliability of the MSLT in N2 and IH, with either "diagnosis" being retained only 30%–47% upon serial MSLT testing. This is in contrast to N1, where repeat MSLT

confirms the diagnosis in 81% of cases [7–9]. One potential explanation for poor reliability of the MSLT is that the test is vulnerable to a host of environmental and behavioral/physiological influencing factors [10,11]. For example, it has long been surmised that patients are least likely to fall asleep and have REM on the last nap of an MSLT series because of anticipation to leave the laboratory after an often protracted stay. However, this hypothesis has never been systematically tested in a large patient data set.

Performance of the MSLT is also influenced by basic demographic factors, some better examined than others. For example, young healthy adolescents often exhibit high sleep pressure and SOREMPs on the MSLT due to developmental changes in circadian biology [12]. This contrasts with what occurs in the aged, where sleep latency increases and REM tendency is markedly reduced [13,14]. Other than studies on confirmed pediatric narcolepsy cases, very little is available on MSLT outcomes in young children.

A small amount of information is available on gender differences in MSLT outcomes. Population-based studies have found conflicting results, with a robust effect of male gender on REM tendency in the Wisconsin Sleep Cohort (WSC; men > women) [10,11] but no significant effect in the tricounty Detroit population-based study [15]. Although narcolepsy risk appears to vary across race/ethnicity [16,17], very little has been systematically studied with regard to racial/ethnic variance in MSLT outcomes. The purpose of this study is to expand our understanding of basic trends and influencing factors in clinical MSLT outcomes in a large, diverse clinical sample. Because the differentiation of N2 from IH relies solely on the MSLT, basic trends affecting MSLT results may have important diagnostic implications.

2. Methods

Data for this study were retrieved from SleepMed's repository of deidentified sleep studies conducted at various sleep disorders clinics in the United States between 2007 and 2015. All patients were sent for a PSG/MSLT because of physician-suspected hypersomnia. All patients consented to the use of their anonymized data for research and the study protocol was approved by Advarra IRB for the protection of human subjects.

2.1. Measures

This study utilized information acquired from the patient's scored and interpreted PSG/MSLT (eg REM and sleep onset latency, sleep indices, etc.) as well as data acquired from the patient's medical intake form and pre-sleep questionnaire. The intake form is a self-reported measure that inquiries about basic demographic information, previous diagnoses, sleep/wake habit and complaints, and symptoms of a variety sleep disorders. Sleepiness was assessed with the Epworth Sleepiness Scale (ESS) [18]. Although some physicians employ actigraphy and sleep diaries prior to PSG/MSLTs, actigraphy data were not available for this study. Habitual sleep duration was instead gleaned from the self-reported intake form (difference between self-reported "typical" bedtime and waketime on weekdays and weekends).

Medications were measured via a pre-sleep questionnaire administered the night of the sleep study. Two questions assessed medications (1) "*please list all medications you are currently taking or will take during the sleep study*" and (2) "*what other medications have you taken over the past month*"? Antidepressants (SSRIs/SNRIs, tricyclic antidepressants, monoamine oxidase inhibitors, and atypical agents) and antipsychotics were categorized as "REM suppressants" [19]. Each patient with a history of REM-suppressant use was categorized as either (1) active or (2) refrained based on the answers to the above questions. That is, if a patient noted use of sertraline as

"*other...taken over the past month*" but NOT as "*currently taking*", the patient was coded as 'refrained' from that compound. Likewise, if a subject noted use of paroxetine as "*currently taking*" but NOT "*other...taken over the past month*", they were coded as 'active' for that compound. Additionally, a patient was noted as 'active' if the technologist specifically noted the use of a REM suppressant at the time of the sleep study. To assess age-related changes in MSLT values, age was categorized as follows- ≤ 12 years = early childhood, 13–20 years = adolescence, 21–59 years = early/middle adulthood, and age≥ 60 years = older adulthood.

2.2. Data acquisition and scoring

Raw PSG and MSLT data were acquired using a variety of native sleep systems collected from various sleep disorder laboratories across the United States. Nocturnal PSG and MSLT data were acquired using traditional electrode placement, preparation, and study execution as per the AASM guidelines [20,21]. Per our standard protocol, raw PSGs and MSLTs underwent automated signal processing using Morpheus™, which decomposes EEG data into a 4-frequency state model using adaptive segmentation with fuzzy clustering and feature extraction [22]. The four frequency states include high frequency, low-frequency, mixed frequency with high energy, and mixed frequency with low energy.

The following rules are applied when EEG is processed with Morpheus: (1) wakefulness is scored when obvious movements are present and/or if EEG membership is predominant in the high frequency domain; (2) N1 is scored when EEG frequencies are predominant in the low energy mixed frequency domain in the presence of relatively high EMG; (3) REM is scored when EEG is similar to N1, rapid eye movements are present, and EMG tone is at the lowest point of the recording; (4) N2 is scored when membership domain is predominant in the high energy mixed frequency state along with the presence of K-complexes and spindles; and (5) N3 is scored when EEG frequencies are predominant in the low-frequency domain with high EEG peak-to-peak amplitude. The EEG sleep stage scoring algorithm has shown good agreement compared to manual scoring of sleep staging using AASM criteria (k = 0.61–0.67) and fair agreement for REM (intraclass correlation coefficient; ICC = 0.72–0.76) [22]. After automated signal processing, a registered sleep technologist edited all autoscored data on a 30-s epoch-by-epoch basis (eg sleep staging, respiratory data, limb movements, etc.). Technologists also validate each REM period, including the start and end times and any stage transitions during each period.

2.3. Data selection and MSLT quality assurance

Fig. 1 represents the data selection process for study inclusion. Initially, all baseline PSGs (no split nights or PAP titrations) with consecutive MSLTs (the following morning) were exported from the Morpheus database (n = 4836). Studies were excluded if the patient reported working swing or night shift at the time of the PSG/MSLT (n = 171). Each MSLT study was scrutinized to ensure adherence to AASM criteria. Studies were excluded if (1) the start of first nap occurred earlier than 1.5 h or later than 3 h from PSG lights on; (2) any of the naps were terminated prematurely (ie < 15 min from sleep onset; or <20 min after lights out if no sleep observed), or late (>15 min from sleep onset or > 20 min after lights out if no sleep observed), or (3) if any nap was spaced more or less than 2 h apart (+/− 5 min). A margin of error of ±30 s was allowed due to natural variance in where notations were tagged in the epoch. Studies that did not meet all the aforementioned criteria were noted as "inadequate quality" (n = 525). Next, studies were excluded if moderate to severe OSA was noted on the preceding PSG (RDI≥ 15, n = 421) or if
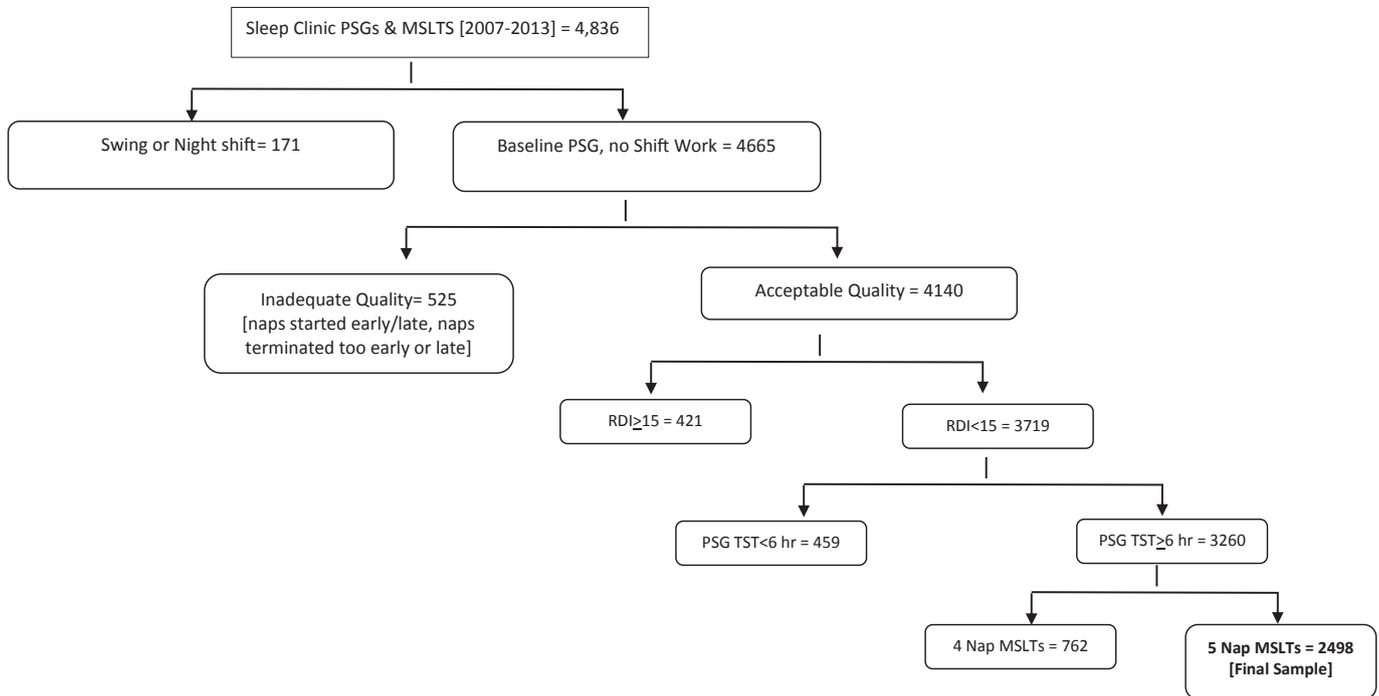
Fig. 1. Study selection and Quality Criteria for inclusion.

the subject had less than 6 h of TST on the PSG (n = 459). The RDI is defined as the average number of apneas, hypopneas (4% desaturation), and flow limited events per hour of sleep [periods of reduced flow or increased effort ($\geq$10 s) that terminate in either an EEG arousal or a 3% desaturation]. For analytic consistency, records with only four naps (which often occurs when the patient has either 0 or 2 REM periods) instead of five were excluded from final analyses (23% of MSLTs; n = 762). This yielded a N = 2498 MSLTs for analyses. Medication data were available on a subsample of 1033 subjects (41%). Please refer to table/figure footnotes for reduced sample size and degrees of freedom for analyses when medication data were analyzed, accordingly.

2.4. Statistical analyses

Analyses were performed using SPSS 23.0 (SPSS Inc., Chicago, IL). For continuous variables, descriptive analyses were completed to analyze the shape, central tendency, and dispersion to ensure parametric testing appropriateness. Group differences in demographics, PSG, and MSLT outcomes were assessed with either ANOVA with Bonferroni post-hoc tests or Chi Square analysis depending on the nature of data. Nap-related variability in sleep/REM propensity was assessed with a Chochran's Q test. Logistic multiple regression was employed to evaluate predictors of key MSLT outcomes (eg $\geq$2 REMs, MSL$\leq$8 min). All comparisons were two-tailed and significance was set at the 0.05 level.

3. Results

3.1. Demographics

The largest group of patients who underwent a five-nap MSLT were White women (50.3%); only 10.1% were Black men. Age ranged from 4 to 89 (Fig. 2) but varied across race/gender groups (Table 1). Despite the recommendation that patients should "ideally" stop REM suppressants for at least two weeks prior to testing [21], only 5.9% of patients taking $\geq$1 REM suppressant agent suggested that they

refrained from said compound(s) prior to the MSLT. Approximately 1/3rd of patients (34.9%) were noted by the technologist to have taken their REM-suppressant(s) the evening of the sleep study. Because of the small sample of individuals in the "refrained" group (and no statistical differences in sleep due to low power), we combined the "active" and "refrained" individuals for all analyses (hereafter referred to as "use of REM-suppressants"). Compared to White individuals, Black individuals had elevated WASO/reduced sleep efficiency and a higher proportion of a PSG SOREMP (especially Black men; Table 1). White women were older and reported the greatest
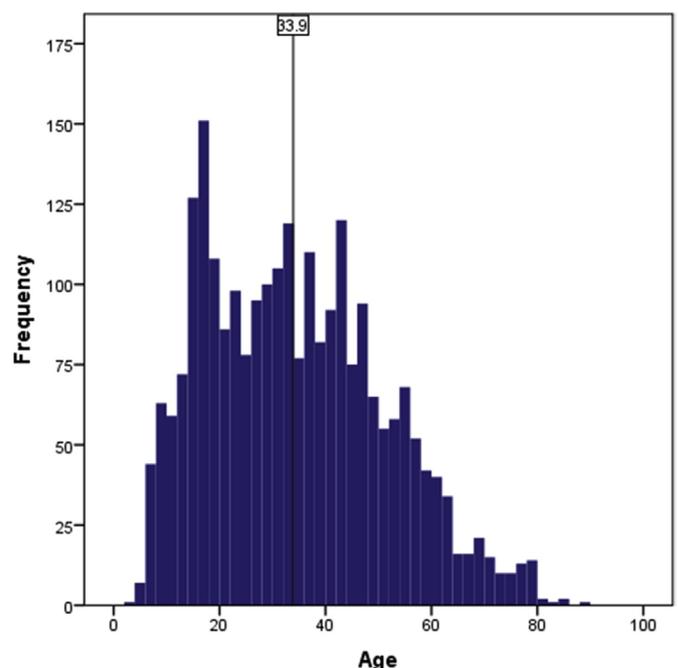

Fig. 2. Age Distribution (N = 2498); vertical line represents mean of the distribution.

Table 1
Demographics.

5 Nap MSLT	Overall	White Female	White Male	Black Female	Black Male	Analyses[1]
Sample Size	n = 2343~	n = 1179 (50.3%)	560 (23.9%)	367 (15.7%)	237 (10.1%)	
DEMOGRAPHICS						
Age (yrs.) [min−max]	33.9 [4−89]	36.9 [4−85]	32.5 [5−89]	32.9 [5−78]	25.3 [4−78]	$F_{(3,2343)} = 36.6$; $p < 0.001$ $\eta^2 = 0.05$; BM<all; WF>all
BMI (kg/m²)	27.6 ± 9.9	27.4 ± 7.2	26.1 ± 5.6	30.5 ± 8.7	28.1 ± 22.9	$F_{(3,2343)} = 14.7$; $p < 0.001$ $\eta^2 = 0.02$; BF > all
Obese (BMI≥30 kg/m²)	33%	33%	26%	50%	30%	$X^2_{(3,2343)} = 60.6$ $p < 0.001$
SLEEPINESS						
Epworth Sleepiness Scale (ESS)	12.0 ± 6.8	13.0 ± 5.9	11.5 ± 6.4	11.5 ± 7.9	9.1 ± 8.4	$F_{(3,2343)} = 25.0$; $p < 0.001$ $\eta^2 = 0.03$; WF > all; BM < all
%Sleepy (ESS≥10)	67%	74%	64%	63%	50%	$X^2_{(3,2343)} = 66.9$ $p < 0.001$
Sleepiness-accidents [y]	8%	7%	13%	9%	3%	$X^2_{(3,2343)} = 16.4$ $p = 0.001$
Problem tiredness [y]	83%	88%	82%	76%	59%	$X^2_{(3,2343)} = 61.5$ $p < 0.001$
COMORBIDITIES						
Depression	32%	43%	21%	29%	14%	$X^2_{(3,2343)} = 81.9$; $p < 0.001$
Hypertension	20%	17%	17%	38%	23%	$X^2_{(3,2343)} = 56.8$ $p < 0.001$
Chronic Pain	23%	28%	14%	22%	13%	$X^2_{(3,2343)} = 37.4$ $p < 0.001$
Diabetes	9%	9%	6%	16%	13%	$X^2_{(3,2343)} = 22.9$ $p < 0.001$
MEDICATION USE (n = 1033)[#]						
REM-Suppressant*	38%	53%	20%	33%	14%	$X^2_{(3,1033)} = 94.8$ $p < 0.001$
Antidepressant	36%	52%	30%	19%	13%	$X^2_{(3,1033)} = 96.9$ $p < 0.001$
Antipsychotic	5%	6%	7%	2%	5%	$p = 0.153$
Antiseizure	17%	25%	13%	15%	6%	$X^2_{(3,1033)} = 23.7$ $p < 0.001$
CNS Stimulant	16%	19%	18%	11%	12%	$p = 0.063$
Anxiolytic	13%	21%	10%	5%	1%	$X^2_{(3,1033)} = 48.2$; $p < 0.001$
Sedative Hypnotic	9%	12%	6%	8%	6%	$X^2_{(3,1033)} = 8.0$; $p = 0.046$
Opiates	9%	12%	8%	6%	6%	$X^2_{(3,1033)} = 8.2$; $p = 0.042$
NARCOLEPSY SYMPTOMS						
Drop attacks[2]	7%	8%	6%	9%	5%	$p = 0.400$
Sleep paralysis[3]	12%	11%	8%	19%	14%	$X^2_{(3,2343)} = 15.4$ $p = 0.001$
Hypnogogic hallucinations[4]	24%	24%	22%	30%	21%	$p = 0.112$
NOCTURNAL POLYSOMNOGRAM						
Apnea Hypopnea Index[5]	2.3 ± 2.5	2.1 ± 2.5	2.5 ± 2.5	2.2 ± 2.5	2.5 ± 2.7	$p = 0.053$
Respiratory Disturbance index[6]	4.1 ± 3.6	4.1 ± 3.6	4.6 ± 3.6	3.7 ± 3.5	3.9 ± 3.7	$F_{(3,2340)} = 4.9$; $p = 0.002$ $\eta^2 <0.01$; WM>all
Sleep Onset Latency (min)	27.6 ± 27.6	28.5 ± 25.5	26.0 ± 28.9	25.5 ± 27.1	27.6 ± 33.0	$p = 0.326$
Sleep Efficiency (%)	91.1 ± 6.6	91.5 ± 6.2	91.7 ± 6.3	90.3 ± 7.1	89.9 ± 7.9	$F_{(3,2340)} = 7.3$; $p < 0.001$ $\eta^2 < 0.01$; BM/BF < WF/WM
Arousal Index	8.2 ± 5.3	8.7 ± 5.6	8.2 ± 4.8	7.2 ± 4.5	7.0 ± 4.9	$F_{(3,2340)} = 12.1$; $p < 0.001$ $\eta^2 = 0.02$; WF/WM > BM/BF
Wake after Sleep Onset (min)	41.5 ± 31.9	38.7 ± 28.9	38.7 ± 30.3	47.6 ± 34.8	50.0 ± 40.0	$F_{(3,2340)} = 13.0$; $p < 0.001$ $\eta^2 = 0.02$; BM/BF > WF/WM
Nocturnal SOREMP[7]	1.1%	0.6%	0.7%	1.9%	3.0%	$X^2_{(3,2340)} = 13.2$ $p = 0.004$
SLEEP HABITS						
Weekday Sleep Duration[8]	8.2 ± 2.0	8.4 ± 2.0	8.0 ± 1.8	8.0 ± 1.7	8.1 ± 2.7	$F_{(1,2340)} = 6.2$; $p < 0.001$ $\eta^2 = 0.01$
Weekend Sleep Duration[8]	9.5 ± 2.5	9.7 ± 2.5	9.1 ± 2.3	9.6 ± 2.8	9.1 ± 3.0	$F_{(1,2340)} = 4.9$; $p = 0.010$ $\eta^2 = 0.01$
Weekday Naps[9]	45%	49%	37%	47%	40%	$X^2_{(3,2340)} = 16.5$ $p = 0.004$
Weekend Naps[9]	52%	59%	39%	55%	44%	$X^2_{(3,2340)} = 44.7$ $p < 0.001$

Total study N = 2498, Table n = 2343 because we removed "other races" from analyses (n = 62 Hispanic, n = 23 Asian, n = 14 American Indian, n = 6 Pacific Islander, 50 'other'); [1] Chi Square analysis for dichotomous variables and ANOVA for continuous variables, missing data were removed from analyses on a case-wise basis; [2] *do you experience drop or paralysis attacks?* [y/n]; [3] *"when falling asleep, how often do you feel unable to move or paralyzed?* [sometimes or more]"; [4] *"when falling asleep, how often do you experience vivid, dreamlike scenes or hallucinations even though you are awake?* [sometimes or more]"; [5] the average number of apneas and hypopneas [4% desaturation] per hour of sleep; [6] the average number of apneas, hypopneas, and flow limited events per/hr [events with reduced flow or increased effort that terminate in an EEG arousal or a 3% desaturation]; [7] SOREMP = short onset REM period (REM latency ≤ 15 min); [#] subsample of data (n = 1033; 41% of sample); * % yes taking either an antidepressant or antipsychotic [including those that 'refrained' for the study]; [8] duration from typical bedtime to wake time; [9] *"Do you take naps on the weekday/weekend?* [yes/no]".

severity of sleepiness compared to all other groups. White women also reported high rates of depression and chronic pain and endorsed the highest proportion of most psychotropic compounds, specifically antidepressants (52%), antiseizure compounds (25%), anxiolytics (21%), sedative hypnotics and opiates (both 12%). In contrast, Black men were younger than other groups (~12 years younger than White women), reported the least severe sleepiness, and were least likely to report depression and psychotropic medication use (antidepressants, antiseizure compounds, and anxiolytics). Black women had the highest proportion of obesity, hypertension, diabetes, and sleep paralysis. White males had a low proportion of obesity and comorbidities but were notably sleepy as per the high proportion of sleepiness-related accidents (13%). Small differences were found in sleep habits, but effect sizes were negligible.

3.2. Sleep propensity and latency

Sleep propensity on the MSLT was defined as the presence of any scored stage of sleep (at least 1 epoch; ≥ 15 s) for any given nap.

Data were analyzed on a per-nap basis to quantify nap-related variation. Sleep propensity decreased as the MSLT progressed Q = (4, N = 2497) = 603.8, $p < 0.001$, but the function was more pronounced for those under 21 years of age, most remarkably so for young children (Fig. 3A). In adults 21 years or older, sleep propensity was generally high (~90%) for the first four naps and reduced to 80% at the fifth nap Q = ((4, N = 1819) = 435.2, $p < 0.001$. A Chi square test of sleep propensity at each MSLT nap revealed that falling asleep was generally less-common for children and adolescents than those older (age≤20 vs. age≥21; all p < 0.001), with only 55% of young children falling asleep at the fifth nap.

Latency to sleep increased, albeit slightly, over the course of the MSLT (from 5.4 ± 5.3 min at nap 1 to 6.8 ± 5.0 min at nap 5; F (4, 9972) = 200.3; p < 0.001). Most of the variance in latency to sleep was accounted for by age. A u-shaped distribution between age and MSL was evident, with notably high MSLs in young children (Fig. 3B; F (3,2494) = 86.6, p < 0.001; all pairwise comparisons p < 0.01). Females were more likely than males to have MSLs<8 min

(A)

Nap & Age Variation in Sleep Propensity

(B)

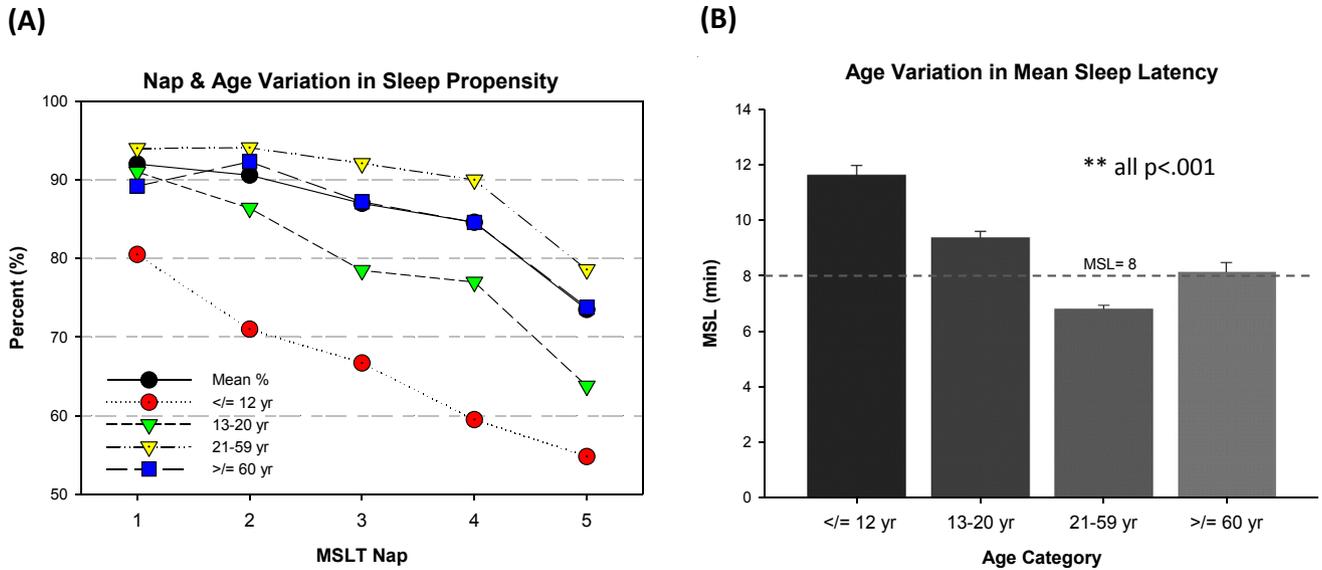### Age Variation in Mean Sleep Latency

Fig. 3. (A) Sleep propensity on successive MSLT Naps by age category; sleep propensity reduced across successive multiple sleep latency test (MSLT) naps (Cochran's Q = $p < 0.001$), sleep propensity was consistently reduced for children and adolescents (age<21) compared to adults (age>21 yr; all p < 0.001) (B) Age-related differences in MSLT mean sleep latency; error bars = standard error; U-shaped function between age and mean sleep latency **univariate ANOVA with Bonferroni correction, all pairwise comparisons p < 0.01; N = 2498 for both figures.

(male OR: 0.72), but this gender effect was no longer significant when controlling for age, race, and use of REM-suppressants (Table 2). Race accounted for a small amount of variance in MSL in the fully adjusted model, but effect size was small (Table 2).

3.3. REM tendency, latency, and MSLT outcomes

REM tendency was quantified as the presence of scored REM (at least 1 epoch; ≥ 15 s) on any given nap where sleep occurred. A Cochran's Q test, including only those that fell asleep at every nap, revealed statistical changes in REM propensity over the course of the MSLT (Q = (4, N = 67.1) = 67.1, $p < 0.001$). Aggregating across age, REM tendency peaked at 23% for nap 2 and troughed at 13% for nap 4, but as Fig. 4A illustrates, the rate of REM at any given nap

varied with age (all χ^2 p < 0.05). Latency to REM remained stable across the MSLT, at approximately 7 min from sleep onset.

A u-shaped function was observed between age and proportion of 2 or more REMs (Fig. 4B; $\chi^2 = (3, 2498) = 30.9$, $p < 0.001$), with advanced age (≥60 years) being associated with greatly reduced proportion of ≥2 REMs in adjusted logistic models (OR: 0.28 p < 0.001; Table 2). Thus, combining sleep latency and REM data, older adults (age≥60) had a relatively low proportion of MSLTs consistent with narcolepsy (10.8% vs. 16.8% in younger adults), and instead demonstrated the highest proportion of MSLTs consistent with IH due to 0 REMs (35.4%; Table 3). Young children (age≤12) also had a reduced proportion of ≥2 REMs, often only demonstrating 1 REM on the MSLT (28%; Fig. 4B), many times on the first nap (Fig. 4A). Thus, because young children demonstrated generally

Table 2
Logistic regression analysis predicting key multiple sleep latency test (MSLT) outcomes.

Predictor	Outcome 1: MSL≤8 Min on The MSLT[a]							
	Unadjusted				**Adjusted**[1]			
	B	OR	95% CI	p	*B*	OR	95% CI	p
Gender (Male)	−0.32	0.72	0.61−0.86	<0.001	—	—	—	0.953
Age	0.02	1.02	1.01−1.02	<0.001	0.02	1.02	1.01−1.03	<0.001
Race (Black)	—	—	—	0.30	−0.34	0.72	0.52−0.97	0.031
REM Suppressants	—	—	—	0.56	—	—	—	0.291
	Outcome 2: ≥2 REMs on The MSLT[b]							
	Unadjusted				**Adjusted**[2]			
	B	OR	95% CI	p	*B*	OR	95% CI	p
Gender (Male)	0.23	1.26	1.03−1.54	0.022	0.40	1.49	1.09−2.02	0.012
Age ≥ 60	−0.62	0.54	0.38−0.76	<0.001	−1.26	0.28	0.13−0.63	0.002
Race (Black)	0.29	1.34	1.08−1.67	0.009	—	—	—	0.282
REM Suppressants	−0.72	0.49	0.35−0.68	<0.001	−0.66	0.52	0.37−0.74	<0.001
	Outcome 3: MSLT Suggestive of Narcolepsy (≥ 2 REM & MSL ≤ 8 min)[c]							
	Unadjusted				**Adjusted**[3]			
	B	OR	95% CI	p	*B*	OR	95% CI	p
Gender (Male)	—	—	—	0.158	0.44	1.55	1.12−2.16	0.009
Age ≥ 60	−0.57	0.57	0.39−0.83	0.004	−0.98	0.38	0.17−0.84	0.016
Race (Black)	0.32	1.37	1.08−1.74	0.009	—	—	—	0.115
REM Suppressants	−0.68	0.51	0.36−0.72	<0.001	−0.51	0.60	0.41−0.88	0.008

[a] Versus MSL>8 min; Unadjusted models N = 2498 except REM-suppressant factor (N = 1033), adjusted models N = 1033; [1] Overall model $R^2 = 0.032$, χ^2 (4, 1033) = 23.1, p < 0.001; [b] Versus <2 REM onsets, [2] Overall model $R^2 = 0.068$, χ^2 (4, n = 1033) = 44.4, p < 0.001; [c] Versus all other outcomes; [3] Overall adjusted model $R^2 = 0.052$, χ2 (4, n = 1033) = 31.4, p < 0.001; [b,c] Number of SOREMPs between PSG and MSLT.
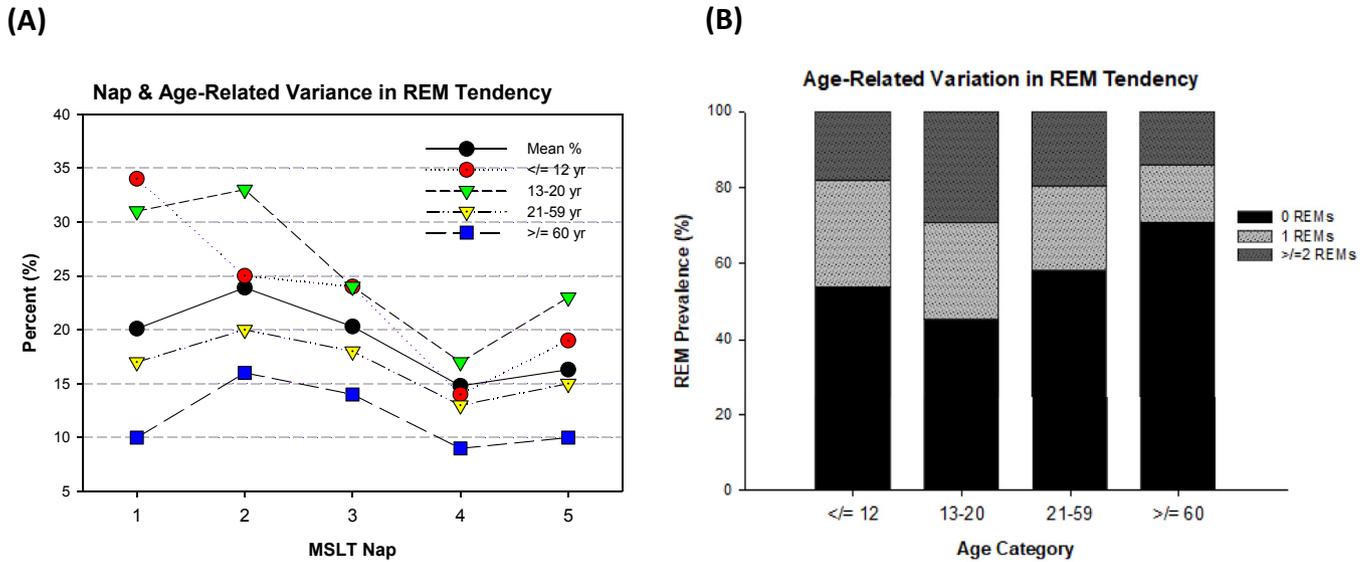
(A)

(B)

Fig. 4. (A) REM tendency on successive MSLT naps by age group; REM tendency decreased over the course of the MSLT (Nap 2 highest, Nap 4 lowest Cochran's Q = $p < 0.001$, but varied with age group ($\chi2 = p < 0.001$); (B) Number of REM onsets on the MSLT by age group; u-shaped function between age and rate of REM ($\chi2 = p < 0.001$); N = 2498 for both figures.

long sleep latencies and only 1 REM, the proportion of MSLTs consistent with narcolepsy was fairly low (11.0%), similar to the proportion found in our older adult group (Table 3). Moreover, 20% of young children had "ambiguous" MSLTs as per 1 REM episode and a MSL >8 min (Table 3). The highest proportion of ≥2 REMs (29%; Fig. 4B) and MSLTs consistent with narcolepsy was observed for adolescents (13−20 years; 19.2%; Table 3). However, adolescents also demonstrated the highest proportion of ambiguous MSLTs due to high REM pressure (≥2 REMs) without high sleep pressure (MSL > 8 min; 10.2%; Table 3).

Besides the robust effect of age on diurnal REM, use of REM-suppressants and gender also contributed unique variance in diurnal REM tendency (Table 2). Although Black race was associated with increased REM tendency in the unadjusted model, effects were no longer statistically significant in the adjusted model (Table 2). Controlling for age, gender, and race, REM-suppressant use was associated with reduced odds of ≥2 REMs (OR: 0.52, $p < 0.001$; Table 2) and MSLTs consistent with narcolepsy (OR: 0.60, $p = 0.008$; Table 2). Likewise, controlling for REM-suppressant use, race, and age, women were less likely to have ≥2 REMs (Men OR: 1.49, $p = 0.012$) or an MSLT consistent with narcolepsy (Men OR:

1.55, $p = 0.009$) than men (Table 2). Thus, because women displayed short MSLs but a low rate of ≥2 REMs, women were more likely than men to have MSLTs consistent with IH in adjusted logistic models (45% vs. 34%; OR: 0.58; $X^2_{(4,601)} = 52.3$, $p < 0.001$).

4. Discussion

The current study represents the largest database of clinical PSGs/MSLTs published to date, limited to patients being evaluated for central disorders of hypersomnolence. Prior to this, the largest published clinical sample of MSLTs was that of Chervin and Aldrich [23−25]. In our large patient sample, we have confirmed and extended several previously reported observations. We confirm that sleep indeed becomes less likely as the MSLT progresses, most dramatically at the last nap, and most prominently for young children. This differs somewhat from the pattern seen in the WSC, in which sleep was least likely on the first nap [26]. However, our finding supports clinical wisdom and early guidelines [27] regarding the "last-test" effect. Ultimately, these data highlight the importance of making the MSLT environment, especially for the

Table 3
Multiple sleep latency test (MSLT) outcomes by age group (n = 2498).

	Mean	Age Category				Analyses[1]
		≤12 yrs.	13−20 yrs.	21−59 yrs.	≥60 yrs.	
MSLT FINDINGS^						
Consistent with Narcolepsy						
≥2 SOREMPs & MSL≤8 min	16.3%	11.0%	19.2%	16.8%	10.8%	$X^2_{(3,2498)} = 14.3$ p = 0.003
Consistent with IH						
Sum	*41.0%*	*18.1%*	*24.1%*	*48.5%*	*44.6%*	$X^2_{(3,2498)} = 138.4$ p < 0.001
1 SOREMP & MSL≤8 min	12.8%	8.1%	8.7%	15.0%	9.2%	
0 SOREMP & MSL≤8 min	28.3%	10.0%	15.4%	33.5%	35.4%	
Normal Result						
0 SOREMP & MSL>8 min	28.0%	43.8%	29.9%	24.5%	35.4%	$X^2_{(3,2498)} = 41.3$ p < 0.001
Ambiguous						
Sum	*14.7%*	*27.1%*	*26.9%*	*10.2%*	*9.2%*	$X^2_{(3,2498)} = 112.4$ p < 0.001
≥2 SOREMPs & MSL>8 min	4.8%	7.1%	10.2%	3.0%	3.6%	
1 SOREMP & MSL>8 min	9.9%	20.0%	16.6%	7.2%	5.6%	

^ Number of SOREMPs between PSG and MSLT; [1] Chi Square analysis-age category vs. dummy coded MSLT outcome variable [% present vs. 'other'].

last nap, conducive for sleep with minimal environmental stimuli, as per the AASM Practice Parameters [21].

The likelihood of REM also decreased across the MSLT naps, most notably at the fourth nap. This temporal pattern differed somewhat from that seen in other population-based studies but is consistent with similar reports on smaller clinical samples [28]. Presumably, this trough in REM tendency reflects the expected circadian variation with a time course inverse to that of core body temperature (ie REM propensity low when core body temperature is relatively high) [29].

The effect of REM-suppressant medications on MSLT SOREMPs is not well-defined within clinical populations. While the WSC investigators initially reported an effect of antidepressants on MSLT SOREMPs [11], no association was found when the WSC sample size was increased [10]. Similarly, no association between nocturnal REM-suppressants and MSLT SOREMPs was found in the tricounty Detroit population-based study [15], suggesting that, at least at the population level where SOREMPs are a rare finding, classic 'REM-suppressant' medications do not have a meaningful effect on MSLT results. Yet, these medications are known to suppress REM sleep on nocturnal polysomnography, either by reducing REM sleep amount or by prolonging REM latency, and recommendations about cessation of REM-suppressants prior to MSLT testing are part of current clinical guidelines [21]. In keeping with this, we have now demonstrated a substantial association between REM suppressant use (specifically antidepressants and antipsychotics) and reduced MSLT SOREMPs/MSLT consistent with narcolepsy. This may represent the expected effect of these medications which was not seen at the population level due to the much lower proportion of SOREMPs in those without hypersomnia disorders (ie low statistical power). However, these data may also reflect a higher risk of depression, and thus use of antidepressants, among clinical participants with non-narcoleptic hypersomnia disorders [30].

We have also demonstrated robust effects of age on MSLT outcomes, with a U-shaped distribution of age for both MSL and number of SOREMPs. Population-based studies have excluded children [10,15] and have yielded conflicting results about the absence/presence of an association between age and SOREMPs during adulthood and a modest correlation between age and MSL (partial correlation r = 0.10, p < 0.001) [31]. Our finding that early childhood was associated with longer sleep onset latencies (often time not falling asleep at all), high REM pressure at the first nap compared to mature adults, and thus a high proportion of ambiguous MSLT results, questions the appropriateness of the MSLT as currently designed and implemented for young children.

In part, the short sleep duration obtained by children studied in a sleep clinic model (6.9 h vs. the National Sleep Foundation recommended >9 h [32]) reflects how the timing of the PSG is often not appropriated for young children's developmental needs. That is, in a typical sleep clinic setting, children are often studied on the same schedule as adults, with a lights out between 10–10:30 PM and a wake time around 6:00 AM. At a minimum, a wake time of 6:00 AM is premature for this stage of development, especially for pubertal children, where there is a natural desire/need to sleep later in the mornings due to changes in circadian biology [12]. Thus, the high proportion of SOREMPs on the first nap (and second nap in the 13–20 age group) may reflect circadian phase and/or REM rebound rather than an abnormality in REM, per se.

Like that found from Dauvilliers and colleagues [13], our study found that older adults were less likely to exhibit MSLTs consistent with narcolepsy compared to younger adults due to reduced proportion of REM and a longer latency to sleep compared to younger adults. Although this may represent age-related remission in narcolepsy symptoms [33], it also highlights the reduced sensitivity of the MSLT in detecting narcolepsy in older individuals. This conclusion is based on the growing literature substantiating age-related decline in nocturnal [34] and diurnal [13,28] REM amount.

Several other important, novel findings from this study require special attention. In our population, women were more likely to have MSLTs consistent with IH than men due to short sleep onset latencies without ≥2 SOREMPs. These findings persisted even after controlling for the fact that women were older and more likely to be taking psychotropic compound(s) at the time of MSLT evaluation. The finding of fewer SOREMPs among women seeking evaluation for hypersomnolence, compared to men seeking assessment for the same symptoms, could reflect a true biological difference in the prevalence of presenting sleepiness without SOREMPs, ie, idiopathic hypersomnia. Indeed, clinical case series suggest a female predominance to idiopathic hypersomnia. However, the other possible explanation for the finding of fewer SOREMPs in women is a biological difference in REM propensity that is independent of diagnosis. This latter hypothesis is more consistent with what is seen at the population-level in the WSC, ie, that women have fewer SOREMPs [11,15]. Because the diagnostic criteria for N2 and IH differ only in number of MSLT SOREMPs, an underlying gender difference in REM propensity would tend to systematically increase the percentage of sleepy women, relative to men, diagnosed with idiopathic hypersomnia.

Our finding that women (in unadjusted models) had a higher proportion of MSLs ≤8 min compared to men was somewhat surprising considering prior literature suggesting that men generally fall asleep quicker than women. Male tendency for short sleep latency has been observed in population-based studies [11,26,31] as well as some [35], but not all [36,37], clinical subgroups. This has been suggested to reflect the diagnostic delay for narcolepsy, which is significantly longer in women [37], such that women who are referred for testing have more disease progression prior to first test [36]. Whether such an explanation accounts for this finding in our study cannot be determined from our data set, but the marked difference between population-based gender effects and our clinical population gender effects requires further elucidation.

The potential for racial differences in MSLT findings has been largely unexplored. As of the 2017 census, 76.6% of the US population identified as White and 13.4% as black or African-American. In our sample, 24.2% were Black, suggesting that there was not a systematic bias against performing MSLT testing based on African-American race. The higher proportion of African-Americans in our MSLT sample compared to US census data may be partially reflective of the geographical distribution of sleep laboratories where data were acquired (densely-populated in the South/Southeastern region of the US). Although race initially appeared to correlate with diurnal SOREMPs in our data, this finding did not persist when analyses were adjusted for age, gender, and medications. Based on our sample, race does not appear to be a predictor of either mean sleep latency or number of SOREMPs on the MSLT.

Our study has several limitations. Primarily, because actigraphy and sleep diary data were not available, sleep restriction and circadian misalignment could not be thoroughly examined or controlled for. Additionally, data from this study represent that from treatment-seeking sleep clinic patients, with a variety of comorbidities and many that use various medications to treat such comorbidities. Although these data have high external validity and reflect the richness of 'real-world' sleep clinic data, one must keep the nature of the sample in mind when drawing conclusions. Furthermore, a large portion of data were acquired from the patient's intake questionnaire and was self-report in nature. Thus, these data are inherently subject to reporting bias, omissions, and errors like any self-report instrument. This is especially the case for medication data and the accuracy of our derived measure of "active vs. refrained" (from REM-suppressants).

Despite these limitations, our very large, diverse clinical MSLT sample provides clear insights into some of the key factors that may influence MSLT outcomes, including those test-related and those specific/unique to certain demographic groups. Our data reinforce good wisdom and practice guidelines to make conducive the sleep environment for an accurate MSLT, especially for the last nap period, where sleep is often more elusive. These data also highlight the need for research on the sensitivity and specificity of the MSLT across the lifespan and age-appropriated (and perhaps other demographic factors, like gender) MSLT guidelines (eg > 20 min nap period for young children, etc.). Until these data are available, it appears critical to allow developmentally-appropriate sleep duration and timing the night prior to/morning of the MSLT for children and adolescents. Lastly, this study reinforces the essential role of the skilled sleep medicine practitioner and a thorough clinical evaluation in the diagnosis and management of hypersomnia disorders. This is especially the case for older individuals, where REM may simply be unlikely and/or more difficult to observe/score.

5. Conclusions

Findings from this study suggest that outcomes from the MSLT, the current gold standard for the assessment of narcolepsy and idiopathic hypersomnia, vary greatly across age and gender groups as well as use of REM-suppressing psychotropic compounds. That is, females were uniquely at-risk for MSLTs that did not support narcolepsy, suggesting a potential sex bias and/or female-risk for idiopathic hypersomnia. These data also highlight the need for age-appropriated MSLT guidelines because diurnal latency to sleep and REM tendency appear to vary across the lifespan. The finding that REM-suppressant psychotropic compound use was associated with reduced REM reinforces AASM guidelines (although it appears infrequently done in routine sleep clinic samples), however more data is needed to understand the magnitude of effect and ideal time course for elimination prior to the MSLT.

Author contributions

AC — (1) conception and design of the study, (2) data analysis, and (3) data interpretation, (4) full manuscript drafting.

LMT — (1) conception and design of the study, (2) data analysis, and (3) data interpretation and manuscript drafting.

RB — (1) conception and design of the study and (2) data interpretation.

Conflict of interest

Drs. Bogan and Cairns have received research support from Jazz Pharmaceuticals, a company that owns therapeutic compounds for Narcolepsy. Dr. Bogan also serves on the speakers bureau for and is a consultant to Jazz Pharmaceuticals. Drs. Bogan and Cairns and are employed by SleepMed, Inc., a commercial diagnostic, research, and therapy company.

The ICMJE Uniform Disclosure Form for Potential Conflicts of Interest associated with this article can be viewed by clicking on the following link: https://doi.org/10.1016/j.sleep.2018.12.013.

References

[1] American Academy of Sleep Medicine. International classification of sleep disorders diagnostic and coding manual. 3rd ed. Westchester: American Academy of Sleep Medicine; 2014.

[2] Nishino S, Ripley B, Overeem S, et al. Hypocretin (orexin) deficiency in human narcolepsy. Lancet 2000;355(9197):39—40.

[3] Khan Z, Trotti LM. Central disorders of hypersomnolence: focus on the narcolepsies and idiopathic hypersomnia. Chest 2015;148:262—73.

[4] Arand D, Bonnet M, Hurwitz T, et al. Review by the MSLT and MWT task force of the standards of practice committee of the american academy of sleep medicine. Sleep 2005;28(1):123—44.

[5] Mignot E, Lammers GJ, Ripley B, et al. The role of cerebrospinal fluid hypocretin measurement in the diagnosis of narcolepsy and other hypersomnias. Arch Neurol 2002;59(10):1553—62.

[6] Aldrich MS, Chervin RD, Malow BA. Value of the multiple sleep latency test (MSLT) for the diagnosis of narcolepsy. Sleep 1997;20(8):620—9.

[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] Ruoff C, Pizza F, Trotti LM, et al. The MSLT is repeatable in narcolepsy type 1 but not narcolepsy type 2: a retrospective patient study. J Clin Sleep Med 2018;14(1):65—74.

[9] Lopez R, Doukkali A, Barateau L, et al. Test-retest reliability of the multiple sleep latency test in central disorders of hypersomnolence. Sleep 2017;40(12).

[10] Goldbart A, Peppard P, Finn L, et al. Narcolepsy and predictors of positive MSLTs in the Wisconsin sleep cohort. Sleep 2014;37(6):1043—51.

[11] Mignot E, Lin L, Finn L, et al. Correlates of sleep-onset REM periods during the multiple sleep latency test in community adults. Brain 2006;129(Pt 6):1609—23.

[12] Carskadon MA, Wolfson AR, Acebo C, et al. Adolescent sleep patterns, circadian timing, and sleepiness at a transition to early school days. Sleep 1998;21(8):871—81.

[13] Dauvilliers Y, Gosselin A, Paquet J, et al. Effect of age on MSLT results in patients with narcolepsy-cataplexy. Neurology 2004;62(1):46—50.

[14] Chakravorty SS, Rye DB. Narcolepsy in the older adult: epidemiology, diagnosis and management. Drugs Aging 2003;20:361—76.

[15] Singh M, Drake CL, Roth T. The prevalence of multiple sleep-onset REM periods in a population-based sample. Sleep 2006;29(7):890—5.

[16] Andlauer O, Moore H, Hong SC, et al. Predictors of hypocretin (orexin) deficiency in narcolepsy without cataplexy. Sleep 2012;35(9):1247—55.

[17] Kawai M, O'Hara R, Einen M, et al. Narcolepsy in african americans. Sleep 2015;38(11):1673—81.

[18] Johns M. A new method for measuring sleepiness: the epworth sleepiness scale. Sleep 1991;14:540—5.

[19] Pagel J, Parnes B. Medications for the treatment of sleep disorders: an overview. J Clin Psychol 2001;3(3):118—25.

[20] Iber C, Ancoli-Israel S, Chesson A, et al. The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications. 1st ed. IL: American Academy of Sleep Medicine; 2007.

[21] Littner MR, Kushida C, Wise M, et al. Practice parameters for clinical use of the multiple sleep latency test and the maintenance of wakefulness test. Sleep 2005;28(1):113—21.

[22] Pittman SD, MacDonald MM, Fogel RB, et al. Assessment of automated scoring of polysomnographic recordings in a population with suspected sleep-disordered breathing. Sleep 2004;27(7):1394—403.

[23] Chervin RD, Aldrich MS. Sleep onset REM periods during multiple sleep latency tests in patients evaluated for sleep apnea. Am J Respir Crit Care Med 2000;161(2 Pt 1):426—31.

[24] Chervin RD, Aldrich MS. Characteristics of apneas and hypopneas during sleep and relation to excessive daytime sleepiness. Sleep 1998;21(8):799—806.

[25] Chervin RD, Aldrich MS. The relation between multiple sleep latency test findings and the frequency of apneic events in REM and non-REM sleep. Chest 1998;113(4):980—4.

[26] Punjabi NM, Bandeen-Roche K, Young T. Predictors of objective sleep tendency in the general population. Sleep 2003;26(6):678—83.

[27] Roehrs T, Roth T. Multiple Sleep Latency Test: technical aspects and normal values. J Clin Neurophysiol 1992;9(1):63—7.

[28] Sansa G, Falup-Pecurariu C, Salamero M, et al. Non-random temporal distribution of sleep onset REM periods in the MSLT in narcolepsy. J Neurol Sci 2014;341(1—2):136—8.

[29] Czeisler CA, Zimmerman JC, Ronda JM, et al. Timing of REM sleep is coupled to the circadian rhythm of body temperature in man. Sleep 1980;2(3):329—46.

[30] Ohayon MM, Carskadon MA, Guilleminault C, et al. Meta-analysis of quantitative sleep parameters from childhood to old age in healthy individuals: developing normative sleep values across the human lifespan. Sleep 2004;27(7):1255—73.

[31] Morrell MJ, Finn L, McMillan A, et al. The impact of ageing and sex on the association between sleepiness and sleep disordered breathing. Eur Respir J 2012;40(2):386—93.

[32] Hirshkowitz M, Whiton K, Albert SM, et al. National sleep foundation's updated sleep duration recommendations: final report. Sleep Health 2015;4:233—43.

[33] Büchele F, Baumann CR, Poryazova R, et al. Remitting narcolepsy? Longitudinal observations in a hypocretin-deficient cohort. Sleep 2018;41(9).

[34] Ohayon MM. Narcolepsy is complicated by high medical and psychiatric comorbidities: a comparison with the general population. Sleep Med 2013;14(6):488—92.

[35] Ali M, Auger RR, Slocumb NL, et al. Idiopathic hypersomnia: clinical features and response to treatment. J Clin Sleep Med 2009;5(6):562—8.

[36] Won C, Mahmoudi M, Qin L, et al. The impact of gender on timeliness of narcolepsy diagnosis. J Clin Sleep Med 2014;10(1):89—95.

[37] Luca G, Haba-Rubio J, Dauvilliers Y, et al. Clinical, polysomnographic and genome-wide association analyses of narcolepsy with cataplexy: a european narcolepsy network study. J Sleep Res 2013;22(5):482—95.